

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Testes bayesianos em ensaios clínicos

Josimara Tatiane da Silva

Tese de Doutorado do Programa Interinstitucional de Pós-Graduação em Estatística (PIPGEs)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Josimara Tatiane da Silva

Testes bayesianos em ensaios clínicos

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Doutora em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística.
VERSÃO REVISADA

Área de Concentração: Estatística

Orientadora: Profa. Dra. Juliana Cobre

Coorientador: Prof. Dr. Mário de Castro Andrade Filho

USP – São Carlos
Março de 2022

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

SS586t Silva, Josimara Tatiane da
Testes bayesianos em ensaios clínicos / Josimara
Tatiane da Silva; orientador Juliana Cobre;
coorientador Mário de Castro Andrade Filho. -- São
Carlos, 2021.
94 p.

Tese (Doutorado - Programa Interinstitucional de
Pós-graduação em Estatística) -- Instituto de Ciências
Matemáticas e de Computação, Universidade de São
Paulo, 2021.

1. Fator de Bayes. 2. Robustez bayesiana. 3.
Sobredispersão. 4. Tamanho amostral. 5. Teste de
hipóteses. I. Cobre, Juliana, orient. II. de Castro
Andrade Filho, Mário, coorient. III. Título.

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Josimara Tatiane da Silva

Testes bayesianos em ensaios clínicos

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP e ao Departamento de Estatística – DEs-UFSCar, como parte dos requisitos para obtenção do título de Doutora em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística.
VERSÃO REVISADA

Área de Concentração: Estatística

Orientadora: Profa. Dra. Juliana Cobre

Coorientador: Prof. Dr. Mário de Castro Andrade Filho

USP – São Carlos
Março de 2022



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Defesa de Tese de Doutorado da candidata Josimara Tatiane da Silva, realizada em 22/02/2022.

Comissão Julgadora:

Prof. Dr. Marinho Gomes de Andrade Filho (USP)

Prof. Dr. Carlos Alberto de Bragança Pereira (IME-USP)

Prof. Dr. Alexandre Galvão Patriota (IME-USP)

Profa. Dra. Rosângela Helena Loschi (UFMG)

Prof. Dr. Edson Zangiacomi Martinez (FMRP-USP)

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

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa Interinstitucional de Pós-Graduação em Estatística.

Josimara Tatiane da Silva

Bayesian tests in clinical trials

Doctoral dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP and to the Department of Statistics – DEs-UFSCar, in partial fulfillment of the requirements for the degree of the Doctorate Interagency Program Graduate in Statistics. *FINAL VERSION*

Concentration Area: Statistics

Advisor: Profa. Dra. Juliana Cobre

Co-advisor: Prof. Dr. Mário de Castro Andrade Filho

USP – São Carlos

March 2022

*A Deus, para toda a sua Honra e a sua Glória.
Aos meus pais e a minha irmã.*

AGRADECIMENTOS

A Deus, pelo dom da vida, por todo o seu cuidado e amor incondicional comigo e com minha família, meus pais e minha irmã. A minha família maravilhosa, gratidão por tudo.

As minhas tias Elaine e Aparecida e aos pastores Cirilo, William e Sandro, por todas as conversas edificantes. As minhas avós, por todo o carinho e por toda sabedoria de vida compartilhada comigo. Aos meus avôs (*in memoriam*), grata pelos momentos vividos.

Aos meus amigos, por todos os momentos compartilhados durante o curso: Adalton, Alex M., Caio M., Claudia M., Dira, Gilson S., Glauciema, Guilherme S., Jardel, Kelly, Oilson, Marcílio, Marina G., Michele, Milena N., Milton M., Rafael P., Vitor A. e Yuri S. Além dos meus amigos, que mesmo distantes, sempre se fizeram presentes nesta jornada: Ana Paula X., Anderson A., Assunção, Bárbara, Edione, Elianne A., Elaine P., Elizama M., Geilson, Jovanira, Karmem, Luz, Marcelo, Paulo Sérgio, Priscila, Silvana, Thaís, Wênia e Zélia.

Aos professores Adriano P., Gustavo, Rafael Stern, Sandro, Milan, Vera T. e Vicente, pelos conhecimentos acadêmicos compartilhados. Agradeço à professora Juliana por ter aceitado a minha orientação e pela valiosa sugestão para trabalhar com inferência bayesiana. Agradeço ao meu coorientador Mário por toda paciência e dedicação ao trabalho e por todo o incentivo à pesquisa científica. Além de todos os conhecimentos humildemente compartilhados; um exemplo de excelente profissional como pesquisador.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradeço a essa instituição pelo apoio financeiro.

A todos, que direta ou indiretamente, participaram desta conquista.

*“Ora, a fé é a certeza daquilo que esperamos e a prova das coisas que não vemos.”
(BÍBLIA, Hebreus 11, 1)*

RESUMO

SILVA, J. T. **Testes bayesianos em ensaios clínicos**. 2022. 99 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

Nesta tese propomos duas novas abordagens bayesianas para testes de hipóteses de equivalência para proporções e provamos que esses testes de hipóteses bayesianos são equivalentes. Essas metodologias bayesianas aplicadas aos testes de equivalência combinam ferramentas frequentistas e bayesianas sob a ótica de teoria da decisão, nas quais regras de decisões ótimas são usadas para minimizar a combinação linear das probabilidades dos erros dos tipos I e II, além do uso de um nível de significância em função do tamanho amostral. Dessa forma, contornando algumas limitações da abordagem frequentista, predominantemente usada para testes de equivalência. Além disso, realizamos um estudo de robustez bayesiana para os testes de significância. Propomos um índice de sensibilidade por meio do uso da derivada dos funcionais estatísticos construídos. Os resultados da análise de dados reais mostram que esse índice de sensibilidade mede satisfatoriamente a sensibilidade local das quantidades *a priori* com respeito à escolha da distribuição *a priori*. Por fim, realizamos um estudo de simulação a fim de avaliar as taxas da probabilidade do erro do tipo I e do poder dos testes considerando diferentes cenários. Também avaliamos dados de contagem sobredispersos por meio de uma análise de má especificação. Apresentamos uma tabela para determinar, de forma prática, o tamanho amostral ótimo. As metodologias desenvolvidas são ilustradas para testes de bioequivalência no contexto de ensaios clínicos.

Palavras-chave: Fator de Bayes. Robustez bayesiana. Sobredispersão. Tamanho amostral. Teste de hipóteses.

ABSTRACT

SILVA, J. T. **Bayesian tests in clinical trials**. 2022. 99 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2022.

In this thesis, we propose two new Bayesian approaches for equivalence hypotheses testing for proportions and prove that these Bayesian hypotheses tests are equivalent. These Bayesian methodologies applied to equivalence tests combine frequentist and Bayesian tools from the perspective of decision theory, in which optimal decision rules are used to minimize the linear combination of the probabilities of type I and II errors, in addition to the use of a significance level as a function of sample size. Thus, these Bayesian methodologies overcome some limitations of the frequentist approach, predominantly used for equivalence testing. Moreover, we perform a Bayesian robustness study for the significance tests. We propose a sensitivity index by using the derivative of the constructed statistical functionals. The results of real data analysis show that this sensitivity index satisfactorily measures the local sensitivity of prior quantities with respect to the choice of the prior distribution. A simulation study was carried out to assess the rate of the type I error probability and the power of the tests under different scenarios. We also evaluate overdispersed count data by means of a misspecification analysis. We present a table to determine, in a practical way, the optimal sample size. The developed methodologies are illustrated for bioequivalence tests in the context of clinical trials.

Keywords: Bayes factor, Bayesian robust, Hypotheses testing, Overdispersion, Sample size.

LISTA DE ILUSTRAÇÕES

Figura 1 – Regiões associadas às hipóteses $H_0 : p - q \geq \delta$ e $H_1 : p - q < \delta$. Diagrama adaptado de Williamson (2007).	29
Figura 2 – Gráficos das funções densidade de probabilidade <i>a priori</i> correspondentes às distribuições Beta(1, 1), Beta(0,5;0,5), Beta(2,4) e Beta(2,2).	39
Figura 3 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,005$	46
Figura 4 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,01$	46
Figura 5 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,1$	47
Figura 6 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,3$	47
Figura 7 – Diagrama adaptado de Lima (2016).	65
Figura 8 – Distribuições dos valores P para o tamanho amostral $m = 50$	83
Figura 9 – Distribuições dos valores P para o tamanho amostral $m = 300$	83
Figura 10 – Regiões associadas às hipóteses $H_0 : p - q = \delta$ e $H_1 : p - q < \delta$	86
Figura 11 – Decomposição de Θ_1 nas regiões $\mathcal{R}_1, \mathcal{R}_2$ e \mathcal{R}_3	90

LISTA DE TABELAS

Tabela 1	– Descrição da função de perda l .	30
Tabela 2	– Proporção de vezes em que H_0 é rejeitada quando $ p - q = \delta$ (H_0 é verdadeira) para o nível de significância fixado $\alpha(\phi_f) = 0,05$ para o teste z e para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017, com a distribuição <i>a priori</i> de Bayes-Laplace para todos os testes.	40
Tabela 3	– Proporção de vezes em que H_0 é rejeitada quando $ p - q = \delta$ (H_0 é verdadeira) para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017 com $l_0 = l_1 = 1$.	41
Tabela 4	– Níveis de significância adaptativo ótimo ($\alpha(\phi_B^*)$) e probabilidades do erro do tipo II adaptativa ótima ($\beta(\phi_B^*)$), poderes ($1 - \beta(\phi_B^*)$) e riscos ($r(\phi_B^*)$) para o teste B2017, para tamanhos amostrais diferentes.	42
Tabela 5	– Proporção de vezes em que H_0 é rejeitada quando $ p - q < \delta$ (H_0 é falsa). Para o teste B2017, a distribuição <i>a priori</i> é a de Bayes-Laplace para todos os testes.	43
Tabela 6	– Proporção de vezes em que H_0 é rejeitada quando $ p - q < \delta$ (H_0 é falsa) para o teste B2017 com $l_0 = l_1 = 1$.	44
Tabela 7	– Valores verdadeiros dos parâmetros no estudo de simulação.	45
Tabela 8	– Proporção de vezes em que H_0 é rejeitada quando $ p - q = \delta$ (H_0 é verdadeira) para o nível de significância fixado $\alpha(\phi_f) = 0,05$ para o teste z e para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017, com a distribuição <i>a priori</i> de Bayes-Laplace para todos os testes.	48
Tabela 9	– Proporção de vezes em que H_0 é rejeitada quando $ p - q = \delta$ (H_0 é verdadeira) para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017 com $l_0 = l_1 = 1$.	49
Tabela 10	– Valores P , níveis de significância adaptativos ótimos ($\alpha(\phi_B^*)$) e probabilidades do erro do tipo II adaptativas ótimas ($\beta(\phi_B^*)$) para o teste B2017.	51
Tabela 11	– Valores das derivadas de Gâteaux dos funcionais \mathbb{T}_P , \mathbb{T}_α e \mathbb{T}_β em g na direção h .	67
Tabela 12	– Valores dos índices de sensibilidade com base na Tabela 11.	67
Tabela 13	– Valores P_ϵ , $\alpha_\epsilon(\phi_B)$ e $\beta_\epsilon(\phi_B)$ considerando a função densidade de probabilidade <i>a priori</i> g_ϵ .	68
Tabela 14	– Proporção de vezes em que H_0 é rejeitada (H_0 verdadeira).	82

SUMÁRIO

1	INTRODUÇÃO	23
2	NOVAS ABORDAGENS BAYESIANAS PARA TESTES DE EQUI- VALÊNCIA	27
2.1	Modelo frequentista	28
2.2	Modelo bayesiano	31
2.3	Comparação dos testes bayesianos	35
2.4	Estudo de simulação	38
2.4.1	<i>Taxa do erro do tipo I</i>	38
2.4.2	<i>Determinação do tamanho amostral</i>	39
2.4.3	<i>Poder</i>	42
2.4.4	<i>Análise de má especificação</i>	42
2.5	Análise de dados	49
2.6	Considerações finais	51
3	ÍNDICE DE SENSIBILIDADE PARA TESTES DE EQUIVALÊNCIA	53
3.1	Medida de robustez	54
3.1.1	<i>Propriedades e construção dos funcionais estatísticos</i>	61
3.2	Análise de dados	66
3.3	Considerações finais	68
	REFERÊNCIAS	71
APÊNDICE A	DIMENSÃO DE ESPAÇOS PARAMÉTRICOS E ELI- MINAÇÃO DE PARÂMETRO DE INCÔMODO	77
APÊNDICE B	AVALIAÇÕES PARA UMA PROPORÇÃO	81
APÊNDICE C	MEDIDAS DE EVIDÊNCIA f_{H_0} E f_{H_1}	85
C.1	f_{H_0}	85
C.2	f_{H_1}	90
APÊNDICE D	MEDIDAS DE EVIDÊNCIA $f_{H_0,\varepsilon}$ E $f_{H_1,\varepsilon}$	93
D.1	$f_{H_0,\varepsilon}$	93
D.2	$f_{H_1,\varepsilon}$	95

APÊNDICE E	OUTRA FORMA DE OBTENÇÃO DE $\mathbb{T}'_p(g)h$, REFERENTE AO TEOREMA 3	97
-------------------	--	-----------

INTRODUÇÃO

O problema de testar a equivalência de dois tratamentos surge em diversas áreas do conhecimento (ARBORETTI; PESARIN; SALMASO, 2020). No presente trabalho estudamos este teste de equivalência para variáveis binomiais, no qual o objetivo é testar a hipótese nula $H_0 : |p - q| \geq \delta$ versus a hipótese alternativa $H_1 : |p - q| < \delta$, em que p e q são as probabilidades de indivíduos curados após a utilização de uma nova droga e de uma droga padrão e $\delta > 0$ é um valor especificado. Assim, verificamos se existe evidência de que as duas drogas exibem pouca diferença na sua eficácia e sobre a escolha dos valores da margem, para uma breve revisão bibliográfica, *vide* Williamson (2007). Dessa forma, podemos concluir a equivalência dos dois tratamentos quando H_0 é rejeitada. Como aplicação da metodologia explorada no presente trabalho ao contexto de ensaios clínicos, consideramos o problema de teste de bioequivalência com dados reais.

Na tomada de decisões, existem dois tipos de erros. O erro do tipo I (II) é caracterizado pela decisão de rejeitar H_0 (não rejeitar H_0) quando, na verdade, H_0 é verdadeira (falsa). Quanto ao erro do tipo I, na prática, isso equivale a dizer que os tratamentos são equivalentes quando na verdade eles não são. Como uma consequência do erro do tipo I, podemos liberar medicamentos de qualidade inferior aos já disponíveis, sendo prejudiciais aos pacientes. O erro do tipo II raramente é considerado nas atuais revisões regulamentares (RUBERG *et al.*, 2019). Na abordagem frequentista, H_0 é configurada tal que o erro do tipo I é mais grave do que o erro do tipo II. Se for desejado que o erro do tipo II seja mais grave do que o erro do tipo I, as hipóteses devem ser alteradas, o que é o caso nos testes de equivalência (WILLIAMSON, 2007). Na abordagem bayesiana da perspectiva de teoria da decisão, se for desejado que o erro do tipo II seja mais grave do que o erro do tipo I, as hipóteses não precisam ser alteradas, porque esta gravidade pode ser facilmente alterada equilibrando as probabilidades dos erros dos tipos I e II.

Ruberg *et al.* (2019) discutiram inferência e tomada de decisão para o desenvolvimento e aprovação de drogas do século XXI. Apontaram que o equilíbrio destes erros é um desafio

difícil no processo de tomada de decisões regulatórias. Afirmaram também que as abordagens bayesianas oferecem a próxima geração de pensamento inferencial para a tomada de decisão.

De acordo com [Ruberg et al. \(2019\)](#), os custos e o tempo aumentaram no desenvolvimento de medicamentos farmacêuticos, com implicações relativamente aos tratamentos que os pacientes podem estar a receber. Assim, é interessante para o investigador e/ou para as agências de desenvolvimento trabalhar com um pequeno número de participantes. Contudo, podemos ter vários problemas com os testes estatísticos utilizados, em particular, com o erro do tipo I e o poder. Além disso, sob o panorama frequentista, os testes intersecção-união (IU) baseados em dois testes unilaterais (*two one-sided tests - TOSTs*) apresentam graves armadilhas dependendo do tamanho da amostra e/ou da margem δ escolhida ([ARBORETTI et al., 2018](#)).

Na abordagem frequentista, quando o tamanho da amostra aumenta é difícil usar um nível de significância fixado em 0,05, digamos, porque podemos rejeitar qualquer hipótese com mais frequência ([PERICCHI; PEREIRA, 2016](#); [PEREIRA et al., 2017](#); [GANNON; PEREIRA; POLPO, 2019](#)). Neste contexto, sob a abordagem bayesiana, existem propostas alternativas na literatura que envolvem a utilização de um nível de significância adaptativo (ou seja, um nível que muda com o tamanho amostral) em [Varuzza e Pereira \(2010\)](#), [Pérez e Pericchi \(2014\)](#), [Pericchi e Pereira \(2016\)](#), [Pereira et al. \(2017\)](#) e [Gannon, Pereira e Polpo \(2019\)](#). O teste bayesiano em [Pereira et al. \(2017\)](#) está em conformidade com o princípio da verossimilhança para qualquer espaço amostral discreto ([GANNON; PEREIRA; POLPO, 2019](#)). [Wasserstein, Schirm e Lazar \(2019\)](#), em uma edição especial do *The American Statistician* sobre o valor p , mencionaram diversos artigos inovadores e instigantes devido aos diversos problemas associados a esse valor, entre esses artigos, mencionaram a proposta sugerida por [Gannon, Pereira e Polpo \(2019\)](#) na seção sobre como complementar ou substituir o valor p , no qual se destacam outros pontos vantajosos como a compatibilidade com o princípio da verossimilhança. Além disso, para uma abordagem sobre índices de significância para tabelas de contingência sob tamanhos amostrais pequenos, *vide* [Oliveira et al. \(2018\)](#).

No contexto bayesiano, uma possibilidade de realizar um teste de hipóteses a um nível de significância especificado é rejeitar a hipótese nula se a probabilidade *a posteriori* da hipótese nula for menor do que o nível de significância especificado. A probabilidade exata do erro do tipo I nos desenhos bayesianos adaptativos pode ser extremamente difícil de calcular ([BERRY et al., 2010](#)). Além disso, estabelecer uma regra sobre a escolha de um nível de probabilidade *a posteriori* a ser alcançado e tendo em conta a correspondência do nível de significância fixo entre os contextos frequentista e bayesiano podem ser questões relevantes relativamente às dificuldades encontradas na utilização de métodos bayesianos pelas agências ([BERRY et al., 2010](#)). Adicionalmente, [Berry \(2006\)](#) apresentou uma discussão sobre o potencial dos ensaios clínicos bayesianos para melhorar a eficácia do desenvolvimento de drogas.

Devido às implicações de custos elevados e à possível necessidade de um maior número de pacientes envolvidos, além de outras dificuldades relatadas relativamente ao equilíbrio de

probabilidades dos erros dos tipos I e II e à utilização do nível de significância, motivamos a procurar por outras abordagens inferenciais a fim de aplicá-las aos testes de equivalência. Assim, as abordagens inferenciais consideradas baseiam-se em [Pericchi e Pereira \(2016\)](#) e [Pereira *et al.* \(2017\)](#), que combinam ferramentas frequentistas e bayesianas de um ponto de vista da teoria da decisão. A perspectiva em [Pericchi e Pereira \(2016\)](#) é modificada para acomodar tais configurações das hipóteses dos testes de equivalência. Em [Pericchi e Pereira \(2016\)](#) existem restrições tais como a hipótese nula é apenas estabelecida para hipótese simples e os espaços paramétricos são unidimensionais. Estas restrições no que diz respeito à formulação de hipóteses e à dimensão dos espaços de parâmetros foram contornadas com base em [Pereira *et al.* \(2017\)](#). Ressaltamos que [Pereira \(1985\)](#) também abordou a minimização da soma ponderada das probabilidades do erros do tipo I e do tipo II e o uso do nível de significância adaptativo.

O presente trabalho tem dois objetivos principais: (i) aplicar e comparar as perspectivas abordadas por [Pericchi e Pereira \(2016\)](#) e [Pereira *et al.* \(2017\)](#) aos testes de equivalência para proporções e (ii) analisar a robustez bayesiana dos testes de significância em relação à escolha da distribuição *a priori* e propor um índice de sensibilidade local usando a derivada dos funcionais estatísticos construídos.

Como contribuições, sugerimos duas abordagens bayesianas alternativas para testes de equivalência para duas proporções com base nas metodologias inferenciais apresentadas em [Pericchi e Pereira \(2016\)](#) (com modificações) e em [Pereira *et al.* \(2017\)](#). Também realizamos uma análise de robustez bayesiana com respeito à escolha da distribuição *a priori*, na qual construímos índices de sensibilidade para o valor P bayesiano, o nível de significância adaptativo e para o poder dos testes de significância bayesianos. Até o presente momento, não encontramos na literatura trabalhos tratando sobre robustez bayesiana para os testes de equivalência para proporções com a abordagem de [Pereira *et al.* \(2017\)](#).

A tese está organizada em dois capítulos, descritos a seguir. No Capítulo 2 apresentamos o estado da arte sobre testes de equivalência para duas proporções. Exploramos as metodologias frequentista e bayesiana sob a ótica de teoria da decisão aplicadas a testes de equivalência. Na sequência, comparamos os testes bayesianos. Também reportamos os resultados numéricos de um estudo de simulação realizado, incluindo uma análise de má especificação. Por fim, ilustramos a aplicação dessas metodologias inferenciais aos testes de bioequivalência com um conjunto de dados reais.

No Capítulo 3 expomos uma breve revisão de literatura sobre robustez bayesiana com ênfase em classes especiais de distribuições *a priori*. Realizamos um estudo sobre a sensibilidade local das quantidades de interesse imprescindíveis nos testes de significância bayesianos apresentados no Capítulo 2, utilizando derivada funcional. Por fim, apresentamos resultados numéricos a fim de ilustrar a metodologia proposta.

No final dos Capítulos 2 e 3 há suas respectivas considerações finais, nas quais expomos algumas possibilidades de pesquisas a partir do que conseguimos estudar até o presente momento

com o propósito de dar continuidade a este trabalho.

NOVAS ABORDAGENS BAYESIANAS PARA TESTES DE EQUIVALÊNCIA

Neste capítulo propomos novas abordagens bayesianas para testes de equivalência para duas proporções. Apresentamos as abordagens frequentista e bayesiana sob a ótica de teoria da decisão, também comparamos os testes bayesianos. Por fim, realizamos um estudo de simulação e uma análise com dados reais. Parte dos principais resultados deste capítulo foi publicada em [Silva, Cobre e de Castro \(2021\)](#).

Sobre o histórico do problema, [Schuirmann \(1987\)](#) propôs o uso de dois testes unilaterais - *two one-sided tests* (TOSTs), [Berger, Hsu et al. \(1996\)](#) abordaram os TOSTs como um caso particular de testes de intersecção-união - *intersection-union tests* (IU). [Barker et al. \(2001\)](#) trataram de testes de equivalência para variáveis aleatórias binomiais. [Williamson \(2007\)](#) propôs uma abordagem bayesiana para estes testes para variáveis aleatórias binomiais utilizando o fator de Bayes. [Kawasaki et al. \(2016\)](#) propuseram um novo índice simplificado que sugere a equivalência de proporções sob a abordagem bayesiana, mas sem utilizar o fator de Bayes. No presente trabalho propomos duas novas abordagens bayesianas aplicadas a estes testes de equivalência e provamos que elas são equivalentes.

Um dos objetivos principais do nosso estudo é aplicar as novas perspectivas introduzidas por [Pericchi e Pereira \(2016\)](#) com modificações e [Pereira et al. \(2017\)](#) aos testes de equivalência para variáveis binomiais, a fim de melhorar a segurança dos testes, para analisar a relação entre o tamanho da amostra e a variação dos valores da margem δ e os níveis de significância desejados. Além disso, o poder dos testes é também avaliado e o tamanho amostral requerido é calculado por meio de um estudo de simulação.

2.1 Modelo frequentista

Nesta seção descrevemos o problema e as estatísticas de teste sob a abordagem frequentista, sob a ótica de teoria da decisão, baseadas nos procedimentos TOSTs (SCHUIRMANN, 1987). O modelo estatístico é dado pelo espaço $(\mathcal{X}, \mathcal{A}, \mathcal{P})$, em que \mathcal{A} é a σ -álgebra de subconjuntos do espaço amostral \mathcal{X} e $\mathcal{P} = \{P_{\theta}; \theta \in \Theta\}$ é a família de distribuições de probabilidade sobre \mathcal{X} , indexada por $\theta \in \Theta = [0, 1] \times [0, 1]$. Consideramos os testes de equivalência para a diferença de proporções. Então, supomos que $X|p \sim \text{binomial}(m, p)$ e $Y|q \sim \text{binomial}(n, q)$ são variáveis aleatórias condicionalmente independentes representando as contagens de indivíduos curados referentes aos grupos tratamento e placebo, respectivamente, em que p e q são as probabilidades de que os indivíduos sejam curados após usar a nova droga e a droga padrão. Além disso, (x, y) denota um vetor de observações do espaço amostral \mathcal{X} e f denota a função massa de probabilidade conjunta, dada por

$$f(x, y | \theta) = \binom{m}{x} \binom{n}{y} p^x (1-p)^{m-x} q^y (1-q)^{n-y} I_{\{0,1,\dots,m\}}(x) I_{\{0,1,\dots,n\}}(y), \quad (2.1)$$

em que $\theta = (p, q) \in \Theta$. Dada uma amostra (x, y) , $L(\theta; x, y)$ denota a função de verossimilhança correspondente a (2.1).

Na estrutura de TOSTs, o objetivo é testar as hipóteses

$$H_{01} : p - q \geq \delta, \quad H_{11} : p - q < \delta \quad \text{e} \quad H_{02} : p - q \leq -\delta, \quad H_{12} : p - q > -\delta. \quad (2.2)$$

Neste trabalho, os limites de equivalência em (2.2) são simétricos em torno de 0, o que significa que utilizamos apenas um único valor $\delta > 0$. Os limites de equivalência podem ser escolhidos de forma simétrica (e.g., $\delta_i = -0,1$ e $\delta_s = 0,1$) ou antissimétrica (e.g., $\delta_i = -0,15$ e $\delta_s = 0,2$ e assim, $-0,15 < p - q < 0,2$), em que os índices i e s indicam os limites inferior e superior do intervalo (δ_i, δ_s) . Sabemos que para valores pequenos de δ , existe uma grande dificuldade em provar a equivalência quando ela existe, além da possível necessidade de tamanhos amostrais maiores, o que por sua vez causa custos mais elevados e pode implicar em problemas éticos (PINTO, 2010; ARBORETTI *et al.*, 2018). Os espaços associados às hipóteses H_{01} , H_{02} , H_{11} e H_{12} são dados, respectivamente, por $\Theta_{01} = \{\theta \in \Theta; p - q \geq \delta\}$, $\Theta_{02} = \{\theta \in \Theta; p - q \leq -\delta\}$, $\Theta_{11} = \{\theta \in \Theta; p - q < \delta\}$ e $\Theta_{12} = \{\theta \in \Theta; p - q > -\delta\}$. Os subespaços paramétricos Θ_{01} , Θ_{02} , Θ_{11} e Θ_{12} estão apresentados na Figura 1.

Consideramos o teste z usual para comparar proporções sob o desenho paralelo para duas amostras independentes. Dessa forma, os TOSTs são baseados nas duas estatísticas em Chow *et al.* (2017) dadas por

$$T_i(X, Y) = \frac{\hat{p} - \hat{q} - \delta}{SE} \quad \text{e} \quad T_s(X, Y) = \frac{\hat{p} - \hat{q} + \delta}{SE}, \quad (2.3)$$

em que $\hat{p} = \frac{X}{m}$ e $\hat{q} = \frac{Y}{n}$ representam as proporções amostrais e $SE = [\hat{p}(1-\hat{p})/m + \hat{q}(1-\hat{q})/n]^{1/2}$ é o erro padrão da diferença entre duas proporções. Como um teste de IU, a região rejeição de

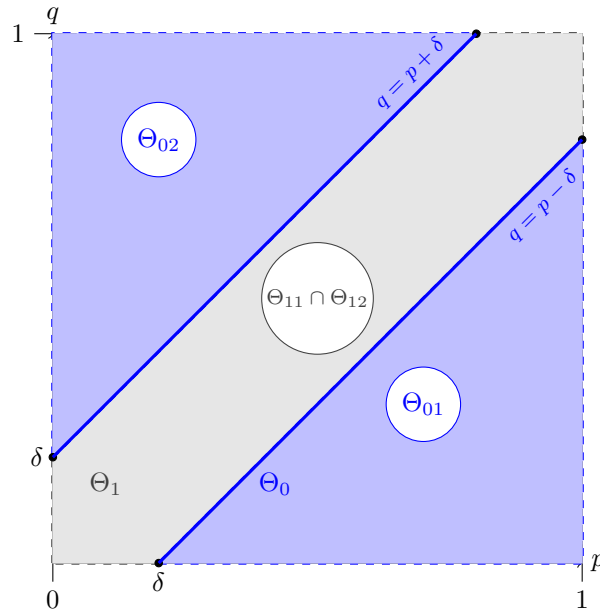


Figura 1 – Regiões associadas às hipóteses $H_0 : |p - q| \geq \delta$ e $H_1 : |p - q| < \delta$. Diagrama adaptado de Williamson (2007).

H_0 a um nível de significância α é dada por

$$C_\alpha = \{(x, y) \in \mathcal{X}; T_i(x, y) < -z_\alpha\} \cap \{(x, y) \in \mathcal{X}; T_s(x, y) > z_\alpha\}, \quad (2.4)$$

em que T_i e T_s denotam as estatísticas inferior e superior em (2.3), respectivamente, e z_α é o 100α -ésimo percentil da distribuição normal padrão. Sob a ótica de teoria da decisão, $\mathbb{A} = \{0, 1\}$ denota o conjunto de ações possíveis, em que 0 corresponde a ação de “não rejeitar H_0 ” e 1 a de “rejeitar H_0 ”. Sendo assim, a função de teste ou a regra de decisão $\phi_{fr} : \mathcal{X} \rightarrow \mathbb{A}$ é dada por $\phi_{fr}(x, y) = I_{C_\alpha}(x, y)$, em que o índice *fr* é usado para indicar a abordagem frequentista e $I_{C_\alpha}(\cdot)$ é a função indicadora sobre o conjunto C_α em (2.4). Portanto, fixado um valor para α , calculamos as estatísticas T_i e T_s e realizamos o teste ϕ_{fr} . Se $\phi_{fr}(x, y) = 1$, o teste rejeita H_0 , *i.e.*, indicando equivalência entre os tratamentos a um nível de significância nominal α . No presente trabalho, usamos o teste z apenas a um nível de significância nominal de 0,05, denotado por $\alpha(\phi_{fr})$.

Observação 1. Podemos verificar que a região de rejeição (2.4) com as estatísticas em (2.3) coincide com o intervalo assintótico simples (SAI) descrito por Barker *et al.* (2001). De fato, nos TOSTs, o teste rejeita H_0 se, e somente se, rejeita H_{01} e H_{02} a um nível de significância α . Sendo assim, dada a estatística T_i expressa em (2.3), rejeitamos H_{01} quando $T_i(x, y) < -z_\alpha$, ou seja, $(\hat{p} - \hat{q} - \delta)/SE < -z_\alpha$, o que implica que $\hat{p} - \hat{q} + z_\alpha SE < \delta$. De modo análogo, dada a estatística T_s em (2.3), rejeitamos H_{02} quando $T_s(x, y) > z_\alpha$, ou seja, $(\hat{p} - \hat{q} + \delta)/SE > z_\alpha$, o que implica que $\hat{p} - \hat{q} - z_\alpha SE > -\delta$. Então, $-\delta < \hat{p} - \hat{q} - z_\alpha SE < \hat{p} - \hat{q} + z_\alpha SE < \delta$. Portanto, $(\hat{p} - \hat{q} - z_\alpha SE, \hat{p} - \hat{q} + z_\alpha SE) \subseteq [-\delta, \delta]$ é um intervalo de $100(1 - 2\alpha)\%$ de confiança para $p - q$, ou seja, conclui-se equivalência ao intervalo de $100(1 - 2\alpha)\%$ confiança tal que esse intervalo esteja completamente contido em $[-\delta, \delta]$. Como $|p - q| \geq \delta$, no TOSTs, concluir equivalência

entre os tratamentos corresponde a usar um intervalo de $100(1 - 2\alpha)\%$ de confiança para $p - q$ tal que esse intervalo esteja inteiramente contido em $[-\delta, \delta]$ (SCHUIRMANN, 1987). No Apêndice A em Williamson (2007) há um explicação análoga ao que fizemos, mas de forma genérica, para o uso desse intervalo de $100(1 - 2\alpha)\%$ de confiança.

Além do mais, dada uma realização amostral, o verdadeiro valor do parâmetro e a regra de decisão ϕ_{fr} , podemos tomar uma decisão correta ou incorreta. Assim, quando a decisão for correta, atribuímos perda igual a 0. Em contrapartida, se for incorreta, atribuímos uma perda igual a 1. Posto isso, definamos a função de perda 0–1, $l : \mathbb{A} \times \Theta \rightarrow \{0, 1\}$, descrita na Tabela 1. Os subespaços paramétricos considerados na Tabela 1 são dados por $\Theta_0 = \{\theta \in \Theta; |p - q| \geq \delta\}$ e $\Theta_1 = \Theta \setminus \Theta_0$.

Tabela 1 – Descrição da função de perda l .

