

**Risco de Bayes e determinação de tamanho amostral em testes de hipóteses agnósticos: Um estudo comparativo**

Bruno Lazarini

DISSERTAÇÃO APRESENTADA  
AO  
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA  
DA  
UNIVERSIDADE DE SÃO PAULO  
PARA  
OBTENÇÃO DO TÍTULO  
DE  
MESTRE EM CIÊNCIAS

Programa: Estatística

Orientador: Prof. Dr. Luís Gustavo Esteves

São Paulo, novembro de 2022

# Risco de Bayes e determinação de tamanho amostral em testes de hipóteses agnósticos: Um estudo comparativo

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por Bruno Lazarini e aprovada pela Comissão Julgadora.

Comissão Julgadora:

- Prof. Dr. Luís Gustavo Esteves (orientador) - IME-USP
- Prof. Dr. Luís Ernesto Bueno Salazar - UFSCar
- Prof. Dr. Marcio Alves Diniz - UFSCar

# Agradecimentos

Inicialmente, agradeço aos meus pais, Lucia e Silvio, por todo o apoio e incentivo aos estudos durante toda minha vida.

Agradeço imensamente ao meu orientador Professor Luís Gustavo Esteves por todos os aprendizados, pela paciência e pela grande ajuda na construção deste trabalho. Agradeço também à todos os professores do IME com os quais tive oportunidade de cursar disciplinas e expandir meus conhecimentos em Estatística.

Por fim, agradeço à minha companheira Debora por todo incentivo, apoio e compreensão durante esse período por vezes árduo, porém bastante recompensador.



# Resumo

Testes de hipóteses constituem um dos principais problemas estudados na Estatística e são amplamente aplicados nas ciências em geral. Na sua formulação tradicional, temos duas possibilidades de decisão: rejeitar ou não rejeitar uma hipótese  $H$  de interesse. A possibilidade de considerar em testes de hipóteses uma alternativa adicional de decisão—a de se abster de uma decisão em relação a  $H$ , também denominada de ficar agnóstico—tem sido estudada recentemente por possuir vantagens em termos de consistência lógica, particularmente quando desejamos testar diferentes hipóteses de forma simultânea para um mesmo conjunto de dados observados. O presente trabalho tem o objetivo de explorar outras vantagens dos chamados testes de hipóteses agnósticos, porém em um cenário de hipótese única e não de hipóteses simultâneas. Mais especificamente, foi demonstrado sob a abordagem bayesiana que testes desse tipo possuem risco de Bayes menor que os testes de duas decisões, sob certas condições. Em decorrência disso, se utilizarmos um método particular de determinação de tamanho de amostra ótimo, existe a possibilidade de obtermos tamanhos amostrais ótimos menores no caso agnóstico. Dadas essas vantagens, diversos problemas de testes de hipóteses foram estudados e implementados tanto no cenário tradicional quanto no agnóstico, considerando alguns modelos estatísticos importantes. Adicionalmente, foram elaborados exemplos numéricos para ilustrar tais vantagens. Para todos os exemplos estudados, foi observada uma redução do risco de Bayes e também dos tamanhos amostrais ótimos, quando comparamos o teste de três decisões com o teste de duas decisões, o que poderia resultar em uma redução de custo no planejamento amostral de experimentos.

**Palavras-chave:** testes de hipóteses, inferência bayesiana, teoria da decisão, risco de Bayes, testes de hipóteses agnósticos, determinação de tamanho amostral



# Abstract

Hypothesis testing is one of the main problems studied in Statistics and is widely applied in the sciences in general. In its traditional formulation, we have two decision possibilities: to reject or not a hypothesis  $H$  of interest. The possibility of considering in hypothesis testing an additional alternative of decision—that of abstaining from a decision with respect to  $H$ , also called to remain agnostic—has been studied recently because of its advantages in terms of logical consistency, particularly when we wish to test different hypotheses simultaneously for the same observed data set. The present work aims to explore other advantages of the so-called agnostic hypothesis test, but in a single hypothesis scenario and not in a simultaneous hypotheses scenario. More specifically, we show under the Bayesian approach that tests of this type have a lower Bayes risk than the two-decision tests, under certain conditions. As a result, if we use a particular method of optimal sample size determination, it is possible to obtain smaller optimal sample sizes in the agnostic case. Given these advantages, several hypothesis testing problems were studied and implemented both in the traditional and agnostic scenarios, considering some important statistical models. Additionally, numerical examples were elaborated to illustrate such advantages. For all the examples studied, a reduction in risk and also in optimal sample sizes were observed when comparing the three-decision test with the two-decision test, which could result in a cost reduction in sample design for experiments.

**Keywords:** hypothesis testing, bayesian inference, decision theory, Bayes risk, agnostic hypothesis tests, sample size determination





# Sumário

<b>Lista de Abreviaturas</b>	<b>ix</b>
<b>Lista de Símbolos</b>	<b>xi</b>
<b>Lista de Figuras</b>	<b>xiii</b>
<b>Lista de Tabelas</b>	<b>xv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Considerações Preliminares . . . . .	1
1.2 Objetivos . . . . .	2
1.3 Organização do Trabalho . . . . .	2
<b>2 Teoria da Decisão</b>	<b>3</b>
2.1 Elementos de um problema de decisão . . . . .	3
2.1.1 Probabilidade subjetiva . . . . .	4
2.1.2 Função de utilidade ou perda . . . . .	5
2.2 Princípio da Maximização da Utilidade Esperada e Risco de Bayes . . . . .	6
2.2.1 Obtenção do risco de Bayes em problemas com observações . . . . .	6
<b>3 Testes de Hipóteses</b>	<b>9</b>
3.1 Testes de Hipóteses Bayesianos . . . . .	9
3.1.1 Testes baseados em Probabilidade a Posteriori . . . . .	11
3.1.2 <i>Full Bayesian Significance Test</i> . . . . .	12
3.1.3 Testes baseados em Estimador por Região . . . . .	13
3.2 Testes de Hipóteses Agnósticos . . . . .	15
3.2.1 Testes Agnósticos baseados em Probabilidade a Posteriori . . . . .	17
3.2.2 <i>Agnostic Full Bayesian Significance Test</i> . . . . .	19
3.2.3 Testes Agnósticos baseados em Estimador por Região e GFBST . . . . .	19
<b>4 Redução do Risco de Bayes e Determinação do Tamanho Amostral</b>	<b>23</b>
4.1 Redução do Risco de Bayes . . . . .	23
4.2 Determinação de tamanho amostral . . . . .	25
4.3 Estudo comparativo: Testes de Hipóteses Unilaterais . . . . .	27
4.3.1 Teste baseado em Probabilidade a Posteriori: Modelo Normal com Variância conhecida . . . . .	27

4.3.2	Teste de Bayes com Perda baseada em Norma $L_1$ : Modelo Normal com Variância conhecida . . . . .	34
4.3.3	Teste baseado em Probabilidade a Posteriori: Modelo Bernoulli . . . . .	40
4.4	Estudo comparativo: Testes de Hipóteses Intervalares . . . . .	46
4.4.1	Teste baseado em Probabilidade a Posteriori: Modelo Normal com Variância conhecida . . . . .	46
4.4.2	Teste baseado em Probabilidade a Posteriori: Modelo Poisson . . . . .	50
4.5	Estudo comparativo: Testes de Hipóteses Precisas . . . . .	54
4.5.1	Teste baseado em estimador por região: Modelo Normal Univariado com Variância conhecida . . . . .	55
4.5.2	Teste baseado em estimador por região: Modelo Normal Bivariado com Variância conhecida . . . . .	60
<b>5</b>	<b>Conclusões</b>	<b>69</b>
5.1	Considerações Finais . . . . .	69
5.2	Sugestões para Pesquisas Futuras . . . . .	71
<b>A</b>	<b>Demonstrações adicionais</b>	<b>73</b>
A.1	Distribuição marginal do modelo Beta-Bernoulli . . . . .	73
A.2	Distribuição marginal do modelo Poisson-Gama . . . . .	74
	<b>Referências Bibliográficas</b>	<b>75</b>

# Lista de Abreviaturas

iid	independentes e identicamente distribuídos.
FBST	Teste de Significância Totalmente Bayesiano ( <i>Full Bayesian Significance Test</i> ).
GFBST	Teste de Significância Totalmente Bayesiano Generalizado ( <i>Generalized Full Bayesian Significance Test</i> ).
HPD	Maior Densidade a Posteriori ( <i>Highest Posterior Density</i> ).



# Lista de Símbolos

$\theta$	Parâmetro.
$\Theta$	Espaço paramétrico.
$\boldsymbol{x}$	Vetor de observações.
$\mathcal{X}$	Espaço amostral de observações.
$d$	Decisão.
$\mathcal{D}$	Espaço de decisões.
$\varphi$	Função de decisão ou teste.
$\Delta$	Espaço de funções de decisão.
$\mathcal{P}$	Medida de probabilidade a priori.
$\bar{\mathcal{P}}$	Medida de probabilidade que representa a incerteza sobre $(\theta, X)$ .
$\mathcal{P}_x$	Medida de probabilidade a posteriori.
$\mathcal{F}$	$\sigma$ -álgebra gerada pelo espaço amostral.
$\bar{\mathcal{F}}$	$\sigma$ -álgebra gerada por $(\Theta, \mathcal{X})$ .
$L$	Função de perda.
$U$	Função de utilidade.
$\rho$	Perda esperada ou risco.
$H_0$	Hipótese nula.
$H_1$	Hipótese alternativa.
$\mathbb{R}$	Conjunto dos números reais.
$\mathbb{N}$	Conjunto dos números naturais.
$\mathbb{N}^*$	Conjunto dos números naturais estritamente positivos.
$\mathbb{P}(\Theta)$	Conjunto das partes de $\Theta$ .



# Lista de Figuras

3.1	Teste agnóstico baseado em estimador por região. . . . .	20
4.1	Riscos de Bayes dos testes unilaterais para modelo Normal - Priori $\mathcal{N}(0, 100)$ . . . . .	31
4.2	Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes unilaterais para modelo Normal. . . . .	31
4.3	Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes unilaterais para modelo Normal. . . . .	32
4.4	Diferença absoluta entre os riscos de Bayes para diferentes perdas - Testes unilaterais para modelo Normal. . . . .	32
4.5	Riscos de Bayes dos testes unilaterais para modelo Normal com função de perda $L_1$ - Priori $\mathcal{N}(3, 100)$ . . . . .	38
4.6	Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes unilaterais com perda $0 - L_1$ para modelo Normal. . . . .	39
4.7	Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes unilaterais com perda $0 - L_1^{(p)}$ para modelo Normal. . . . .	39
4.8	Riscos de Bayes dos testes unilaterais para modelo Bernoulli - Priori $Beta(1, 1)$ . . . . .	44
4.9	Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias semelhantes - Testes unilaterais para modelo Bernoulli. . . . .	44
4.10	Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes unilaterais para modelo Bernoulli. . . . .	45
4.11	Riscos de Bayes dos testes intervalares para modelo Normal - Priori $\mathcal{N}(0, 100)$ . . . . .	48
4.12	Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes intervalares para modelo Normal. . . . .	49
4.13	Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes intervalares para modelo Normal. . . . .	49
4.14	Riscos de Bayes dos testes intervalares para modelo Poisson - Priori $\mathcal{G}(4, 1)$ . . . . .	53
4.15	Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes intervalares para modelo Poisson. . . . .	53
4.16	Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes intervalares para modelo Poisson. . . . .	54
4.17	Riscos dos testes baseados em estimador por região para modelo Normal univariado - Priori $\mathcal{N}(0, 100)$ . . . . .	58
4.18	Diferença absoluta entre os riscos para prioris com médias diferentes e variâncias iguais - Testes baseados em estimador por região para modelo Normal univariado. . . . .	58

4.19 Diferença absoluta entre os riscos para prioris com mesma média e variâncias diferentes - Testes baseados em estimador por região para modelo Normal univariado. . . . . 59

4.20 Figura ilustrativa da hipótese  $H_0$  no caso bivariado. . . . . 60

4.21 Figura ilustrativa do teste agnóstico baseado em estimador por região no caso bivariado. 63

4.22 Riscos dos testes baseados em estimador por região para modelo Normal bivariado - Prioris  $\mathcal{N}(0, 100)$ . . . . . 66

4.23 Diferença absoluta entre os riscos para prioris com médias diferentes e variâncias iguais - Testes baseados em estimador por região para modelo Normal bivariado. . . . . 66

4.24 Diferença absoluta entre os riscos para prioris com mesma média e variâncias diferentes - Testes baseados em estimador por região para modelo Normal bivariado. . . . . 67