$l(\phi_{fr}(X, Y), \theta)$		Situações	
		$\theta \in \Theta_0$	$\theta \in \Theta_1$
Rejeitar H_0	$\phi_{fr}(X, Y) = 1$	1	0
Não rejeitar H_0	$\phi_{fr}(X, Y) = 0$	0	1

Na abordagem frequentista da perspectiva de teoria da decisão, na qual $l_1 > 0$ ($l_0 > 0$) denota a perda por rejeitar H_0 (não rejeitar H_0) quando H_0 é verdadeira (falsa), somente $l_0 = l_1 = 1$ ocorre considerando a função de perda 0–1. Ademais, as hipóteses são formuladas de forma que o erro do tipo I seja mais grave do que o erro do tipo II. Além disso, para o valor do parâmetro $\theta \in \Theta$, a perda esperada ou o risco associado a uma decisão ϕ_{fr} baseada nos dados (X, Y) , de acordo com a Tabela 1, é dado por

$$R_{\phi_{fr}}(\theta) = \mathbb{E}_{\theta}[l(\theta, \phi_{fr})] = \begin{cases} P_{\theta}(\phi_{fr}(X, Y) = 1) = \alpha_{\theta}(\phi_{fr}), & \text{se } \theta \in \Theta_0, \\ P_{\theta}(\phi_{fr}(X, Y) = 0) = \beta_{\theta}(\phi_{fr}), & \text{se } \theta \in \Theta_1, \end{cases}$$

em que $\alpha_{\theta}(\phi_{fr})$ e $\beta_{\theta}(\phi_{fr})$ são, respectivamente, as probabilidades dos erros do tipo I e do tipo II para o teste ϕ_{fr} para o valor do parâmetro θ , dadas por

$$\alpha_{\theta}(\phi_{fr}) = P_{\theta}(\text{Rejeitar } H_0) = P_{\theta}(\phi_{fr}(X, Y) = 1), \quad \text{se } \theta \in \Theta_0, \quad (2.5)$$

$$\text{e } \beta_{\theta}(\phi_{fr}) = P_{\theta}(\text{Não rejeitar } H_0) = P_{\theta}(\phi_{fr}(X, Y) = 0), \quad \text{se } \theta \in \Theta_1. \quad (2.6)$$

Por fim, considerando o modelo frequentista $(\mathcal{X}, \mathcal{A}, \mathcal{P})$, ressaltamos que a medida de probabilidade $P_{\theta} \in \mathcal{P}$ necessariamente precisa ser indexada em θ e que $\theta : \Theta \rightarrow \mathcal{P}$ é uma correspondência um-a-um, em que $\Theta = [0, 1] \times [0, 1]$, com a σ -álgebra de Borel $\sigma(\Theta)$. Para mais detalhes, vide Pereira (1993, Cap. 1). Também as probabilidades dos erros dos tipos I e II devem ser definidas conforme (2.5) e (2.6), sem o uso da medida de probabilidade condicionada no evento $\{\theta \in \Theta_i\}$, com $i = 0, 1$. Primeiro, observe que $C_{\alpha} = \{(x, y) \in \mathcal{X}; \phi_{fr}(x, y) = 1\} = \phi_{fr}^{-1}(\{1\}) \subset \mathcal{X}$, então

C_α é mensurável na σ -álgebra \mathcal{A} . Dessa forma, no contexto frequentista, por uma questão de mensurabilidade, não podemos definir $P_\theta \in \mathcal{P}$ como uma medida de probabilidade condicional. Em outras palavras, no contexto frequentista, dada $P_\theta \in \mathcal{P}$ uma medida de probabilidade, não definimos, por exemplo, a probabilidade do erro do tipo I como $\alpha(\phi_{fr}) = P_\theta(\text{Rejeitar } H_0 | \theta \in \Theta_0)$ porque se $P_\theta(\theta \in \Theta_0) > 0$, então $P_\theta(\cdot | \theta \in \Theta_0)$ é uma medida de probabilidade em \mathcal{A} , mas $\theta \in \Theta_0$.

2.2 Modelo bayesiano

Nesta seção tratamos de alguns conceitos imprescindíveis aos procedimentos dos testes bayesianos abordados em [Pericchi e Pereira \(2016\)](#) e em [Pereira et al. \(2017\)](#). Por fim, apresentamos os testes de hipóteses bayesianos. Consideramos o modelo estatístico como apresentado em [Pereira et al. \(2017\)](#). Sejam \mathcal{X} , \mathcal{A} e Θ os conjuntos apresentados na Seção 2.1. Na abordagem bayesiana, as hipóteses são

$$H_0 : |p - q| = \delta \quad \text{e} \quad H_1 : |p - q| < \delta. \quad (2.7)$$

Além do mais, Θ denota o espaço paramétrico e $\Theta_0 \subset \Theta$ o espaço paramétrico correspondente à hipótese nula. Com relação às hipóteses em (2.7), dizemos que H_0 é uma hipótese precisa quando a dimensão de Θ_0 é menor do que a dimensão de Θ . Usamos a formulação das hipóteses do teste em [Patel e Gupta \(1984\)](#), no qual H_0 é precisa.

Como os subespaços paramétricos Θ_0 e Θ_1 possuem dimensões diferentes, devemos considerar duas distribuições *a priori*, a saber, uma sobre as hipóteses e outra sobre todo o espaço paramétrico Θ . Uma vez que a medida de Lebesgue de Θ_0 é nula (H_0 é precisa), precisamos definir uma probabilidade positiva de H_0 ser verdadeira. Equivalentemente, suponhamos que o analista possa ou deseje expressar sua crença com relação às hipóteses ([IRONY; PEREIRA, 1995](#)). A seguir definimos a variável aleatória ξ por

$$\xi(\theta) = \begin{cases} 0, & \text{se } H_0 \text{ é verdadeira } (\theta \in \Theta_0), \\ 1, & \text{se } H_1 \text{ é verdadeira } (\theta \in \Theta_1), \end{cases} \quad (2.8)$$

em que $P(\xi = 0) = P(H_0) > 0$ é a probabilidade *a priori* de a hipótese nula H_0 ser verdadeira e $P(\xi = 1) = P(H_1) = 1 - P(H_0)$ é a probabilidade *a priori* de a hipótese alternativa H_1 ser verdadeira. A saber, ξ equivale à função λ em [Pereira et al. \(2017\)](#) e em [Gannon, Pereira e Polpo \(2019\)](#).

Agora ξ torna-se o parâmetro de interesse, representando a decisão, e o parâmetro original θ é tratado como um parâmetro de incômodo. Uma vez que o novo parâmetro ξ é binário e as hipóteses envolvendo-o são simples *versus* simples, podemos aplicar o Lema 1 ([PEREIRA et al., 2017](#)). No Lema 1, f_{H_i} é a medida de evidência ponderada para os dados (x, y) sob cada hipótese, para $i = 0, 1$, definida posteriormente em (2.13) e (2.14).

Lema 1 (Lema de Neyman-Pearson bayesiano). Sejam $\mathbf{z} = (x, y)$ e ϕ_B^* um teste tais que (i) rejeita H_0 em favor de H_1 se $\varpi_1 f_{H_0}(\mathbf{z}) < \varpi_2 f_{H_1}(\mathbf{z})$, (ii) não rejeita H_0 se $\varpi_1 f_{H_0}(\mathbf{z}) > \varpi_2 f_{H_1}(\mathbf{z})$ e (iii) é indiferente se $\varpi_1 f_{H_0}(\mathbf{z}) = \varpi_2 f_{H_1}(\mathbf{z})$. Então, para qualquer outro teste ϕ_B , temos que

$$\varpi_1 \alpha(\phi_B) + \varpi_2 \beta(\phi_B) \geq \varpi_1 \alpha(\phi_B^*) + \varpi_2 \beta(\phi_B^*), \quad (2.9)$$

para $\varpi_1 = l_1 P(H_0) > 0$ e $\varpi_2 = l_0 P(H_1) > 0$ especificados, em que $l_1 > 0$ é a perda por rejeitar H_0 quando H_0 é verdadeira e $l_0 > 0$ é a perda por não rejeitar H_0 quando H_0 é falsa.

Em (2.9), $\alpha(\cdot)$ e $\beta(\cdot)$ representam as probabilidades ponderadas dos erros dos tipos I e II, respectivamente, definidas mais adiante em (2.18) e (2.19). Note que no contexto do Lema 1, os testes para as hipóteses em (2.7) podem ser reformulados usando ideias de teoria da decisão como uma combinação de ferramentas frequentistas e bayesianas. Do Lema 1, obtemos o teste ótimo ϕ_B^* para testar as hipóteses em (2.7). Para mais detalhes sobre inferência de uma perspectiva de teoria da decisão, *vide* Young e Smith (2005).

Com respeito ao espaço Θ , consideramos p e q com distribuições *a priori* beta independentes tais que $p \sim \text{Beta}(a, b)$ e $q \sim \text{Beta}(c, d)$ com $a, b, c, d > 0$. Como p e q são independentes, a função densidade de probabilidade conjunta *a priori* de θ é dada por

$$g(\theta) = \frac{p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1}}{\text{Be}(a, b)\text{Be}(c, d)} I_{[0,1] \times [0,1]}(\theta), \quad (2.10)$$

em que $\text{Be}(\cdot, \cdot)$ é a função beta. Em particular, a distribuição *a priori* é a (i) de Jeffreys para $a = b = 0,5$ e coincide com a distribuição *a priori* de referência, (ii) de Bayes-Laplace para $a = b = 1$ e (iii) de Haldane para $a = b = 0$, que são distribuições *a priori* não informativas. Além do mais, a distribuição *a posteriori* de θ dado (x, y) é expressa como $\theta|(x, y) \sim \text{Beta}(A, B)\text{Beta}(C, D)$, em que $A = x + a$, $B = m - x + b$, $C = y + c$ e $D = n - y + d$.

Como $\Theta_0 \subset \Theta$ tem menor dimensão do que Θ , a função densidade *a priori* condicional de θ dada a hipótese H_0 (ou dado $\xi = 0$), seguindo a notação de Pereira *et al.* (2017), é calculada por

$$g(\theta|H_0) = \frac{g(\theta)}{\int_{\gamma} g(\theta) ds} I_{\Theta_0}(\theta), \quad (2.11)$$

em que o denominador é a integral de linha de g sobre a curva γ , para mais detalhes *vide* Apêndice C. Para a hipótese alternativa H_1 , a função densidade *a priori* condicional de θ sob H_1 é dada por

$$g(\theta|H_1) = \frac{g(\theta)}{\int_{\Theta_1} g(\theta) d\theta} I_{\Theta_1}(\theta). \quad (2.12)$$

Para mais detalhes sobre a técnica de condicionamento de função densidade *a priori*, *vide* Apêndice A.

Em Berger e Pericchi (2001), com adaptações para nosso trabalho, as medidas de evidência ponderadas para os dados (x, y) sob cada hipótese são

$$f_{H_0}(x, y) = f(x, y|H_0) = \int_{\gamma} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}|H_0) ds \quad (2.13)$$

$$\text{e } f_{H_1}(x, y) = f(x, y|H_1) = \int_{\Theta_1} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}|H_1) d\boldsymbol{\theta}, \quad (2.14)$$

em que $g(H_0|\boldsymbol{\theta})$ e $g(H_1|\boldsymbol{\theta})$ vêm de (2.11) e de (2.12), respectivamente. Então, substituindo (2.11) em (2.13), reescrevemos f_{H_0} como

$$f_{H_0}(x, y) = \int_{\gamma} L(\boldsymbol{\theta}; x, y) \frac{g(\boldsymbol{\theta})}{\int_{\gamma} g(\boldsymbol{\theta}) ds} ds = \frac{\int_{\gamma} \varphi(x, y, \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds}, \quad (2.15)$$

em que $\varphi(x, y, \boldsymbol{\theta}) = L(\boldsymbol{\theta}; x, y)g(\boldsymbol{\theta})$. De modo semelhante, substituindo (2.12) em (2.14), reescrevemos f_{H_1} como

$$f_{H_1}(x, y) = \frac{\int_{\Theta_1} \varphi(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta}}. \quad (2.16)$$

Com respeito às hipóteses em (2.7), as medidas de evidência ponderadas para os dados (x, y) sob cada hipótese são

$$f_{H_0}(x, y) = \binom{m}{x} \binom{n}{y} \frac{\eta_{H01} + \eta_{H02}}{\eta_{H03} + \eta_{H04}} \quad \text{e} \quad f_{H_1}(x, y) = \binom{m}{x} \binom{n}{y} \frac{\eta_{H11} + \eta_{H12} + \eta_{H13}}{\eta_{H14} + \eta_{H15} + \eta_{H16}}, \quad (2.17)$$

em que

$$\begin{aligned} \eta_{H01} &= \int_{\delta}^1 t^{A-1} (1-t)^{B-1} (t-\delta)^{C-1} [1-(t-\delta)]^{D-1} dt, \\ \eta_{H02} &= \int_0^{1-\delta} t^{A-1} (1-t)^{B-1} (t+\delta)^{C-1} [1-(t+\delta)]^{D-1} dt, \\ \eta_{H03} &= \int_{\delta}^1 t^{a-1} (1-t)^{b-1} (t-\delta)^{c-1} [1-(t-\delta)]^{d-1} dt, \\ \eta_{H04} &= \int_0^{1-\delta} t^{a-1} (1-t)^{b-1} (t+\delta)^{c-1} [1-(t+\delta)]^{d-1} dt, \\ \eta_{H11} &= \int_0^{\delta} \int_0^{q+\delta} p^{A-1} (1-p)^{B-1} q^{C-1} (1-q)^{D-1} dpdq, \\ \eta_{H12} &= \int_{\delta}^{1-\delta} \int_{q-\delta}^{q+\delta} p^{A-1} (1-p)^{B-1} q^{C-1} (1-q)^{D-1} dpdq, \\ \eta_{H13} &= \int_{1-\delta}^1 \int_{q-\delta}^1 p^{A-1} (1-p)^{B-1} q^{C-1} (1-q)^{D-1} dpdq, \\ \eta_{H14} &= \int_0^{\delta} \int_0^{q+\delta} p^{a-1} (1-p)^{b-1} q^{c-1} (1-q)^{d-1} dpdq, \\ \eta_{H15} &= \int_{\delta}^{1-\delta} \int_{q-\delta}^{q+\delta} p^{a-1} (1-p)^{b-1} q^{c-1} (1-q)^{d-1} dpdq, \\ \text{e } \eta_{H16} &= \int_{1-\delta}^1 \int_{q-\delta}^1 p^{a-1} (1-p)^{b-1} q^{c-1} (1-q)^{d-1} dpdq. \end{aligned}$$

No Apêndice C encontram-se mais detalhes sobre a obtenção das medidas de evidência em (2.17). Com as medidas de evidências em (2.15) e (2.16) e a notação em Gannon, Pereira e Polpo (2019), para dados (x, y) , o fator de Bayes sob H_0 é dado por $FB_{01}(x, y) = f_{H_0}(x, y)/f_{H_1}(x, y)$.

Com relação à terminologia, essas medidas de evidência também são chamadas de marginais de (X, Y) com respeito ao subespaço Θ_j para $j = 0, 1$ (BERGER; PERICCHI, 2001) ou ainda funções de verossimilhanças ponderadas com as distribuições *a priori* sendo as “funções de peso” ($g(\boldsymbol{\theta}|H_0)$ e $g(\boldsymbol{\theta}|H_1)$); de densidades preditivas *a priori* sob cada hipótese (PEREIRA *et al.*, 2017) ou ainda, de densidades preditivas *a priori* para uma variável aleatória observada $(X, Y) = (x, y)$ ponderada pela função densidade *a priori* g condicionada sob cada hipótese, H_0 e H_1 . Também observe que a parametrização de entrada do modelo estatístico passa a depender do par $(\xi, \boldsymbol{\theta})$, no entanto, agora nos interessa apenas ξ . Para detalhes sobre eliminação de parâmetros de incômodo, *vide* Fraser (1956), Berger, Liseo e Wolpert (1999). Quanto à terminologia, a função densidade *a priori* é condicionada a cada hipótese, sendo denotada por $g(\boldsymbol{\theta}|\xi = i)$, e simplificada para $g(\boldsymbol{\theta}|H_i)$, a função de probabilidade $f(x, y|\boldsymbol{\theta}) = f(x, y|\boldsymbol{\theta}, \xi)$, a função de verossimilhanças por $L(\boldsymbol{\theta}; x, y) = L(\boldsymbol{\theta}, \xi; x, y)$ para $i = 0, 1$, também valendo para as medidas de evidência.

Seguindo Gannon, Pereira e Polpo (2019), as probabilidades ponderadas dos erros do tipo I e do tipo II são definidas, respectivamente, como

$$\alpha(\phi_B) = \sum_{(x, y) \in B_k} f_{H_0}(x, y) \quad (2.18)$$

$$\text{e } \beta(\phi_B) = \sum_{(x, y) \notin B_k} f_{H_1}(x, y), \quad (2.19)$$

em que $B_k = \{(x, y) \in \mathcal{X}; FB_{01}(x, y) < k\}$ e $k = \omega_2/\omega_1$ é a razão ou taxa de perda. Alternativamente, as probabilidades ponderadas dos erros do tipo I e do tipo II são definidas em Pericchi e Pereira (2016) como

$$\alpha(\phi_B) = \int_{\gamma} P(\phi_B = 1 | H_0 \text{ é verdadeira}) g(\boldsymbol{\theta}|H_0) ds \quad (2.20)$$

$$\text{e } \beta(\phi_B) = \int_{\Theta_1} P(\phi_B = 0 | H_0 \text{ é falsa}) g(\boldsymbol{\theta}|H_1) d\boldsymbol{\theta}, \quad (2.21)$$

em que $P(\phi_B = \cdot | \cdot)$ é a probabilidade de erro para o teste ϕ_B para o parâmetro $\boldsymbol{\theta}$ na tradição de Neyman-Pearson. Em nosso trabalho, a definição (2.20) de Pericchi e Pereira (2016) é modificada usando uma integral de linha. Finalmente, como em Gannon, Pereira e Polpo (2019), o valor P é calculado por

$$\text{valor } P = \sum_{(x, y) \in B_{\text{obs}}} f_{H_0}(x, y), \quad (2.22)$$

em que $B_{\text{obs}} = \{(x, y) \in \mathcal{X}; FB_{01}(x, y) \leq FB_{01}(x, y)_{\text{obs}}\}$, em que $(x, y)_{\text{obs}} \in \mathcal{X}$ é um vetor observado. Nos procedimentos de teste usando o valor P (PEREIRA *et al.*, 2017; GANNON; PEREIRA; POLPO, 2019), o fator de Bayes é usado para induzir uma ordem no espaço amostral,

ou seja, considere $\{FB_{01}(x,y); (x,y) \in \mathcal{X}\} \subset \mathbb{R}$ para estabelecer a ordem para cada $(x,y) \in \mathcal{X}$. Nesse sentido, \mathcal{X}' denota o espaço amostral ordenado pelo fator de Bayes.

Exemplo 1. Considerando as hipóteses em (2.7), suponhamos que o espaço amostral seja $\mathcal{X} = \{(0,0), (0,1), (1,0), (1,1)\}$ e que os fatores de Bayes são $FB_{01}(0,0) = 2$, $FB_{01}(0,1) = 1,5$, $FB_{01}(1,0) = 0,6$ e $FB_{01}(1,1) = 0,01$. Observamos que esses valores dos fatores de Bayes apresentados são ilustrativos. Dessa forma, podemos ordenar o espaço amostral \mathcal{X} da seguinte forma: $\mathcal{X}' = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, com $\omega_1 = (1,1)$, $\omega_2 = (1,0)$, $\omega_3 = (0,1)$ e $\omega_4 = (0,0)$, em que o índice indica a ordem. Assim, passamos a usar \mathcal{X}' o espaço amostral com ordem induzida pelo fator de Bayes.

O teste ótimo proposto em [Pericchi e Pereira \(2016\)](#) é definido como $\phi_1(x,y) = I_{B_k}(x,y)$, em que B_k é a região de rejeição. No teste ϕ_1 , dada uma amostra observada $(x,y)_{\text{obs}} \in \mathcal{X}'$, decidimos por rejeitar H_0 se $\varpi_1 f_{H_0}(x,y) < \varpi_2 f_{H_1}(x,y)$, ou seja, $FB_{01}(x,y)_{\text{obs}} < k$, com $(x,y) \in \mathcal{X}'$. Note que minimizar o risco dado por $r(\phi_B) = \varpi_1 \alpha(\phi_B) + \varpi_2 \beta(\phi_B)$ equivale a minimizar a soma das probabilidades ponderadas dos erros ([PERICCHI; PEREIRA, 2016](#)). Além disso, sob as abordagens em [Pericchi e Pereira \(2016\)](#) com modificações, tais modificações consistiram em calcular o fator de Bayes como em [Pereira et al. \(2017\)](#), considerando a eliminação de um parâmetro de incômodo e a ordem induzida pelo fator de Bayes no espaço amostral para simplificar o cálculo das probabilidades dos erros.

O teste ótimo proposto em [Pereira et al. \(2017\)](#), denotado por ϕ_2 , é dado por $\phi_2(x,y) = I_{C_P}(x,y)$, em que $C_P = \{(x,y) \in \mathcal{X}'; \text{valor } P < \alpha(\phi_B^*)\}$ é a região de rejeição desse teste e $\alpha(\phi_B^*)$ é o nível de significância adaptativo ótimo. Dada uma amostra observada, $(x,y)_{\text{obs}} \in \mathcal{X}'$, calculamos o valor P e decidimos por rejeitar H_0 se valor $P < \alpha(\phi_B^*)$. Note que o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ depende do tamanho amostral, enquanto que o valor P depende da amostra observada $(x,y)_{\text{obs}}$, conforme (2.18) e (2.22).

2.3 Comparação dos testes bayesianos

Em [Pericchi e Pereira \(2016\)](#), [Pereira et al. \(2017\)](#), [Gannon, Pereira e Polpo \(2019\)](#), é esperado que o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ e a probabilidade do erro do tipo II adaptativa ótima tenderão a decrescer com o aumento do tamanho amostral. Na sequência, estabelecemos que as definições (2.18) e (2.19) coincidem com (2.20) e (2.21), respectivamente. Note que ambas abordagens bayesianas são formuladas sob as suposições de que a função densidade *a priori* g é própria e a família de funções de verossimilhança indexadas por (x,y) , $\{L(\boldsymbol{\theta}; x,y) = f(x,y|\boldsymbol{\theta}); (x,y) \in \mathcal{X}\}$, deve ser mensurável na σ -álgebra *a priori*, usando a notação em ([PEREIRA et al., 2017](#)). A seguinte proposição garante que os testes ϕ_1 e ϕ_2 rejeitam H_0 ao mesmo nível de significância adaptativo ótimo $\alpha(\phi_B^*)$. Na Proposição 1, uma curva suave é uma curva regular de classe \mathcal{C}^1 ([CARMO, 2005](#), Cap. 1).

Proposição 1. Sejam γ uma curva suave em $[0, 1] \times [0, 1]$ e f e g funções mensuráveis não negativas. Suponha que $f(\cdot, \cdot | \boldsymbol{\theta})$ e $g(\cdot)$ são funções contínuas quando definidas sobre a imagem de γ . Se \mathcal{X} é finito, então

$$\begin{aligned} \int_{\gamma} P(\phi_B = 1 | H_0 \text{ é verdadeira}) g(\boldsymbol{\theta} | H_0) ds &= \sum_{(x,y) \in B_k} f_{H_0}(x, y) \\ \text{e } \int_{\Theta_1} P(\phi_B = 0 | H_0 \text{ é falsa}) g(\boldsymbol{\theta} | H_1) d\boldsymbol{\theta} &= \sum_{(x,y) \notin B_k} f_{H_1}(x, y). \end{aligned} \quad (2.23)$$

Demonstração. Note que $\sum_{(x,y) \in B_k} f(x, y | \boldsymbol{\theta}) = \int_{B_k} f(x, y | \boldsymbol{\theta}) d\mu$ porque $B_k \subset \mathcal{X}$ é finito, em que μ denota a medida de contagem e $f(x, y | \boldsymbol{\theta}) \geq 0$ é dada em (2.1). Considere uma curva suave γ em $[0, 1] \times [0, 1]$, sem perda de generalidade, parametrizada pelo comprimento de arco t dada por $\gamma: [t_1, t_2] \rightarrow [0, 1] \times [0, 1]$, em que $[t_1, t_2] \subset [0, 1]$. Além do mais, assumindo que $f(x, y | \boldsymbol{\theta})$ e $g(\cdot)$ são funções contínuas quando definidas sobre a imagem de γ , o que implica que o produto $f(x, y | \boldsymbol{\theta}) g(\boldsymbol{\theta} | H_0)$ é também uma função contínua, uma vez que $g(\cdot | H_0)$ é função contínua, pois $g(\cdot | H_0)$ é uma restrição de $g(\cdot)$. Definamos $\Phi(\boldsymbol{\theta}) = \int_{B_k} f(x, y | \boldsymbol{\theta}) g(\boldsymbol{\theta} | H_0) d\mu$. Então,

$$\begin{aligned} \int_{\gamma} P(\phi_B = 1 | H_0 \text{ é verdadeira}) g(\boldsymbol{\theta} | H_0) ds &= \int_{\gamma} \left(\int_{B_k} f(x, y | \boldsymbol{\theta}) d\mu \right) g(\boldsymbol{\theta} | H_0) ds \\ &= \int_{\gamma} \int_{B_k} f(x, y | \boldsymbol{\theta}) g(\boldsymbol{\theta} | H_0) d\mu ds \\ &= \int_{\gamma} \Phi(\boldsymbol{\theta}) ds \\ &= \int_{t_1}^{t_2} \Phi(\gamma(t)) \|\gamma'(t)\| dt. \end{aligned}$$

Segue da definição da parametrização por comprimento de arco que $\|\gamma'(t)\| = 1$. Então,

$$\begin{aligned} \int_{\gamma} P(\phi_B = 1 | H_0 \text{ é verdadeira}) g(\boldsymbol{\theta} | H_0) ds &= \int_{t_1}^{t_2} \Phi(\gamma(t)) dt \\ &= \int_{t_1}^{t_2} \int_{B_k} f(x, y | \gamma(t)) g(\gamma(t) | H_0) d\mu dt. \end{aligned}$$

Fazendo $\varphi(x, y, \gamma(t)) = f(x, y | \gamma(t)) g(\gamma(t) | H_0)$, temos que φ é não negativa, pois f e g também são não negativas. Então,

$$\begin{aligned} \int_{\gamma} \int_{B_k} f(x, y | \boldsymbol{\theta}) g(\boldsymbol{\theta} | H_0) d\mu ds &= \int_{t_1}^{t_2} \int_{B_k} \varphi(x, y, t) d\mu dt \\ &\stackrel{(*)}{=} \int_{B_k} \int_{t_1}^{t_2} \varphi(x, y, t) dt d\mu \\ &= \int_{B_k} \int_{t_1}^{t_2} f(x, y | \gamma(t)) g(\gamma(t) | H_0) dt d\mu \\ &= \int_{B_k} \int_{t_1}^{t_2} f(x, y | \gamma(t)) g(\gamma(t) | H_0) \|\gamma'(t)\| dt d\mu, \end{aligned}$$

em que a mudança em (*) foi possível com a aplicação do Teorema de Tonelli. Portanto,

$$\begin{aligned} \int_{\gamma} \int_{B_k} f(x, y | \boldsymbol{\theta}) g(\boldsymbol{\theta} | H_0) d\mu ds &= \int_{B_k} \int_{\gamma} f(x, y | \boldsymbol{\theta}) g(\boldsymbol{\theta} | H_0) ds d\mu \\ &\stackrel{(**)}{=} \int_{B_k} \int_{\gamma} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta} | H_0) ds d\mu \\ &= \int_{B_k} f_{H_0}(x, y) d\mu. \end{aligned}$$

Como $B_k \subset \mathcal{X}$ é finito, temos que $\int_{B_k} f_{H_0}(x, y) d\mu = \sum_{(x, y) \in B_k} f_{H_0}(x, y)$. Além disso, a mudança em (**) é válida porque para todos os pontos $(x, y) \in \mathcal{X}$, $f(x, y | \boldsymbol{\theta}) = L(\boldsymbol{\theta}; x, y)$, para todos os pontos $\boldsymbol{\theta} \in \Theta_0 \subset \Theta$ e portanto, as integrais são iguais. De modo análogo, obtemos (2.23). \square

Da Proposição 1, vemos que o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ é o mesmo para os dois testes, ϕ_1 and ϕ_2 . Para testes de equivalência, em (2.7) temos as curvas γ_1 e γ_2 , para mais detalhes *vide* Apêndice C. Na prova da Proposição 1 é suficiente considerar somente um intervalo $[t_1, t_2] \subset [0, 1]$ devido ao fato de que a propriedade aditiva para integrais de linha é válida porque os traços das curvas são disjuntos e desconexos. De acordo com a Proposição 2, esses testes bayesianos são equivalentes.

Proposição 2. Sejam \mathcal{X}' finito e $f_{H_0} > 0$. Então, $\phi_1 = \phi_2$.

Demonstração. Note que ϕ_1 e ϕ_2 são funções indicadoras, sendo assim é suficiente mostrar a igualdade $B_k = C_P$. Dado $(x, y)_{\text{obs}} \in B_k$ uma amostra observada, isto é, $FB_{01}(x, y)_{\text{obs}} < k$. Observe que $FB_{01}(x, y) \leq FB_{01}(x, y)_{\text{obs}} < k$, para todo $(x, y) \in B_{\text{obs}}$, o que implica que $FB_{01}(x, y) < k$, para todo $(x, y) \in B_{\text{obs}}$, o que equivale a $(x, y) \in B_k$, para todo $(x, y) \in B_{\text{obs}}$, então $B_{\text{obs}} \subset B_k$. Como $f_{H_0} > 0$, segue que valor $P < \alpha(\phi_B^*)$ e concluímos que $(x, y)_{\text{obs}} \in C_P$, então $B_k \subset C_P$. Por outro lado, considere $(x, y)_{\text{obs}} \in C_P$, o que significa que valor $P < \alpha(\phi_B^*)$. Para prosseguir, precisamos provar que $B_{\text{obs}} \subset B_k$. De fato, suponha que $B_{\text{obs}} \not\subset B_k$. Então, existe $(x, y) \in B_{\text{obs}} \setminus B_k$, o que implica que $k \leq FB_{01}(x, y) \leq FB_{01}(x, y)_{\text{obs}}$. Assim, $B_k \subset B_{\text{obs}}$ e como B_{obs} e B_k são diferentes, uma vez que $B_{\text{obs}} \not\subset B_k$, concluímos que $\sum_{(x, y) \in B_{\text{obs}}} f_{H_0}(x, y) > \sum_{(x, y) \in B_k} f_{H_0}(x, y)$ porque $f_{H_0} > 0$, o que contradiz a hipótese valor $P < \alpha(\phi_B^*)$. Portanto, a afirmação $B_{\text{obs}} \subset B_k$ é válida. Então, $(x, y)_{\text{obs}} \in B_k$, pois $(x, y)_{\text{obs}} \in B_{\text{obs}}$, e portanto, $B_k = C_P$. \square

Na abordagem bayesiana, quando tomamos $l_0 = l_1 = 1$ isso significa que os dois tipos de erros têm a mesma gravidade relativa (PEREIRA *et al.*, 2017). Além do mais, nas Seções 2.4 e 2.5 diferentes valores são configurados para l_0 e l_1 tais que $l_1 > l_0$, o que significa que o erro do tipo I é visto como mais grave do que o erro do tipo II. Para o teste em Pereira *et al.* (2017), não encontramos na literatura um estudo de simulação considerando a proporção de vezes em que H_0 é rejeitada quando $|p - q| = \delta$ (H_0 é verdadeira). Nos resultados do estudo de simulação para as taxas do erro do tipo I esperamos que o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ tenda a

se aproximar da proporção de vezes em que H_0 é rejeitada quando $|p - q| = \delta$ (H_0 é verdadeira) calculada com base nas 3000 réplicas.

2.4 Estudo de simulação

Nesta seção reportamos os resultados numéricos e as conclusões do estudo de simulação, analisando a relação entre o tamanho amostral e os níveis de significância bem como o poder para diferentes valores da margem δ . As rotinas implementadas para realizar as simulações foram elaboradas na linguagem de programação R (R Core Team, 2021). Quanto à nomenclatura dos métodos, B2017 denota o teste bayesiano apresentado em Pereira *et al.* (2017) e teste z denota o teste frequentista. Da Proposição 2, apresentamos somente os resultados para o teste B2017.

Obtemos os dados simulando valores das variáveis aleatórias $X|p \sim \text{binomial}(m, p)$ e $Y|q \sim \text{binomial}(n, q)$ para tamanhos amostrais $m = n = 50, 100, 200$ e 300 . Ao considerar também a escolha do tamanho da amostra utilizando a teoria da decisão para testes de equivalência, alguns tamanhos de amostra diferentes ($m \neq n$) foram escolhidos. Tomamos 3000 réplicas para cada valor de p , q e δ , que são alguns dos valores em Barker *et al.* (2001). Para calcular as integrais nas medidas de evidência em (2.17), utilizamos o método de Monte Carlo com 2000 réplicas.

Na abordagem bayesiana consideramos os valores dos parâmetros para as distribuições *a priori* beta conforme Williamson (2007). Em Williamson (2007), as escolhas das distribuições *a priori* Beta(2,4) ou Beta(2,2) são arbitrárias a fim de modelar uma certa crença subjetiva vaga de que as proporções se situavam entre 0 e 0,5 ou próximas de 0,5, respectivamente. Para a distribuição *a priori* de Haldane, não foi possível calcular o fator Bayes. Na Figura 2 podemos ver as diferentes formas de funções densidade de probabilidade g da família Beta para estes valores dos parâmetros. Em todas as simulações, consideramos $P(H_0) = P(H_1) = 0,5$, ou seja, as duas hipóteses têm probabilidades *a priori* iguais.

2.4.1 Taxa do erro do tipo I

Nesta seção analisamos as taxas do erro do tipo I dos testes z e B2017. A Tabela 2 mostra as taxas do erro do tipo I dos testes z e B2017. De acordo com os resultados, o teste z tem uma taxa do erro do tipo I próxima ao nível de significância nominal fixado em 0,05 à medida que o tamanho amostral aumenta para uma margem menos conservadora $\delta \geq 0,2$. Na Tabela 2, com o aumento dos valores da margem δ os resultados apontam para taxas do erro do tipo I menores do teste B2017 e de $\alpha(\phi_B^*)$ em comparação com o cenário com $l_0 = l_1 = 1$. Na sequência, quando $\delta \geq 0,2$ e $n = 300$, vemos que as taxas do erro do tipo I do teste B2017 e $\alpha(\phi_B^*)$ estão muito próximas (ou menores do que) de 0,03. Na abordagem frequentista, para um nível de significância nominal de 0,03, um tamanho de amostra maior pode ser requerido

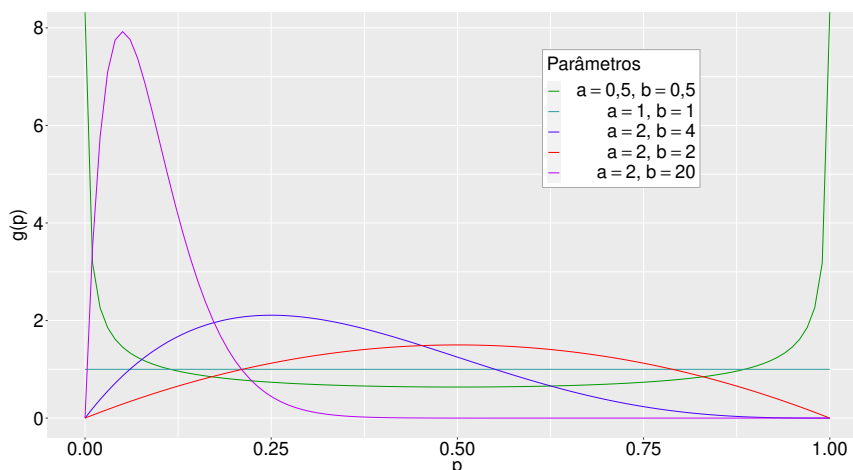


Figura 2 – Gráficos das funções densidade de probabilidade *a priori* correspondentes às distribuições Beta(1, 1), Beta(0,5;0,5), Beta(2, 4) e Beta(2, 2).

para indicar a equivalência entre os dois tratamentos. Finalmente, considerando as Tabelas 2 e 3, quando as perdas são iguais os resultados sugerem um melhor desempenho para a distribuição *a priori* de Jeffreys em comparação com as outras distribuições *a priori* porque os resultados apontam para taxas do erro do tipo I menores do teste B2017 com a distribuição *a priori* de Jeffreys. Quanto às taxas do erro do tipo I, em geral, os resultados indicam que o teste B2017 é pouco sensível à escolha da distribuição *a priori*. Para distribuições *a priori* mais informativas como a distribuição Beta(2, 20), os resultados são melhores, no entanto, nem sempre é simples obter uma distribuição *a priori* mais informativa. Distribuições *a priori* mais informativas como a Beta(20, 2), mas longe da informação correta, podem levar a uma decisão contrária quando se considera Beta(2, 20).

2.4.2 Determinação do tamanho amostral

A Tabela 4 mostra o decaimento dos dois tipos de probabilidades de erro com o aumento dos tamanhos amostrais, como esperado, para o teste B2017. Conseqüentemente, o risco para o fabricante e os pacientes é reduzido à medida que mais dados se tornam disponíveis. É de interesse determinar o tamanho amostral para alcançar o nível de significância fixado e o poder. Na Tabela 4, considerando a distribuição *a priori* de Bayes-Laplace e $l_0 = l_1 = 1$, consideramos o cenário com proporções $p = 0,5$ e $q = 0,3$, com valor de margem $\delta = 0,2$. Para o nível de significância adaptativo ótimo de 0,0809 e o poder de 0,7723, os resultados indicam que os tamanhos de amostra sugeridos são $m = 300$ e $n = 300$. Note que quando esse nível de significância adaptativo ótimo é usado, o tamanho da amostra é naturalmente considerado no planejamento amostral. Além disso, para um tamanho amostral total fixado ($m + n$), os resultados apontam para um maior desequilíbrio (desbalanceamento) nas probabilidades de erro quando $m \neq n$. Para exemplificar, para $m = 300$ e $n = 100$, os resultados indicam que os valores de $\alpha(\phi_B^*)$ e $\beta(\phi_B^*)$ aumentam quando comparados com o teste com $m = n = 200$. Uma situação

Tabela 2 – Proporção de vezes em que H_0 é rejeitada quando $|p - q| = \delta$ (H_0 é verdadeira) para o nível de significância fixado $\alpha(\phi_f) = 0,05$ para o teste z e para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017, com a distribuição *a priori* de Bayes-Laplace para todos os testes.

p	q	δ	n	$l_0 = l_1 = 1$			$l_0 = 1$ e $l_1 = 2$	
				teste z	B2017	$\alpha(\phi_B^*)$	B2017	$\alpha(\phi_B^*)$
0,10	0,05	0,05	50	0,0080	0,5347	0,5178	0,0100	0,0118
			100	0,0047	0,4563	0,5062	0,0520	0,0157
			200	0,0487	0,3323	0,4316	0,1420	0,0312
			300	0,0477	0,2483	0,3776	0,1160	0,0552
0,20	0,10	0,10	50	0,0117	0,3923	0,4399	0,0637	0,0305
			100	0,0510	0,2977	0,3388	0,1363	0,0845
			200	0,0507	0,1910	0,2373	0,0887	0,1022
			300	0,0487	0,1510	0,1883	0,0647	0,0824
0,25	0,10	0,15	50	0,0607	0,2937	0,3243	0,1377	0,1057
			100	0,0617	0,2030	0,2187	0,0930	0,0978
			200	0,0533	0,1213	0,1487	0,0537	0,0653
			300	0,0490	0,0843	0,1154	0,0350	0,0509
0,40	0,20	0,20	50	0,0540	0,2607	0,2363	0,1230	0,1064
			100	0,0557	0,1670	0,1572	0,0877	0,0711
			200	0,0483	0,1003	0,1042	0,0480	0,0465
			300	0,0490	0,0847	0,0816	0,0350	0,0363
0,45	0,20	0,25	50	0,0520	0,1990	0,1820	0,0920	0,0859
			100	0,0507	0,1237	0,1198	0,0483	0,0532
			200	0,0427	0,0617	0,0777	0,0257	0,0349
			300	0,0430	0,0480	0,0607	0,0220	0,0272
0,50	0,20	0,30	50	0,0540	0,1383	0,1451	0,0673	0,0657
			100	0,0523	0,0880	0,0934	0,0387	0,0424
			200	0,0463	0,0497	0,0611	0,0250	0,0273
			300	0,0447	0,0387	0,0478	0,0177	0,0217
0,50	0,30	0,20	50	0,0480	0,2767	0,2363	0,1397	0,1064
			100	0,0440	0,1687	0,1580	0,0807	0,0706
			200	0,0497	0,1143	0,1041	0,0510	0,0466
			300	0,0430	0,0850	0,0809	0,0387	0,0362
0,60	0,30	0,30	50	0,0347	0,1237	0,1451	0,0580	0,0657
			100	0,0470	0,0970	0,0951	0,0463	0,0416
			200	0,0503	0,0657	0,0610	0,0323	0,0274
			300	0,0450	0,0483	0,0477	0,0247	0,0216
0,65	0,40	0,25	50	0,0413	0,1890	0,1820	0,0917	0,0859
			100	0,0437	0,1330	0,1200	0,0527	0,0533
			200	0,0423	0,0857	0,0774	0,0353	0,0347
			300	0,0540	0,0713	0,0607	0,0357	0,0269

semelhante é também relatada no Exemplo 3 em [Pereira et al. \(2017\)](#) para uma diferença de duas proporções.

Tabela 3 – Proporção de vezes em que H_0 é rejeitada quando $|p - q| = \delta$ (H_0 é verdadeira) para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017 com $l_0 = l_1 = 1$.

p	q	δ	n	$p, q \sim \text{Beta}(0, 5, 0, 5)$		$p, q \sim \text{Beta}(2, 4)$		$p, q \sim \text{Beta}(2, 2)$	
				B2017	$\alpha(\phi_B^*)$	B2017	$\alpha(\phi_B^*)$	B2017	$\alpha(\phi_B^*)$
0,10	0,05	0,05	50	0,1910	0,1957	0,6310	0,6492	0,6180	0,6254
			100	0,2927	0,2045	0,5040	0,5968	0,4920	0,5983
			200	0,2617	0,2244	0,3390	0,4485	0,3543	0,4896
			300	0,1750	0,2115	0,2637	0,4013	0,2643	0,4253
0,20	0,10	0,10	50	0,2667	0,2370	0,4577	0,4931	0,4633	0,5043
			100	0,2200	0,2232	0,3423	0,3812	0,3423	0,3957
			200	0,1343	0,1508	0,1937	0,2510	0,2090	0,2751
			300	0,1097	0,1177	0,1587	0,2037	0,1553	0,2148
0,25	0,10	0,15	50	0,2237	0,2273	0,3137	0,3586	0,3127	0,3735
			100	0,1457	0,1503	0,2213	0,2455	0,2213	0,2553
			200	0,0870	0,1038	0,1313	0,1609	0,1313	0,1723
			300	0,0623	0,0778	0,0903	0,1227	0,0863	0,1311
0,40	0,20	0,20	50	0,1713	0,1766	0,2737	0,2632	0,2767	0,2701
			100	0,1113	0,1045	0,1720	0,1643	0,1910	0,1811
			200	0,0780	0,0715	0,1007	0,1051	0,1060	0,1184
			300	0,0590	0,0546	0,0857	0,0815	0,0893	0,0921
0,45	0,20	0,25	50	0,1433	0,1300	0,2053	0,1869	0,2057	0,2014
			100	0,0797	0,0821	0,1223	0,1185	0,1337	0,1328
			200	0,0440	0,0540	0,0617	0,0753	0,0637	0,0854
			300	0,0337	0,0438	0,0480	0,0587	0,0483	0,0667
0,50	0,20	0,30	50	0,1077	0,1096	0,1530	0,1515	0,1693	0,1642
			100	0,0607	0,0675	0,0853	0,0909	0,1000	0,1034
			200	0,0383	0,0451	0,0480	0,0564	0,0513	0,0651
			300	0,0320	0,0344	0,0383	0,0432	0,0397	0,0511
0,50	0,30	0,20	50	0,1907	0,1766	0,2983	0,2632	0,3093	0,2701
			100	0,1303	0,1048	0,1740	0,1633	0,1793	0,1808
			200	0,0823	0,0720	0,1190	0,1058	0,1243	0,1176
			300	0,0600	0,0549	0,0850	0,0808	0,0890	0,0919
0,60	0,30	0,30	50	0,1043	0,1096	0,1360	0,1515	0,1580	0,1642
			100	0,0697	0,0676	0,0933	0,0937	0,0977	0,1051
			200	0,0460	0,0451	0,0560	0,0563	0,0667	0,0652
			300	0,0373	0,0347	0,0440	0,0427	0,0507	0,0510
0,65	0,40	0,25	50	0,1400	0,1300	0,1890	0,1869	0,1997	0,2014
			100	0,0960	0,0821	0,1330	0,1216	0,1473	0,1357
			200	0,0670	0,0548	0,0827	0,0753	0,0900	0,0849
			300	0,0523	0,0439	0,0683	0,0584	0,0773	0,0668

Tabela 4 – Níveis de significância adaptativo ótimo ($\alpha(\phi_B^*)$) e probabilidades do erro do tipo II adaptativa ótima ($\beta(\phi_B^*)$), poderes ($1 - \beta(\phi_B^*)$) e riscos ($r(\phi_B^*)$) para o teste B2017, para tamanhos amostrais diferentes.

m	n	$\alpha(\phi_B^*)$	$\beta(\phi_B^*)$	$1 - \beta(\phi_B^*)$	$r(\phi_B^*)$
200	200	0,1041	0,2556	0,7444	0,1799
300	100	0,1217	0,2734	0,7266	0,1975
300	200	0,0918	0,2431	0,7569	0,1675
300	300	0,0809	0,2277	0,7723	0,1543

2.4.3 Poder

A Tabela 5 reporta os resultados relativos ao poder dos testes considerando a distribuição *a priori* de Bayes-Laplace para o teste B2017. Quando as perdas são iguais, para $\delta = 0,15$ o poder simulado do teste B2017 está mais próximo do poder ($1 - \beta(\phi_B^*)$) e para $\delta \neq 0,15$, os resultados indicam que a diferença absoluta entre o poder simulado e o poder tende a aumentar à medida que o tamanho da amostra aumenta, sendo maiores para as margens $\delta = 0,05; 0,25$ e $0,3$. Quando os valores de perda são diferentes, os resultados indicam um comportamento semelhante ao encontrado quando os valores da margem são $\delta = 0,15$ ou $\delta \neq 0,15$ e indicam também um decaimento maior do poder simulado e do poder com o aumento de l_1 . Em geral, estes resultados mostram um melhor desempenho para um tamanho de amostra menor com o teste B2017 versus o teste z . Finalmente, considerando as Tabelas 5 e 6, para $\delta = 0,1, 0,15$ e $0,2$, para $n = 300$, quando as perdas são iguais, o valor do poder do teste B2017 é maior quando a distribuição *a priori* é a de Jeffreys e o valor do poder simulado do teste B2017 é maior quando as distribuições *a priori* são de Bayes-Laplace, Beta(2,4) e Beta(2,2). Quanto à diferença absoluta entre o poder simulado e o poder, para $\delta = 0,1$ e $n = 300$, os resultados sugerem um melhor desempenho para a distribuição *a priori* de Jeffreys em comparação com as outras distribuições *a priori* porque os resultados apontam para diferenças absolutas menores do teste B2017 com a distribuição *a priori* de Jeffreys. Da mesma forma, para $\delta = 0,15$, a distribuição *a priori* de Bayes-Laplace mostrou o melhor desempenho e para $\delta = 0,2$, as distribuições *a priori* de Jeffreys e Beta(2,4) mostraram os melhores desempenhos. No sentido pontual, os resultados indicam que o teste B2017 não é muito sensível à escolha da distribuição *a priori*.

2.4.4 Análise de má especificação

Nesta seção analisamos dados de contagens com sobredispersão. O objetivo é avaliar os testes z e B2017 sob a suposição de que os dados possuem maior variabilidade do que o previsto pelo modelo binomial teórico. Assim, consideramos o modelo beta-binomial, com $X_{BB} \sim \text{BetaBinomial}(m, a_{BB}, b_{BB})$ e $Y_{BB} \sim \text{BetaBinomial}(n, c_{BB}, d_{BB})$ variáveis aleatórias condicionalmente independentes, como distribuição alternativa ao modelo binomial, com as variáveis aleatórias $X|p \sim \text{Binomial}(m, p)$ e $Y|q \sim \text{Binomial}(n, q)$. Para uma breve revisão do histórico

Tabela 5 – Proporção de vezes em que H_0 é rejeitada quando $|p - q| < \delta$ (H_0 é falsa). Para o teste B2017, a distribuição *a priori* é a de Bayes-Laplace para todos os testes.

p	q	δ	n	$l_0 = l_1 = 1$			$l_0 = 1$ and $l_1 = 2$	
				teste z	B2017	$1 - \beta(\phi_B^*)$	B2017	$1 - \beta(\phi_B^*)$
0,05	0,05	0,05	50	0,0420	0,7793	0,5946	0,0723	0,0288
			100	0,0827	0,8213	0,6359	0,3350	0,0507
			200	0,5010	0,9087	0,6430	0,7050	0,1112
			300	0,7393	0,9500	0,6499	0,8600	0,1893
0,10	0,05	0,10	50	0,1813	0,7780	0,6402	0,4110	0,1049
			100	0,4193	0,7207	0,6472	0,5540	0,2801
			200	0,6120	0,8193	0,6643	0,6613	0,4805
			300	0,7627	0,8693	0,6813	0,7607	0,5367
0,15	0,05	0,15	50	0,2523	0,5880	0,6467	0,4217	0,3385
			100	0,3590	0,6303	0,6629	0,4173	0,4947
			200	0,5330	0,6737	0,7053	0,5227	0,5928
			300	0,6707	0,7757	0,7319	0,6263	0,6418
0,40	0,20	0,25	50	0,1517	0,3687	0,6848	0,2193	0,5545
			100	0,2073	0,3347	0,7294	0,2017	0,6379
			200	0,2927	0,3500	0,7768	0,2157	0,7190
			300	0,3833	0,3940	0,8040	0,2697	0,7577
0,45	0,20	0,30	50	0,1527	0,2907	0,7103	0,1830	0,6015
			100	0,2047	0,2970	0,7583	0,1693	0,6874
			200	0,2853	0,2980	0,8048	0,1960	0,7591
			300	0,3817	0,3690	0,8296	0,2327	0,7941
0,45	0,30	0,20	50	0,1203	0,4783	0,6595	0,2713	0,4793
			100	0,1710	0,4123	0,6979	0,2523	0,5781
			200	0,2703	0,4300	0,7444	0,2700	0,6651
			300	0,3553	0,4807	0,7729	0,3267	0,7108
0,50	0,30	0,25	50	0,1410	0,3917	0,6848	0,2163	0,5545
			100	0,1720	0,3600	0,7278	0,2040	0,6375
			200	0,2680	0,3723	0,7775	0,2267	0,7190
			300	0,3530	0,4050	0,8042	0,2723	0,7575
0,55	0,40	0,20	50	0,0767	0,4353	0,6595	0,2347	0,4793
			100	0,1680	0,4043	0,6988	0,2543	0,5793
			200	0,2600	0,4373	0,7433	0,2760	0,6650
			300	0,3607	0,4833	0,7728	0,3233	0,7107
0,65	0,40	0,30	50	0,1017	0,2923	0,7103	0,1533	0,6015
			100	0,1637	0,3013	0,7589	0,1637	0,6861
			200	0,2557	0,3267	0,8052	0,1930	0,7589
			300	0,3443	0,3613	0,8296	0,2357	0,7939

do modelo beta-binomial e de suas aplicações em diversas áreas do conhecimento, *vide* [Martinez, Achcar e Aragon \(2015\)](#). Então, fixamos $p = a_{BB}/(a_{BB} + b_{BB})$ e $q = c_{BB}/(c_{BB} + d_{BB})$. Para a variável X_{BB} , fazendo $\lambda_x = p$ e $\rho_x = 1/(a_{BB} + b_{BB} + 1)$, obtemos a parametrização do modelo beta-binomial como em [Hinde e Demétrio \(1998\)](#), em que $\mathbb{E}(X_{BB}) = m\lambda_x$ e $\text{var}(X_{BB}) =$

Tabela 6 – Proporção de vezes em que H_0 é rejeitada quando $|p - q| < \delta$ (H_0 é falsa) para o teste B2017 com $l_0 = l_1 = 1$.

p	q	δ	n	$p, q \sim \text{Beta}(0, 5, 0, 5)$		$p, q \sim \text{Beta}(2, 4)$		$p, q \sim \text{Beta}(2, 2)$	
				B2017	$1 - \beta(\phi_B^*)$	B2017	$1 - \beta(\phi_B^*)$	B2017	$1 - \beta(\phi_B^*)$
0,05	0,05	0,05	50	0,4937	0,4439	0,8577	0,7088	0,8570	0,6881
			100	0,6787	0,4893	0,8383	0,6990	0,8377	0,7000
			200	0,8610	0,6001	0,9067	0,6484	0,9107	0,6647
			300	0,9163	0,6548	0,9557	0,6551	0,9557	0,6652
0,10	0,05	0,10	50	0,6380	0,5740	0,7933	0,6672	0,7933	0,6714
			100	0,6710	0,6638	0,7660	0,6584	0,7507	0,6598
			200	0,7517	0,7089	0,8227	0,6629	0,8370	0,6602
			300	0,8290	0,7257	0,8747	0,6758	0,8747	0,6699
0,15	0,05	0,15	50	0,5327	0,6707	0,6247	0,6563	0,5890	0,6600
			100	0,5067	0,6986	0,6617	0,6681	0,6617	0,6596
			200	0,6227	0,7502	0,7107	0,6979	0,7067	0,6895
			300	0,7037	0,7793	0,7810	0,7293	0,7970	0,7169
0,40	0,20	0,25	50	0,2910	0,7247	0,3840	0,6885	0,3840	0,6770
			100	0,2497	0,7723	0,3347	0,7386	0,3523	0,7201
			200	0,2863	0,8173	0,3447	0,7887	0,3567	0,7683
			300	0,3337	0,8401	0,3940	0,8185	0,4017	0,7963
0,45	0,20	0,30	50	0,2453	0,7501	0,3210	0,7300	0,3440	0,7088
			100	0,2323	0,7972	0,2917	0,7762	0,3067	0,7536
			200	0,2613	0,8372	0,2923	0,8253	0,3067	0,8015
			300	0,3107	0,8592	0,3553	0,8497	0,3750	0,8266
0,45	0,30	0,20	50	0,3550	0,7064	0,4937	0,6728	0,5087	0,6604
			100	0,3293	0,7399	0,4127	0,6982	0,4287	0,6843
			200	0,3570	0,7859	0,4280	0,7516	0,4340	0,7317
			300	0,4107	0,8159	0,4817	0,7813	0,4860	0,7611
0,50	0,30	0,25	50	0,3173	0,7247	0,3917	0,6885	0,4137	0,6770
			100	0,2827	0,7671	0,3583	0,7396	0,3790	0,7162
			200	0,3117	0,8148	0,3680	0,7896	0,3823	0,7694
			300	0,3473	0,8412	0,3993	0,8186	0,4143	0,7955
0,55	0,40	0,20	50	0,3443	0,7064	0,4763	0,6728	0,4880	0,6604
			100	0,3420	0,7407	0,4047	0,6998	0,4297	0,6819
			200	0,3650	0,7867	0,4463	0,7510	0,4470	0,7332
			300	0,4147	0,8173	0,4810	0,7803	0,4937	0,7607
0,65	0,40	0,30	50	0,2337	0,7501	0,3093	0,7300	0,3250	0,7088
			100	0,2387	0,7938	0,2837	0,7775	0,3070	0,7509
			200	0,2790	0,8353	0,3113	0,8262	0,3313	0,8021
			300	0,3010	0,8609	0,3313	0,8496	0,3657	0,8261

$m\lambda_x(1 - \lambda_x)[1 + \rho_x(m - 1)]$. Para a variável Y_{BB} , temos $\lambda_y = q$ e $\rho_y = 1/(c_{BB} + d_{BB} + 1)$, com $\mathbb{E}(Y_{BB}) = m\lambda_y$ e $\text{var}(Y_{BB}) = m\lambda_y(1 - \lambda_y)[1 + \rho_y(n - 1)]$. Nas simulações, consideramos apenas o parâmetro $0 < \rho < 1$, com $\rho = \rho_x = \rho_y$, o qual representa o parâmetro de sobredispersão e no limite, quando $\rho = 0$, as variâncias das variáveis binomiais e beta-binomiais são iguais. Como espaço amostral \mathcal{X} do vetor das variáveis aleatórias $(X|p, Y|q)$ é o mesmo para o vetor (X_{BB}, Y_{BB}) com distribuição beta-binomial bivariada, então o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ é o mesmo ou próximo para os diferentes valores de ρ , ou seja, os dados sobredispersos não influenciam de forma nos valores de $\alpha(\phi_B^*)$ quando o tamanho amostral é menor, por conta do espaço amostral \mathcal{X} .

Tabela 7 – Valores verdadeiros dos parâmetros no estudo de simulação.

ρ	p	q	a_{BB}	b_{BB}	c_{BB}	d_{BB}
0,005			39,80	159,20	19,90	179,10
0,01	0,2	0,1	19,80	79,20	9,90	89,10
0,1			1,80	7,20	0,90	8,10
0,3			0,47	1,87	0,23	2,10
0,005			99,50	99,50	59,70	139,30
0,01	0,5	0,3	49,50	49,50	29,70	69,30
0,1			4,50	4,50	2,70	6,30
0,3			1,17	1,17	0,70	1,63

Dados os valores de ρ e de p , fazendo $\rho_x = \rho$, calculamos os valores de a_{BB} e b_{BB} . Como $\rho = 1/(a_{BB} + b_{BB} + 1)$, temos que $a_{BB} + b_{BB} = (1 - \rho)/\rho$. Por outro lado, $p = a_{BB}/(a_{BB} + b_{BB})$, então $p = a_{BB}/[(1 - \rho)/\rho]$. Logo, $a_{BB} = p(1 - \rho)\rho$ e substituindo o valor de a_{BB} em $b_{BB} = [(1 - \rho)/\rho] - a_{BB}$, obtemos b_{BB} . De forma análoga, obtemos os valores de c_{BB} e d_{BB} referentes à variável aleatória Y_{BB} . Para simular os valores de X_{BB} e Y_{BB} , utilizamos o pacote `rmutil` (SWIHART; LINDSEY, 2020). Na Tabela 7 apresentamos os valores verdadeiros dos parâmetros correspondentes às distribuições binomial e beta-binomial. As Figuras 3–6 apresentam os gráficos das distribuições binomial e beta-binomial considerando $\rho = 0,005; 0,01; 0,1$ e $0,3$, respectivamente.

Na Tabela 8, considerando a distribuição *a priori* de Bayes-Laplace, os resultados mostram que a taxa do erro do tipo I para o teste z aumenta à medida que o tamanho amostral aumenta, independentemente da margem escolhida, a um nível de significância $\alpha(\phi_{fr}) = 0,05$. Para $l_0 = l_1 = 1$, $\delta = 0,1$ e $\rho \leq 0,01$, a taxa do erro do tipo I para o teste B2017 tende a diminuir à medida que o tamanho amostral aumenta, mas quando $\rho > 0,01$, isto é, a dispersão dos dados é maior, a taxa do erro do tipo I do teste B2017 tende a oscilar com o aumento do tamanho amostral. Para uma margem maior, $\delta = 0,2$ e com $\rho = 0,005$, ocorre a mesma situação quando $\delta = 0,1$ e $\rho \leq 0,01$; para $\rho = 0,01$, a taxa do erro do tipo I do teste B2017 tende a oscilar com o aumento do tamanho amostral e para $\rho \geq 0,1$, a taxa do erro do tipo I do teste B2017 aumenta com o aumento do tamanho amostral. Para perdas diferentes, a taxa do erro do tipo I do teste

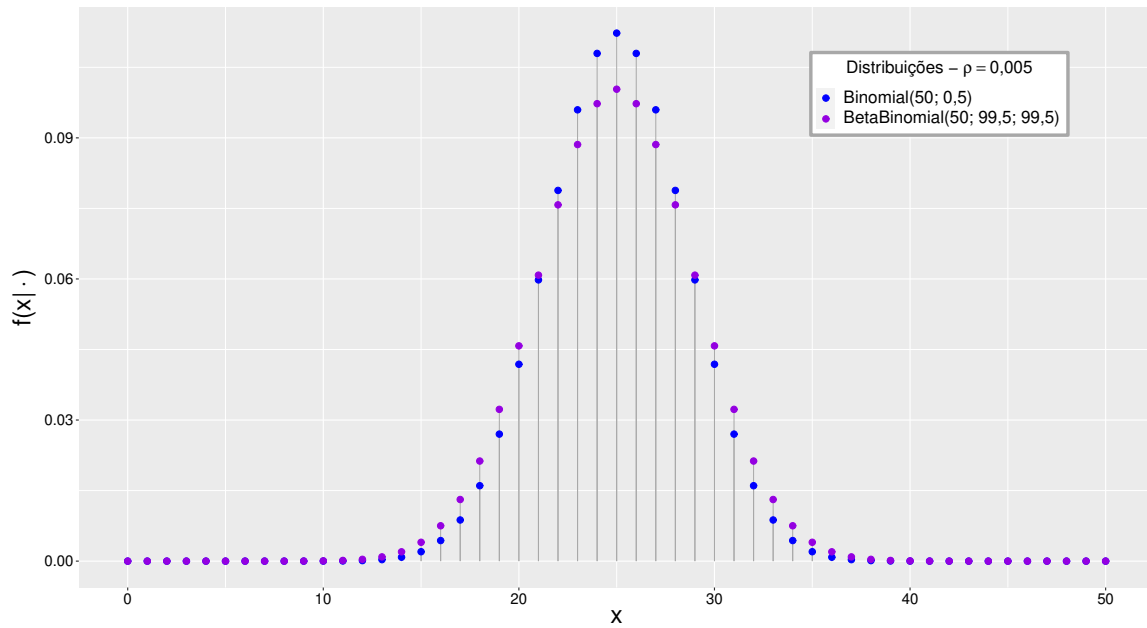


Figura 3 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,005$.

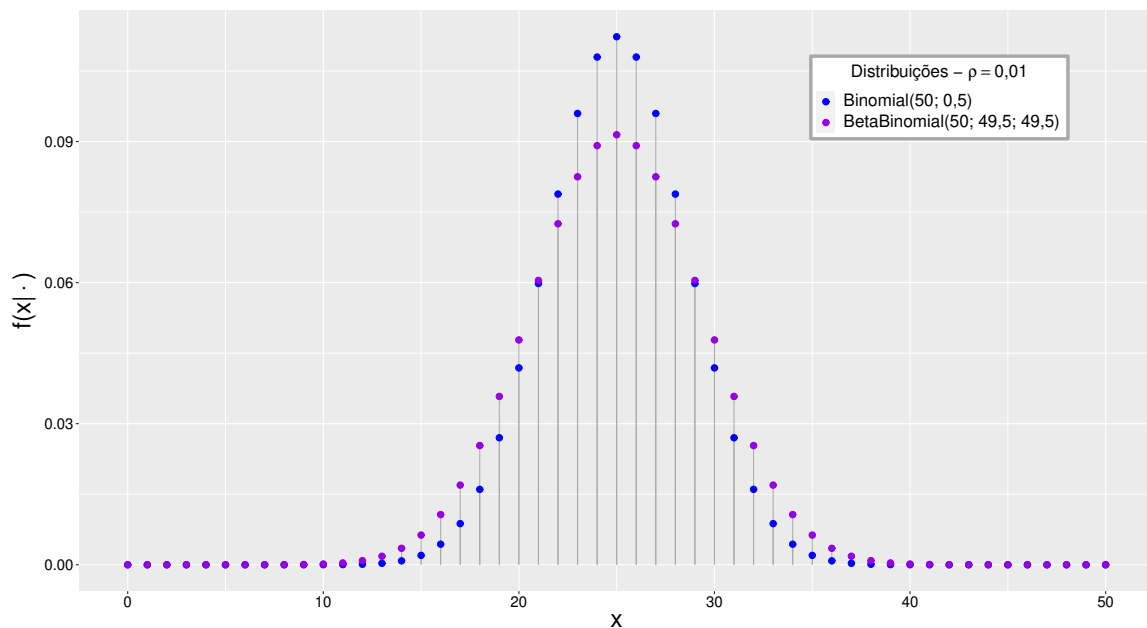


Figura 4 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,01$.

B2017 tende a aumentar com o aumento do tamanho amostral. Nas Tabelas 8 e 9, para $\rho = 0,05$, os resultados mostram que a taxa do erro do tipo I para o teste B2017 diminui à medida que o tamanho amostral aumenta, independentemente da margem escolhida; para $\rho = 0,01$, a taxa do erro do tipo I para o teste B2017 tende a diminuir à medida que o tamanho amostral aumenta, independentemente da margem escolhida, exceto quando $\delta = 0,2$ considerando a distribuição

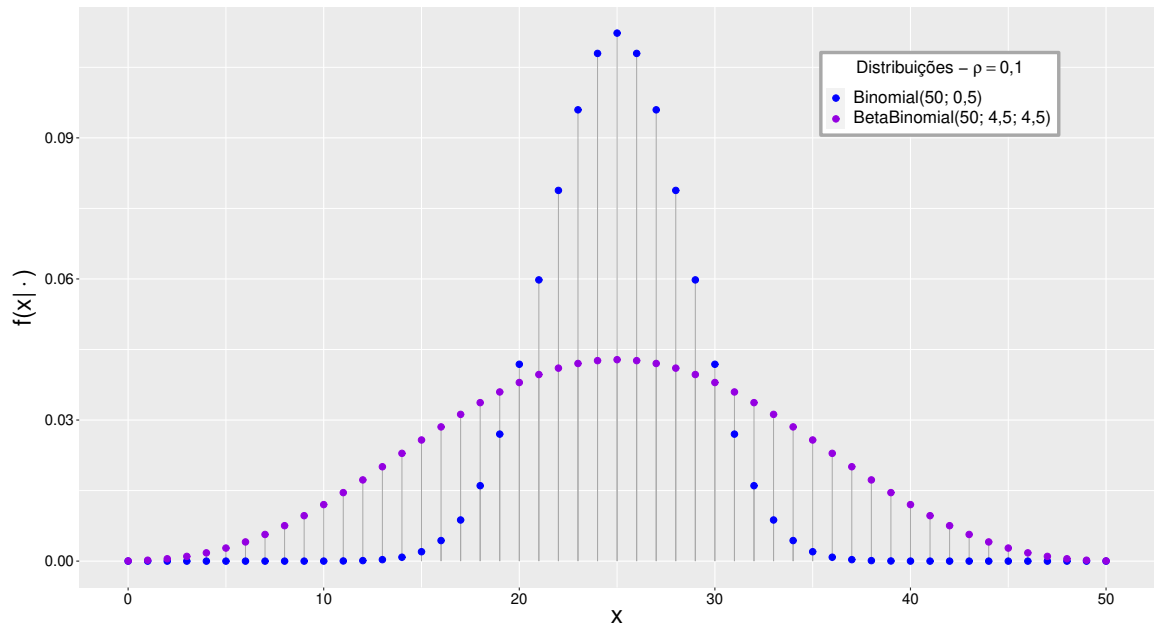


Figura 5 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,1$.

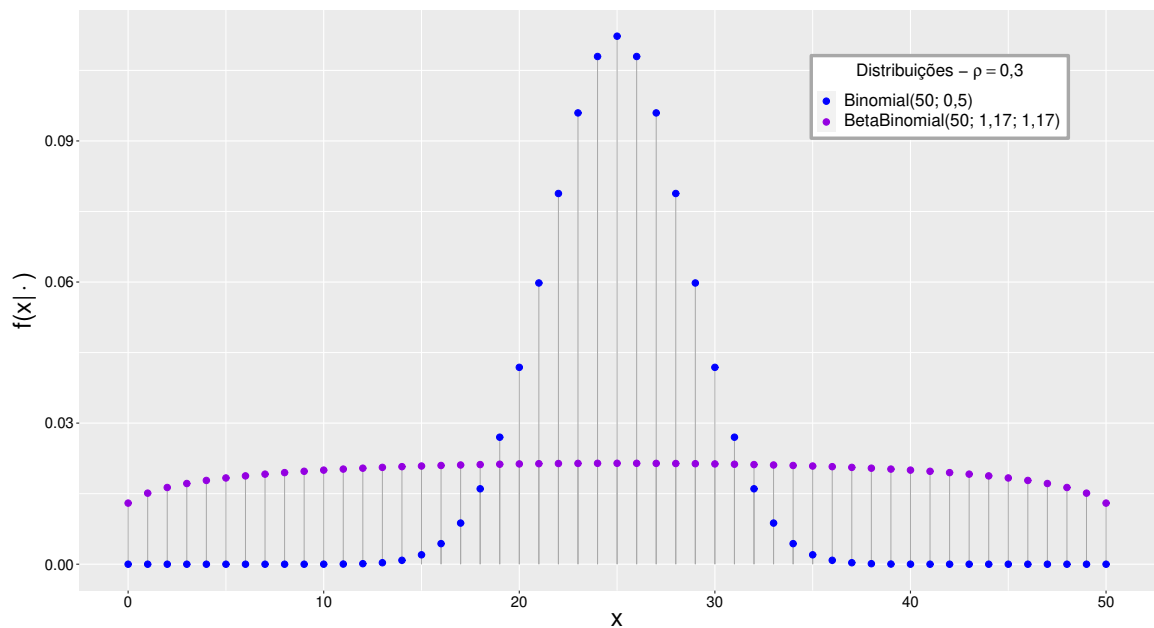


Figura 6 – Gráficos das funções massa de probabilidade correspondendo às distribuições binomial e beta-binomial, com $\rho = 0,3$.

a priori de Jeffreys; para $\rho \geq 0,1$ e $\delta = 0,1$, a taxa do erro do tipo I para o teste B2017 tende diminuir à medida que o tamanho amostral aumenta, com exceção quando a distribuição *a priori* é a de Jeffreys, e $\delta = 0,2$, a taxa do erro do tipo I para o teste B2017 aumenta à medida que o tamanho amostral aumenta, independentemente da distribuição *a priori* elicitada.