# Lista de Tabelas

3.1	Função de perda $0 - a_0 - a_1$ . . . . .	10
3.2	Função de perda $0 - a_0 - a_1 - b_0 - b_1$ . . . . .	17
4.1	Tamanhos amostrais e riscos para os testes de Bayes unilaterais - Modelo Normal. . .	33
4.2	Função de perda $0 - L_1$ . . . . .	34
4.3	Função de perda $0 - L_1^{(p)}$ . . . . .	35
4.4	Tamanhos amostrais e riscos para os testes de Bayes unilaterais com perda $L_1$ - Modelo Normal. . . . .	40
4.5	Exemplo de cálculo dos riscos de Bayes de cada teste para $n = 10$ . . . . .	43
4.6	Tamanhos amostrais e riscos para os testes de Bayes unilaterais - Modelo Bernoulli. .	45
4.7	Tamanhos amostrais e riscos para os testes de Bayes intervalares - Modelo Normal. .	50
4.8	Tamanhos amostrais e riscos para os testes de Bayes intervalares - Modelo Poisson. .	54
4.9	Tamanhos amostrais e riscos para os testes baseados em estimador por região - Mo- delo Normal univariado. . . . .	59
4.10	Tamanhos amostrais e riscos para os testes baseados em estimador por região - Mo- delo Normal bivariado. . . . .	67



# Capítulo 1

## Introdução

### 1.1 Considerações Preliminares

Testes de hipóteses compõem um dos pilares da inferência estatística juntamente com estimação pontual e intervalar. Apesar de extenso debate na literatura sobre o abuso na utilização de testes de hipóteses—com problemas que incluem: ênfase indevida em atingir significância estatística gerando vieses, dicotomização desnecessária ao tirar conclusões sobre teorias científicas (Meehl [32] e McShane et al. [31]) e fonte de má interpretação de resultados em trabalhos científicos por pesquisadores (Greenland et al. [18])—seguem sendo uma das principais ferramentas da Estatística para se realizar inferências na prática da investigação científica e mais recentemente no ambiente corporativo, principalmente em empresas de tecnologia (ver Kohavi [23]).

O objetivo dos testes de hipóteses consiste em testar afirmações a respeito de um estado da natureza desconhecido, aqui representado pelo parâmetro  $\theta$ , cujo conhecimento sobre seu valor numérico pode ser importante para fins científicos, informativos ou para tomadas de decisão. Tal parâmetro pode representar numericamente a eficácia de um medicamento, a porcentagem da população de um país com intenção de voto para um determinado candidato em uma eleição ou uma grandeza física de interesse em um experimento, por exemplo.

A formulação mais difundida de testes de hipóteses (Neyman e Pearson [34]) consiste em abordá-los como um problema de decisão em que temos duas possibilidades de ação: rejeitar uma hipótese  $H$  de interesse ou não rejeitar  $H$ . Porém, uma outra abordagem ao mesmo tipo de problema envolve adicionar uma terceira possibilidade de ação, a de ficar agnóstico, ou seja, se abster de uma decisão definitiva em relação à hipótese  $H$ . Testes desse tipo são denominados por Esteves et al. [17] de testes de hipóteses agnósticos.

Como será visto adiante, a abordagem de utilizar três possíveis decisões em testes de hipóteses é mencionada em algumas das referências clássicas de Inferência Estatística, porém não é muito explorada e nem contrastada com a abordagem tradicional de duas decisões, em termos de diferenças de resultados e propriedades. Mais recentemente, Izbicki e Esteves [19] mostram algumas limitações de testes de duas decisões e Esteves et al. [17] exploram algumas vantagens dos testes agnósticos bastante interessantes, principalmente em termos lógicos, quando desejamos testar diferentes hipóteses de forma simultânea para o mesmo conjunto de dados observados—por exemplo, quando desejamos testar se os coeficientes de um modelo linear são todos iguais a zero ( $\beta_1 = \beta_2 = 0$ ) e, simultaneamente, testar se apenas um deles é igual a zero ( $\beta_1 = 0$ ). Tendo isso em vista, o foco deste trabalho é explorar outras potenciais vantagens dos testes de hipóteses agnósticos. Mais precisamente, vamos derivá-los em alguns problemas usuais de testes de hipóteses e implementar computacionalmente o cálculo (estimativa) do risco de Bayes de tais testes, confrontando-o com o correspondente risco de testes tradicionais de duas decisões, sob certas condições. Finalmente, os riscos calculados são utilizados no problema de determinação do tamanho de amostra ótimo.

## 1.2 Objetivos

Os objetivos do presente trabalho englobam:

- demonstrar que, dadas algumas restrições, o risco de Bayes de um teste de hipótese agnóstico para uma hipótese nula específica é menor que o risco de Bayes do teste de hipótese tradicional correspondente;
- estudar alguns exemplos de diferentes tipos de testes de hipóteses para alguns modelos estatísticos, implementando-os através de rotinas computacionais com o propósito de calcular e comparar os riscos de Bayes para os cenários tradicional e agnóstico;
- utilizar os riscos de Bayes calculados para resolver um problema de determinação do tamanho amostral ótimo, a partir de um método específico, a fim de contrastar os testes tradicionais e agnósticos.

## 1.3 Organização do Trabalho

No Capítulo 2, serão revisadas algumas ideias fundamentais da Teoria da Decisão. Serão apresentados os principais elementos que compõem um problema de decisão, o Princípio da Maximização da Utilidade Esperada e a definição de Risco de Bayes. Testes de hipóteses serão definidos no Capítulo 3 e, na sequência, o foco consistirá em apresentar e discutir alguns tipos de testes sob o paradigma bayesiano. Adicionalmente, testes de hipóteses agnósticos serão definidos de forma geral, juntamente com as versões agnósticas dos testes bayesianos tradicionais apresentados anteriormente. No Capítulo 4, será demonstrado o principal resultado deste trabalho em termos teóricos e um método de cálculo de tamanho amostral será exposto. Também serão elaborados estudos abordando diferentes tipos de testes bayesianos para alguns modelos estatísticos importantes, com a finalidade de comparar os testes tradicionais e agnósticos em termos de risco de Bayes e tamanhos amostrais ótimos. O impacto da escolha de diferentes prioris também será explorado nos mesmos exemplos. Por fim, no Capítulo 5, serão destacadas as conclusões deste trabalho e algumas possibilidades de temas a serem explorados no futuro. O Apêndice A expõe duas demonstrações adicionais cujos resultados são utilizados no trabalho.

## Capítulo 2

# Teoria da Decisão

Como indivíduos tomamos decisões diariamente, sejam elas triviais, como escolher qual roupa vestir para ir ao trabalho, ou complexas, como aceitar ou não uma nova proposta de trabalho. Similarmente, instituições e empresas precisam tomar, com frequência, decisões extremamente complexas que geram consequências na sociedade como um todo, como a decisão de um Banco Central sobre o aumento da taxa de juros ou a aprovação de um novo medicamento por um órgão regulador de um determinado país.