Em síntese, quando a dispersão dos dados é menor, a taxa do erro do tipo I para o teste

B2017 diminui à medida que o tamanho amostral aumenta. Nas Tabelas 8 e 9, considerando a distribuição beta-binomial, quando a margem é menor, a dispersão dos dados é menor, para tamanhos amostrais $n = 50$ e 100 , os resultados numéricos são próximos aos valores obtidos nas Tabelas 2 e 3, quando a distribuição é a binomial, e quando a sobredispersão dos dados é maior, para tamanhos amostrais $n = 200$ e 300 , as taxas do erro do tipo I tendem a ser maiores se comparadas com as taxas do erro do tipo I apresentadas nas Tabelas 2 e 3.

Tabela 8 – Proporção de vezes em que H_0 é rejeitada quando $|p - q| = \delta$ (H_0 é verdadeira) para o nível de significância fixado $\alpha(\phi_f) = 0,05$ para o teste z e para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017, com a distribuição *a priori* de Bayes-Laplace para todos os testes.

ρ	δ	n	$l_0 = l_1 = 1$			$l_0 = 1$ e $l_1 = 2$	
			teste z	B2017	$\alpha(\phi_B^*)$	B2017	$\alpha(\phi_B^*)$
0,005	0,1	50	0,0157	0,4060	0,4399	0,0723	0,0305
		100	0,0720	0,3310	0,3388	0,1813	0,0845
		200	0,1420	0,2850	0,2369	0,1817	0,1033
		300	0,1680	0,2677	0,1877	0,1837	0,0826
0,01	0,1	50	0,0153	0,4070	0,4399	0,0817	0,0305
		100	0,0870	0,3467	0,3388	0,2013	0,0845
		200	0,1800	0,3187	0,2369	0,2287	0,1033
		300	0,2257	0,3090	0,1877	0,2410	0,0826
0,1	0,1	50	0,0723	0,3583	0,4399	0,1367	0,0305
		100	0,1473	0,3217	0,3388	0,2213	0,0845
		200	0,2283	0,3230	0,2369	0,2563	0,1033
		300	0,2640	0,3267	0,1877	0,2753	0,0826
0,3	0,1	50	0,2250	0,3897	0,4399	0,2490	0,0305
		100	0,2727	0,3560	0,3388	0,3033	0,0845
		200	0,3120	0,3630	0,2369	0,3247	0,1033
		300	0,3330	0,3620	0,1877	0,3403	0,0826
0,005	0,2	50	0,0620	0,2940	0,2363	0,1653	0,1064
		100	0,0960	0,2453	0,1581	0,1380	0,0696
		200	0,1363	0,2143	0,1032	0,1410	0,0460
		300	0,1630	0,2180	0,0816	0,1533	0,0362
0,01	0,2	50	0,0727	0,3093	0,2363	0,1853	0,1064
		100	0,1330	0,2703	0,1581	0,1740	0,0696
		200	0,1853	0,2593	0,1032	0,1890	0,0460
		300	0,2233	0,2703	0,0816	0,2067	0,0362
0,1	0,2	50	0,1043	0,3173	0,2363	0,2030	0,1064
		100	0,2093	0,3180	0,1581	0,2480	0,0696
		200	0,2900	0,3303	0,1032	0,2917	0,0460
		300	0,3213	0,3470	0,0816	0,3103	0,0362
0,3	0,2	50	0,1217	0,2393	0,2363	0,1727	0,1064
		100	0,1823	0,2413	0,1581	0,1970	0,0696
		200	0,2247	0,2510	0,1032	0,2217	0,0460
		300	0,2457	0,2577	0,0816	0,2367	0,0362

Tabela 9 – Proporção de vezes em que H_0 é rejeitada quando $|p - q| = \delta$ (H_0 é verdadeira) para o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ para o teste B2017 com $l_0 = l_1 = 1$.

ρ	δ	n	$p, q \sim \text{Beta}(0.5, 0.5)$		$p, q \sim \text{Beta}(2, 4)$		$p, q \sim \text{Beta}(2, 2)$	
			B2017	$\alpha(\phi_B^*)$	B2017	$\alpha(\phi_B^*)$	B2017	$\alpha(\phi_B^*)$
0,005	0,1	50	0,2763	0,2370	0,4640	0,4931	0,4697	0,5043
		100	0,2610	0,2232	0,3653	0,3812	0,3653	0,3957
		200	0,2410	0,1546	0,2890	0,2444	0,2967	0,2727
		300	0,2310	0,1175	0,2760	0,2002	0,2830	0,2210
0,01	0,1	50	0,2800	0,2370	0,4593	0,4931	0,4663	0,5043
		100	0,2840	0,2232	0,3780	0,3812	0,3777	0,3957
		200	0,2787	0,1546	0,3220	0,2444	0,3293	0,2727
		300	0,2750	0,1175	0,3200	0,2002	0,3280	0,2210
0,1	0,1	50	0,2780	0,2370	0,3850	0,4931	0,3883	0,5043
		100	0,2833	0,2232	0,3410	0,3812	0,3377	0,3957
		200	0,2937	0,1546	0,3267	0,2444	0,3303	0,2727
		300	0,3057	0,1175	0,3320	0,2002	0,3340	0,2210
0,3	0,1	50	0,3490	0,2370	0,3987	0,4931	0,4000	0,5043
		100	0,3367	0,2232	0,3610	0,3812	0,3610	0,3957
		200	0,3510	0,1546	0,3670	0,2444	0,3687	0,2727
		300	0,3557	0,1175	0,3633	0,2002	0,3637	0,2210
0,005	0,2	50	0,2183	0,1766	0,3130	0,2632	0,3240	0,2701
		100	0,1983	0,1159	0,2483	0,1708	0,2487	0,1797
		200	0,1760	0,0731	0,2127	0,1039	0,2270	0,1168
		300	0,1877	0,0555	0,2180	0,0823	0,2270	0,0926
0,01	0,2	50	0,2343	0,1766	0,3270	0,2632	0,3403	0,2701
		100	0,2340	0,1159	0,2760	0,1708	0,2763	0,1797
		200	0,2283	0,0731	0,2577	0,1039	0,2673	0,1168
		300	0,2450	0,0555	0,2703	0,0823	0,2743	0,0926
0,1	0,2	50	0,2593	0,1766	0,3297	0,2632	0,3347	0,2701
		100	0,2907	0,1159	0,3220	0,1708	0,3237	0,1797
		200	0,3137	0,0731	0,3300	0,1039	0,3347	0,1168
		300	0,3307	0,0555	0,3467	0,0823	0,3517	0,0926
0,3	0,2	50	0,2093	0,1766	0,2487	0,2632	0,2520	0,2701
		100	0,2227	0,1159	0,2463	0,1708	0,2480	0,1797
		200	0,2387	0,0731	0,2517	0,1039	0,2550	0,1168
		300	0,2513	0,0555	0,2583	0,0823	0,2607	0,0926

2.5 Análise de dados

Analizamos os dados em [Buchem et al. \(1997\)](#) de um estudo para investigar a melhora associada ao tratamento com o antibiótico amoxicilina para pacientes com sinusite maxilar aguda. Neste ensaio aleatorizado placebo-controlado, 68 pacientes foram curados de 105 no grupo tratamento (com amoxicilina) e 53 de 101 no grupo placebo.

O resultado observado é $(x, y)_{\text{obs}} = (68, 53)$ e as proporções amostrais observadas são

$\hat{p} = x/m = 68/105 \approx 0,65$ e $\hat{q} = y/n = 53/101 \approx 0,52$. O teste z fornece $T_i(68, 53) = 0,3356$ e $T_s(68, 53) = 3,2709$, com $z_\alpha = 1,6449$, a um nível de significância de 5%. Como $T_i(68, 53) > -z_\alpha$, então não temos evidência suficiente para concluir a bioequivalência entre os tratamentos a um nível de significância $\alpha(\phi_{fr}) = 0,05$.

Para o teste B2017, consideramos $P(H_0) = 0,5$ e para calcular as integrais nas medidas de evidência utilizamos o método Monte Carlo com 2000 réplicas. A Tabela 10 mostra os resultados do teste B2017 e também apresenta a situação em que o erro do tipo I é mais grave do que o erro do tipo II com $k = 0,5$, ou seja, é muito mais grave liberar o tratamento com amoxicilina para este tipo de doença quando, de fato, não é bioequivalente ao placebo. Da Tabela 10, considerando $\delta = 0,1$ e $0,15$, temos o valor $P > \alpha(\phi_B^*)$. Portanto, não temos evidência suficiente para rejeitar H_0 com estes valores da margem. O teste B2017 é considerado a um nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ e com uma probabilidade do erro do tipo II adaptativa ótima $\beta(\phi_B^*)$, os valores de $\alpha(\phi_B^*)$ e $\beta(\phi_B^*)$ sendo menores quando a distribuição *a priori* é a de Jeffreys.

Por outro lado, para $\delta = 0,2$ e $l_0 = l_1 = 1$, o teste B2017 rejeita H_0 . Para exemplificar, quando a distribuição *a priori* é a de Bayes-Laplace, obtemos o valor $P = 0,0989 < 0,1543 = \alpha(\phi_B^*)$. Assim, há evidência suficiente para concluir a bioequivalência ao nível de significância adaptativo ótimo $\alpha(\phi_B^*) = 0,1543$ e com probabilidade do erro do tipo II adaptativa ótima $\beta(\phi_B^*) = 0,3008$. Para a distribuição *a priori* de Jeffreys, o teste B2017 rejeita H_0 a um nível de significância adaptativo ótimo inferior a $\alpha(\phi_B^*) = 0,1063$, enquanto que para $\delta = 0,2$ e com diferentes valores da perda, não temos evidência suficiente para concluir a bioequivalência, quaisquer que sejam os parâmetros especificados para as distribuições *a priori*.

Williamson (2007) realizou uma abordagem bayesiana e alguns testes frequentistas apresentados em Barker *et al.* (2001) para os valores $\delta = 0,10$; $0,15$ e $0,20$, concluindo que não haveria evidência de bioequivalência em todas as abordagens discutidas para os dados em Buchem *et al.* (1997). Na abordagem bayesiana em Williamson (2007), ambos os subespaços paramétricos são regiões. Trabalhando com fatores de Bayes, calculando-os utilizando distribuições *a priori* e *a posteriori* e também não levando em conta a minimização do risco r , Williamson (2007) apontou apenas dois valores possíveis $\delta = 0,15$ e $0,20$ considerando a distribuição *a priori* de Jeffreys que poderia levar à rejeição de H_0 , com respeito às orientações dadas por Jeffreys (1961) e Kass e Raftery (1995). No nosso trabalho, quando os valores de perda são iguais e $\delta = 0,2$, o teste Bayesiano B2017 rejeita H_0 , qualquer que seja a distribuição *a priori* Beta com os parâmetros considerados. Note que fixada uma margem δ , o teste B2017 rejeita H_0 dentro de conjuntos de distribuições *a priori* escolhidas nesta análise de dados, ou seja, independentemente dos valores atribuídos para os parâmetros das distribuições *a priori* Beta, o teste rejeita H_0 , sendo mais robusto neste sentido da decisão a ser tomada não sofrer alteração.

Tabela 10 – Valores P , níveis de significância adaptativos ótimos ($\alpha(\phi_B^*)$) e probabilidades do erro do tipo II adaptativas ótimas ($\beta(\phi_B^*)$) para o teste B2017.

Distribuição <i>a priori</i>	δ	valor P	$l_0 = l_1 = 1$		$l_0 = 1$ e $l_1 = 2$	
			$\alpha(\phi_B^*)$	$\beta(\phi_B^*)$	$\alpha(\phi_B^*)$	$\beta(\phi_B^*)$
Bayes-Laplace	0,10	0,4416	0,3353	0,3524	0,0879	0,7074
Jeffreys		0,3676	0,2059	0,3332	0,0594	0,5365
Beta(2,4)		0,4785	0,3582	0,3453	0,0944	0,7212
Beta(2,2)		0,4798	0,3960	0,3404	0,1099	0,7480
Bayes-Laplace	0,15	0,2437	0,2184	0,3330	0,0961	0,5023
Jeffreys		0,2285	0,1459	0,2891	0,0619	0,4052
Beta(2,4)		0,2733	0,2344	0,3268	0,1027	0,5097
Beta(2,2)		0,2623	0,2507	0,3434	0,1121	0,5375
Bayes-Laplace	0,20	0,0989	0,1543	0,3008	0,0689	0,4188
Jeffreys		0,0906	0,1063	0,2557	0,0494	0,3337
Beta(2,4)		0,1118	0,1608	0,2967	0,0732	0,4191
Beta(2,2)		0,1048	0,1775	0,3127	0,0786	0,4504

2.6 Considerações finais

Neste capítulo estudamos duas novas abordagens bayesianas alternativas para os testes de equivalência de proporções. Na verdade, provamos que tais testes bayesianos são equivalentes. No estudo de simulação, a probabilidade do erro do tipo I e o poder são avaliados sob diferentes cenários, considerando diferentes valores dos parâmetros das distribuições *a priori*, valores para perdas e tamanhos amostrais. No contexto frequentista, o nível de significância nominal é, geralmente, fixado em 0,05. Sob o panorama bayesiano, em alguns cenários os resultados indicaram que podem ser necessárias amostras menores a um nível de significância adaptativo ótimo menor do que 0,05. Além disso, semelhante a [Pereira et al. \(2017\)](#), apresentamos uma tabela com os tamanhos das amostras variando para testes de equivalência, na qual é possível determinar o tamanho amostral necessário para alcançar o nível de significância adaptativo ótimo e o poder. No contexto do estudo de simulação e da aplicação a dados reais, em geral, os resultados sugerem pouca sensibilidade à escolha da distribuição *a priori* relativamente às decisões a serem tomadas. Além disso, os testes se comportaram melhor quando as distribuições *a priori* são Jeffreys ou Bayes-Laplace e mais informativas, levando a resultados numéricos muito próximos.

Também analisamos a inflação da variância induzida pela sobredispersão dos dados teóricos binomiais com a distribuição alternativa beta-binomial. Realizamos um estudo de simulação, no qual os resultados mostram um pior desempenho do teste z com o aumento do tamanho amostral. Para o teste bayesiano, conforme esperado, o nível de significância adaptativo não é alterado com a sobredispersão dos dados com tamanhos amostrais menores e a taxa do erro do tipo I diminui com o aumento do tamanho amostral para valores menores do parâmetro

de sobredispersão dos dados binomiais, principalmente, para tamanhos amostrais menores.

Com a minimização do risco e a utilização de um nível de significância adaptativo ótimo, os testes bayesianos foram muito atrativos em comparação com a abordagem frequentista, indicando uma maior segurança para os ensaios clínicos. Finalmente, analisamos um conjunto de dados sobre a sinusite aguda maxilar. Os resultados indicaram que os testes bayesianos podem ser alternativas mais flexíveis e robustas no que diz respeito à escolha da distribuição *a priori* relativamente às decisões a serem tomadas.

Uma investigação do comportamento dos valores p dos testes de equivalência para duas proporções dos pontos de vista frequentista e bayesiano pode ser objeto de pesquisa futura. Gostaríamos de estudar as classes de testes de hipóteses de [Pericchi e Pereira \(2016\)](#) e de [Pereira et al. \(2017\)](#) com o objetivo de avaliar as propriedades lógicas como a monotonicidade, a invertibilidade e as consonâncias da união e da intersecção, apresentadas em [Izbicki \(2010\)](#). Um estudo com uma revisão de literatura mais estudo de simulação sobre diferentes abordagens bayesianas e frequentistas para os testes de superioridade e de não inferioridade poderia ser uma proposta de trabalho, além das abordagens bayesianas apresentadas no presente trabalho.

ÍNDICE DE SENSIBILIDADE PARA TESTES DE EQUIVALÊNCIA

Neste capítulo propomos um índice de sensibilidade local para analisar o valor P , o nível de significância adaptativo e o poder com respeito à escolha da distribuição *a priori*. Consideramos essas quantidades de interesse apresentadas na Seção 2.2, considerando as hipóteses em (2.7) dadas por $H_0 : |p - q| = \delta$ e $H_1 : |p - q| < \delta$. Também realizamos um estudo de sensibilidade com os dados reais apresentados na Seção 2.5, avaliando essas quantidades de interesse *a priori* presentes nos testes de significância bayesiano. Por fim, apresentamos as considerações finais referentes a este capítulo.

Ao especificar uma função densidade de probabilidade *a priori* sobre o espaço paramétrico, requerida no paradigma bayesiano, na prática esse conhecimento sobre o modelo é impreciso (BASU; JAMMALAMADAKA; LIU, 1996). No nosso trabalho, o problema consiste em como medir a sensibilidade local das quantidades *a priori* referente à escolha da distribuição *a priori*, com ênfase nos testes de equivalência para o modelo beta-binomial bivariado.

Neste capítulo o conceito de robustez bayesiana é usado, com ênfase na análise de sensibilidade local com respeito à escolha da distribuição *a priori*, e portanto, estudamos sobre o quanto que pequenas perturbações infinitesimais na distribuição *a priori* podem provocar maiores perturbações nas quantidades de interesse. Berger *et al.* (1994) apresentaram uma visão geral sobre a análise bayesiana robusta, em particular, no contexto de análise de sensibilidade local notaram um rápido crescimento da literatura sobre a utilização de derivadas funcionais. Clarke (2000) relatou um crescimento paralelo entre as áreas de estudo de robustez e de diferenciabilidade na literatura estatística, também fez uma extensa revisão de literatura sobre o uso de diversos tipos de derivadas (Gâteaux, compacta ou Haadamard e Fréchet), entre os anos de 1947 e 2000, na área de Estatística. Sendo assim, para quantificar o grau de imprecisão dessas quantidades de interesse *a priori* ou o quanto elas são influenciadas pela escolha da distribuição *a priori*, propomos um índice de sensibilidade usando a derivada de Gâteaux. Quanto à interpretação deste índice,

considerando quantidades de interesse *a posteriori*, de acordo com [Gustafson \(1996\)](#), quando a derivada ou mais precisamente a sua norma é grande, a inferência é sensível à distribuição *a priori* elicitada, ou seja, perturbações infinitesimais podem causar grandes mudanças ou perturbações na média *a posteriori*, por exemplo.

[Berger e Moreno \(1994\)](#) abordaram questões sobre robustez bayesiana em modelos bidimensionais envolvendo a esperança *a posteriori* e ressaltaram dificuldades para lidar com espaços paramétricos multidimensionais. Neste trabalho, consideramos a independência entre as coordenadas do vetor paramétrico $\boldsymbol{\theta} = (p, q)$ e a classe ε -contaminada é construída com as funções densidades de probabilidade conjuntas *a priori*. Dessa forma, dada $g \in \mathcal{G}$ uma função densidade de probabilidade conjunta *a priori* elicitada, em que \mathcal{G} denota a classe de todas as possíveis funções densidade de probabilidade conjunta *a priori* sobre o espaço paramétrico Θ , a classe ε -contaminada é expressa por

$$\Gamma_{\varepsilon, g} = \{g_{\varepsilon} = (1 - \varepsilon)g + \varepsilon g_* ; g_* \in \mathcal{G}_* \subset \mathcal{G}\}, \quad \text{com } 0 \leq \varepsilon \leq 1, \quad (3.1)$$

em que g é a função densidade de probabilidade *a priori* elicitada, \mathcal{G}_* é uma subclasse de distribuições de \mathcal{G} e ε é uma constante dada refletindo a quantidade de contaminação.

Recentemente, [Al-Labadi, Asl e Wang \(2021\)](#) revisaram o uso da classe em (3.1), relatando diversos estudos com diferentes objetivos. Por exemplo, há trabalhos nos quais usaram a classe em (3.1) fazendo várias escolhas da subclasse \mathcal{G}_* . Além de outros trabalhos em que a usaram em estudos sobre a robustez de regiões de verossimilhança, a robustez baseada em curvatura e a robustez de taxas de crenças relativas. [Rai, Tomer e Chaturvedi \(2021\)](#) abordaram a classe em (3.1) no contexto de análise de sobrevivência. Diferentes quantidades *a posteriori* são analisadas com o uso de índices baseados em derivadas funcionais. Neste contexto, construímos funcionais estatísticos definidos no espaço das funções densidades de probabilidade conjuntas *a priori* possíveis para elicitamos a derivada de Gâteaux como derivada funcional. Dessa forma, o principal objetivo é propor um índice de sensibilidade usando a derivada de Gâteaux com a finalidade de analisar a sensibilidade local das quantidades de interesse *a priori*, presentes nos testes de significância bayesianos para os testes de equivalência, referente à escolha da distribuição *a priori*.

3.1 Medida de robustez

Neste trabalho assumimos uma estrutura ε -contaminada para um modelo bivariado, supondo independência entre as variáveis aleatórias p e q . Antes de definirmos os funcionais estatísticos e suas respectivas derivadas de Gâteaux aplicadas à função densidade de probabilidade *a priori* elicitada, precisamos estabelecer alguns resultados fundamentais sobre as medidas de evidência sob cada hipótese e referentes aos conjuntos B_k e B_{obs} .

Lema 2. Sejam $g \in \mathcal{G}$ e $g \in \mathcal{G}_* \subset \mathcal{G}$ funções densidade de probabilidade *a priori*. As medidas de evidência ponderadas para os dados (x, y) sob cada hipótese, considerando a função densidade de probabilidade $g_\varepsilon = (1 - \varepsilon)g + \varepsilon g_*$, com $0 \leq \varepsilon \leq 1$, são dadas por

$$f_{H_0, \varepsilon}(x, y) = \frac{(1 - \varepsilon) \int_{\gamma} \varphi(x, y, \boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} \varphi_*(x, y, \boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} g_*(\boldsymbol{\theta}) ds} \quad (3.2)$$

$$\text{e } f_{H_1, \varepsilon}(x, y) = \frac{(1 - \varepsilon) \int_{\Theta_1} \varphi(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta} + \varepsilon \int_{\Theta_1} \varphi_*(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta}}{(1 - \varepsilon) \int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta} + \varepsilon \int_{\Theta_1} g_*(\boldsymbol{\theta}) d\boldsymbol{\theta}}, \quad (3.3)$$

em que $\varphi(x, y, \boldsymbol{\theta}) = L(\boldsymbol{\theta}; x, y)g(\boldsymbol{\theta})$ e $\varphi_*(x, y, \boldsymbol{\theta}) = L(\boldsymbol{\theta}; x, y)g_*(\boldsymbol{\theta})$.

Demonstração. Temos que

$$f_{H_0, \varepsilon}(x, y) = \int_{\gamma} L(\boldsymbol{\theta}; x, y) g_\varepsilon(\boldsymbol{\theta} | H_0) ds,$$

em que $g_\varepsilon(\boldsymbol{\theta} | H_0) = \frac{g_\varepsilon(\boldsymbol{\theta})}{\int_{\gamma} g_\varepsilon(\boldsymbol{\theta}) ds}$, com $g_\varepsilon(\boldsymbol{\theta}) = (1 - \varepsilon)g(\boldsymbol{\theta}) + \varepsilon g_*(\boldsymbol{\theta})$. Então, reescrevemos $f_{H_0, \varepsilon}$

como

$$\begin{aligned} f_{H_0, \varepsilon}(x) &= \int_{\gamma} L(\boldsymbol{\theta}; x, y) \frac{g_\varepsilon(\boldsymbol{\theta})}{\int_{\gamma} g_\varepsilon(\boldsymbol{\theta}) ds} ds \\ &= \frac{\int_{\gamma} L(\boldsymbol{\theta}; x, y) g_\varepsilon(\boldsymbol{\theta}) ds}{\int_{\gamma} g_\varepsilon(\boldsymbol{\theta}) ds} \\ &= \frac{\int_{\gamma} L(\boldsymbol{\theta}; x, y) [(1 - \varepsilon)g(\boldsymbol{\theta}) + \varepsilon g_*(\boldsymbol{\theta})] ds}{\int_{\gamma} [(1 - \varepsilon)g(\boldsymbol{\theta}) + \varepsilon g_*(\boldsymbol{\theta})] ds} \\ &= \frac{\int_{\gamma} [(1 - \varepsilon)\varphi(x, y, \boldsymbol{\theta}) + \varepsilon\varphi_*(x, y, \boldsymbol{\theta})] ds}{\int_{\gamma} [(1 - \varepsilon)g(\boldsymbol{\theta}) + \varepsilon g_*(\boldsymbol{\theta})] ds}. \end{aligned}$$

De modo semelhante, obtemos

$$f_{H_1, \varepsilon}(x) = \frac{\int_{\Theta_1} \varphi_\varepsilon(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_1} g_\varepsilon(\boldsymbol{\theta}) d\boldsymbol{\theta}} = \frac{(1 - \varepsilon) \int_{\Theta_1} \varphi(x, y, \boldsymbol{\theta}) ds + \varepsilon \int_{\Theta_1} \varphi_*(x, y, \boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_{\Theta_1} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\Theta_1} g_*(\boldsymbol{\theta}) ds},$$

em que $\varphi_\varepsilon(x, y) = (1 - \varepsilon)\varphi(x, y, \boldsymbol{\theta}) + \varepsilon\varphi_*(x, y, \boldsymbol{\theta})$. □

Proposição 3. Sejam \mathcal{X} finito e $f_{H_1, \varepsilon}$ e $f_{H_1} > 0$. Então, para cada amostra $(x, y) \in \mathcal{X}$, $\lim_{\varepsilon \rightarrow 0} FB_{01, \varepsilon}(x, y) = FB_{01}(x, y)$, em que $FB_{01, \varepsilon}(\mathbf{z}) = f_{H_0, \varepsilon}(\mathbf{z})/f_{H_1, \varepsilon}(\mathbf{z})$.

Demonstração. Para cada amostra $\mathbf{z} = (x, y) \in \mathcal{X}$, considerando (3.2), temos que $\lim_{\varepsilon \rightarrow 0} f_{H_0, \varepsilon}(\mathbf{z}) = \frac{\int_{\gamma} \varphi(\mathbf{z}, \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} = f_{H_0}(\mathbf{z})$. De forma análoga, obtemos $\lim_{\varepsilon \rightarrow 0} f_{H_1, \varepsilon}(\mathbf{z}) = f_{H_1}(\mathbf{z})$. Logo, $\lim_{\varepsilon \rightarrow 0} FB_{01, \varepsilon}(\mathbf{z}) = FB_{01}(\mathbf{z})$. \square

Os Teoremas 1 e 2 são usados na demonstração do Teorema 3, no qual obtemos as derivadas de Gâteaux dos funcionais estatísticos aplicados à função densidade de probabilidade *a priori* elicitada.

Teorema 1. Sejam \mathcal{X} finito, com $(x, y)_{\text{obs}} \in \mathcal{X}$ tal que $FB_{01}^{-1}(\{FB_{01}(x, y)_{\text{obs}}\}) = \{(x, y)_{\text{obs}}\}$, em que $FB_{01}^{-1}(\{FB_{01}(x, y)_{\text{obs}}\})$ é a imagem inversa de $\{FB_{01}(x, y)_{\text{obs}}\}$ por FB_{01} , e $B_{\text{obs}, \varepsilon} = \{(x, y) \in \mathcal{X}; FB_{01, \varepsilon}(x, y) \leq FB_{01, \varepsilon}(x, y)_{\text{obs}}\}$. Existe $\varepsilon_0 > 0$ tal que $\varepsilon \in (0, \varepsilon_0)$ implica $B_{\text{obs}, \varepsilon} = B_{\text{obs}}$.

Demonstração. Primeiro mostramos que existe $\varepsilon_0 > 0$ tal que $\varepsilon \in (0, \varepsilon_0)$ implica $B_{\text{obs}, \varepsilon} \subset B_{\text{obs}}$. Suponha por contradição que não existe $\varepsilon_0 > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0)$, temos a inclusão $B_{\text{obs}, \varepsilon} \subset B_{\text{obs}}$, equivalentemente,

$$\text{para todo } \varepsilon_0 > 0, \text{ existe } \varepsilon \in (0, \varepsilon_0) \text{ tal que } B_{\text{obs}, \varepsilon} \not\subset B_{\text{obs}}. \quad (3.4)$$

A seguir, provamos que existe uma sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ tal que vale a não inclusão $B_{\text{obs}, \varepsilon_N} \not\subset B_{\text{obs}}$ e vale a desigualdade $\varepsilon_{N+1} < \varepsilon_N/2$, para todo $N \in \mathbb{N}$. De fato, definimos a sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ por indução sobre N . Para $N = 1$, dado $\varepsilon_0 > 0$ qualquer, logo, por (3.4), existe $\varepsilon_1 \in (0, \varepsilon_0)$ tal que $B_{\text{obs}, \varepsilon_1} \not\subset B_{\text{obs}}$. Supomos que ε_N está definido. Se ε_N está definido, definimos ε_{N+1} da seguinte forma: como $\varepsilon_N/2 > 0$ então, por (3.4), existe $\varepsilon_{N+1} < \varepsilon_N/2$ tal que $B_{\text{obs}, \varepsilon_{N+1}} \not\subset B_{\text{obs}}$. Assim, a sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ está definida por indução. Portanto, existe uma sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ tal que vale a não inclusão $B_{\text{obs}, \varepsilon_N} \not\subset B_{\text{obs}}$ e vale a desigualdade $\varepsilon_{N+1} < \varepsilon_N/2$, para todo $N \in \mathbb{N}$. Além disso, observe que $\varepsilon_N \rightarrow 0$ quando $N \rightarrow \infty$, pois

$$\varepsilon_{N+1} < \frac{\varepsilon_N}{2} < \frac{\varepsilon_{N-1}}{2^2} < \frac{\varepsilon_{N-2}}{2^3} < \dots < \frac{\varepsilon_1}{2^N} \rightarrow 0, \text{ quando } N \rightarrow \infty.$$

Consideramos \mathbf{z}_{obs} para denotar o ponto $(x, y)_{\text{obs}} \in \mathcal{X}$. Note que para todo $N \in \mathbb{N}$, existe $\mathbf{z}_N = (x_N, y_N) \in B_{\text{obs}, \varepsilon_N} \setminus B_{\text{obs}}$ e como $\mathbf{z}_N \in \mathcal{X}$, o qual é um conjunto finito, existe uma subsequência constante $(\mathbf{z}_{N_j})_{j \in \mathbb{N}}$. Logo, existe $\mathbf{z} \in \mathcal{X}$ tal que $\mathbf{z}_{N_j} = \mathbf{z}$, para todo $j \in \mathbb{N}$. Como $\mathbf{z} = \mathbf{z}_{N_j} \in B_{\text{obs}, \varepsilon_{N_j}} \setminus B_{\text{obs}}$, para todo $j \in \mathbb{N}$, então $FB_{01, \varepsilon_{N_j}}(\mathbf{z}) \leq FB_{01, \varepsilon_{N_j}}(\mathbf{z}_{\text{obs}})$, para todo $j \in \mathbb{N}$, e $FB_{01}(\mathbf{z}) > FB_{01}(\mathbf{z}_{\text{obs}})$. Então, fazendo $j \rightarrow \infty$, temos que $\varepsilon_{N_j} \rightarrow 0$, e pela Proposição 3 e pelo Corolário 2 (LIMA, 2020, Cap. 5), segue que $\lim_{\varepsilon_{N_j} \rightarrow 0} FB_{01, \varepsilon_{N_j}}(\mathbf{z}) \leq \lim_{\varepsilon_{N_j} \rightarrow 0} FB_{01, \varepsilon_{N_j}}(\mathbf{z}_{\text{obs}})$, ou seja, $FB_{01}(\mathbf{z}) \leq FB_{01}(\mathbf{z}_{\text{obs}})$ e $FB_{01}(\mathbf{z}) > FB_{01}(\mathbf{z}_{\text{obs}})$, o que é um absurdo. Portanto, existe $\varepsilon_0 > 0$ tal que, se $\varepsilon \in (0, \varepsilon_0)$, então $B_{\text{obs}, \varepsilon} \subset B_{\text{obs}}$.

Agora, provamos que existe $\varepsilon_0 > 0$ tal que $\varepsilon \in (0, \varepsilon_0)$ implica $B_{\text{obs}} \subset B_{\text{obs}, \varepsilon}$. Supomos por absurdo que

$$\text{para todo } \varepsilon > 0, \text{ existe } \varepsilon_1 \in (0, \varepsilon) \text{ tal que } B_{\text{obs}} \not\subset B_{\text{obs}, \varepsilon_1}. \quad (3.5)$$

Dado $\mathbf{z}_1 = (x_1, y_1) \in B_{\text{obs}} \setminus B_{\text{obs}, \varepsilon_1}$, obtemos $FB_{01, \varepsilon_1}(\mathbf{z}_1) > FB_{01, \varepsilon_1}(\mathbf{z}_{\text{obs}})$ e $FB_{01}(\mathbf{z}_1) \leq FB_{01}(\mathbf{z}_{\text{obs}})$, o que implica que $\mathbf{z}_1 \neq \mathbf{z}_{\text{obs}}$. De fato, se $\mathbf{z}_1 = \mathbf{z}_{\text{obs}}$, então $FB_{01, \varepsilon_1}(\mathbf{z}_1) = FB_{01, \varepsilon_1}(\mathbf{z}_{\text{obs}})$, o que é um absurdo, pois $FB_{01, \varepsilon_1}(\mathbf{z}_1) > FB_{01, \varepsilon_1}(\mathbf{z}_{\text{obs}})$. Logo, $\mathbf{z}_1 \neq \mathbf{z}_{\text{obs}}$. Como $FB_{01}^{-1}(\{FB_{01}(x, y)_{\text{obs}}\}) = \{(x, y)_{\text{obs}}\}$, então $FB_{01}(\mathbf{z}_1) \neq FB_{01}(\mathbf{z}_{\text{obs}})$. Portanto, $FB_{01}(\mathbf{z}_1) < FB_{01}(\mathbf{z}_{\text{obs}})$. Pela Proposição 3 e pelo Teorema 1 (LIMA, 2020, Cap. 5), existe $\zeta_1 > 0$ tal que para todo $\zeta \in (0, \zeta_1)$, $FB_{01, \zeta}(\mathbf{z}_1) < FB_{01, \zeta}(\mathbf{z}_{\text{obs}})$. Por (3.5), existe $\varepsilon_2 \in (0, \varepsilon_1/2)$ tal que $B_{\text{obs}} \not\subset B_{\text{obs}, \varepsilon_2}$. Logo, existe $\mathbf{z}_2 \in B_{\text{obs}} \setminus B_{\text{obs}, \varepsilon_2}$, o que implica que $FB_{01, \varepsilon_2}(\mathbf{z}_2) > FB_{01, \varepsilon_2}(\mathbf{z}_{\text{obs}})$. Portanto, $\mathbf{z}_2 \neq \mathbf{z}_{\text{obs}}$, o que implica que $FB_{01}(\mathbf{z}_2) < FB_{01}(\mathbf{z}_{\text{obs}})$. Portanto, pela Proposição 3 e pelo Teorema 1 (LIMA, 2020, Cap. 5), existe $\zeta_2 > 0$ tal que para todo $\zeta \in (0, \zeta_2)$, temos a desigualdade $FB_{01, \zeta}(\mathbf{z}_2) < FB_{01, \zeta}(\mathbf{z}_{\text{obs}})$. Continuando com $\varepsilon_3 < \varepsilon_2/2$, obtemos as sequências $(\varepsilon_N)_{N \in \mathbb{N}}$, $(\mathbf{z}_N)_{N \in \mathbb{N}}$ e $(\zeta_N)_{N \in \mathbb{N}}$ que satisfazem as seguintes condições:

$$\varepsilon_{N+1} < \varepsilon_N/2, \text{ com } \varepsilon_N \rightarrow 0, \text{ quando } N \rightarrow \infty; \quad (3.6)$$

$$FB_{01, \varepsilon_N}(\mathbf{z}_N) > FB_{01, \varepsilon_N}(\mathbf{z}_{\text{obs}}); \quad (3.7)$$

$$FB_{01, \zeta}(\mathbf{z}_N) < FB_{01, \zeta}(\mathbf{z}_{\text{obs}}), \text{ para todo } \zeta \in (0, \zeta_N) \text{ e para todo } N \in \mathbb{N}. \quad (3.8)$$

Note que $(\mathbf{z}_N)_{N \in \mathbb{N}}$ possui subsequência $(\mathbf{z}_{N_j})_{j \in \mathbb{N}}$ constante \mathbf{z}_* , ou seja, $\mathbf{z}_{N_j} = \mathbf{z}_*$ para todo $j \in \mathbb{N}$ e por (3.7), segue que $FB_{01, \varepsilon_{N_j}}(\mathbf{z}_*) > FB_{01, \varepsilon_{N_j}}(\mathbf{z}_{\text{obs}})$ e por (3.8), temos que $FB_{01, \zeta}(\mathbf{z}_*) < FB_{01, \zeta}(\mathbf{z}_{\text{obs}})$, para todo $\zeta \in (0, \zeta_{N_j})$ e para todo $j \in \mathbb{N}$. Em particular, considere

$$FB_{01, \zeta}(\mathbf{z}_*) < FB_{01, \zeta}(\mathbf{z}_{\text{obs}}), \text{ para todo } \zeta \in (0, \zeta_{N_1}). \quad (3.9)$$

Logo, por (3.6), como $\varepsilon_N \rightarrow 0$ quando $N \rightarrow \infty$, escolhemos $J_0 \in \mathbb{N}$ tal que $\varepsilon_{N_{J_0}} \in (0, \zeta_{N_1})$, o que é possível, pois $\varepsilon_{N_j} \rightarrow 0$ quando $j \rightarrow \infty$. Logo, $FB_{01, \varepsilon_{N_{J_0}}}(\mathbf{z}_*) > FB_{01, \varepsilon_{N_{J_0}}}(\mathbf{z}_{\text{obs}})$ e por (3.9), $FB_{01, \varepsilon_{N_{J_0}}}(\mathbf{z}_*) < FB_{01, \varepsilon_{N_{J_0}}}(\mathbf{z}_{\text{obs}})$, o que é um absurdo. Portanto, existe $\varepsilon_0 > 0$, $\varepsilon \in (0, \varepsilon_0)$ implica $B_{\text{obs}} \subset B_{\text{obs}, \varepsilon}$. Logo, existe $\varepsilon_0 > 0$ tal que quando $\varepsilon \in (0, \varepsilon_0)$, então $B_{\text{obs}, \varepsilon} = B_{\text{obs}}$. \square

No Teorema 1, não é necessário supor $FB_{01}^{-1}(\{FB_{01}(x, y)_{\text{obs}}\}) = \{(x, y)_{\text{obs}}\}$ se o fator de Bayes FB_{01} for uma função injetiva. Dessa forma, propomos a seguinte conjectura.

Conjectura 1. Seja \mathcal{X} finito. Existe $\varepsilon_0 > 0$ tal que, se $\varepsilon \in (0, \varepsilon_0)$, então $B_{\text{obs}, \varepsilon} = B_{\text{obs}}$.

Observamos que, na prática, o cálculo de cada fator de Bayes para cada ponto do espaço amostral envolve diversas simulações pelo método de Monte Carlo, ou seja, são valores aproximados e pode ser verificado que todos os fatores de Bayes são diferentes. Além disso, se o fator de Bayes FB_{01} não for uma função injetiva, o Teorema 1 continua válido.

Teorema 2. Sejam \mathcal{X} um conjunto finito, k uma razão de perdas não pertencendo à imagem da função FB_{01} e $B_{k,\varepsilon} = \{(x,y) \in \mathcal{X}; FB_{01,\varepsilon}(x,y) < k\}$. Existe $\varepsilon_0 > 0$ tal que $\varepsilon \in (0, \varepsilon_0)$ implica $B_{k,\varepsilon} = B_k$.

Demonstração. Primeiro mostramos que existe $\varepsilon_0 > 0$ tal que $\varepsilon \in (0, \varepsilon_0)$ implica $B_{k,\varepsilon} \subset B_k$. Suponha por contradição que não existe $\varepsilon_0 > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0)$, temos a inclusão $B_{k,\varepsilon} \subset B_k$, equivalentemente,

$$\text{para todo } \varepsilon_0 > 0, \text{ existe } \varepsilon \in (0, \varepsilon_0) \text{ tal que } B_{k,\varepsilon} \not\subset B_k. \quad (3.10)$$

A seguir, provamos que existe uma sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ tal que vale a não inclusão $B_{k,\varepsilon_N} \not\subset B_k$ e vale a desigualdade $\varepsilon_{N+1} < \varepsilon_N/2$, para todo $N \in \mathbb{N}$. De fato, definimos a sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ por indução sobre N . Para $N = 1$, dado $\varepsilon_0 > 0$ qualquer, logo, por (3.10), existe $\varepsilon_1 \in (0, \varepsilon_0)$ tal que $B_{k,\varepsilon_1} \not\subset B_k$. Supomos que ε_N está definido. Se ε_N está definido, definimos ε_{N+1} da seguinte forma: como $\varepsilon_N/2 > 0$ então, por (3.10), existe $\varepsilon_{N+1} < \varepsilon_N/2$ tal que $B_{k,\varepsilon_{N+1}} \not\subset B_k$. Assim, a sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ está definida por indução. Portanto, existe uma sequência $(\varepsilon_N)_{N \in \mathbb{N}}$ tal que vale a não inclusão $B_{k,\varepsilon_N} \not\subset B_k$ e vale a desigualdade $\varepsilon_{N+1} < \varepsilon_N/2$, para todo $N \in \mathbb{N}$. Além disso, observe que $\varepsilon_N \rightarrow 0$ quando $N \rightarrow \infty$, pois

$$\varepsilon_{N+1} < \frac{\varepsilon_N}{2} < \frac{\varepsilon_{N-1}}{2^2} < \frac{\varepsilon_{N-2}}{2^3} < \dots < \frac{\varepsilon_1}{2^N} \rightarrow 0, \text{ quando } N \rightarrow \infty.$$

Consideramos \mathbf{z} para denotar o ponto $(x,y) \in \mathcal{X}$. Note que para todo $N \in \mathbb{N}$, existe $\mathbf{z}_N = (x_N, y_N) \in B_{k,\varepsilon_N} \setminus B_k$ e como $\mathbf{z}_N \in \mathcal{X}$, o qual é um conjunto finito, existe uma subsequência constante $(\mathbf{z}_{N_j})_{j \in \mathbb{N}}$. Logo, existe $\mathbf{z} \in \mathcal{X}$ tal que $\mathbf{z}_{N_j} = \mathbf{z}$, para todo $j \in \mathbb{N}$. Como $\mathbf{z} = \mathbf{z}_{N_j} \in B_{k,\varepsilon_{N_j}} \setminus B_k$, para todo $j \in \mathbb{N}$, então $FB_{01,\varepsilon_{N_j}}(\mathbf{z}) < k$, para todo $j \in \mathbb{N}$, e $FB_{01}(\mathbf{z}) \geq k$. Assim, fazendo $j \rightarrow \infty$, então $\varepsilon_{N_j} \rightarrow 0$ e portanto, $FB_{01,\varepsilon_{N_j}}(\mathbf{z}) \rightarrow FB_{01}(\mathbf{z})$, pela Proposição 3. Consequentemente, pelo teorema do confronto, $FB_{01}(\mathbf{z}) = k$, o que contraria o fato de $k \notin Im_{FB_{01}}$, em que $Im_{FB_{01}}$ denota o conjunto imagem da função FB_{01} . Logo, para $k \notin Im_{FB_{01}}$, existe $\varepsilon_0 > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0)$, temos que $B_{k,\varepsilon} \subset B_k$.

Agora, mostramos que existe $\varepsilon_0 > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0)$, temos a inclusão $B_k \subset B_{k,\varepsilon}$. Assim, dado $(x,y) \in B_k$, temos que $FB_{01}(x,y) < k$. Pela Proposição 3, segue que $\lim_{\varepsilon \rightarrow 0} FB_{01,\varepsilon}(x,y) = FB_{01}(x,y) < k$. Então, pelo Corolário 1 (LIMA, 2020, Cap. 5), existe $\varepsilon_0 > 0$ tal que $FB_{01,\varepsilon}(x,y) < k$, para todo $\varepsilon \in [0, 1]$ com $0 < \varepsilon < \varepsilon_0$, o que implica que existe $\varepsilon_0 > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0)$, obtemos $(x,y) \in B_{k,\varepsilon}$. Portanto, existe $\varepsilon_0 > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0)$, temos que $B_k \subset B_{k,\varepsilon}$. Logo, existe $\varepsilon_0 > 0$ tal que $\varepsilon \in (0, \varepsilon_0)$ implica $B_{k,\varepsilon} = B_k$. \square

O valor P_ε denota o valor P calculado considerando a função densidade de probabilidade *a priori* g_ε . A Proposição 4 apresenta o valor P_ε como a média ponderada dos valores P e P_* modificados, no sentido do somatório das medidas de evidência serem aplicadas aos pontos $(x,y) \in B_{\text{obs},\varepsilon}$. As Proposições 4–6 são usadas na demonstração do Teorema 3.

Proposição 4. Valor $P_\varepsilon = \frac{\tau_1 \text{valor } P^\varepsilon + \tau_2 \text{valor } P_*^\varepsilon}{\tau_1 + \tau_2}$, em que $\tau_1 = (1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds$, $\tau_2 = \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds$, valor $P^\varepsilon = \sum_{(x,y) \in B_{\text{obs},\varepsilon}} f_{H_0}(x,y)$ e valor $P_*^\varepsilon = \sum_{(x,y) \in B_{\text{obs},\varepsilon}} f_{H_{0,*}}(x,y)$.

Demonstração. Dada (x,y) uma amostra observada, temos, pelo Lema 2 que

$$\begin{aligned}
\text{valor } P_\varepsilon &= \sum_{(x,y) \in B_{\text{obs},\varepsilon}} f_{H_{0,\varepsilon}}(x,y) \\
&= \sum_{(x,y) \in B_{\text{obs},\varepsilon}} \frac{(1 - \varepsilon) \int_\gamma \varphi(x,y, \boldsymbol{\theta}) ds + \varepsilon \int_\gamma \varphi_*(x,y, \boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds} \\
&= \sum_{(x,y) \in B_{\text{obs},\varepsilon} } \left[\frac{(1 - \varepsilon) \int_\gamma \varphi(x,y, \boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds} + \frac{\varepsilon \int_\gamma \varphi_*(x,y, \boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds} \right] \\
&= \frac{(1 - \varepsilon) \sum_{(x,y) \in B_{\text{obs},\varepsilon}} \int_\gamma \varphi(x,y, \boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds} + \frac{\varepsilon \sum_{(x,y) \in B_{\text{obs},\varepsilon}} \int_\gamma \varphi_*(x,y, \boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds} \\
&= \frac{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds} \sum_{(x,y) \in B_{\text{obs},\varepsilon}} f_{H_0}(x,y) \\
&\quad + \frac{\varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds} \sum_{(x,y) \in B_{\text{obs},\varepsilon}} f_{H_{0,*}}(x,y),
\end{aligned}$$

em que $f_{H_0}(x,y) = \frac{\int_\gamma \varphi(x,y, \boldsymbol{\theta}) ds}{\int_\gamma g(\boldsymbol{\theta}) ds}$ e $f_{H_{0,*}}(x,y) = \frac{\int_\gamma \varphi_*(x,y, \boldsymbol{\theta}) ds}{\int_\gamma g_*(\boldsymbol{\theta}) ds}$. Assim,

$$\text{valor } P_\varepsilon = \frac{(1 - \varepsilon) \left(\int_\gamma g(\boldsymbol{\theta}) ds \right) \text{valor } P^\varepsilon + \varepsilon \left(\int_\gamma g_*(\boldsymbol{\theta}) ds \right) \text{valor } P_*^\varepsilon}{(1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds}.$$

Portanto, fazendo $\tau_1 = (1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds$ e $\tau_2 = \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds$, obtemos

$$\text{valor } P_\varepsilon = \frac{\tau_1 \text{valor } P^\varepsilon + \tau_2 \text{valor } P_*^\varepsilon}{\tau_1 + \tau_2}.$$

□

Proposição 5. $\alpha_\varepsilon(\phi_B) = \frac{\tau_1 \alpha^\varepsilon(\phi_B) + \tau_2 \alpha_*^\varepsilon(\phi_B)}{\tau_1 + \tau_2}$, em que $\alpha^\varepsilon(\phi_B) = \sum_{(x,y) \in B_{k,\varepsilon}} f_{H_0}(x,y)$ e $\alpha_*^\varepsilon(\phi_B) = \sum_{(x,y) \in B_{k,\varepsilon}} f_{H_{0,*}}(x,y)$.

Demonstração. De forma análoga à Proposição 4. □

Proposição 6. $\beta_\varepsilon(\phi_B) = \frac{\tau_1 \beta^\varepsilon(\phi_B) + \tau_2 \beta_*^\varepsilon(\phi_B)}{\tau_1 + \tau_2}$, em que $\tau_1 = (1 - \varepsilon) \int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta}$, $\tau_2 = \varepsilon \int_{\Theta_1} g_*(\boldsymbol{\theta}) d\boldsymbol{\theta}$, $\beta^\varepsilon(\phi_B) = \sum_{(x,y) \notin B_{k,\varepsilon}} f_{H_1}(x,y)$ e $\beta_*^\varepsilon(\phi_B) = \sum_{(x,y) \notin B_{k,\varepsilon}} f_{H_{1,*}}(x,y)$.

Demonstração. De forma análoga à Proposição 4. □

Observamos que a função g_ε pode ser reescrita como $g_\varepsilon = (1 - \varepsilon)g + \varepsilon g_* = g + \varepsilon h$, em que $h = g_* - g \in \mathcal{M}$. Seja \mathcal{X} um espaço amostral finito, consideramos $\mathcal{L}(\mathcal{G}, [0, 1])$, com $\mathcal{G} \subset \mathcal{M}$, o espaço dos funcionais estatísticos $\mathbb{T}_P, \mathbb{T}_\alpha$ e \mathbb{T}_β dados por $\mathbb{T}_P(g)$, em que $\mathbb{T}_P(g)$ é o valor P para cada $(x,y)_{\text{obs}} \in \mathcal{X}$, $\mathbb{T}_\alpha(g) = \alpha(\phi_B)$ e $\mathbb{T}_\beta(g) = \beta(\phi_B)$, para toda função densidade de probabilidade *a priori* elicitada $g \in \mathcal{G}$. A seguir, apresentamos a derivada de Gâteaux dos funcionais na função densidade de probabilidade *a priori* elicitada. Usamos a definição da derivada de Gâteaux apresentada em [Schwartz \(1969\)](#) com adaptações, para mais detalhes *vide* Seção 3.1.1. Nos referimos à derivada de Gâteaux no Teorema 3 no sentido da derivada de Gâteaux da identificação (3.14) apresentada na Seção 3.1.1.

Teorema 3. As derivadas de Gâteaux dos funcionais $\mathbb{T}_P, \mathbb{T}_\alpha, \mathbb{T}_\beta \in \mathcal{L}(\mathcal{G}, [0, 1])$ em $g \in \mathcal{G}$ na direção $h = g_* - g \in \mathcal{M}$ são dadas por

$$\begin{aligned} \mathbb{T}'_P(g)h &= \frac{\int_\gamma g_*(\boldsymbol{\theta}) ds}{\int_\gamma g(\boldsymbol{\theta}) ds} [\text{valor } P_*^{\text{obs}} - \mathbb{T}_P(g)], \text{ em que } \text{valor } P_*^{\text{obs}} = \sum_{(x,y) \in B_{\text{obs}}} f_{H_{0,*}}(x,y), \\ \mathbb{T}'_\alpha(g)h &= \frac{\int_\gamma g_*(\boldsymbol{\theta}) ds}{\int_\gamma g(\boldsymbol{\theta}) ds} [\alpha_*^k(\phi_B) - \mathbb{T}_\alpha(g)], \text{ em que } \alpha_*^k(\phi_B) = \sum_{(x,y) \in B_k} f_{H_{0,*}}(x,y) \\ \text{e } \mathbb{T}'_\beta(g)h &= \frac{\int_{\Theta_1} g_*(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta}} [\beta_*^k(\phi_B) - \mathbb{T}_\beta(g)], \text{ em que } \beta_*^k(\phi_B) = \sum_{(x,y) \notin B_k} f_{H_{1,*}}(x,y). \end{aligned}$$

Demonstração. Dados $(x,y)_{\text{obs}} \in \mathcal{X}$ uma amostra observada e $g \in \mathcal{G}$ uma função densidade de probabilidade *a priori* elicitada na direção $h = g_* - g \in \mathcal{M}$, a derivada de Gâteaux de \mathbb{T}_P em g na direção de h é dada por

$$\mathbb{T}'_P(g)h = \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{T}(g + \varepsilon h) - \mathbb{T}(g)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{T}(g_\varepsilon) - \mathbb{T}(g)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{\text{valor } P_\varepsilon - \text{valor } P}{\varepsilon}. \quad (3.11)$$

Pela Proposição 4, temos que

$$\begin{aligned}\mathbb{T}'_P(g)h &= \lim_{\varepsilon \rightarrow 0^+} \frac{\tau_1 \text{valor } P^\varepsilon + \tau_2 \text{valor } P_*^\varepsilon - \text{valor } P}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\tau_1 \text{valor } P^\varepsilon + \tau_2 \text{valor } P_*^\varepsilon - (\tau_1 + \tau_2) \text{valor } P}{\varepsilon(\tau_1 + \tau_2)}.\end{aligned}$$

Pelo Teorema 1, existe $\varepsilon_0 > 0$ tal que quando $\varepsilon \in (0, \varepsilon_0)$, então $B_{\text{obs}, \varepsilon} = B_{\text{obs}}$. Logo,

$$\begin{aligned}\mathbb{T}'_P(g)h &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon < \varepsilon_0}} \frac{\tau_1 \text{valor } P^\varepsilon + \tau_2 \text{valor } P_*^\varepsilon - (\tau_1 + \tau_2) \text{valor } P}{\varepsilon(\tau_1 + \tau_2)} \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon < \varepsilon_0}} \frac{\tau_1 \text{valor } P + \tau_2 \text{valor } P_* - (\tau_1 + \tau_2) \text{valor } P}{\varepsilon(\tau_1 + \tau_2)} \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon < \varepsilon_0}} \frac{\tau_2(\text{valor } P_* - \text{valor } P)}{\varepsilon(\tau_1 + \tau_2)} \\ &= \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon < \varepsilon_0}} \frac{\varepsilon \int_{\gamma} g_*(\boldsymbol{\theta}) ds (\text{valor } P_* - \text{valor } P)}{\varepsilon \left\{ \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} [g_*(\boldsymbol{\theta}) - g(\boldsymbol{\theta})] ds \right\}} \\ &= \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} (\text{valor } P_*^{\text{obs}} - \text{valor } P),\end{aligned}$$

em que $\text{valor } P_*^{\text{obs}} = \sum_{(x,y) \in B_{\text{obs}}} f_{H_{0,*}}(x,y)$. Portanto, $\mathbb{T}'_i(g)h = \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} [\text{valor } P_*^{\text{obs}} - \mathbb{T}_P(g)]$.

Para $i = \alpha$, o resultado segue da Proposição 5 e do Teorema 2 e para $i = \beta$, o resultado segue da Proposição 6 e do Teorema 2, com

$$\mathbb{T}'_{\alpha}(g)h = \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} [\alpha_*^k(\phi_B) - \mathbb{T}_{\alpha}(g)] \quad \text{e} \quad \mathbb{T}'_{\beta}(g)h = \frac{\int_{\Theta_1} g_*(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta}} [\beta_*^k(\phi_B) - \mathbb{T}_{\beta}(g)].$$

□

Outra forma de obtenção de $\mathbb{T}'_P(g)h$, referente ao Teorema 3, sem usar a Proposição 4 na sua demonstração, é apresentada no Apêndice E. A seguir, apresentamos algumas propriedades dos funcionais estatísticos.

3.1.1 Propriedades e construção dos funcionais estatísticos

A Observação 2 é útil na Observação 3.

Observação 2. Note que

$$\sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} = \sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} L(\boldsymbol{\theta}; x, y) [-g(\boldsymbol{\theta})] ds}{\int_{\gamma} -g(\boldsymbol{\theta}) ds},$$

o que significa que a expressão

$$\sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} \quad (3.12)$$

é par na variável g . Observe que a expressão (3.12) é a mesma usada para calcular o va-

lor P . Além disso, o mesmo também vale para as expressões $\sum_{(x,y) \in B_k} \frac{\int_{\gamma} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds}$ e

$$\sum_{(x,y) \notin B_k} \frac{\int_{\Theta_1} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta}},$$

as quais estão associadas a $\alpha(\phi_B)$ e $\beta(\phi_B)$, respectivamente.

Observação 3. Usando a Observação 2, note que as derivadas de Gâteaux (em termos do limite) da expressão (3.12) são ímpares na variável g .

Lema 3. Seja $(V_2, \|\cdot\|)$ um espaço vetorial normado. Se $\mathbb{T} : V_1 \rightarrow V_2$ é uma transformação linear não nula, então \mathbb{T} é uma transformação ilimitada.

Demonstração. Como \mathbb{T} é uma transformação linear não nula, então existe algum elemento $g \in V_1$ tal que $\mathbb{T}(g) \neq 0$ e portanto, $\|\mathbb{T}(g)\| \neq 0$. Sendo assim, fazendo $g_N = (N/\|\mathbb{T}(g)\|)g$, para todo $N \in \mathbb{N}$, e portanto, $\|\mathbb{T}(g_N)\| = N$. Logo, \mathbb{T} não é limitado. De fato, suponha que \mathbb{T} é limitado, então existe uma constante $M > 0$ tal que $\|\mathbb{T}(g)\| \leq M$ para todo $g \in V_1$, assim obtemos $\|\mathbb{T}(g_N)\| \leq M$ para todo $N \in \mathbb{N}$, ou seja, $N \leq M$ para todo $N \in \mathbb{N}$, o que é uma contradição por que o conjunto dos números naturais \mathbb{N} é ilimitado superiormente. Portanto, \mathbb{T} é uma transformação ilimitada. \square

Proposição 7. Para qualquer função densidade de probabilidade *a priori* elicitada $g \in \mathcal{G}$, o funcional $\mathbb{T}'_i(g)$ não é linear, para $i \in \{P, \alpha, \beta\}$.

Demonstração. De fato, dada $g \in \mathcal{G}$ uma função densidade de probabilidade *a priori* elicitada, temos que $|\mathbb{T}'_i(g)h| \leq \left(\int_{\gamma} g(\boldsymbol{\theta}) ds\right)^{-1}$, para $i \in \{P, \alpha\}$, e $|\mathbb{T}'_{\beta}(g)h| \leq \left(\int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta}\right)^{-1}$. Portanto, $\mathbb{T}'_i(g)$ é um funcional limitado. Então, pelo Lema 3, como $\mathbb{T}'_i(g)$ não é um funcional nulo, logo $\mathbb{T}'_i(g)$ é um funcional não linear em h , para $i \in \{P, \alpha, \beta\}$. \square

Na Proposição 8, observamos que os funcionais estatísticos são transformações não lineares, *i.e.*, são quantidades de interesse *a priori* não lineares.

Proposição 8. O funcional \mathbb{T}_i não é linear, para $i \in \{P, \alpha, \beta\}$.

Demonstração. Supomos, por absurdo, que o funcional estatístico \mathbb{T}_P é linear. Então, dados $(x, y)_{\text{obs}} \in \mathcal{X}$ uma amostra observada e $g \in \mathcal{G}$ uma função densidade de probabilidade *a priori* elicitada na direção $h = g_* - g \in \mathcal{M}$, a derivada de Gâteaux de \mathbb{T}_P em g na direção de h é dada por

$$\mathbb{T}'_P(g)h = \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{T}_P(g + \varepsilon h) - \mathbb{T}_P(g)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{T}_P(g) + \varepsilon \mathbb{T}_P(h) - \mathbb{T}_P(g)}{\varepsilon} = \mathbb{T}_P(h),$$

ou seja, $\mathbb{T}'_P(g)h = \mathbb{T}_P(h)$, para todo $h \in \mathcal{M}$. Logo, $\mathbb{T}'_P(g)$ é linear, o que é um absurdo porque $\mathbb{T}'_P(g)$ não é linear, pela Proposição 7. Portanto, \mathbb{T}_P não é linear. Analogamente mostramos para $i \in \{\alpha, \beta\}$. \square

Carota (1996) mencionou que frequentemente derivadas funcionais como a de Gâteaux e a de Fréchet de quantidades *a posteriori* são usadas para medir suas taxas de mudança e mencionou duas formas comuns na literatura sobre quais conjuntos os funcionais estatísticos são definidos. Uma delas é primeiro estender artificialmente o funcional estatístico ao espaço de todas as medidas com sinal, para então aplicar os conceitos de derivadas funcionais para quantificar a sensibilidade local dessas quantidades *a posteriori*. Neste sentido, Carota (1996) citou alguns autores como Diaconis e Freedman (1986), Srinivasan e Trusczyńska (1990), Ruggeri e Wasserman (1993). Outra solução alternativa mencionada por Carota (1996) foi sugerida em Huber (1981), Clarke (1983), Basu (1996), que generalizaram o conceito dessas derivadas funcionais para englobar o caso no qual o funcional estatístico é definido somente sobre um conjunto convexo. Notamos que Basu (1996) se baseiou em Huber (1981) para poder definir o funcional estatístico dessa forma, usando as derivadas de Fréchet e de Gâteaux. Basu (1996) também estabeleceu a diferenciação de quantidades *a posteriori* não lineares por meio da derivada de Gâteaux, no entanto, as definições dessas derivadas funcionais apresentadas em Basu (1996) requerem a propriedade de linearidade do funcional derivada.

No nosso trabalho, como o funcional $\mathbb{T}'_i(g)$ não é linear (Proposição 7), não podemos usar a derivada de Fréchet. Em Schwartz (1969), na definição da derivada de Gâteaux, o funcional é definido em um conjunto aberto U de um espaço vetorial normado a fim de que o limite faça sentido porque não teríamos como aplicar \mathbb{T}_i em g_ε , com $g_\varepsilon \notin U$. Neste trabalho, adaptamos a definição apresentada em Schwartz (1969, Cap. 1), retirando a suposição de que o funcional estatístico precise estar definido sobre um conjunto aberto, pois g_ε pertence ao domínio do funcional \mathbb{T}_i e portanto, faz sentido calcular o limite em (3.13).

Definição 1. Sejam $(V_1, \|\cdot\|)$ e $(V_2, \|\cdot\|)$ espaços vetoriais normados e $\mathbb{T} : E_1 \subset V_1 \rightarrow E_2 \subset V_2$ um funcional. Se $g \in E_1$ e $h \in V_1$ são tais que $g + \varepsilon h \in E_1$, para todo $\varepsilon > 0$, então \mathbb{T} possui derivada de Gâteaux em g na direção h se existe o limite

$$\left. \frac{d\mathbb{T}(g + \varepsilon h)}{d\varepsilon} \right|_{\varepsilon=0} = \mathbb{T}'(g)h. \quad (3.13)$$

Se existe o limite (3.13) para todo $h \in V_1$, então dizemos que \mathbb{T} possui derivada de Gâteaux em $g \in E_1$.

Para poder usar a Definição 1, construímos um novo funcional estatístico $\tilde{\mathbb{T}}_i$ aplicado à medida de probabilidade *a priori* associando-o ao funcional estatístico \mathbb{T}_i aplicado à função densidade de probabilidade *a priori*, para $i \in \{P, \alpha, \beta\}$. Com base em Diaconis e Freedman (1986), consideramos a ideia de usar o conjunto das medidas com sinal finitas. Dessa forma, dados $G \in \tilde{\mathcal{G}}$ e $G_* \in \tilde{\mathcal{G}}_* \subset \tilde{\mathcal{G}}$ medidas de probabilidade *a priori* sobre o espaço mensurável $(\Theta, \mathcal{B}(\Theta))$ e suas respectivas funções densidade de probabilidade associadas $g \in \mathcal{G}$ e $g_* \in \mathcal{G}_* \subset \mathcal{G}$, para cada função $H = G_* - G \in \tilde{\mathcal{M}}$, em que $\tilde{\mathcal{M}}$ é o conjunto das medidas com sinal finitas, obtemos $h = g_* - g \in \mathcal{M}$. Pelo Exercício 13 (FERNANDEZ, 2015, Cap. 6), se G e G_* são medidas de probabilidade *a priori*, então a métrica é dada por $\tilde{d}(G_*, G) = \|G_* - G\|_* = \sup_{W \in \mathcal{B}(\Theta)} |G_*(W) - G(W)|$. Além disso, $\tilde{\mathcal{M}}$ é um espaço vetorial, pelo Exercício 7 (FERNANDEZ, 2015, Cap. 6). Sendo assim, definimos o funcional estatístico dado por

$$\begin{aligned} \tilde{\mathbb{T}}_i : \tilde{\mathcal{G}} \subset \tilde{\mathcal{M}} &\rightarrow [0, 1] \subset \mathbb{R} \\ G &\mapsto \tilde{\mathbb{T}}_i(G) = \mathbb{T}_i(g), \end{aligned} \quad (3.14)$$

em que $\tilde{\mathbb{T}}_i$ identifica cada medida de probabilidade *a priori* com sua respectiva função densidade de probabilidade, para $i \in \{P, \alpha, \beta\}$. Portanto, podemos usar a Definição 1 pelo fato de $(\tilde{\mathcal{M}}, \|\cdot\|_*)$ ser um espaço vetorial normado e assumindo que $G + \varepsilon H \in \tilde{\mathcal{G}}$, com $G \in \tilde{\mathcal{G}}$ e $H \in \tilde{\mathcal{M}}$, para todo $\varepsilon > 0$.

Note que na Definição 1 não há a necessidade de que $\tilde{\mathbb{T}}'_i(G)$ seja linear, ou melhor, que $\mathbb{T}'_i(g)$ seja linear. Também observamos que

$$\begin{aligned} \tilde{\mathbb{T}}'_i(G)H &= \left. \frac{d\tilde{\mathbb{T}}_i(G + \varepsilon H)}{d\varepsilon} \right|_{\varepsilon=0^+} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\tilde{\mathbb{T}}(G + \varepsilon H) - \tilde{\mathbb{T}}(G)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{T}(g + \varepsilon h) - \mathbb{T}(g)}{\varepsilon} \\ &= \left. \frac{d\mathbb{T}_i(g + \varepsilon h)}{d\varepsilon} \right|_{\varepsilon=0^+} \\ &= \mathbb{T}'_i(g)h, \end{aligned} \quad (3.15)$$

e assim, calculamos a derivada de Gâteaux como em (3.11), para $i \in \{P, \alpha, \beta\}$. Observamos que as derivadas de Gâteaux nesta área de robustez bayesiana são calculadas como derivadas laterais à direita, *vide* Ruggeri e Wasserman (1993), por exemplo.

Fazendo $g_\varepsilon = \kappa(\varepsilon)$, construímos um caminho $\kappa : [0, 1] \rightarrow \mathcal{G}$ entre as funções densidades de probabilidade *a priori* g e g_* no espaço das funções densidade de probabilidade *a priori* \mathcal{G} , dado por $\varepsilon \mapsto \kappa(\varepsilon) = g + \varepsilon h$, com $\kappa(0) = g$ e $\frac{d\kappa}{d\varepsilon}(0) = h$. A Figura 7 apresenta um diagrama

retratando $\mathbb{T}_i(g_\epsilon)$ como a função composta $\mathbb{T}_i \circ \kappa$ e ilustrando $\mathbb{T}'_i(g)$ conforme (3.15), para $i \in \{P, \alpha, \beta\}$.

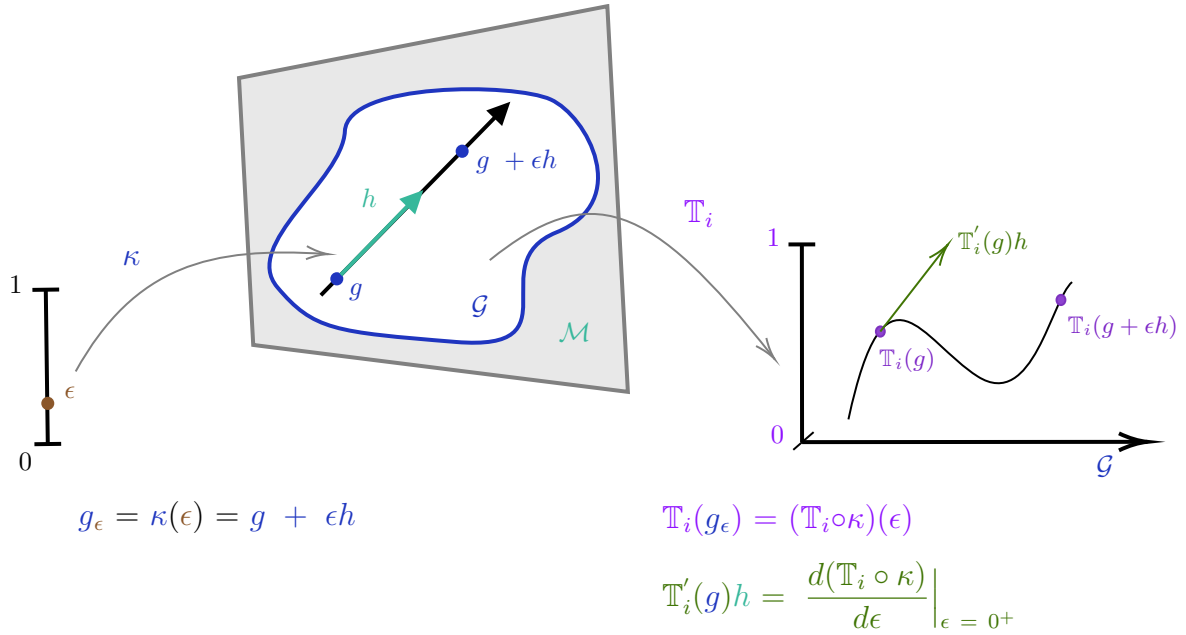


Figura 7 – Diagrama adaptado de Lima (2016).