Um problema de decisão pode ser definido informalmente como o processo em que um ou mais agentes decisores precisam escolher uma ação dentre duas ou mais possíveis, em que cada ação pode gerar consequências diferentes, muitas vezes dependentes de fatores desconhecidos pelos agentes. Tais escolhas podem ser feitas de forma racional, intuitiva ou até inconsciente. Porém, para decisões importantes de grande impacto, é de interesse dos agentes abordá-las de forma racional. Uma questão que surge é como definir se uma decisão é racional ou não, questão essa que será formalizada e detalhada mais adiante.

Um dos principais focos da Teoria da Decisão é estudar e formalizar como um agente, diante de uma tomada de decisão, deve escolher a melhor ação possível de forma racional e coerente, dadas todas as restrições e incertezas definidas pelo próprio agente e impostas pela realidade. É uma área do conhecimento que faz uso de conceitos e métodos principalmente dos campos da Estatística e da Economia, com destaque para os trabalhos de Ramsey [38] com uma abordagem mais conceitual e von Neumann e Morgenstern [33] com uma abordagem mais formal e sob uma perspectiva econômica. No campo da Estatística, possui uma relação importante, dentre outras, com o pensamento Bayesiano e, conseqüentemente, com a utilização dos conceitos de probabilidade subjetiva e utilidade, que será a abordagem utilizada neste trabalho.

### 2.1 Elementos de um problema de decisão

Problemas de decisão frequentemente envolvem uma complexidade grande, com diversas possibilidades de ações, consequências difíceis de serem avaliadas e dependências de diversos fatores desconhecidos e incertos. Assim como na Estatística e em todas as ciências de forma geral, o responsável pelo estudo do problema em questão pode se utilizar de construções matemáticas convenientes e formular um modelo de forma a simplificar o problema e tentar representar de forma aproximada a realidade, tornando-o solucionável.

O problema de decisão central deste trabalho trata de situações em que as consequências de cada ação possível dependem de um estado da natureza, que é desconhecido a priori, mas sobre o qual é possível obter informações, sob certas condições, através de um ou mais experimentos por um custo pré-determinado, o que permitirá que o agente escolha de forma lógica a melhor ação (Raiffa e Schlaifer [37]). Matematicamente, o modelo que descreve tal problema de decisão pode ser formulado a partir de alguns elementos centrais, definidos a seguir.

Definimos um estado da natureza de interesse, representado pelo parâmetro  $\theta$ , que pode ser unidimensional ou multidimensional. Este parâmetro é desconhecido e pertence a um espaço pa-

ramétrico  $\Theta$  adequado, que contém todos os possíveis valores para esse estado da natureza. Uma possível decisão a ser escolhida pelo agente decisor é denotada por  $d$ . Todas as possíveis decisões compõem o conjunto  $\mathcal{D}$ . Os dados obtidos nos experimentos, que trarão informação sobre  $\theta$ , são denotados pelo vetor  $\mathbf{X} = (X_1, X_2, \dots, X_n)$ , que representa uma amostra de tamanho  $n$ . O espaço amostral, que contém todas as possíveis realizações do experimento, é denotado por  $\mathcal{X}$ , podendo ser discreto ou contínuo. Para lidar com a incerteza do problema, introduzimos um espaço de probabilidade representado pela tripla  $(\Theta \times \mathcal{X}, \bar{\mathcal{F}}, \bar{\mathcal{P}})$ , que tem o intuito de representar numericamente a incerteza a respeito de  $\theta$  e  $\mathbf{X}$  antes e depois da realização do experimento. E finalmente, temos uma função  $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$ , que representa numericamente as consequências—seja esta uma recompensa ou prejuízo—de cada decisão  $d \in \mathcal{D}$  e estado da natureza  $\theta \in \Theta$ . Essa função é chamada de função utilidade ou perda na literatura, dependendo do contexto em que é utilizada. A seguir, dois dos conceitos centrais na Teoria da Decisão Bayesiana são detalhados.

### 2.1.1 Probabilidade subjetiva

Um dos pilares da Teoria da Decisão Bayesiana é o conceito de probabilidade subjetiva. “A ideia central da probabilidade subjetiva é que a probabilidade de um evento reflita numericamente a crença pessoal de um indivíduo sobre a chance de esse evento ocorrer” (Berger [3]). Essa ideia contrasta com a definição de probabilidade sob a perspectiva frequentista, que define probabilidade de um evento como sendo o limite da frequência relativa deste evento em uma sequência infinita de repetições de experimentos independentes, sempre sob as mesmas condições.

Para a especificação das probabilidades dos eventos de interesse de forma subjetiva, parte-se do pressuposto que um indivíduo seja capaz de avaliar e realizar comparações sobre as crenças de cada evento. Ao comparar dois eventos  $A$  e  $B$ , por exemplo, o indivíduo deve ter uma crença tal que  $A$  é mais, menos ou igualmente provável à  $B$ . A partir dessa premissa e de outros axiomas, DeGroot [16] e Savage [42], entre outros, formalizam e constroem a ideia de probabilidade subjetiva de forma a garantir que exista uma única medida de probabilidade que corresponde à determinada relação de crença entre eventos, determinada pelo indivíduo. Outras construções são propostas por outros autores, com destaque para de Finetti [15], que formaliza o conceito e a quantificação de probabilidades subjetivas através de uma associação com cenários hipotéticos de apostas envolvendo dinheiro, nos quais o indivíduo tem interesse de ser coerente e para tanto, quantifica suas preferências entre eventos de forma a não perder dinheiro, ou seja, suas probabilidades são definidas de forma a evitar que o indivíduo não perca dinheiro com certeza.

Uma das motivações da construção da probabilidade subjetiva é formalizar como um indivíduo deve agir de forma consistente diante de incertezas. As construções citadas acima viabilizam a utilização do cálculo de probabilidades usual para a representação de incerteza de cada indivíduo. No contexto dos problemas de decisão, definimos o espaço de probabilidade  $(\Theta \times \mathcal{X}, \bar{\mathcal{F}}, \bar{\mathcal{P}})$ , no qual estamos interessados na distribuição conjunta do parâmetro  $\theta$  e da amostra  $\mathbf{X}$ . Para realizarmos inferências a respeito do parâmetro desconhecido  $\theta$ , utilizamos o teorema de Bayes decompondo a função de densidade conjunta  $f_{\theta, \mathbf{X}}(\theta, \mathbf{x})$  nas funções priori  $f_{\theta}(\theta)$  e verossimilhança  $f_{\mathbf{X}|\theta}(\mathbf{x}|\theta)$ , com o objetivo de calcular a função posteriori  $f_{\theta|\mathbf{X}}(\theta|\mathbf{x})$ —que sob a perspectiva bayesiana contém toda a informação necessária sobre  $\theta$  após a revelação da amostra  $\mathbf{x}$ . Doravante, a fim de simplificar a notação, omitiremos os subscritos das funções acima, de modo que o argumento de  $f$  descreva de qual distribuição estamos nos referindo. Assim, o cálculo da função posteriori  $f(\theta|\mathbf{x})$  é dado por

$$f(\theta|\mathbf{x}) = \frac{f(\mathbf{x}|\theta)f(\theta)}{f(\mathbf{x})} = \frac{f(\mathbf{x}|\theta)f(\theta)}{\int_{\Theta} f(\mathbf{x}|\theta)f(\theta)d\theta}. \quad (2.1)$$