Como medida de sensibilidade de quantidades *a posteriori*, Diaconis e Freedman (1986) usaram a norma da derivada do funcional estatístico, a qual é a norma espectral pelo fato de o funcional derivada ser linear. Neste trabalho, sugerimos o índice de sensibilidade como na definição a seguir, o qual usamos na análise de dados para avaliar quantidades não lineares *a priori*.

Definição 2. Nas condições anteriores, seja \mathcal{G}_* finito, o índice de sensibilidade é dado por $IS_i(g) = \max_{g_* \in \mathcal{G}_*} |\mathbb{T}'_i(g)h|$, para $i \in \{P, \alpha, \beta\}$.

Por fim, discorremos sobre as relações e propriedades entre as classes definidas \mathcal{G} e $\Gamma_{\epsilon, g}$. Além disso, dada uma função densidade de probabilidade *a priori* $g \in \mathcal{G}_{\epsilon, g}$, para cada $g_* \in \Gamma_{\epsilon, g}$, o segmento que une g a g_* é dado por $\Gamma_{g, g_*} = \{g_\epsilon \in \mathcal{G}; g_\epsilon = (1 - \epsilon)g + \epsilon g_*, 0 \leq \epsilon \leq 1\}$ e dessa forma, observamos que a classe ϵ -contaminada é diferente do segmento de funções densidade de probabilidade *a priori*.

Proposição 9. Valem as seguintes inclusões de conjuntos: (i) $\Gamma_{\epsilon, g} \subset \mathcal{G}$ e (ii) $\Gamma_{g, g_*} \subset \mathcal{G}$.

Demonstração. (i) De fato, dado $g_\epsilon \in \Gamma_{\epsilon, g}$, temos que $g_\epsilon = (1 - \epsilon)g + \epsilon g_*$, com $g_* \in \mathcal{G}_* \subset \mathcal{G}$. Como $g_* \in \mathcal{G}$ é uma função densidade de probabilidade *a priori*, portanto $g, g_* \in \mathcal{G}$, pois g e g_* são funções densidade de probabilidade *a priori*. Então, g_ϵ é uma função densidade de probabilidade *a priori*, o que implica que $g_\epsilon \in \mathcal{G}$. Logo, $\Gamma_{\epsilon, g} \subset \mathcal{G}$.

(ii) Dado $g_* \in \mathcal{G}_* \subset \mathcal{G}$, temos que $g_* \in \mathcal{G}$ e portanto, g_* é uma função densidade de probabilidade *a priori*, o que implica que g_ε é uma função densidade de probabilidade *a priori*, para todo $\varepsilon \in [0, 1]$. Logo, $g_\varepsilon \in \mathcal{G}$, para todo $\varepsilon \in [0, 1]$. Portanto, o segmento de funções $\Gamma_{g, g_*} \subset \mathcal{G}$. \square

Ademais, a Proposição 10 trata sobre a convexidade da classe $\Gamma_{\varepsilon, g}$.

Proposição 10. Se \mathcal{G}_* é convexo, a classe ε -contaminada $\Gamma_{\varepsilon, g}$ é convexa.

Demonstração. Uma vez elicitada a função densidade de probabilidade *a priori* $g \in \mathcal{G}$ e fixado $\varepsilon \in (0, 1)$, dadas as funções densidade de probabilidade *a priori* $g_{\varepsilon_1}, g_{\varepsilon_2} \in \Gamma_{\varepsilon, g}$, então $g_{\varepsilon_1} = g + \varepsilon(g_{*,1} - g)$ e $g_{\varepsilon_2} = g + \varepsilon(g_{*,2} - g)$, com $g_{*,1}, g_{*,2} \in \mathcal{G}_*$. Tomando $\zeta \in (0, 1)$, temos

$$\begin{aligned} (1 - \zeta)g_{\varepsilon_1} + \zeta g_{\varepsilon_2} &= g_{\varepsilon_1} - \zeta g_{\varepsilon_1} + \zeta g_{\varepsilon_2} \\ &= g + \varepsilon(g_{*,1} - g) + \zeta[g + \varepsilon(g_{*,1} - g)] + \zeta[g + \varepsilon(g_{*,2} - g)] \\ &= g + \varepsilon g_{*,1} - \varepsilon g + \zeta g + \zeta \varepsilon g_{*,1} - \zeta \varepsilon g + \zeta g + \zeta \varepsilon g_{*,2} - \zeta \varepsilon g \\ &= (1 - \varepsilon)g + \varepsilon g_{*,3}, \end{aligned}$$

em que $g_{*,3} = (1 - \zeta)g_{*,1} + \zeta g_{*,2}$. Como \mathcal{G}_* é convexo, segue que $g_{*,3} \in \mathcal{G}_*$, o que implica que $(1 - \zeta)g_{\varepsilon_1} + \zeta g_{\varepsilon_2} \in \Gamma_{\varepsilon, g}$, com $g_{*,3} \in \mathcal{G}_*$. Logo, $\Gamma_{\varepsilon, g}$ é uma classe convexa. \square

3.2 Análise de dados

Nesta seção apresentamos uma análise de sensibilidade utilizando os dados apresentados na Seção 2.5. Para a análise de dados, usamos a classe ε -contaminada $\Gamma_{\varepsilon, g}$ com funções densidades de probabilidade *a priori* beta, reduzindo o número de possibilidades de escolhas das funções densidades de probabilidade *a priori* g e g_* , o que é possível de acordo com o item (i) da Proposição 9. Consideramos as subclasses $\mathcal{G}_i \subset \Gamma_{\varepsilon, g}$, com $i \in \{J, BL, 24, 42\}$, dadas por

- (i) $\mathcal{G}_J = \{g \text{ é a função densidade de probabilidade } a \text{ priori de Jeffreys}\}$,
- (ii) $\mathcal{G}_{BL} = \{g \text{ é a função densidade de probabilidade } a \text{ priori de Bayes-Laplace}\}$,
- (iii) $\mathcal{G}_{24} = \{g \text{ é a função densidade de probabilidade } a \text{ priori com } p, q \sim B(2, 4)\}$
- e (iv) $\mathcal{G}_{42} = \{g \text{ é a função densidade de probabilidade } a \text{ priori com } p, q \sim B(4, 2)\}$.

Na Tabela 11 os cenários são identificados por meio dos pares ordenados $(g, g_*) \in \mathcal{G}_i \times \mathcal{G}_j$, com $i \neq j$, em que g é sempre a função densidade de probabilidade *a priori* elicitada, para $i, j \in \{J, BL, 24, 42\}$. Na Tabela 11, para $\delta = 0, 1$, elicitada a função densidade de probabilidade *a priori* de Jeffreys, os resultados indicam que os funcionais \mathbb{T}_P e \mathbb{T}_β variam mais rápido na direção de $h = g_* - g$, com $g_* \in \mathcal{G}_{24}$, o que indica maior sensibilidade do valor P e do poder referente à escolha da função densidade de probabilidade *a priori*, ou seja, pequenas perturbações na função densidade de probabilidade *a priori* de Jeffreys podem causar maiores perturbações no valor P e no poder. Com o aumento da margem, os valores $|\mathbb{T}'_P(g)h|$, $|\mathbb{T}'_\alpha(g)h|$ e $|\mathbb{T}'_\beta(g)h|$ diminuem, o que indica menor sensibilidade dessas quantidades de interesse referentes

Tabela 11 – Valores das derivadas de Gâteaux dos funcionais \mathbb{T}_P , \mathbb{T}_α e \mathbb{T}_β em g na direção h .

(g, g_*)	$\delta = 0,1$			$\delta = 0,15$			$\delta = 0,2$		
	$\mathbb{T}'_P(g)h$	$\mathbb{T}'_\alpha(g)h$	$\mathbb{T}'_\beta(g)h$	$\mathbb{T}'_P(g)h$	$\mathbb{T}'_\alpha(g)h$	$\mathbb{T}'_\beta(g)h$	$\mathbb{T}'_P(g)h$	$\mathbb{T}'_\alpha(g)h$	$\mathbb{T}'_\beta(g)h$
$\mathcal{G}_{BL} \times \mathcal{G}_J$	-0,0519	-0,0421	-0,0807	-0,0183	-0,0162	-0,0841	-0,0037	-0,0040	-0,0714
$\mathcal{G}_{BL} \times \mathcal{G}_{24}$	0,0505	0,0333	0,0249	0,0185	0,0210	0,0134	0,0066	0,0069	0,0070
$\mathcal{G}_{BL} \times \mathcal{G}_{22}$	0,0594	0,0455	0,0251	0,0254	0,0227	0,0369	0,0086	0,0119	0,0345
$\mathcal{G}_J \times \mathcal{G}_{BL}$	0,0855	0,0350	0,1263	0,0196	0,0182	0,1095	0,0101	0,0114	0,0934
$\mathcal{G}_J \times \mathcal{G}_{24}$	0,2006	0,0761	0,2477	0,0505	0,0405	0,1990	0,0232	0,0221	0,1537
$\mathcal{G}_J \times \mathcal{G}_{22}$	0,1828	0,0752	0,2006	0,0534	0,0449	0,1784	0,0229	0,0259	0,1541
$\mathcal{G}_{24} \times \mathcal{G}_{BL}$	-0,0082	-0,0053	-0,0027	-0,0033	-0,0033	-0,0012	-0,0013	-0,0001	0,0006
$\mathcal{G}_{24} \times \mathcal{G}_J$	-0,0369	-0,0291	-0,0533	-0,0154	-0,0131	-0,0536	-0,0053	-0,0066	-0,0435
$\mathcal{G}_{24} \times \mathcal{G}_{22}$	0,0244	0,0200	0,0106	0,0126	0,0090	0,0202	0,0041	0,0071	0,0212
$\mathcal{G}_{22} \times \mathcal{G}_{BL}$	-0,0280	-0,0226	-0,0126	-0,0092	-0,0091	-0,0224	-0,0034	-0,0054	-0,0203
$\mathcal{G}_{22} \times \mathcal{G}_J$	-0,0692	-0,0522	-0,0775	-0,0190	-0,0182	-0,0917	-0,0070	-0,0119	-0,0733
$\mathcal{G}_{22} \times \mathcal{G}_{24}$	-0,0047	-0,0086	-0,0074	-0,0020	0,0018	-0,0283	-0,0001	-0,0020	-0,0294

à escolha da distribuição *a priori*. Comparando os funcionais estatísticos \mathbb{T}_i , independentemente do valor da margem, os resultados da Tabela 11 também indicam que o funcional \mathbb{T}_β , ou melhor, o poder é mais sensível à escolha da distribuição *a priori* do que os funcionais \mathbb{T}_P e \mathbb{T}_α , e consequentemente, o valor P e o nível de significância adaptativo ótimo $\alpha(\phi_B^*)$ são menos sensíveis com respeito à escolha dessa distribuição *a priori*. Ademais, suponhamos que g , a função densidade de probabilidade *a priori* elicitada, é a de Jeffreys na direção $h = g_* - g$, em que g_* é a função densidade de probabilidade *a priori* de Bayes-Laplace, e com $\mathbb{T}'_P(g)h > 0$. Por outro lado, suponhamos que g , a função densidade de probabilidade *a priori* elicitada, é a de Bayes-Laplace na direção $h = g_* - g$, em que g_* é a função densidade de probabilidade *a priori* de Jeffreys, com $\mathbb{T}'_P(g_*)(-h) < 0$. Na prática, os resultados sugerem a existência da relação $\mathbb{T}'_i(g)h \propto \mathbb{T}'_i(g_*)(-h)$, com constante de proporcionalidade negativa, para $i \in \{P, \alpha, \beta\}$. Na Tabela

Tabela 12 – Valores dos índices de sensibilidade com base na Tabela 11.

g	$\delta = 0,1$			$\delta = 0,15$			$\delta = 0,2$		
	$IS_P(g)$	$IS_\alpha(g)$	$IS_\beta(g)$	$IS_P(g)$	$IS_\alpha(g)$	$IS_\beta(g)$	$IS_P(g)$	$IS_\alpha(g)$	$IS_\beta(g)$
\mathcal{G}_{BL}	0,0594	0,0455	0,0807	0,0254	0,0227	0,0841	0,0086	0,0119	0,0345
\mathcal{G}_J	0,2006	0,0761	0,2477	0,0534	0,0449	0,1990	0,0232	0,0259	0,1541
\mathcal{G}_{24}	0,0369	0,0291	0,0533	0,0154	0,0131	0,0536	0,0053	0,0071	0,0435
\mathcal{G}_{22}	0,0692	0,0522	0,0775	0,0190	0,0182	0,0917	0,0070	0,0119	0,0733

12, os maiores valores dos índices de sensibilidade, com base nas direções escolhidas na Tabela 11, ocorrem quando g é a função densidade de probabilidade *a priori* de Jeffreys. Além disso, note na Tabela 10 que os menores valores de valor P , $\alpha(\phi_B^*)$ e $\beta(\phi_B^*)$ ocorrem quando a distribuição *a priori* é a de Jeffreys. Dessa forma, dada a margem δ , os resultados sugerem a existência de uma relação inversamente proporcional entre as quantidades $IS_i(g)$ e mínimo de $\mathbb{T}_i(g)$ com respeito as distribuições *a priori*, para $i \in \{P, \alpha, \beta\}$. Na Tabela 13 consideramos as misturas com

Tabela 13 – Valores P_ε , $\alpha_\varepsilon(\phi_B)$ e $\beta_\varepsilon(\phi_B)$ considerando a função densidade de probabilidade *a priori* g_ε .

(g, g_*, g_ε)	$\delta = 0, 1$			$\delta = 0, 15$			$\delta = 0, 2$		
	valor P_ε	$\alpha_\varepsilon(\phi_B^*)$	$\beta_\varepsilon(\phi_B^*)$	valor P_ε	$\alpha_\varepsilon(\phi_B^*)$	$\beta_\varepsilon(\phi_B^*)$	valor P_ε	$\alpha_\varepsilon(\phi_B^*)$	$\beta_\varepsilon(\phi_B^*)$
$\mathcal{G}_{BL} \times \mathcal{G}_J \times \mathcal{E}_{01}$	0,5616	0,3368	0,3493	0,3169	0,2190	0,3310	0,1203	0,1547	0,3008
$\mathcal{G}_{BL} \times \mathcal{G}_{24} \times \mathcal{E}_{01}$	0,5628	0,3375	0,3504	0,3176	0,2211	0,3302	0,1202	0,1549	0,3015
$\mathcal{G}_{BL} \times \mathcal{G}_{22} \times \mathcal{E}_{01}$	0,5641	0,3376	0,3504	0,3178	0,2212	0,3304	0,1207	0,1550	0,3018
$\mathcal{G}_{BL} \times \mathcal{G}_J \times \mathcal{E}_1$	0,5560	0,3267	0,3476	0,3146	0,2147	0,3262	0,1190	0,1521	0,2965
$\mathcal{G}_{BL} \times \mathcal{G}_{24} \times \mathcal{E}_1$	0,5707	0,3416	0,3498	0,3198	0,2226	0,3311	0,1232	0,1566	0,3003
$\mathcal{G}_{BL} \times \mathcal{G}_{22} \times \mathcal{E}_1$	0,5772	0,3442	0,3492	0,3246	0,2247	0,3317	0,1216	0,1582	0,3023
$\mathcal{G}_J \times \mathcal{G}_{BL} \times \mathcal{E}_{01}$	0,5013	0,2257	0,3298	0,3006	0,1473	0,2975	0,1187	0,1138	0,2583
$\mathcal{G}_J \times \mathcal{G}_{24} \times \mathcal{E}_{01}$	0,5023	0,2273	0,3291	0,2980	0,1482	0,2979	0,1188	0,1142	0,2585
$\mathcal{G}_J \times \mathcal{G}_{22} \times \mathcal{E}_{01}$	0,5028	0,2263	0,3303	0,3010	0,1479	0,2979	0,1189	0,1141	0,2589
$\mathcal{G}_J \times \mathcal{G}_{BL} \times \mathcal{E}_1$	0,5165	0,2340	0,3341	0,3051	0,1561	0,2995	0,1208	0,1175	0,2631
$\mathcal{G}_J \times \mathcal{G}_{24} \times \mathcal{E}_1$	0,5234	0,2455	0,3346	0,3037	0,1613	0,3028	0,1209	0,1207	0,2649
$\mathcal{G}_J \times \mathcal{G}_{22} \times \mathcal{E}_1$	0,5301	0,2404	0,3373	0,3096	0,1610	0,3031	0,1207	0,1205	0,2668
$\mathcal{G}_{24} \times \mathcal{G}_{BL} \times \mathcal{E}_{01}$	0,5804	0,3719	0,3471	0,3257	0,2365	0,3364	0,1234	0,1650	0,3001
$\mathcal{G}_{24} \times \mathcal{G}_J \times \mathcal{E}_{01}$	0,5793	0,3707	0,3468	0,3254	0,2364	0,3359	0,1226	0,1650	0,2997
$\mathcal{G}_{24} \times \mathcal{G}_{22} \times \mathcal{E}_{01}$	0,5815	0,3721	0,3472	0,3259	0,2368	0,3358	0,1243	0,1651	0,3003
$\mathcal{G}_{24} \times \mathcal{G}_{BL} \times \mathcal{E}_1$	0,5876	0,3691	0,3474	0,3295	0,2347	0,3365	0,1265	0,1647	0,3003
$\mathcal{G}_{24} \times \mathcal{G}_J \times \mathcal{E}_1$	0,5778	0,3642	0,3444	0,3238	0,2324	0,3320	0,1242	0,1623	0,2976
$\mathcal{G}_{24} \times \mathcal{G}_{22} \times \mathcal{E}_1$	0,6003	0,3720	0,3478	0,3364	0,2377	0,3367	0,1282	0,1659	0,3019
$\mathcal{G}_{22} \times \mathcal{G}_{BL} \times \mathcal{E}_{01}$	0,6071	0,3866	0,3486	0,3409	0,2511	0,3416	0,1263	0,1783	0,3112
$\mathcal{G}_{22} \times \mathcal{G}_J \times \mathcal{E}_{01}$	0,6065	0,3858	0,3484	0,3404	0,2502	0,3418	0,1261	0,1778	0,3107
$\mathcal{G}_{22} \times \mathcal{G}_{24} \times \mathcal{E}_{01}$	0,6070	0,3865	0,3488	0,3410	0,2509	0,3418	0,1264	0,1782	0,3111
$\mathcal{G}_{22} \times \mathcal{G}_{BL} \times \mathcal{E}_1$	0,6070	0,3827	0,3489	0,3402	0,2488	0,3408	0,1257	0,1761	0,3110
$\mathcal{G}_{22} \times \mathcal{G}_J \times \mathcal{E}_1$	0,5993	0,3756	0,3459	0,3379	0,2427	0,3392	0,1252	0,1735	0,3073
$\mathcal{G}_{22} \times \mathcal{G}_{24} \times \mathcal{E}_1$	0,6077	0,3849	0,3480	0,3396	0,2476	0,3424	0,1276	0,1762	0,3100

$\varepsilon = 0, 01$ ($\mathcal{E}_{01} = \{g_\varepsilon = 0, 99g + 0, 01g_*\}$) e com $\varepsilon = 0, 1$ ($\mathcal{E}_1 = \{g_\varepsilon = 0, 9g + 0, 1g_*\}$), em que $g \in \mathcal{G}_i$ e $g_* \in \mathcal{G}_j$, para $i, j \in \{J, BL, 24, 42\}$, com $i \neq j$. As funções densidades de probabilidade *a priori* e a mistura dessas funções densidades de probabilidade *a priori* são indicadas pelos pontos ordenados $(g, g_*, g_\varepsilon) \in \mathcal{G}_i \times \mathcal{G}_j \times \mathcal{E}_{01}$ ou $(g, g_*, g_\varepsilon) \in \mathcal{G}_i \times \mathcal{G}_j \times \mathcal{E}_1$. Na Tabela 13, quando $\delta = 0, 1$, os resultados mostram que quando aumentamos o valor de $\varepsilon = 0, 01$ para $\varepsilon = 0, 1$, os valores P_ε , $\alpha_\varepsilon(\phi_B^*)$ e $\beta_\varepsilon(\phi_B^*)$ tendem a aumentar, o que não ocorre quando $\delta = 0, 2$. Ao compararmos os resultados da Tabela 13 com a Tabela 10, quando $\delta = 0, 1$, os resultados mostram que o valor P_ε é maior do que o valor P em todos os casos e observamos que a diferença absoluta entre os valores P e P_ε tendem a ser maiores do que quando comparadas as diferenças absoluta entre $\alpha(\phi_B^*)$ e $\alpha_\varepsilon(\phi_B^*)$, $\beta(\phi_B^*)$ e $\beta_\varepsilon(\phi_B^*)$. Quando $\delta = 0, 2$ essas diferenças absolutas referentes ao valores P , $\alpha(\phi_B^*)$ e $\beta(\phi_B^*)$ tendem a se menores, mesmo para os casos em que $\varepsilon = 0, 1$.

3.3 Considerações finais

Neste capítulo propomos um índice de sensibilidade baseado na derivada de Gâteaux dos funcionais estatísticos aplicados à função densidade de probabilidade *a priori* elicitada para avaliar a sensibilidade local do valor P , do nível de significância adaptativo e do poder. Para

valores menores da margem δ , os resultados da análise de dados indicam que o valor P e o poder tendem a ser mais sensíveis à escolha da distribuição *a priori* do que nível de significância adaptativo. Para valores maiores da margem δ , os valores P , o nível de significância adaptativo e o poder tendem a ser menos sensíveis à escolha da distribuição *a priori*. Estudamos algumas propriedades dos funcionais estatísticos definidos e mostramos que essas quantidades de interesse *a priori* e seus respectivos funcionais derivada de Gâteaux aplicados à função densidade de probabilidade *a priori* elicitada são não lineares. Ademais, na análise de dados os índices de sensibilidade propostos refletem de forma resumida os resultados das quantidades de interesse *a priori* sob perturbações infinitésimas na função densidade de probabilidade *a priori* elicitada.

Na análise bayesiana robusta, gostaríamos de estudar normas, se existirem, para a derivada de Gâteaux não linear visando uma análise do comportamento assintótico das normas com relação ao aumento do tamanho amostral, como em [Gustafson \(1996\)](#). Também gostaríamos de implementar um pacote em linguagem R para análise de robustez, além de estender esse estudo de robustez com respeito à escolha da distribuição *a priori* para outros modelos probabilísticos.

REFERÊNCIAS

AL-LABADI, L.; ASL, F. F.; WANG, C. Measuring Bayesian robustness using Rényi divergence. **Stats**, v. 4, p. 251–268, 2021. Citado na página 54.

ARBORETTI, R.; CARROZZO, E.; PESARIN, F.; SALMASO, L. Testing for equivalence: an intersection-union permutation solution. **Statistics in Biopharmaceutical Research**, v. 10, p. 130–138, 2018. Citado nas páginas 24 e 28.

ARBORETTI, R.; PESARIN, F.; SALMASO, L. A unified approach to permutation testing for equivalence. **Statistical Methods & Applications**, p. 1–20, 2020. Citado na página 23.

BARKER, L.; ROLKA, H.; ROLKA, D.; BROWN, C. Equivalence testing for binomial random variables: Which test to use? **The American Statistician**, v. 55, p. 279–287, 2001. Citado nas páginas 27, 29, 38 e 50.

BASU, S. Local sensitivity, functional derivatives and nonlinear posterior quantities. **Statistics and Decisions**, Oldenbourg Wissenschaft Verlag, München, v. 14, p. 405–418, 1996. Citado na página 63.

BASU, S.; JAMMALAMADAKA, S. R.; LIU, W. Local posterior robustness with parametric priors: maximum and average sensitivity. In: HEIDBREDE, G. R. (Ed.). **Maximum Entropy and Bayesian Methods**. Dordrecht: Kluwer Academic Publishers, 1996. p. 97–106. Citado na página 53.

BERGER, J.; MORENO, E. Bayesian robustness in bidimensional models: prior independence. **Journal of Statistical Planning and Inference**, v. 40, p. 161–176, 1994. Citado na página 54.

BERGER, J. O.; DELAMPADY, M. Testing precise hypotheses. **Statistical Science**, Institute of Mathematical Statistics, v. 2, p. 317–352, 1987. Citado na página 79.

BERGER, J. O.; LISEO, B.; WOLPERT, R. L. Integrated likelihood methods for eliminating nuisance parameters. **Statistical Science**, v. 14, p. 1–28, 1999. Citado na página 34.

BERGER, J. O.; MORENO, E.; PERICCHI, L. R.; BAYARRI, M. J.; BERNARDO, J. M.; CANO, J. A.; HORRA, J. De la; MARTÍN, J.; RÍOS-INSÚA, D.; BETRÒ, B.; DASGUPTA, A.; GUSTAFSON, P.; WASSERMAN, L.; KADANE, J. B.; SRINIVASAN, C.; LAVINE, M.; O'HAGAN, A.; POLASEK, W.; ROBERT, C. P.; GOUTIS, C.; RUGGERI, F.; SALINETTI, G.; SIVAGANESAN, S. An overview of robust Bayesian analysis. **Test**, v. 3, p. 5–124, 1994. Citado na página 53.

BERGER, J. O.; PERICCHI, L. R. Objective Bayesian methods for model selection: Introduction and comparison. **Lecture Notes-Monograph Series**, v. 38, p. 135–207, 2001. Citado nas páginas 33 e 34.

BERGER, R. L.; HSU, J. C. *et al.* Bioequivalence trials, intersection-union tests and equivalence confidence sets. **Statistical Science**, v. 11, p. 283–319, 1996. Citado na página 27.

- BERRY, D. A. Bayesian clinical trials. **Nature Reviews Drug Discovery**, v. 5, p. 27–36, 2006. Citado na página 24.
- BERRY, S. M.; CARLIN, B. P.; LEE, J. J.; MULLER, P. **Bayesian Adaptive Methods for Clinical Trials**. Boca Raton: Chapman & Hall/CRC, 2010. Citado nas páginas 24 e 79.
- BUCHEM, F. L. van; KNOTTNERUS, J. A.; SCHRIJNEMAEKERS, V. J. J.; PEETERS, M. F. Primary-care-based randomised placebo-controlled trial of antibiotic treatment in acute maxillary sinusitis. **The Lancet**, Elsevier, v. 349, p. 683–687, 1997. Citado nas páginas 49 e 50.
- CARMO, M. P. D. **Geometria Diferencial de Curvas e Superfícies**. 2^a. ed. Rio de Janeiro: Sociedade Brasileira de Matemática, 2005. Citado na página 35.
- CAROTA, C. Local robustness of Bayes factors for nonparametric alternatives. **Lecture Notes-Monograph Series**, p. 283–291, 1996. Citado na página 63.
- CELLA, L. O. G. **Regressão Ordinal Bayesiana**. Dissertação (Mestrado) — Universidade de Brasília, 2013. Citado na página 78.
- CHOW, S.-C.; SHAO, J.; WANG, H.; LOKHNYGINA, Y. **Sample Size Calculations in Clinical Research**. Boca Raton: Chapman & Hall/CRC, 2017. Citado na página 28.
- CLARKE, B. R. Uniqueness and Fréchet differentiability of functional solutions to maximum likelihood type equations. **The Annals of Statistics**, p. 1196–1205, 1983. Citado na página 63.
- _____. A review of differentiability in relation to robustness with application to seismic data analysis. **Proceedings of the Indian National Science Academy**, Indian National Science Academy, v. 66, p. 467–482, 2000. Citado na página 53.
- CONSONNI, G.; GUTIÉRREZ-PEÑA, E.; VERONESE, P. Compatible priors for Bayesian model comparison with an application to the Hardy–Weinberg equilibrium model. **Test**, v. 17, p. 585–605, 2008. Citado na página 77.
- DAWID, A. P.; LAURITZEN, S. L. Compatible prior distributions. In: GEORGE, E. I. (Ed.). **Bayesian Methods with Applications to Science, Policy and Official Statistics, Selected Papers from ISBA 2000: The Sixth World Meeting of the International Society for Bayesian Analysis**. Luxembourg: Eurostat, 2001. p. 109–118. Citado na página 78.
- DIACONIS, P.; FREEDMAN, D. On the consistency of Bayes estimates. **The Annals of Statistics**, p. 1–26, 1986. Citado nas páginas 63, 64 e 65.
- DICKEY, J. M. The weighted likelihood ratio, linear hypotheses on normal location parameters. **The Annals of Mathematical Statistics**, p. 204–223, 1971. Citado na página 78.
- FERNANDEZ, P. J. **Medida e integração**. 2^a. ed. Rio de Janeiro: Impa, 2015. Citado na página 64.
- FRASER, D. Sufficient statistics with nuisance parameters. **The Annals of Mathematical Statistics**, v. 27, p. 838–842, 1956. Citado na página 34.
- GANNON, M. A.; PEREIRA, C. A. B.; POLPO, A. Blending Bayesian and classical tools to define optimal sample-size-dependent significance levels. **The American Statistician**, v. 73, p. 213–222, 2019. Citado nas páginas 24, 31, 34 e 35.

- GUSTAFSON, P. Local sensitivity of posterior expectations. **The Annals of Statistics**, v. 24, p. 174–195, 1996. Citado nas páginas 54 e 69.
- HINDE, J.; DEMÉTRIO, C. G. Overdispersion: models and estimation. **Computational Statistics & Data Analysis**, v. 27, p. 151–170, 1998. Citado na página 43.
- HUBER, P. **Robust Statistics**. New York: Wiley, 1981. Citado na página 63.
- IRONY, T. Z.; PEREIRA, C. A. B. Bayesian hypothesis test: Using surface integrals to distribute prior information among the hypotheses. **Resenhas do Instituto de Matemática e Estatística da Universidade de São Paulo**, v. 2, p. 27–46, 1995. Citado na página 31.
- IZBICKI, R. **Classes de testes de hipóteses**. Dissertação (Mestrado) — Universidade de São Paulo, 2010. Citado na página 52.
- JEFFREYS, H. **Theory of Probability**. 3^a. ed. Oxford: Clarendon Press, 1961. Citado na página 50.
- KASS, R. E.; RAFTERY, A. E. Bayes factor. **Journal of the American Statistical Association**, v. 90, p. 773–795, 1995. Citado na página 50.
- KAWASAKI, Y.; SHIMOKAWA, A.; YAMADA, H.; MIYAOKA, E. A Bayesian equivalency test for two independent binomial proportions. **Journal of Biopharmaceutical Statistics**, v. 26, p. 781–789, 2016. Citado na página 27.
- LIMA, E. L. **Análise no Espaço \mathbb{R}^n** . 2^a. ed. Rio de Janeiro: Impa, 2016. Citado nas páginas 17 e 65.
- _____. **Análise Real: Funções de Uma Variável**. 13^a. ed. Rio de Janeiro: Impa, 2020. v. 1. Citado nas páginas 56, 57 e 58.
- MARTINEZ, E. Z.; ACHCAR, J. A.; ARAGON, D. C. Parameter estimation of the beta-binomial distribution: an application using the SAS software. **Ciência e Natura**, v. 37, p. 12–19, 2015. Citado na página 43.
- OLIVEIRA, N. L.; PEREIRA, C. A. B.; DINIZ, M. A.; POLPO, A. A discussion on significance indices for contingency tables under small sample sizes. **PLoS One**, v. 13, 2018. Citado na página 24.
- PATEL, H. I.; GUPTA, G. D. A problem of equivalence in clinical trials. **Biometrical Journal**, v. 26, p. 471–474, 1984. Citado na página 31.
- PEREIRA, C. A. **Teste de Hipóteses Definidas em Espaços de Diferentes Dimensões: Visão Bayesiana e Interpretação Clássica**. 1985. Tese (Livre Docência) — Universidade de São Paulo. Citado na página 25.
- PEREIRA, C. A. B.; NAKANO, E. Y.; FOSSALUZA, V.; ESTEVES, L. G.; GANNON, M. A.; POLPO, A. Hypothesis tests for Bernoulli experiments: Ordering the sample space by Bayes factors and using adaptive significance levels for decisions. **Entropy**, v. 19, p. 696, 2017. Citado nas páginas 24, 25, 27, 31, 32, 34, 35, 37, 38, 40, 51, 52, 77, 78 e 81.
- PEREIRA, C. A. d. B. **Teoria Estatística: 19º Colóquio Brasileiro de Matemática**. Rio de Janeiro: Impa, 1993. Citado na página 30.

- PEREIRA, C. A. d. B.; STERN, J. M. Evidence and credibility: full Bayesian significance test for precise hypotheses. **Entropy**, Molecular Diversity Preservation International, v. 1, p. 99–110, 1999. Citado na página 78.
- PÉREZ, M.-E.; PERICCHI, L. R. Changing statistical significance with the amount of information: The adaptive α significance level. **Statistics and Probability Letters**, v. 85, p. 20–24, 2014. Citado na página 24.
- PERICCHI, L.; PEREIRA, C. Adaptive significance levels using optimal decision rules: Balancing by weighting the error probabilities. **Brazilian Journal of Probability and Statistics**, v. 30, p. 70–90, 2016. Citado nas páginas 24, 25, 27, 31, 34, 35, 52, 81 e 82.
- PINTO, V. F. Non-inferiority clinical trials: concepts and issues. **Jornal Vascular Brasileiro**, v. 9, p. 145–151, 2010. Citado na página 28.
- R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2021. Disponível em: <<https://www.R-project.org/>>. Citado na página 38.
- RAI, H.; TOMER, S. K.; CHATURVEDI, A. Robust estimation with variational Bayes in presence of competing risks. **METRON**, p. 1–17, 2021. Citado na página 54.
- RUBERG, S. J.; HARRELL, F. E.; GAMALO-SIEBERS, M.; LAVANGE, L.; LEE, J. J.; PRICE, K.; PECK, C. Inference and decision making for 21st-century drug development and approval. **The American Statistician**, v. 73, p. 319–327, 2019. Citado nas páginas 23 e 24.
- RUGGERI, F.; WASSERMAN, L. Infinitesimal sensitivity of posterior distributions. **Canadian Journal of Statistics**, v. 21, p. 195–203, 1993. Citado nas páginas 63 e 64.
- SCHUIRMANN, D. J. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. **Journal of Pharmacokinetics and Biopharmaceutics**, v. 15, p. 657–680, 1987. Citado nas páginas 27, 28, 30 e 79.
- SCHWARTZ, J. T. **Nonlinear Functional Analysis**. New York: Gordon and Breach, 1969. Citado nas páginas 60 e 63.
- SILVA, J. T. da; COBRE, J.; de Castro, M. New Bayesian approaches to equivalence testing. **Journal of Statistical Computation and Simulation**, v. 92, p. 1–17, 2021. Citado na página 27.
- SRINIVASAN, C.; TRUSZCZYNSKA, H. Approximation to the range of a ratio linear posterior quantity based on frechet derivative. In: **Technical Report No. 289**. University of Kentucky: Department of Statistics, 1990. Citado na página 63.
- SWIHART, B.; LINDSEY, J. **rmutil: Utilities for Nonlinear Regression and Repeated Measurements Models**. [S.l.], 2020. R package version 1.1.5. Disponível em: <<https://CRAN.R-project.org/package=rmutil>>. Acesso em: 8 de dezembro de 2021. Citado na página 45.
- VARUZZA, L.; PEREIRA, C. A. B. Significance test for comparing digital gene expression profiles: Partial likelihood application. **Chilean Journal of Statistics**, v. 1, p. 91–102, 2010. Citado na página 24.
- WASSERSTEIN, R. L.; SCHIRM, A. L.; LAZAR, N. A. Moving to a world beyond “ $p < 0.05$ ”. **The American Statistician**, v. 73, p. 1–19, 2019. Citado na página 24.

WETZELS, R.; GRASMAN, R. P. P. P.; WAGENMAKERS, E.-J. An encompassing prior generalization of the Savage–Dickey density ratio. **Computational Statistics & Data Analysis**, v. 54, p. 2094–2102, 2010. Citado na página 78.

WILLIAMSON, P. P. Bayesian equivalence testing for binomial random variables. **Journal of Statistical Computation and Simulation**, v. 77, p. 739–755, 2007. Citado nas páginas 17, 23, 27, 29, 30, 38 e 50.

YOUNG, G. A.; SMITH, R. L. **Essentials of Statistical Inference**. Cambridge: Cambridge University Press, 2005. Citado na página 32.

DIMENSÃO DE ESPAÇOS PARAMÉTRICOS E ELIMINAÇÃO DE PARÂMETRO DE INCÔMODO

Neste apêndice discutimos alguns conceitos sobre a relação entre espaços paramétricos com dimensões diferentes e parâmetros de incômodo, fundamentais na compreensão do teste em [Pereira *et al.* \(2017\)](#). Suponha que H_0 é uma hipótese precisa, então, o espaço Θ_0 pode ser i) Um conjunto só com um ponto em \mathbb{R} , em que a dimensão de Θ_0 é igual a 0, *e.g.*, $\Theta_0 = \{\theta_0\}$ ou qualquer conjunto enumerável, ii) Uma curva no espaço (em \mathbb{R}^2) em que a dimensão é igual a 1 e iii) Uma superfície (em \mathbb{R}^n , para $n > 2$) em que a dimensão é maior do que 1. Ainda sobre a diferença de dimensões, elaboramos o Exemplo 2 baseando-se em [Consonni, Gutiérrez-Peña e Veronese \(2008\)](#), envolvendo hipóteses para diferença de proporções, o qual é importante para complementar alguns comentários sobre a presença do parâmetro de incômodo posteriormente.

Exemplo 2. Considere o teste das hipóteses para duas proporções p e q dadas por

$$H_0 : p = q \quad e \quad H_1 : p \neq q, \tag{A.1}$$

em que $\Theta = [0, 1] \times [0, 1]$ e assumindo $p > q$. Temos também os conjuntos $\Theta'_0 = \{(p, q) \in [0, 1] \times [0, 1]; p = q\}$ e $\Theta_0 = \{t \in [0, 1]; t = p - q\}$. Note que $\Theta'_0 \subset \Theta$, que $\dim(\Theta'_0) = \dim(\Theta_0) < \dim(\Theta)$ e que a hipótese $p - q = 0$ equivale a $t = 0$.

A questão é se conseguimos definir uma distribuição *a priori* sobre Θ , mas como construir uma distribuição *a priori* “compatível” sobre Θ_0 ? É nesse contexto em que há diferença de dimensões entre o espaço e o subespaço paramétricos que se inserem as distribuições *a priori* compatíveis. Ilustrando com o Exemplo 2, um caminho natural é derivar uma distribuição *a priori* em Θ'_0 pelo condicionamento da distribuição *a priori* para Θ em $p - q = 0$. Observe que a representação geométrica de Θ'_0 corresponde à diagonal de um quadrado unitário. Sendo assim,

podemos estar interessados em outras parametrizações para essa diagonal. Por exemplo, podemos definir $\iota = \frac{p}{q} = 1$ ou $\iota = p - q = 0$, o que pode implicar em diferentes distribuições *a priori* sobre Θ_0 pelo fato de o procedimento de condicionamento não ser invariante sob reparametrizações de modelos. Note que teremos problema quanto à forma como parametrizamos e isso implica na associação do paradoxo de Borel-Kolmogorov, pois estamos lidando com condicionamento sobre um evento com probabilidade nula. Para um exemplo mais didático sobre esse paradoxo, *vide* [Wetzels, Grasman e Wagenmakers \(2010\)](#). Existem várias formas para se construir essas distribuições com o propósito de evitar esse paradoxo e uma dessas propostas é a técnica de condicionamento de Jeffreys sugerida por [Dawid e Lauritzen \(2001\)](#), tendo a vantagem de ser invariante para modelar a reparametrização. Segundo [Wetzels, Grasman e Wagenmakers \(2010\)](#), que mencionaram outras técnicas presentes na literatura, a ideia é desviar-se desse paradoxo, mas cada uma dessas propostas terá seus problemas específicos associados.

No Exemplo 2 temos o problema de construir essa parametrização para testes de hipóteses bayesianos com a presença de hipótese nula precisa para valores fixos do parâmetro p . Após esse procedimento de parametrização é que haverá a necessidade de ora se usar uma integral de linha (ou de superfície) para poder calcular a medida de evidência referente ao subespaço paramétrico da hipótese nula H_0 , ora uma integral dupla para a hipótese alternativa H_1 . Um exemplo desses cálculos por um caminho mais dependente da parametrização pode ser encontrado em [Dickey \(1971\)](#), em que se faz o uso de métodos de integração para conseguir a eliminação do parâmetro de incômodo. Sabemos que o fator de Bayes é sensível à escolha de distribuições *a priori* porque dependendo das parametrizações adotadas para ι em termos de p e q como no Exemplo 2, podemos ter diferentes distribuições sobre Θ_0 . Logo, a proposta de condicionamento de [Dawid e Lauritzen \(2001\)](#) é interessante nesse sentido por conta da propriedade de invariância para reparametrização, assim garantindo mais estabilidade quanto ao fator de Bayes calculado, diante da sensibilidade dos fatores de Bayes com respeito à escolha da distribuição *a priori*.

Sobre a construção do modelo estatístico em [Pereira et al. \(2017\)](#), entre as referências mencionadas, destacamos que a abordagem de condicionamento proposta por [Dawid e Lauritzen \(2001\)](#) é uma abordagem mais geral, enquanto que [Dickey \(1971\)](#) faz uma abordagem mais dependente da parametrização, segundo [Pereira et al. \(2017\)](#). Enfatizando, no trabalho de [Dickey \(1971\)](#) subentende-se melhor a escolha do uso dessas integrais de linha e usuais, a depender das hipóteses tratadas. No tratamento de hipótese precisa, ao fixarmos um valor para H_0 , levando-se em conta a reparametrização em virtude da presença de um parâmetro de incômodo, dessa forma contemplando algum dos objetos geométricos possíveis mencionados nos casos ii) ou iii).

Por fim, também [Pereira e Stern \(1999\)](#) trataram com hipóteses precisas propondo o teste de significância completamente bayesiano (FBST). Algumas referências e discussões importantes sobre hipóteses precisas são encontradas em [Cella \(2013\)](#), que utiliza o FBST. No contexto prático, discussões sobre hipóteses precisas ao testar a significância de algum tratamento, primordialmente, pode implicar em uma formulação tal que nos leve a hipóteses nulas pontuais e

precisas, conforme a formulação apresentada em [Berry *et al.* \(2010\)](#) para introduzir de modo amplo os testes de equivalência. Uma ilustração prática e simples dessas hipóteses pode ser vista em [Berger e Delampady \(1987\)](#) ao introduzir sua Seção 2. Nessa linha, repensando em hipóteses compostas, surge a proposta de [Schuirmann \(1987\)](#).

AVALIAÇÕES PARA UMA PROPORÇÃO

Neste apêndice exemplificamos com um breve estudo de simulação o que ocorre se H_0 não for uma hipótese precisa nos testes em [Pericchi e Pereira \(2016\)](#) e [Pereira *et al.* \(2017\)](#). Sendo assim, o objetivo é testar as hipóteses compostas

$$H_0 : p \leq 0,5 \quad e \quad H_1 : p > 0,5. \quad (\text{B.1})$$

Em (B.1) o objetivo é testar a proporção de sucessos p de uma população de interesse $X|p$ com uma proporção de referência 0,5, em que $X|p \sim \text{binomial}(m, p)$. Além disso, dada uma amostra x , $L(p; x)$ denota a função de verossimilhança com $0 \leq p \leq 1$. A seguir, obtemos as medidas de evidência f_{H_0} e f_{H_1} . Note que a função densidade *a priori* $g(p|H_0)$ passa a ser dada por

$$g(p|H_0) = \frac{g(p)}{\int_{\Theta_0} g(p) dp},$$

em que o denominador é a integral usual de g sobre o espaço $\Theta_0 = [0; 0,5]$. Então, a medida de evidência f_{H_0} passa a ser

$$\begin{aligned} f_{H_0}(x) &= \int_{\Theta_0} L(p; x) g(p|H_0) dp = \frac{\int_{\Theta_0} L(p; x) g(p) dp}{\int_{\Theta_0} g(p) dp} \\ &= \frac{\int_0^{0,5} \binom{m}{x} p^x (1-p)^{m-x} \frac{p^{a-1} (1-p)^{b-1}}{\text{Be}(a, b)} dp}{\int_0^{0,5} \frac{p^{a-1} (1-p)^{b-1}}{\text{Be}(a, b)} dp} \\ &= \binom{m}{x} \frac{\int_0^{0,5} p^{x+a-1} (1-p)^{m-x+b-1} dp}{\int_0^{0,5} p^{a-1} (1-p)^{b-1} dp}. \end{aligned}$$

Para a hipótese $H_1 : p > 0,5$, com $\Theta_1 = (0,5, 1]$, temos que

$$\begin{aligned}
 f_{H_1}(x) &= \int_{\Theta_1} L(p;x)g(p|H_1) dp = \frac{\int_{\Theta_1} L(p;x)g(p) dp}{\int_{\Theta_1} g(p) dp} \\
 &= \frac{\int_{0,5}^1 \binom{m}{x} p^x(1-p)^{m-x} \frac{p^{a-1}(1-p)^{b-1}}{\text{Be}(a,b)} dp}{\int_{0,5}^1 \frac{p^{a-1}(1-p)^{b-1}}{\text{Be}(a,b)} dp} \\
 &= \binom{m}{x} \frac{\int_{0,5}^1 p^{x+a-1}(1-p)^{m-x+b-1} dp}{\int_{0,5}^1 p^{a-1}(1-p)^{b-1} dp} \\
 &= \binom{m}{x} \frac{\int_0^1 p^{x+a-1}(1-p)^{m-x+b-1} dp - \int_0^{0,5} p^{x+a-1}(1-p)^{m-x+b-1} dp}{\int_0^1 p^{a-1}(1-p)^{b-1} dp - \int_0^{0,5} p^{a-1}(1-p)^{b-1} dp} \\
 &= \binom{m}{x} \frac{\text{Be}(x+a, m-x+b) - \text{Be}(0,5; x+a, m-x+b)}{\text{Be}(a,b) - \text{Be}(0,5; a,b)},
 \end{aligned}$$

em que

$$\begin{aligned}
 \text{Be}(0,5; x+a, m-x+b) &= \int_0^{0,5} p^{x+a-1}(1-p)^{m-x+b-1} dp \\
 \text{e } \text{Be}(0,5; a,b) &= \int_0^{0,5} p^{a-1}(1-p)^{b-1} dp
 \end{aligned}$$

são as funções beta incompletas.

Para testar as hipóteses em (B.1), consideramos a distribuição *a priori* de Bayes-Laplace, as perdas são $l_0 = l_1 = 1$ e 3000 réplicas para todos os testes. O teste em Pericchi e Pereira (2016) é denotado por B2016. Na Tabela 14, os resultados mostram um aumento da taxa do erro do tipo I para os métodos B2016 e B2017 para além do nível de significância adaptativo, o que pode ser observado nos gráficos dos valores P nas Figuras 8 e 9. Também notamos que os testes B2016 e B2017 deixam de ser equivalentes.

Tabela 14 – Proporção de vezes em que H_0 é rejeitada (H_0 verdadeira).

m	B2016	B2017	$\alpha(\phi_B^*)$
50	0,4297	0,3160	0,0463
300	0,4867	0,4393	0,0214

Fonte: Elaborada pela autora.

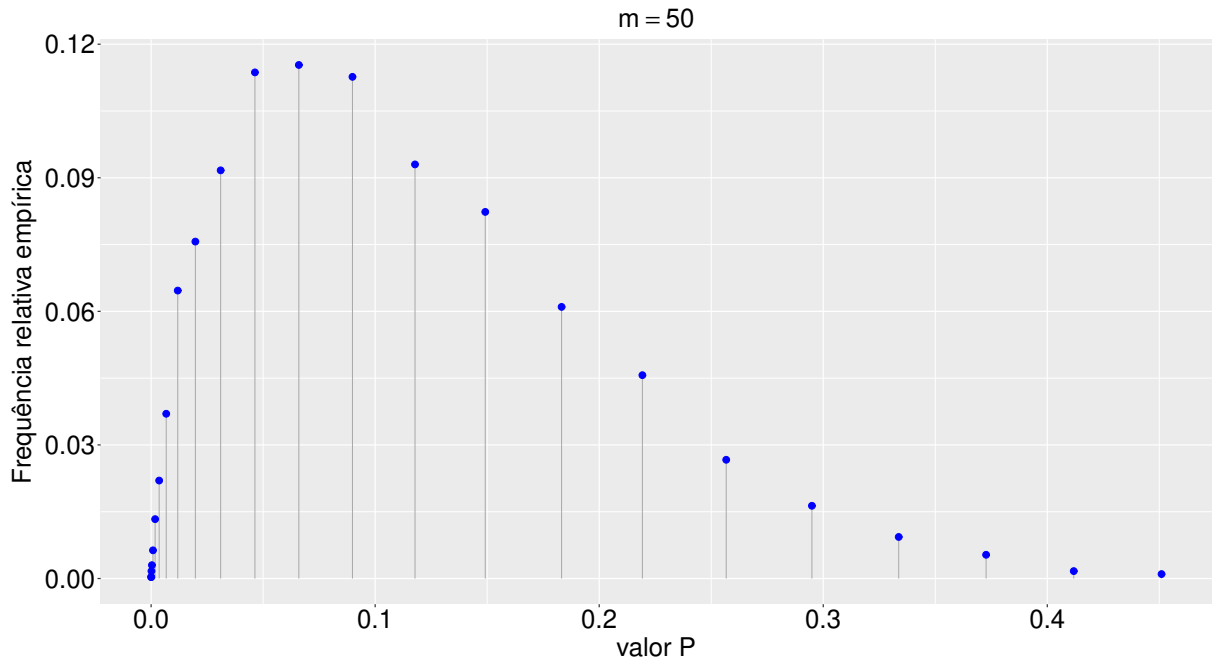


Figura 8 – Distribuições dos valores P para o tamanho amostral $m = 50$.

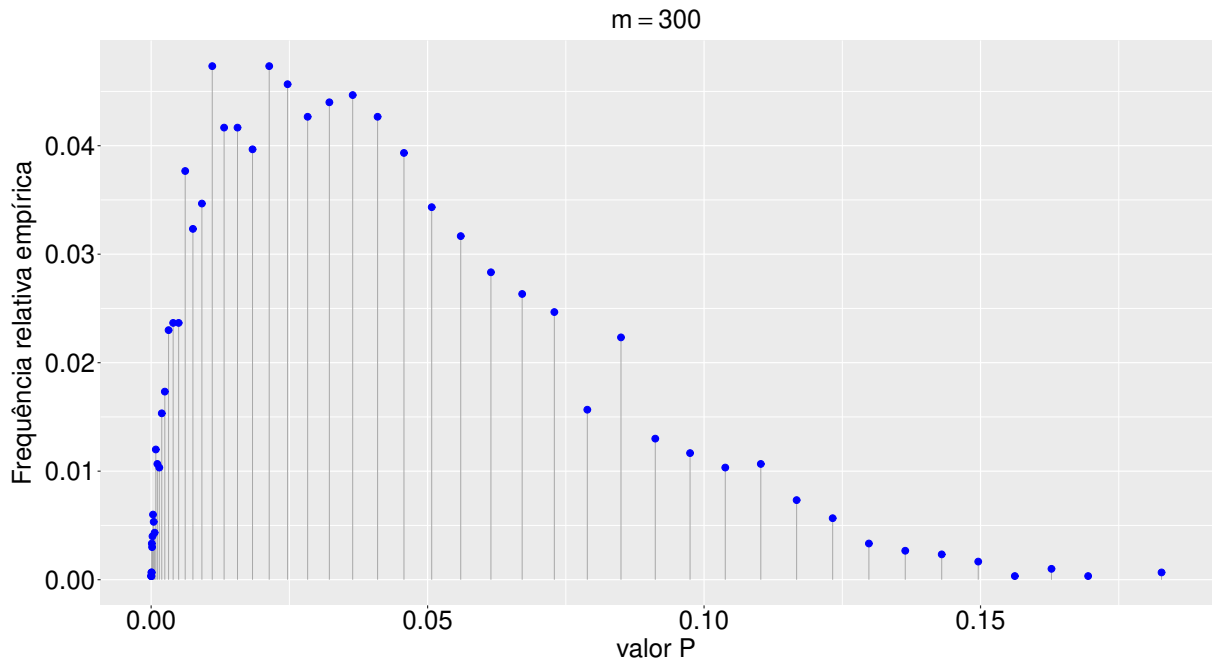


Figura 9 – Distribuições dos valores P para o tamanho amostral $m = 300$.

MEDIDAS DE EVIDÊNCIA f_{H_0} E f_{H_1}

Neste apêndice obtemos as medidas de evidência ponderadas para os dados (x, y) sob cada hipótese, f_{H_0} e f_{H_1} , utilizadas nas abordagens bayesianas referentes ao Capítulo 2.

C.1 f_{H_0}

Os subespaços paramétricos de Θ em (2.7) são apresentados na Figura 10. Note que o subespaço paramétrico Θ_0 corresponde à união disjunta das linhas C_1 e C_2 e o subespaço Θ_1 é representado pela região diagonal. Reescrevemos as hipóteses em (2.7) como

$$H_0 : q = p - \delta \text{ ou } q = p + \delta \quad \text{versus} \quad H_1 : |p - q| < \delta. \quad (\text{C.1})$$

O traço da curva γ_1 corresponde ao segmento de reta no plano pq de vértices $(\delta, 0)$ e $(1, p - \delta)$ orientado no sentido horário, analogamente para o traço da curva C_2 de vértices $(0, \delta)$ e $(1 - \delta, p + \delta)$. Assim, os traços das curvas γ_1 e γ_2 são as retas de $[0, 1] \times [0, 1]$, $C_1 = \gamma_1([\delta, 1])$ e $C_2 = \gamma_2([0, 1 - \delta])$, respectivamente. Então, reescrevemos Θ_0 como $\Theta_0 = \{(p, q) \in [0, 1] \times [0, 1]; q = p - \delta, \delta \leq p \leq 1\} \cup \{(p, q) \in [0, 1] \times [0, 1]; q = p + \delta, 0 \leq p \leq 1 - \delta\}$. Por outro lado, sejam $[\delta, 1]$ e $[0, 1 - \delta]$ intervalos de $[0, 1] \subset \mathbb{R}$. As curvas $\gamma_1 : [\delta, 1] \rightarrow [0, 1] \times [0, 1]$ e $\gamma_2 : [0, 1 - \delta] \rightarrow [0, 1] \times [0, 1]$ são dadas por

$$\gamma_1(t) = (t, t - \delta) \text{ e} \quad (\text{C.2})$$

$$\gamma_2(t) = (t, t + \delta). \quad (\text{C.3})$$

Reparametrizando γ_1 por $\tilde{\gamma}_1(t) = (t + 2, t + 2 - \delta)$ com $[\delta - 2, -1]$, obtemos $\gamma : [\delta - 2, -1] \cup [0, 1 - \delta] \rightarrow [0, 1] \times [0, 1]$ dada por $\gamma(t) = \tilde{\gamma}_1(t)$ se $t \in [\delta - 2, -1]$ e $\gamma(t) = \gamma_2(t)$ se $t \in [0, 1 - \delta]$. Dessa forma, o traço de γ corresponde a Θ_0 , pois $\gamma([\delta - 2, -1] \cup [0, 1 - \delta]) = \gamma([\delta - 2, -1]) \cup \gamma([0, 1 - \delta]) = \tilde{\gamma}_1([\delta - 2, -1]) \cup \gamma_2([0, 1 - \delta]) = \Theta_0 \subset [0, 1] \times [0, 1]$. Recapitulando de (2.15) que

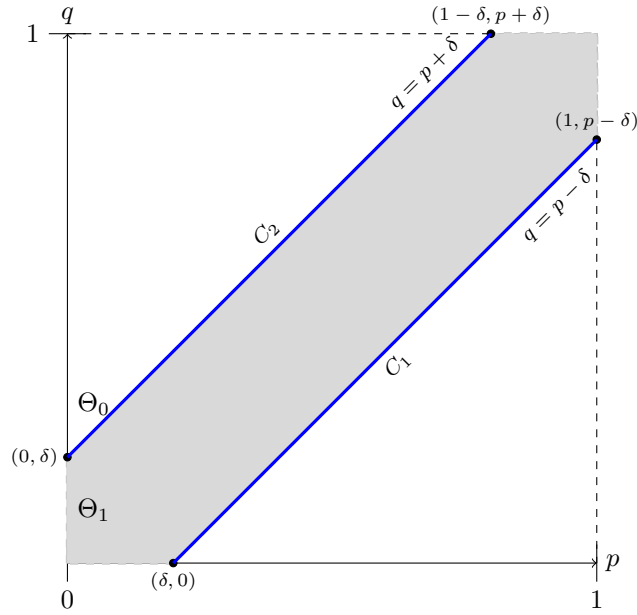


Figura 10 – Regiões associadas às hipóteses $H_0 : |p - q| = \delta$ e $H_1 : |p - q| < \delta$.

f_{H_0} é dada por

$$f_{H_0}(x, y) = \frac{\int_{\gamma} \varphi(x, y, \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds}. \quad (\text{C.4})$$

Assim, a função φ é dada por

$$\begin{aligned} \varphi(x, y, \boldsymbol{\theta}) &= L(x, y; \boldsymbol{\theta}) g(\boldsymbol{\theta}) \\ &= \binom{m}{x} \binom{n}{y} p^x (1-p)^{m-x} q^y (1-q)^{n-y} \\ &\quad \cdot \frac{p^{a-1} (1-p)^{b-1}}{\text{Be}(a, b)} \frac{q^{c-1} (1-q)^{d-1}}{\text{Be}(c, d)} I_{[0,1] \times [0,1]}(p, q) \\ &= \binom{m}{x} \binom{n}{y} \frac{p^{x+a-1} (1-p)^{m-x+b-1} q^{y+c-1} (1-q)^{n-y+d-1}}{\text{Be}(a, b) \text{Be}(c, d)} I_{[0,1] \times [0,1]}(p, q). \end{aligned} \quad (\text{C.5})$$

Fazendo $A = x + a$, $B = m - x + b$, $C = y + c$ e $D = n - y + d$ em (C.5), temos que

$$\varphi(x, y, \boldsymbol{\theta}) = \binom{m}{x} \binom{n}{y} \frac{p^{A-1} (1-p)^{B-1} q^{C-1} (1-q)^{D-1}}{\text{Be}(a, b) \text{Be}(c, d)} I_{[0,1] \times [0,1]}(p, q). \quad (\text{C.6})$$

Como Θ_0 é decomposto em dois traços disjuntos, então

$$\begin{aligned} \int_{\gamma} g(\boldsymbol{\theta}) ds &= \int_{\gamma_1} g(\boldsymbol{\theta}) ds + \int_{\gamma_2} g(\boldsymbol{\theta}) ds \\ &= \int_{\delta}^1 g(\gamma_1(t)) \|\gamma_1'(t)\| dt + \int_0^{1-\delta} g(\gamma_2(t)) \|\gamma_2'(t)\| dt \end{aligned} \quad (\text{C.7})$$

$$\begin{aligned}
e \int_{\gamma} \varphi(x, y, \boldsymbol{\theta}) ds &= \int_{\gamma_1} \varphi(x, y, \boldsymbol{\theta}) ds + \int_{\gamma_2} \varphi(x, y, \boldsymbol{\theta}) ds \\
&= \int_{\delta}^1 \varphi(x, y, \gamma_1(t)) \|\gamma_1'(t)\| dt + \int_0^{1-\delta} \varphi(x, y, \gamma_2(t)) \|\gamma_2'(t)\| dt. \quad (\text{C.8})
\end{aligned}$$

em que $\gamma_j'(t) = \left(\frac{d\psi_{j,x}}{dt}(t), \frac{d\psi_{j,y}}{dt}(t) \right)$ e $\|\gamma_j'(t)\| = \sqrt{\left[\frac{d\psi_{j,x}}{dt}(t) \right]^2 + \left[\frac{d\psi_{j,y}}{dt}(t) \right]^2}$ para $j = 1$ e 2 , com $\psi_{1,x}(t) = t$ e $\psi_{1,y}(t) = t - \delta$ para a curva γ_1 e $\psi_{2,x}(t) = t$ e $\psi_{2,y}(t) = t + \delta$ para γ_2 . Portanto, substituindo (C.7) e (C.8) em (C.4), f_{H_0} é reescrita como

$$f_{H_0}(x, y) = \frac{\int_{\delta}^1 \varphi(x, y, \gamma_1(t)) \|\gamma_1'(t)\| dt + \int_0^{1-\delta} \varphi(x, y, \gamma_2(t)) \|\gamma_2'(t)\| dt}{\int_{\delta}^1 g(\gamma_1(t)) \|\gamma_1'(t)\| dt + \int_0^{1-\delta} g(\gamma_2(t)) \|\gamma_2'(t)\| dt}. \quad (\text{C.9})$$

Observação 4. Para obtermos (C.9), realizamos uma mudança de variável nas integrais referentes à curva $\tilde{\gamma}_1$.

$$\begin{aligned}
f_{H_0}(x, y) &= \frac{\int_{\delta-2}^{-1} \varphi(x, y, \gamma(t)) \|\gamma'(t)\| dt + \int_0^{1-\delta} \varphi(x, y, \gamma(t)) \|\gamma'(t)\| dt}{\int_{\delta-2}^{-1} g(\gamma(t)) \|\gamma'(t)\| dt + \int_0^{1-\delta} g(\gamma(t)) \|\gamma'(t)\| dt} \\
&= \frac{\int_{\delta-2}^{-1} \varphi(x, y, \tilde{\gamma}_1(t)) \|\tilde{\gamma}_1'(t)\| dt + \int_0^{1-\delta} \varphi(x, y, \gamma_2(t)) \|\gamma_2'(t)\| dt}{\int_{\delta-2}^{-1} g(\tilde{\gamma}_1(t)) \|\tilde{\gamma}_1'(t)\| dt + \int_0^{1-\delta} g(\gamma_2(t)) \|\gamma_2'(t)\| dt} \\
&= \frac{\int_{\delta}^1 \varphi(x, y, \tilde{\gamma}_1(t-2)) \|\tilde{\gamma}_1'(t-2)\| dt + \int_0^{1-\delta} \varphi(x, y, \gamma_2(t)) \|\gamma_2'(t)\| dt}{\int_{\delta}^1 g(\tilde{\gamma}_1(t-2)) \|\tilde{\gamma}_1'(t-2)\| dt + \int_0^{1-\delta} g(\gamma_2(t)) \|\gamma_2'(t)\| dt}.
\end{aligned}$$

Note que $\tilde{\gamma}_1(t-2) = (t, t - \delta)$. Então,

$$f_{H_0}(x, y) = \frac{\int_{\delta}^1 \varphi(x, y, \gamma_1(t)) \|\gamma_1'(t)\| dt + \int_0^{1-\delta} \varphi(x, y, \gamma_2(t)) \|\gamma_2'(t)\| dt}{\int_{\delta}^1 g(\gamma_1(t)) \|\gamma_1'(t)\| dt + \int_0^{1-\delta} g(\gamma_2(t)) \|\gamma_2'(t)\| dt}.$$

Inicialmente, para calcularmos (C.7) recapitulamos a função densidade *a priori* beta

expressa em (2.10) e obtemos

$$g(\gamma_1(t)) = g(t, t - \delta) = \frac{t^{a-1}(1-t)^{b-1}}{\text{Be}(a, b)} \frac{(t - \delta)^{c-1}[1 - (t - \delta)]^{d-1}}{\text{Be}(c, d)} I_{[\delta, 1]}(t), \quad (\text{C.10})$$

$$\begin{aligned} \gamma_1'(t) &= (1, 1), \\ \|\gamma_1'(t)\| &= \sqrt{1^2 + 1^2} = \sqrt{2}, \end{aligned} \quad (\text{C.11})$$

$$\begin{aligned} g(\gamma_2(t)) &= g(t, t + \delta) \\ &= \frac{t^{a-1}(1-t)^{b-1}}{\text{Be}(a, b)} \frac{(t + \delta)^{c-1}[1 - (t + \delta)]^{d-1}}{\text{Be}(c, d)} I_{[0, 1-\delta]}(t), \end{aligned} \quad (\text{C.12})$$

$$\begin{aligned} \gamma_2'(t) &= (1, 1) \\ \text{e } \|\gamma_2'(t)\| &= \sqrt{1^2 + 1^2} = \sqrt{2}. \end{aligned} \quad (\text{C.13})$$

Para a curva γ_1 , substituindo (C.10) e (C.11) em (C.7), segue que

$$\begin{aligned} \int_{\gamma_1} g(\boldsymbol{\theta}) ds &= \int_{\delta}^1 g(\gamma_1(t)) \|\gamma_1'(t)\| dt \\ &= \int_{\delta}^1 \frac{t^{a-1}(1-t)^{b-1}}{\text{Be}(a, b)} \frac{(t - \delta)^{c-1}[1 - (t - \delta)]^{d-1}}{\text{Be}(c, d)} \sqrt{2} dt \\ &= \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} \int_{\delta}^1 t^{a-1}(1-t)^{b-1}(t - \delta)^{c-1}[1 - (t - \delta)]^{d-1} dt. \end{aligned} \quad (\text{C.14})$$

Fazendo $\eta_{H_{03}} = \int_{\delta}^1 t^{a-1}(1-t)^{b-1}(t - \delta)^{c-1}[1 - (t - \delta)]^{d-1} dt$, obtemos

$$\int_{\gamma_1} g(\boldsymbol{\theta}) ds = \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{03}}. \quad (\text{C.15})$$

Também para a curva γ_2 , substituindo (C.12) e (C.13) em (C.7), temos

$$\begin{aligned} \int_{\gamma_2} g(\boldsymbol{\theta}) ds &= \int_{t_1}^{t_2} g(\gamma_2(t)) \|\gamma_2'(t)\| dt \\ &= \int_0^{1-\delta} \frac{t^{a-1}(1-t)^{b-1}}{\text{Be}(a, b)} \frac{(t + \delta)^{c-1}[1 - (t + \delta)]^{d-1}}{\text{Be}(c, d)} \sqrt{2} dt \\ &= \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} \int_0^{1-\delta} t^{a-1}(1-t)^{b-1}(t + \delta)^{c-1}[1 - (t + \delta)]^{d-1} dt. \end{aligned} \quad (\text{C.16})$$

Fazendo $\eta_{H_{04}} = \int_0^{1-\delta} t^{a-1}(1-t)^{b-1}(t + \delta)^{c-1}[1 - (t + \delta)]^{d-1} dt$, obtemos

$$\int_{\gamma_2} g(\boldsymbol{\theta}) ds = \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{04}}. \quad (\text{C.17})$$

Logo, substituindo (C.15) e (C.17) em (C.7), resulta em

$$\begin{aligned} \int_{\gamma} g(\boldsymbol{\theta}) ds &= \int_{\gamma_1} g(\boldsymbol{\theta}) ds + \int_{\gamma_2} g(\boldsymbol{\theta}) ds \\ &= \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{03}} + \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{04}} \\ &= \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} (\eta_{H_{03}} + \eta_{H_{04}}). \end{aligned} \quad (\text{C.18})$$

Utilizando (C.2) e (C.6), temos que

$$\begin{aligned}\varphi(x, y, \gamma_1(t)) &= \varphi(x, y, (t, t - \delta)) \\ &= \binom{m}{x} \binom{n}{y} \frac{t^{A-1}(1-t)^{B-1}(t-\delta)^{C-1}[1-(t-\delta)]^{D-1}}{\text{Be}(a, b)\text{Be}(c, d)} I_{[\delta, 1]}(t).\end{aligned}\quad (\text{C.19})$$

Então, substituindo (C.11) e (C.19) em (C.8), segue que

$$\begin{aligned}\int_{\gamma_1} \varphi(x, y, \boldsymbol{\theta}) ds &= \int_{\delta}^1 \varphi(x, y, \gamma_1(t)) \|\gamma_1'(t)\| dt \\ &= \int_{\delta}^1 \binom{m}{x} \binom{n}{y} \frac{t^{A-1}(1-t)^{B-1}(t-\delta)^{C-1}[1-(t-\delta)]^{D-1}}{\text{Be}(a, b)\text{Be}(c, d)} \sqrt{2} dt \\ &= \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \int_{\delta}^1 t^{A-1}(1-t)^{B-1}(t-\delta)^{C-1}[1-(t-\delta)]^{D-1} dt.\end{aligned}\quad (\text{C.20})$$

Fazendo $\eta_{H_{01}} = \int_{\delta}^1 t^{A-1}(1-t)^{B-1}(t-\delta)^{C-1}[1-(t-\delta)]^{D-1} dt$, temos que

$$\int_{\gamma_1} \varphi(x, y, \boldsymbol{\theta}) ds = \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{01}}.\quad (\text{C.21})$$