De forma resumida, este é o ferramental para a representação de incerteza a respeito do estado da natureza nos problemas de decisão sob a perspectiva bayesiana. Através de escolhas adequadas das funções de priori e de verossimilhança para o problema em questão, dados são coletados através de experimentos e realizamos inferências a respeito do estado da natureza aplicando a operação bayesiana descrita acima.

### 2.1.2 Função de utilidade ou perda

Outro conceito fundamental na Teoria da Decisão é o de função de utilidade, que é definida a fim de representar numericamente as preferências de um agente entre diferentes consequências das ações ou decisões. Utilidade tem o intuito de representar tudo que é relevante para o agente em termos de consequências das possíveis ações. Não somente valor financeiro das consequências, mas também outros fatores muitas vezes intangíveis aparentemente como reputação e qualidade de vida, por exemplo. Como escolher uma função que reflita todos esses fatores depende do problema estudado e é uma das dificuldades que surgem na aplicação da Teoria da Decisão, mas tal questão não será o foco deste trabalho. É comum a utilização de funções mais simples, tratáveis do ponto de vista matemático e este será o caso aqui.

A ideia de utilidade como conhecida hoje é bastante antiga, sendo discutida em trabalhos de Daniel Bernoulli no século 18. Bernoulli propõe que o princípio da esperança matemática, em que o indivíduo escolhe a opção com maior valor esperado em termos monetários numa aposta, por exemplo, não consiste na melhor abordagem ao problema. O autor sugere que um princípio mais adequado é o da maximização da utilidade esperada, em que utilidade poderia corresponder ao valor moral da escolha, que englobaria não só o valor financeiro mas também outros fatores importantes para o indivíduo tomando a decisão em questão. Uma discussão sobre as principais ideias propostas por Bernoulli e alguns exemplos relacionados ao conceito de utilidade pode ser encontrada em Savage [42].

As primeiras formalizações modernas de utilidade foram estabelecidas por Ramsey [38] e von Neumann e Morgenstern [33]. A primeira, revolucionária do ponto de vista conceitual e matematicamente incompleta (Bernardo e Smith [6]), desenvolve os conceitos de probabilidade e utilidade subjetivas de forma simultânea. A segunda formaliza de forma axiomática o conceito de utilidade isoladamente, em que as probabilidades são tratadas como conhecidas. Posteriormente, outros autores como DeGroot [16] e Savage [42] também formalizam a construção da função de utilidade e suas propriedades em um contexto bayesiano. A partir de axiomas similares aos da formalização da probabilidade subjetiva, porém agora se referindo a preferências entre ações ao invés de preferências entre eventos, tais construções garantem a existência de uma função de utilidade que teoricamente descreve o comportamento de um indivíduo racional, definido como aquele que age de forma a maximizar sua utilidade esperada.

**Definição 2.1.** *Seja um parâmetro  $\theta \in \Theta$  que representa o estado da natureza de interesse e um conjunto de decisões  $\mathcal{D}$ . A função de utilidade é uma função  $U : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$  que associa ao par  $(d, \theta)$  um valor que determina a recompensa para cada decisão  $d \in \mathcal{D}$  e estado da natureza possível  $\theta \in \Theta$ .*

Na Estatística, de forma geral, é comum a utilização do termo perda ao invés de utilidade, mais frequente no campo da Economia. Tal terminologia é motivada também pelo tipo de problema de decisão enfrentado, ou seja, há decisões em que faz mais sentido tratar de prejuízos em vez de recompensas, e vice-versa. Dependendo do contexto e autor, a função pode ter premissas e propriedades distintas, mas essencialmente o objetivo é o mesmo, quantificar preferências entre ações. Neste trabalho, foi escolhido introduzir o conceito como função de utilidade pois as principais construções axiomáticas são realizadas utilizando essa terminologia. Porém, daqui em diante, o termo de função de perda será o preponderante.

Após a escolha da função de utilidade  $U$ , DeGroot [16] define a função de perda  $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$  como

$$L(\theta, d) = -U(\theta, d). \quad (2.2)$$

## 2.2 Princípio da Maximização da Utilidade Esperada e Risco de Bayes

A partir das construções axiomáticas de probabilidade subjetiva e utilidade—que podem ser interpretados como condições que garantem consistência e coerência na avaliação de decisões que dependem de incertezas—temos que um indivíduo que segue tais axiomas age como se existisse uma função de utilidade (perda) que descreve seu comportamento frente às decisões disponíveis, e sua escolha da melhor decisão é determinada pela maximização (minimização) do valor esperado dessa função.

Para as definições abaixo, adaptadas de DeGroot [16], considere um problema de decisão em que temos possíveis decisões  $d \in \mathcal{D}$ , o espaço de probabilidade  $(\Theta, \mathcal{F}, \mathcal{P})$  que representa a incerteza a respeito do estado da natureza  $\theta$  e uma função de perda  $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$ .

**Definição 2.2** (DeGroot [16]). *A perda esperada ou risco da decisão  $d$  contra a distribuição (a priori)  $\mathcal{P}$  é definida como*

$$\rho(d, \mathcal{P}) = E[L(d, \theta)] = \int_{\Theta} L(d, \theta) d\mathcal{P}(\theta). \quad (2.3)$$

**Definição 2.3** (DeGroot [16]). *O risco de Bayes é definido como o menor risco dentre todos os possíveis riscos  $\rho(d, \mathcal{P})$  associados às decisões  $d \in \mathcal{D}$ . Logo, escrevemos o risco de Bayes como*

$$\rho^*(\mathcal{P}) = \inf_{d \in \mathcal{D}} \rho(d, \mathcal{P}). \quad (2.4)$$

**Definição 2.4** (DeGroot [16]). *Toda decisão  $d^*$  em que o risco  $\rho(d^*, \mathcal{P})$  é igual ao risco de Bayes é denominada decisão de Bayes contra a distribuição  $\mathcal{P}$ .*

### 2.2.1 Obtenção do risco de Bayes em problemas com observações

Agora, considere o cenário em que observações serão coletadas com o propósito de obter informações a respeito do estado da natureza de forma a dar suporte à tomada de decisão. Nesse caso, as decisões dependerão dos dados obtidos denotados por  $\mathbf{x}$  e, assim, define-se uma função de decisão  $\varphi : \mathcal{X} \rightarrow \mathcal{D}$ , que associa cada valor de  $\mathbf{x} \in \mathcal{X}$  a uma decisão  $d \in \mathcal{D}$ . A classe de todas as funções de decisão possíveis  $\varphi$  é denotada por  $\Delta$ . São definidos também um novo espaço de probabilidade  $(\Theta \times \mathcal{X}, \bar{\mathcal{F}}, \bar{\mathcal{P}})$ , em que  $\bar{\mathcal{P}}$  é a medida de probabilidade que representa a incerteza sobre  $(\theta, \mathbf{X})$ , e uma nova função de perda  $\bar{L} : \mathcal{D} \times (\Theta \times \mathcal{X}) \rightarrow \mathbb{R}$  de modo que  $\bar{L}(\varphi, (\theta, \mathbf{x})) = L(\varphi(\mathbf{x}), \theta)$ , onde  $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$  é como discutida na seção anterior. A partir das definições acima e, assumindo que  $(\theta, \mathbf{X})$  é contínuo (para simplificar a notação e sem perda de generalidade), temos que a perda esperada (ou risco)  $\bar{\rho}$  da função de decisão  $\varphi$  é dada por:

$$\bar{\rho}(\varphi, \bar{\mathcal{P}}) = E[L(\varphi(\mathbf{X}), \theta)] = \int_{\Theta} \int_{\mathcal{X}} L(\varphi(\mathbf{x}), \theta) d\bar{\mathcal{P}}(\theta, \mathbf{X}) = \int_{\Theta} \int_{\mathcal{X}} L(\varphi(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta, \quad (2.5)$$

em que  $f(\theta, \mathbf{x})$  é a função densidade de probabilidade conjunta de  $(\theta, \mathbf{X})$ .

Para a obtenção da função de decisão  $\varphi^* \in \Delta$  que minimiza (maximiza) a perda (utilidade) esperada, devemos então encontrar o risco de Bayes. Segundo Raiffa e Schlaifer [37], existem duas formas equivalentes de encontrar a função de decisão que minimiza a perda esperada do agente decisor, a forma normal e a forma extensiva.

Na forma normal, a análise é realizada de forma a se avaliar os riscos de todas as possíveis funções de decisão e determinar aquela que minimiza a perda esperada, ou seja, uma função de decisão de Bayes  $\varphi^*$ . Já na forma extensiva, uma decisão de Bayes é obtida condicionalmente a cada  $\mathbf{x} \in \mathcal{X}$ , fazendo com que não seja necessário avaliar os riscos de todas as possíveis funções de



decisão explicitamente. Com efeito, a partir da equação (2.5) e assumindo que a ordem das integrais pode ser invertida, temos que

$$\begin{aligned}\bar{\rho}(\varphi, \bar{\mathcal{P}}) &= E[L(\varphi(\mathbf{X}), \theta)] = \int_{\Theta} \int_{\mathcal{X}} L(\varphi(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta = \int_{\mathcal{X}} \int_{\Theta} L(\varphi(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\theta d\mathbf{x} \\ &= \int_{\mathcal{X}} \left\{ \int_{\Theta} L(\varphi(\mathbf{x}), \theta) f(\theta|\mathbf{x}) d\theta \right\} f(\mathbf{x}) d\mathbf{x}.\end{aligned}\quad (2.6)$$

A partir daí, a minimização é realizada na integral entre chaves, encontrando a chamada decisão de Bayes a posteriori  $d_x^*$ , condicional à observação de  $\mathbf{x} \in \mathcal{X}$ . Denotando a integral entre chaves por  $\rho_x(\varphi(\mathbf{x}), \mathcal{P}_x)$ , em que  $\mathcal{P}_x$  é a medida de probabilidade (a posteriori) para  $\theta$  obtida após o condicionamento em  $\mathbf{x}$ , temos que,  $\forall \varphi \in \Delta$ ,

$$\begin{aligned}\bar{\rho}(\varphi, \bar{\mathcal{P}}) &= \int_{\mathcal{X}} \left\{ \int_{\Theta} L(\varphi(\mathbf{x}), \theta) f(\theta|\mathbf{x}) d\theta \right\} f(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathcal{X}} \rho_x(\varphi(\mathbf{x}), \mathcal{P}_x) f(\mathbf{x}) d\mathbf{x} \\ &\geq \int_{\mathcal{X}} \rho_x(d_x^*, \mathcal{P}_x) f(\mathbf{x}) d\mathbf{x} \\ &= \int_{\mathcal{X}} \left\{ \int_{\Theta} L(\varphi^*(\mathbf{x}), \theta) f(\theta|\mathbf{x}) d\theta \right\} f(\mathbf{x}) d\mathbf{x} \\ &= \bar{\rho}(\varphi^*, \bar{\mathcal{P}}),\end{aligned}\quad (2.7)$$

onde, para cada  $\mathbf{x} \in \mathcal{X}$ ,  $d_x^*$  é tal que

$$\rho_x(d_x^*, \mathcal{P}_x) = \inf_{d \in \mathcal{D}} \rho_x(d, \mathcal{P}_x) \quad (2.8)$$

e  $\varphi^* : \mathcal{X} \rightarrow \mathcal{D}$  é dada por  $\varphi^*(\mathbf{x}) = d_x^*$ .

Assim, a função  $\varphi^*$  definida acima é uma decisão de Bayes, uma vez que minimiza (2.5). Note que, na construção de  $\varphi^*$  não é necessária a avaliação de todas as funções de decisão possíveis como na forma normal. Logo, a forma extensiva costuma ser a mais utilizada das duas abordagens.



## Capítulo 3

# Testes de Hipóteses

Um dos precursores na formalização e utilização de testes de hipóteses a partir de uma perspectiva frequentista foi Ronald Fisher, com a ideia de teste de significância estatística (Cox [11]) em que o foco é medir a compatibilidade entre uma hipótese nula e os dados observados, sem se preocupar com a formulação de uma hipótese alternativa. Também sob a perspectiva frequentista, porém com algumas diferenças em termos filosóficos e metodológicos, Jerzy Neyman e Egon Pearson propuseram a principal e mais disseminada abordagem para a realização de testes de hipóteses, em que tanto a hipótese nula quanto a hipótese alternativa possuem importância. Já no campo bayesiano, vale mencionar o trabalho pioneiro de Jeffreys [20] com a aplicação de testes utilizando o *Fator de Bayes*, abordagem que não será apresentada aqui, mas que possui importância na literatura bayesiana e que continua sendo aplicada na prática científica. Uma interessante comparação entre as três abordagens acima pode ser encontrada em Berger [4]. Apesar de haver maior disseminação e aplicação dos testes de hipóteses em um contexto frequentista, o foco do presente trabalho será somente no tratamento bayesiano para esse tipo de problema.