De modo análogo, com (C.3) e (C.6) obtemos

$$\begin{aligned}\varphi(x, y, \gamma_2(t)) &= \varphi(x, y, (t, t + \delta)) \\ &= \binom{m}{x} \binom{n}{y} \frac{t^{A-1}(1-t)^{B-1}(t+\delta)^{C-1}[1-(t+\delta)]^{D-1}}{\text{Be}(a, b)\text{Be}(c, d)} I_{[0, 1-\delta]}(t).\end{aligned}\quad (\text{C.22})$$

Então, substituindo (C.13) e (C.22) em (C.8), temos que

$$\begin{aligned}\int_{\gamma_2} \varphi(x, y, \boldsymbol{\theta}) ds &= \int_0^{1-\delta} \varphi(x, y, \gamma_2(t)) \|\gamma_2'(t)\| dt \\ &= \int_0^{1-\delta} \binom{m}{x} \binom{n}{y} \frac{t^{A-1}(1-t)^{B-1}(t+\delta)^{C-1}[1-(t+\delta)]^{D-1}}{\text{Be}(a, b)\text{Be}(c, d)} \sqrt{2} dt \\ &= \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \int_0^{1-\delta} t^{A-1}(1-t)^{B-1}(t+\delta)^{C-1}[1-(t+\delta)]^{D-1} dt.\end{aligned}\quad (\text{C.23})$$

Fazendo $\eta_{H_{02}} = \int_0^{1-\delta} t^{A-1}(1-t)^{B-1}(t+\delta)^{C-1}[1-(t+\delta)]^{D-1} dt$, obtemos

$$\int_{\gamma_2} \varphi(x, y, \boldsymbol{\theta}) ds = \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{02}}.\quad (\text{C.24})$$

Por fim, substituindo (C.21) e (C.24) em (C.8) segue que

$$\begin{aligned}\int_{\gamma} \varphi(x, y, \boldsymbol{\theta}) ds &= \int_{\gamma_1} \varphi(x, y, \boldsymbol{\theta}) ds + \int_{\gamma_2} \varphi(x, y, \boldsymbol{\theta}) ds \\ &= \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{01}} + \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \eta_{H_{02}} \\ &= \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} (\eta_{H_{01}} + \eta_{H_{02}}).\end{aligned}\quad (\text{C.25})$$

Logo, encontramos f_{H_0} substituindo (C.18) e (C.25) em (C.9). Então,

$$f_{H_0}(x, y) = \frac{\int_{\gamma} \varphi(x, y, \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} = \binom{m}{x} \binom{n}{y} \frac{\eta_{H_{01}} + \eta_{H_{02}}}{\eta_{H_{03}} + \eta_{H_{04}}}. \quad (\text{C.26})$$

C.2 f_{H_1}

Nesta seção calculamos f_{H_1} , a medida de evidência ponderada para os dados (x, y) sob a hipótese alternativa H_1 . Para tanto, calculamos integrais duplas das funções φ e g sobre o espaço paramétrico Θ_1 . A Figura 11 mostra o espaço Θ_1 dividido em três regiões disjuntas denotadas por \mathcal{R}_1 , \mathcal{R}_2 e \mathcal{R}_3 , sendo descritas por

$$\mathcal{R}_1 = \{(p, q) \in (0, 1)^2; 0 < p < q + \delta, 0 < q < \delta\}, \quad (\text{C.27})$$

$$\mathcal{R}_2 = \{(p, q) \in (0, 1)^2; q - \delta < p < q + \delta, \delta < q < 1 - \delta\} \quad (\text{C.28})$$

$$\text{e } \mathcal{R}_3 = \{(p, q) \in (0, 1)^2; q - \delta < p < 1, 1 - \delta < q < 1\}. \quad (\text{C.29})$$

Dado f_{H_1} de (2.16) e como $\Theta_1 = \mathcal{R}_1 \cup \mathcal{R}_2 \cup \mathcal{R}_3$, φ e g são integráveis sobre cada região, pela

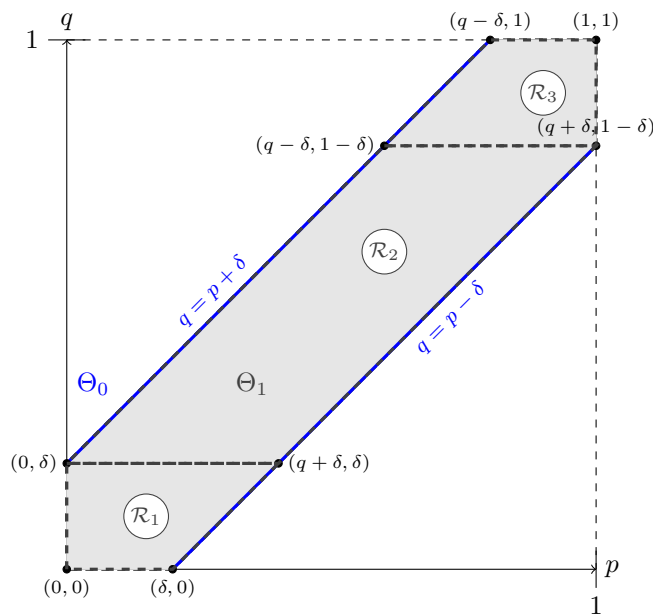


Figura 11 – Decomposição de Θ_1 nas regiões \mathcal{R}_1 , \mathcal{R}_2 e \mathcal{R}_3 .

aditividade, isso implica que

$$\begin{aligned}
 f_{H_1}(x, y) &= \frac{\int_{\Theta_1} \varphi(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta}} \\
 &= \frac{\int_{\mathcal{R}_1} \varphi(x, y, \boldsymbol{\theta}) dpdq + \int_{\mathcal{R}_2} \varphi(x, y, \boldsymbol{\theta}) dpdq + \int_{\mathcal{R}_3} \varphi(x, y, \boldsymbol{\theta}) dpdq}{\int_{\mathcal{R}_1} g(\boldsymbol{\theta}) dpdq + \int_{\mathcal{R}_2} g(\boldsymbol{\theta}) dpdq + \int_{\mathcal{R}_3} g(\boldsymbol{\theta}) dpdq} \\
 &= \frac{\sum_{j=1}^3 \int_{\mathcal{R}_j} \varphi(x, y, \boldsymbol{\theta}) dpdq}{\sum_{j=1}^3 \int_{\mathcal{R}_j} g(\boldsymbol{\theta}) dpdq}. \tag{C.30}
 \end{aligned}$$

Por (2.10) e (C.6) temos que

$$\begin{aligned}
 \int_{\mathcal{R}_j} g(\boldsymbol{\theta}) dpdq &= \int_{\mathcal{R}_j} \frac{p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1}}{\text{Be}(a, b)\text{Be}(c, d)} dpdq \\
 &= \frac{1}{\text{Be}(a, b)\text{Be}(c, d)} \int_{\mathcal{R}_j} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq \tag{C.31}
 \end{aligned}$$

e que

$$\begin{aligned}
 \int_{\mathcal{R}_j} \varphi(x, y, \boldsymbol{\theta}) dpdq &= \int_{\mathcal{R}_j} \frac{\binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq \\
 &= \frac{\binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \int_{\mathcal{R}_j} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq, \tag{C.32}
 \end{aligned}$$

para $j = 1, 2$ e 3 . Então, substituindo (C.32) e (C.31) em (C.30), obtemos

$$\begin{aligned}
 f_{H_1}(x, y) &= \frac{\sum_{j=1}^3 \frac{\binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} \int_{\mathcal{R}_j} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq}{\sum_{j=1}^3 \frac{1}{\text{Be}(a, b)\text{Be}(c, d)} \int_{\mathcal{R}_j} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq} \\
 &= \binom{m}{x} \binom{n}{y} \frac{\sum_{j=1}^3 \int_{\mathcal{R}_j} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq}{\sum_{j=1}^3 \int_{\mathcal{R}_j} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq}. \tag{C.33}
 \end{aligned}$$

Portanto, fixando os limites de integração de acordo com as regiões (C.27), (C.28) e (C.29) em (C.33), temos que

$$f_{H_1}(x, y) = \binom{m}{x} \binom{n}{y} \frac{\eta_{H11} + \eta_{H12} + \eta_{H13}}{\eta_{H14} + \eta_{H15} + \eta_{H16}}, \tag{C.34}$$

em que

$$\begin{aligned}
 \eta_{H11} &= \int_{\mathcal{R}_1} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq \\
 &= \int_0^\delta \int_0^{q+\delta} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq, \\
 \eta_{H12} &= \int_{\mathcal{R}_2} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq \\
 &= \int_\delta^{1-\delta} \int_{q-\delta}^{q+\delta} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq, \\
 \eta_{H13} &= \int_{\mathcal{R}_3} p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq \\
 &= \int_{1-\delta}^1 \int_{q-\delta}^1 p^{A-1}(1-p)^{B-1}q^{C-1}(1-q)^{D-1} dpdq, \\
 \eta_{H14} &= \int_{\mathcal{R}_1} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq \\
 &= \int_0^\delta \int_0^{q+\delta} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq, \\
 \eta_{H15} &= \int_{\mathcal{R}_2} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq \\
 &= \int_\delta^{1-\delta} \int_{q-\delta}^{q+\delta} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq \\
 \text{e } \eta_{H16} &= \int_{\mathcal{R}_3} p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq \\
 &= \int_{1-\delta}^1 \int_{q-\delta}^1 p^{a-1}(1-p)^{b-1}q^{c-1}(1-q)^{d-1} dpdq.
 \end{aligned}$$

MEDIDAS DE EVIDÊNCIA $f_{H_0,\varepsilon}$ E $f_{H_1,\varepsilon}$

Neste apêndice apresentamos as medidas de evidência com a função densidade de probabilidade *a priori* g_ε referente ao Capítulo 3.

D.1 $f_{H_0,\varepsilon}$

Nesta seção calculamos $f_{H_0,\varepsilon}$. Com respeito ao espaço Θ , sejam $p \sim \text{Beta}(a_*, b_*)$ e $q \sim \text{Beta}(c_*, d_*)$ variáveis aleatórias independentes. De (2.10) temos que a função densidade de probabilidade conjunta *a priori* de θ é dada por

$$g_*(\theta) = \frac{p^{a_*-1}(1-p)^{b_*-1}q^{c_*-1}(1-q)^{d_*-1}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} I_{[0,1] \times [0,1]}(\theta). \quad (\text{D.1})$$

Então, de (C.6), a função φ_* é expressa por

$$\varphi_*(x, y, \theta) = \binom{m}{x} \binom{n}{y} \frac{p^{A_*-1}(1-p)^{B_*-1}q^{C_*-1}(1-q)^{D_*-1}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} I_{[0,1] \times [0,1]}(p, q), \quad (\text{D.2})$$

em que $A_* = x + a_*$, $B_* = m - x + b_*$, $C_* = y + c_*$ e $D_* = n - y + d_*$. Pelo Lema 2, de (3.2) temos que

$$f_{H_0,\varepsilon}(x, y) = \frac{\int_{\gamma} \varphi_\varepsilon(x, y, \theta) ds}{\int_{\gamma} g_\varepsilon(\theta) ds} = \frac{(1-\varepsilon) \int_{\gamma} \varphi(x, y, \theta) ds + \varepsilon \int_{\gamma} \varphi_*(x, y, \theta) ds}{(1-\varepsilon) \int_{\gamma} g(\theta) ds + \varepsilon \int_{\gamma} g_*(\theta) ds}, \quad (\text{D.3})$$

em que $\varphi_\varepsilon(x, y, \boldsymbol{\theta}) = (1 - \varepsilon)\varphi(x, y, \boldsymbol{\theta}) + \varepsilon\varphi_*(x, y, \boldsymbol{\theta})$. Considerando a curva $\gamma(t) = \gamma_1(t)I_{[\delta, 1]}(t) + \gamma_2(t)I_{[0, 1-\delta]}(t)$ apresentada no Apêndice C, temos que

$$\begin{aligned} \int_\gamma g_\varepsilon(\boldsymbol{\theta}) ds &= (1 - \varepsilon) \int_\gamma g(\boldsymbol{\theta}) ds + \varepsilon \int_\gamma g_*(\boldsymbol{\theta}) ds, \\ &= (1 - \varepsilon) \left[\int_{\gamma_1} g(\boldsymbol{\theta}) ds + \int_{\gamma_2} g(\boldsymbol{\theta}) ds \right] + \varepsilon \left[\int_{\gamma_1} g_*(\boldsymbol{\theta}) ds + \int_{\gamma_2} g_*(\boldsymbol{\theta}) ds \right] \end{aligned} \quad (\text{D.4})$$

$$\begin{aligned} \text{e } \int_\gamma \varphi_\varepsilon(x, y, \boldsymbol{\theta}) ds &= (1 - \varepsilon) \int_\gamma \varphi(x, y, \boldsymbol{\theta}) ds + \varepsilon \int_\gamma \varphi_*(x, y, \boldsymbol{\theta}) ds \\ &= (1 - \varepsilon) \left[\int_{\gamma_1} \varphi(x, y, \boldsymbol{\theta}) ds + \int_{\gamma_2} \varphi(x, y, \boldsymbol{\theta}) ds \right] \\ &\quad + \varepsilon \left[\int_{\gamma_1} \varphi_*(x, y, \boldsymbol{\theta}) ds + \int_{\gamma_2} \varphi_*(x, y, \boldsymbol{\theta}) ds \right]. \end{aligned} \quad (\text{D.5})$$

De (C.18) obtemos

$$\int_\gamma g_*(\boldsymbol{\theta}) ds = \frac{\sqrt{2}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} (\eta_{H_{03},*} + \eta_{H_{04},*}), \quad (\text{D.6})$$

em que

$$\begin{aligned} \eta_{H_{03},*} &= \int_\delta^1 t^{a_*-1} (1-t)^{b_*-1} (t-\delta)^{c_*-1} [1-(t-\delta)]^{d_*-1} dt \\ \text{e } \eta_{H_{04},*} &= \int_0^{1-\delta} t^{a_*-1} (1-t)^{b_*-1} (t+\delta)^{c_*-1} [1-(t+\delta)]^{d_*-1} dt. \end{aligned}$$

Então, substituindo (C.18) e (D.6) em (D.4) temos que

$$\begin{aligned} \int_\gamma g_\varepsilon(\boldsymbol{\theta}) ds &= (1 - \varepsilon) \frac{\sqrt{2}}{\text{Be}(a, b)\text{Be}(c, d)} (\eta_{H_{03}} + \eta_{H_{04}}) + \varepsilon \frac{\sqrt{2}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} (\eta_{H_{03},*} + \eta_{H_{04},*}) \\ &= \sqrt{2} \left[(1 - \varepsilon) \frac{\eta_{H_{03}} + \eta_{H_{04}}}{\text{Be}(a, b)\text{Be}(c, d)} + \varepsilon \frac{\eta_{H_{03},*} + \eta_{H_{04},*}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} \right]. \end{aligned} \quad (\text{D.7})$$

De forma análoga, de (C.25) obtemos

$$\int_\gamma \varphi_*(x, y, \boldsymbol{\theta}) ds = \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} (\eta_{H_{01},*} + \eta_{H_{02},*}), \quad (\text{D.8})$$

em que

$$\begin{aligned} \eta_{H_{01},*} &= \int_\delta^1 t^{A_*-1} (1-t)^{B_*-1} (t-\delta)^{C_*-1} [1-(t-\delta)]^{D_*-1} dt \\ \text{e } \eta_{H_{02},*} &= \int_0^{1-\delta} t^{A_*-1} (1-t)^{B_*-1} (t+\delta)^{C_*-1} [1-(t+\delta)]^{D_*-1} dt. \end{aligned}$$

Dessa forma, substituindo (C.25) e (D.8) em (D.5) temos que

$$\begin{aligned} \int_\gamma \varphi_\varepsilon(x, y, \boldsymbol{\theta}) ds &= (1 - \varepsilon) \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a, b)\text{Be}(c, d)} (\eta_{H_{01}} + \eta_{H_{02}}) \\ &\quad + \varepsilon \frac{\sqrt{2} \binom{m}{x} \binom{n}{y}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} (\eta_{H_{01},*} + \eta_{H_{02},*}) \\ &= \sqrt{2} \binom{m}{x} \binom{n}{y} \left[(1 - \varepsilon) \frac{\eta_{H_{01}} + \eta_{H_{02}}}{\text{Be}(a, b)\text{Be}(c, d)} + \varepsilon \frac{\eta_{H_{01},*} + \eta_{H_{02},*}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} \right]. \end{aligned} \quad (\text{D.9})$$

Logo, substituindo (D.9) e (D.7) em (D.3) obtemos

$$\begin{aligned} f_{H_0, \varepsilon}(x, y) &= \frac{\sqrt{2} \binom{m}{x} \binom{n}{y} \left[(1 - \varepsilon) \frac{\eta_{H_{01}} + \eta_{H_{02}}}{\text{Be}(a, b)\text{Be}(c, d)} + \varepsilon \frac{\eta_{H_{01,*}} + \eta_{H_{02,*}}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} \right]}{\sqrt{2} \left[(1 - \varepsilon) \frac{\eta_{H_{03}} + \eta_{H_{04}}}{\text{Be}(a, b)\text{Be}(c, d)} + \varepsilon \frac{\eta_{H_{03,*}} + \eta_{H_{04,*}}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} \right]} \\ &= \binom{m}{x} \binom{n}{y} \frac{(1 - \varepsilon) \frac{\eta_{H_{01}} + \eta_{H_{02}}}{\text{Be}(a, b)\text{Be}(c, d)} + \varepsilon \frac{\eta_{H_{01,*}} + \eta_{H_{02,*}}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)}}{(1 - \varepsilon) \frac{\eta_{H_{03}} + \eta_{H_{04}}}{\text{Be}(a, b)\text{Be}(c, d)} + \varepsilon \frac{\eta_{H_{03,*}} + \eta_{H_{04,*}}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)}}. \end{aligned}$$

D.2 $f_{H_1, \varepsilon}$

Pelo Lema 2, de (3.3) temos que

$$f_{H_1, \varepsilon}(x, y) = \frac{\int_{\Theta_1} \varphi_{\varepsilon}(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta}}{\int_{\Theta_1} g_{\varepsilon}(\boldsymbol{\theta}) d\boldsymbol{\theta}} = \frac{(1 - \varepsilon) \int_{\Theta_1} \varphi(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta} + \varepsilon \int_{\Theta_1} \varphi_*(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta}}{(1 - \varepsilon) \int_{\Theta_1} g(\boldsymbol{\theta}) d\boldsymbol{\theta} + \varepsilon \int_{\Theta_1} g_*(\boldsymbol{\theta}) d\boldsymbol{\theta}}. \quad (\text{D.10})$$

De (C.31) e (C.32) temos que

$$\int_{\mathcal{R}_j} \varphi_*(x, y, \boldsymbol{\theta}) dpdq = \frac{\binom{m}{x} \binom{n}{y}}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} \int_{\mathcal{R}_j} p^{A_*-1} (1-p)^{B_*-1} q^{C_*-1} (1-q)^{D_*-1} dpdq \quad (\text{D.11})$$

$$\text{e } \int_{\mathcal{R}_j} g_*(\boldsymbol{\theta}) dpdq = \frac{1}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)} \int_{\mathcal{R}_j} p^{a_*-1} (1-p)^{b_*-1} q^{c_*-1} (1-q)^{d_*-1} dpdq, \quad (\text{D.12})$$

para $j = 1, 2$ e 3 . Como Θ_1 é união disjunta das regiões \mathcal{R}_j , para $j = 1, 2$ e 3 , em (C.27)–(C.29), temos que

$$\int_{\Theta_1} \varphi_{\varepsilon}(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta} = (1 - \varepsilon) \sum_{j=1}^3 \int_{\mathcal{R}_j} \varphi(x, y, \boldsymbol{\theta}) dpdq + \varepsilon \sum_{j=1}^3 \int_{\mathcal{R}_j} \varphi_*(x, y, \boldsymbol{\theta}) dpdq \quad (\text{D.13})$$

$$\text{e } \int_{\Theta_1} g_{\varepsilon}(\boldsymbol{\theta}) d\boldsymbol{\theta} = (1 - \varepsilon) \sum_{j=1}^3 \int_{\mathcal{R}_j} g(\boldsymbol{\theta}) dpdq + \varepsilon \sum_{j=1}^3 \int_{\mathcal{R}_j} g_*(\boldsymbol{\theta}) dpdq. \quad (\text{D.14})$$

Substituindo (D.11) em (D.13), obtemos

$$\int_{\Theta_1} \varphi_{\varepsilon}(x, y, \boldsymbol{\theta}) d\boldsymbol{\theta} = \binom{m}{x} \binom{n}{y} \frac{(1 - \varepsilon)(\eta_{H_{11}} + \eta_{H_{12}} + \eta_{H_{13}})}{\text{Be}(a, b)\text{Be}(c, d)} + \frac{\varepsilon(\eta_{H_{11,*}} + \eta_{H_{12,*}} + \eta_{H_{13,*}})}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)}, \quad (\text{D.15})$$

em que

$$\begin{aligned}
 \eta_{H11,*} &= \int_{\mathcal{R}_1} p^{A_*-1} (1-p)^{B_*-1} q^{C_*-1} (1-q)^{D_*-1} dpdq \\
 &= \int_0^\delta \int_0^{q+\delta} p^{A_*-1} (1-p)^{B_*-1} q^{C_*-1} (1-q)^{D_*-1} dpdq, \\
 \eta_{H12,*} &= \int_{\mathcal{R}_2} p^{A_*-1} (1-p)^{B_*-1} q^{C_*-1} (1-q)^{D_*-1} dpdq \\
 &= \int_\delta^{1-\delta} \int_{q-\delta}^{q+\delta} p^{A_*-1} (1-p)^{B_*-1} q^{C_*-1} (1-q)^{D_*-1} dpdq \\
 \text{e } \eta_{H13,*} &= \int_{\mathcal{R}_3} p^{A_*-1} (1-p)^{B_*-1} q^{C_*-1} (1-q)^{D_*-1} dpdq \\
 &= \int_{1-\delta}^1 \int_{q-\delta}^1 p^{A_*-1} (1-p)^{B_*-1} q^{C_*-1} (1-q)^{D_*-1} dpdq.
 \end{aligned}$$

Analogamente, substituindo (D.12) em (D.14) temos que

$$\int_{\Theta_1} g_\varepsilon(\boldsymbol{\theta}) d\boldsymbol{\theta} = \frac{(1-\varepsilon)(\eta_{H14} + \eta_{H15} + \eta_{H16})}{\text{Be}(a, b)\text{Be}(c, d)} + \frac{\varepsilon(\eta_{H14,*} + \eta_{H15,*} + \eta_{H16,*})}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)}, \quad (\text{D.16})$$

em que

$$\begin{aligned}
 \eta_{H14,*} &= \int_{\mathcal{R}_1} p^{a_*-1} (1-p)^{b_*-1} q^{c_*-1} (1-q)^{d_*-1} dpdq \\
 &= \int_0^\delta \int_0^{q+\delta} p^{a_*-1} (1-p)^{b_*-1} q^{c_*-1} (1-q)^{d_*-1} dpdq, \\
 \eta_{H15,*} &= \int_{\mathcal{R}_2} p^{a_*-1} (1-p)^{b_*-1} q^{c_*-1} (1-q)^{d_*-1} dpdq \\
 &= \int_\delta^{1-\delta} \int_{q-\delta}^{q+\delta} p^{a_*-1} (1-p)^{b_*-1} q^{c_*-1} (1-q)^{d_*-1} dpdq \\
 \text{e } \eta_{H16,*} &= \int_{\mathcal{R}_3} p^{a_*-1} (1-p)^{b_*-1} q^{c_*-1} (1-q)^{d_*-1} dpdq \\
 &= \int_{1-\delta}^1 \int_{q-\delta}^1 p^{a_*-1} (1-p)^{b_*-1} q^{c_*-1} (1-q)^{d_*-1} dpdq.
 \end{aligned}$$

Logo, de (D.15) e (D.16) segue que

$$f_{H_1, \varepsilon}(x, y) = \binom{m}{x} \binom{n}{y} \frac{\frac{(1-\varepsilon)(\eta_{H11} + \eta_{H12} + \eta_{H13})}{\text{Be}(a, b)\text{Be}(c, d)} + \frac{\varepsilon(\eta_{H11,*} + \eta_{H12,*} + \eta_{H13,*})}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)}}{\frac{(1-\varepsilon)(\eta_{H14} + \eta_{H15} + \eta_{H16})}{\text{Be}(a, b)\text{Be}(c, d)} + \frac{\varepsilon(\eta_{H14,*} + \eta_{H15,*} + \eta_{H16,*})}{\text{Be}(a_*, b_*)\text{Be}(c_*, d_*)}}. \quad (\text{D.17})$$

OUTRA FORMA DE OBTENÇÃO DE $\mathbb{T}'_P(g)h$, REFERENTE AO TEOREMA 3

Teorema 4. Seja \mathcal{X} finito. A derivada de Gâteaux de \mathbb{T}_P de $g \in \mathcal{G}$ na direção $h \in \mathcal{M}$ é dada por

$$\mathbb{T}'_P(g)h = \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} (\text{valor } P_*^{\text{obs}} - \text{valor } P),$$

em que $\text{valor } P_*^{\text{obs}} = \sum_{(x,y) \in B_{\text{obs}}} f_{H_0,*}(x,y)$.

Demonstração. Dados $(x,y)_{\text{obs}} \in \mathcal{X}$ uma amostra observada e $g \in \mathcal{G}$ uma função densidade de probabilidade *a priori* elicitada na direção $h = g_* - g \in \mathcal{M}$, a derivada de Gâteaux de \mathbb{T}_P em g na direção h é dada por

$$\begin{aligned} \mathbb{T}'_P(g)h &= \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{T}_P(g + \varepsilon h) - \mathbb{T}_P(g)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\mathbb{T}_P(g\varepsilon) - \mathbb{T}_P(g)}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\text{valor } P_{\varepsilon} - \text{valor } P}{\varepsilon} \\ &= \lim_{\varepsilon \rightarrow 0^+} \frac{\sum_{(x,y) \in B_{\text{obs},\varepsilon}} f_{H_0,\varepsilon}(x,y) - \sum_{(x,y) \in B_{\text{obs}}} f_{H_0}(x,y)}{\varepsilon} \end{aligned}$$

Pelo Teorema 1, existe $\varepsilon_0 > 0$ tal que $\varepsilon \in (0, \varepsilon_0)$ implica $B_{\text{obs},\varepsilon} = B_{\text{obs}}$. Portanto,

$$\mathbb{T}'_P(g)h = \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon < \varepsilon_0}} \frac{\sum_{(x,y) \in B_{\text{obs}}} [f_{H_0,\varepsilon}(x,y) - f_{H_0}(x,y)]}{\varepsilon}$$

$$\stackrel{(*)}{=} \sum_{(x,y) \in B_{\text{obs}}} \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon < \varepsilon_0}} \frac{f_{H_0,\varepsilon}(x,y) - f_{H_0}(x,y)}{\varepsilon}, \quad (\text{E.1})$$

em que a mudança (*) é válida, pois $B_{\text{obs}} \subset \mathcal{X}$ é finito e $f_{H_0,\varepsilon}(x,y) = \int_{\gamma} L(\boldsymbol{\theta}; x,y) g_{\varepsilon}(\boldsymbol{\theta}|H_0) ds$,

em que $g_{\varepsilon}(\boldsymbol{\theta}|H_0) = \frac{g_{\varepsilon}(\boldsymbol{\theta})}{\int_{\gamma} g_{\varepsilon}(\boldsymbol{\theta}) ds}$. Note que o numerador em (E.1) é dado por

$$\begin{aligned} f_{H_0,\varepsilon}(x,y) - f_{H_0}(x,y) &= \int_{\gamma} L(\boldsymbol{\theta}; x,y) g_{\varepsilon}(\boldsymbol{\theta}|H_0) ds - \int_{\gamma} L(\boldsymbol{\theta}; x,y) g(\boldsymbol{\theta}|H_0) ds \\ &= \int_{\gamma} L(\boldsymbol{\theta}; x,y) [g_{\varepsilon}(\boldsymbol{\theta}|H_0) ds - g(\boldsymbol{\theta}|H_0)] ds, \end{aligned} \quad (\text{E.2})$$

em que a diferença das funções densidades *a priori* condicionadas na hipótese H_0 em (E.2) é expressa por

$$\begin{aligned} g_{\varepsilon}(\boldsymbol{\theta}|H_0) - g(\boldsymbol{\theta}|H_0) &= \frac{g_{\varepsilon}(\boldsymbol{\theta})}{\int_{\gamma} g_{\varepsilon}(\boldsymbol{\theta}) ds} I_{\Theta_0}(\boldsymbol{\theta}) - \frac{g(\boldsymbol{\theta})}{\int_{\gamma} g(\boldsymbol{\theta}) ds} I_{\Theta_0}(\boldsymbol{\theta}) \\ &= \frac{g_{\varepsilon}(\boldsymbol{\theta}) \int_{\gamma} g(\boldsymbol{\theta}) ds - g(\boldsymbol{\theta}) \int_{\gamma} g_{\varepsilon}(\boldsymbol{\theta}) ds}{\int_{\gamma} g_{\varepsilon}(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds} I_{\Theta_0}(\boldsymbol{\theta}). \end{aligned} \quad (\text{E.3})$$

A seguir reescrevemos o numerador e o denominador em (E.3). Note que $g_{\varepsilon} = (1 - \varepsilon)g + \varepsilon g_*$ = $g + \varepsilon(g_* - g)$. Então, o numerador e o denominador são dados, respectivamente, por

$$\begin{aligned} g_{\varepsilon}(\boldsymbol{\theta}) \int_{\gamma} g(\boldsymbol{\theta}) ds - g(\boldsymbol{\theta}) \int_{\gamma} g_{\varepsilon}(\boldsymbol{\theta}) ds &= [(1 - \varepsilon)g(\boldsymbol{\theta}) + \varepsilon g_*(\boldsymbol{\theta})] \int_{\gamma} g(\boldsymbol{\theta}) ds \\ &\quad - g(\boldsymbol{\theta}) \int_{\gamma} [(1 - \varepsilon)g(\boldsymbol{\theta}) + \varepsilon g_*(\boldsymbol{\theta})] ds \\ &= (1 - \varepsilon)g(\boldsymbol{\theta}) \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon g_*(\boldsymbol{\theta}) \int_{\gamma} g(\boldsymbol{\theta}) ds \\ &\quad - (1 - \varepsilon)g(\boldsymbol{\theta}) \int_{\gamma} g(\boldsymbol{\theta}) ds - \varepsilon g(\boldsymbol{\theta}) \int_{\gamma} g_*(\boldsymbol{\theta}) ds \\ &= \varepsilon \left[g_*(\boldsymbol{\theta}) \int_{\gamma} g(\boldsymbol{\theta}) ds - g(\boldsymbol{\theta}) \int_{\gamma} g_*(\boldsymbol{\theta}) ds \right] \end{aligned} \quad (\text{E.4})$$

$$\begin{aligned} \text{e } \int_{\gamma} g_{\varepsilon}(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds &= \int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon (g_*(\boldsymbol{\theta}) - g(\boldsymbol{\theta})) ds \\ &= \int_{\gamma} g(\boldsymbol{\theta}) ds \left\{ \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} [g_*(\boldsymbol{\theta}) - g(\boldsymbol{\theta})] ds \right\} \\ &= \int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} [g_*(\boldsymbol{\theta}) - g(\boldsymbol{\theta})] ds. \end{aligned} \quad (\text{E.5})$$

Substituindo (E.4) e (E.5) em (E.3), temos que

$$g_{\varepsilon}(\boldsymbol{\theta}|H_0) - g(\boldsymbol{\theta}|H_0) = \frac{\varepsilon \left[g_*(\boldsymbol{\theta}) \int_{\gamma} g(\boldsymbol{\theta}) ds - g(\boldsymbol{\theta}) \int_{\gamma} g_*(\boldsymbol{\theta}) ds \right]}{\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} [g_*(\boldsymbol{\theta}) - g(\boldsymbol{\theta})] ds}. \quad (\text{E.6})$$

Novamente, substituindo (E.6) em (E.2), obtemos

$$f_{H_0, \varepsilon}(x, y) - f_{H_0}(x, y) = \frac{\varepsilon \left[\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} L(\boldsymbol{\theta}; x, y) g_*(\boldsymbol{\theta}) ds - \int_{\gamma} g_*(\boldsymbol{\theta}) ds \int_{\gamma} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}) ds \right]}{\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} [g_*(\boldsymbol{\theta}) - g(\boldsymbol{\theta})] ds}.$$

(E.7)

Logo, substituindo (E.7) em (E.1) temos que

$$\begin{aligned} \mathbb{T}'_P(g)h &= \sum_{(x,y) \in B_{\text{obs}}} \lim_{\substack{\varepsilon \rightarrow 0 \\ \varepsilon < \varepsilon_0}} \frac{\varepsilon \left[\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} L(\boldsymbol{\theta}; x, y) g_*(\boldsymbol{\theta}) ds - \int_{\gamma} g_*(\boldsymbol{\theta}) ds \int_{\gamma} L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta}) ds \right]}{\varepsilon \left\{ \int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds + \varepsilon \int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} [g_*(\boldsymbol{\theta}) - g(\boldsymbol{\theta})] ds \right\}} \\ &= \sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} \varphi_*(x, y; \boldsymbol{\theta}) ds - \int_{\gamma} g_*(\boldsymbol{\theta}) ds \int_{\gamma} \varphi(x, y; \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds} \\ &= \sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} \varphi_*(x, y; \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds} - \sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds \int_{\gamma} \varphi(x, y; \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds \int_{\gamma} g(\boldsymbol{\theta}) ds} \\ &= \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} \sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} \varphi_*(x, y; \boldsymbol{\theta}) ds}{\int_{\gamma} g_*(\boldsymbol{\theta}) ds} - \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} \sum_{(x,y) \in B_{\text{obs}}} \frac{\int_{\gamma} \varphi(x, y; \boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds}, \end{aligned}$$

em que $\varphi_*(x, y; \boldsymbol{\theta}) = L(\boldsymbol{\theta}; x, y) g_*(\boldsymbol{\theta})$ e $\varphi(x, y; \boldsymbol{\theta}) = L(\boldsymbol{\theta}; x, y) g(\boldsymbol{\theta})$. Logo,

$$\mathbb{T}'_P(g)h = \frac{\int_{\gamma} g_*(\boldsymbol{\theta}) ds}{\int_{\gamma} g(\boldsymbol{\theta}) ds} (\text{valor } P_*^{\text{obs}} - \text{valor } P).$$

□