Um teste de hipóteses pode ser considerado como um caso particular de um problema de decisão, no qual as possíveis ações se referem a afirmar algo sobre o parâmetro desconhecido  $\theta$ . Para a formulação do problema de teste, definimos um espaço paramétrico  $\Theta$  e duas hipóteses (afirmações sobre o parâmetro)  $H_0$  e  $H_1$ —denominadas de hipótese nula e hipótese alternativa, respectivamente—da seguinte forma:  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$ , em que  $\Theta_0$  e  $\Theta_1$  formam uma partição de  $\Theta$ , isto é,  $\Theta = \Theta_0 \cup \Theta_1$  e  $\Theta_0 \cap \Theta_1 = \emptyset$ . Os conjuntos  $\Theta_0$  e  $\Theta_1$  podem ambos conter um único ponto ( $H_0 : \theta = \theta_0$  e  $H_1 : \theta = \theta_1$ ), caso que chamamos  $H_0$  e  $H_1$  de hipóteses simples. Caso contrário, as hipóteses são denominadas compostas, por exemplo, com hipóteses da forma  $H_0 : \theta \leq \theta_0$  e  $H_1 : \theta > \theta_0$ . Apesar de a partição do espaço paramétrico em dois subconjuntos ser a mais utilizada, vale mencionar que não é uma restrição e que é possível particionar  $\Theta$  em mais de dois subconjuntos.

Tradicionalmente, os testes de hipóteses possibilitam duas decisões: rejeição de  $H_0$  e não-rejeição (ou aceitação) de  $H_0$ . Tais testes serão revisados na seção 3.1. Porém, na seção 3.2 será apresentada uma outra alternativa de teste de hipóteses em que será adicionada mais uma possibilidade de decisão: ficar agnóstico sobre  $H_0$ . Neste trabalho, testes de hipóteses com duas opções de decisão serão denominados testes de hipóteses tradicionais ou apenas testes de hipóteses. Os testes com três alternativas serão denominados de testes de hipóteses agnósticos.

### 3.1 Testes de Hipóteses Bayesianos

Conforme visto no capítulo anterior, o paradigma bayesiano faz uso do conceito de probabilidade subjetiva e tem como ideia central combinar a incerteza pré-experimento sobre  $\theta$ , através da priori, com a informação dos dados observados por meio da função de verossimilhança para atualizar a incerteza sobre  $\theta$ . Devido ao fato de nessa abordagem tratarmos o parâmetro  $\theta$  como uma variável aleatória, ao contrário do paradigma frequentista que o trata como fixo, faz sentido calcularmos e nos referirmos à probabilidade de uma hipótese  $H_0$  ser verdadeira ou falsa. Segundo Schervish [43], outra diferença grande entre as abordagens frequentista e bayesiana é a importância da função

de perda para a formulação dos testes de hipóteses. No tratamento frequentista, o foco em geral é encontrar testes que atendam a um determinado critério (por exemplo, que simultaneamente minimizam a probabilidade de erro de tipo II e não excedem certo nível quanto à probabilidade do erro de tipo I), muitas vezes não importando qual a função de perda escolhida. Já no tratamento bayesiano, a função de perda é essencial pois influencia diretamente o teste resultante, uma vez que a perda esperada (ou risco) posterior depende desta.

A definição a seguir, adaptada de Schervish [43], formaliza o problema de teste de hipóteses sob a perspectiva da Teoria da Decisão.

**Definição 3.1** (Schervish [43]). *Sejam o parâmetro  $\theta \in \Theta$  e as hipóteses  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$  em que  $\Theta_0$  e  $\Theta_1$  formam uma partição de  $\Theta$ . Um problema de decisão é chamado de teste de hipóteses quando temos um espaço de decisão  $\mathcal{D} = \{0, 1\}$  e uma função de perda  $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$  tal que:  $L(1, \theta) > L(0, \theta)$  para  $\theta \in \Theta_0$  e  $L(1, \theta) < L(0, \theta)$  para  $\theta \in \Theta_1$ . A decisão  $d = 1$  é entendida como o ato de rejeitar a hipótese  $H_0$  e a decisão  $d = 0$  como o ato de aceitar, ou não rejeitar, a hipótese  $H_0$ .*

Analogamente, em um contexto em que observamos dados, representados por  $\mathbf{x}$ , definimos um teste como uma função de decisão  $\varphi : \mathcal{X} \rightarrow \{0, 1\}$  em que  $\varphi(\mathbf{x}) = 1$  representa rejeitar a hipótese  $H_0$  e  $\varphi(\mathbf{x}) = 0$  representa aceitar, ou não rejeitar, a hipótese  $H_0$ . Existem pontos de vista divergentes quanto à interpretação da segunda possibilidade de decisão: se a hipótese  $H_0$  deve ser aceita ou simplesmente não rejeitada. A fim de simplificar, vamos nos ater à terminologia de aceitar a hipótese  $H_0$  daqui em diante. O conjunto de pontos do espaço amostral que resultam na rejeição de  $H_0$  é chamado de região crítica ou região de rejeição do teste. Consequentemente, o conjunto complementar a este é chamado de região de aceitação do teste.

A função de perda  $L$  para um teste de hipótese pode ser definida de diversas formas. Como temos duas possíveis decisões e duas afirmações ou hipóteses sobre o estado da natureza, existem quatro possíveis consequências ao tomarmos uma decisão. Podemos tomar decisões corretas, ao rejeitarmos a hipótese  $H_0$  quando esta é de fato falsa ou ao aceitarmos  $H_0$  quando esta é de fato verdadeira. Usualmente, atribui-se o valor zero para a função de perda para esses dois cenários, já que não cometemos erros ao tomarmos tais decisões. Adicionalmente, temos outros dois cenários possíveis, a saber: rejeitamos a hipótese  $H_0$  quando esta é verdadeira ou aceitamos a hipótese  $H_0$  quando esta é falsa. Tradicionalmente, no primeiro cenário, dizemos que cometemos um erro do tipo I e, no segundo, um erro do tipo II. Para esses tipos de erros associa-se uma perda maior que zero, pois pela definição, perdas associadas a erros são maiores que perdas associadas a decisões corretas. Uma possível função de perda simples, chamada de perda  $0 - a_0 - a_1$  (com  $a_0 > 0$  e  $a_1 > 0$ ), é dada por

$$L(d, \theta) = \begin{cases} 0, & \text{se } d = \mathbb{1}_{\Theta_1}(\theta) \\ a_0, & \text{se } d = 1 \text{ e } \theta \in \Theta_0 \\ a_1, & \text{se } d = 0 \text{ e } \theta \in \Theta_1 \end{cases}, \quad (3.1)$$

em que  $\mathbb{1}_A$  representa a função indicadora  $\mathbb{1}_A(\theta) = \begin{cases} 1, & \theta \in A \\ 0, & \theta \notin A \end{cases}$ , onde  $A$  é um subconjunto de  $\Theta$ .

Alguns autores chamam esse tipo de perda na literatura é  $0 - 1 - c$ , que é equivalente ao caso acima se tomarmos  $c = \frac{a_1}{a_0}$  e outros preferem escrever  $0 - a_1 - a_2$ , onde  $a_i$  é a penalidade por incorrer no erro do tipo  $i$ ,  $i = 1, 2$ . A função de perda  $0 - a_0 - a_1$  também pode ser representada pela tabela abaixo.

$L(d, \theta)$	$\theta \in \Theta_0$	$\theta \in \Theta_1$
$d = 0$	0	$a_1$
$d = 1$	$a_0$	0

**Tabela 3.1:** Função de perda  $0 - a_0 - a_1$ .

Outras funções de perda mais complexas—que dependem da distância de  $\theta$  à hipótese, por exemplo—podem ser construídas de acordo com o tipo de problema. Um exemplo de função de perda desse tipo será apresentado mais adiante. Porém, aqui o foco será principalmente na perda  $0 - a_0 - a_1$ .

Com a função de perda escolhida, as hipóteses definidas e fixada uma probabilidade para  $(\theta, \mathbf{X})$ , podemos calcular a perda esperada ou risco do teste  $\varphi$  conforme a definição dada no capítulo anterior. O exemplo abaixo ilustra o cálculo do risco para a perda  $0 - a_0 - a_1$ .

**Exemplo 3.1.** *Seja o parâmetro  $\theta \in \Theta$ , os dados observados  $\mathbf{x} \in \mathcal{X}$  e  $f(\theta, \mathbf{x})$  a função de densidade conjunta de  $(\theta, \mathbf{X})$  (associada à probabilidade  $\bar{\mathcal{P}}$ ). Deseja-se testar a hipótese  $H_0 : \theta \in \Theta_0$  contra  $H_1 : \theta \in \Theta_1$ . Para a função de decisão  $\varphi : \mathcal{X} \rightarrow \{0, 1\}$ , sob perda  $0 - a_0 - a_1$ , o risco é dado por*

$$\begin{aligned} \bar{\rho}(\varphi, \bar{\mathcal{P}}) &= \int_{\Theta} \int_{\mathcal{X}} L(\varphi(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\ &= 0 \int_{\Theta} \int_{\mathcal{X}} \mathbb{1}_{\Theta_0}(\theta) \mathbb{1}_{\varphi^{-1}(\{0\})}(\mathbf{x}) f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\ &\quad + a_0 \int_{\Theta} \int_{\mathcal{X}} \mathbb{1}_{\Theta_0}(\theta) \mathbb{1}_{\varphi^{-1}(\{1\})}(\mathbf{x}) f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\ &\quad + a_1 \int_{\Theta} \int_{\mathcal{X}} \mathbb{1}_{\Theta_1}(\theta) \mathbb{1}_{\varphi^{-1}(\{0\})}(\mathbf{x}) f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\ &\quad + 0 \int_{\Theta} \int_{\mathcal{X}} \mathbb{1}_{\Theta_1}(\theta) \mathbb{1}_{\varphi^{-1}(\{1\})}(\mathbf{x}) f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\ &= a_0 \int_{\Theta_0} \int_{\varphi^{-1}(\{1\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta + a_1 \int_{\Theta_1} \int_{\varphi^{-1}(\{0\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta. \end{aligned} \quad (3.2)$$

Como será visto a seguir, uma forma de derivar um teste de hipótese bayesiano é através da forma extensiva já descrita anteriormente, em que o risco a posteriori é minimizado para cada  $\mathbf{x} \in \mathcal{X}$ . Porém, existem outras abordagens bayesianas importantes para obtenção de testes que também serão apresentadas neste capítulo.

### 3.1.1 Testes baseados em Probabilidade a Posteriori

O teste baseado em probabilidade a posteriori é o teste mais intuitivo e natural dentro de um contexto bayesiano. Tal teste é derivado através da minimização do risco posterior sob a perda  $0 - a_0 - a_1$  e informalmente consiste em, após observação dos dados, calcular a probabilidade a posteriori da hipótese  $H_0$  de interesse e verificar se o valor obtido é maior ou menor que um limiar que depende da função de perda  $0 - a_0 - a_1$  escolhida.

**Proposição 3.1.** *Sejam as hipóteses  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$ , em que  $\Theta_0$  e  $\Theta_1$  formam uma partição de  $\Theta$ . A função do teste baseado em probabilidade a posteriori para testar  $H_0$  contra  $H_1$  com função de perda  $0 - a_0 - a_1$  é dada por*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \text{se } \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) > \frac{a_1}{a_0 + a_1} \\ 1, & \text{se } \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq \frac{a_1}{a_0 + a_1} \end{cases}. \quad (3.3)$$

*Demonstração.* Através da forma extensiva de análise, a decisão  $d_x^* = 1$  de rejeitar a hipótese  $H_0$  é tomada ao observar  $\mathbf{x}$  se, e somente se, o risco a posteriori de rejeitar  $H_0$  for menor ou igual ao

risco a posteriori da decisão de aceitar  $H_0$ . Logo, temos que

$$\begin{aligned}
d_x^* = 1 &\iff \rho_x(1, \mathcal{P}_x) \leq \rho_x(0, \mathcal{P}_x) \\
&\iff \int_{\Theta} L(1, \theta) f(\theta | \mathbf{x}) d\theta \leq \int_{\Theta} L(0, \theta) f(\theta | \mathbf{x}) d\theta \\
&\iff \int_{\Theta_0} a_0 f(\theta | \mathbf{x}) d\theta \leq \int_{\Theta_1} a_1 f(\theta | \mathbf{x}) d\theta \\
&\iff a_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq a_1 \mathbb{P}(\theta \in \Theta_1 | \mathbf{x}) \\
&\iff a_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq a_1 [1 - \mathbb{P}(\theta \in \Theta_0 | \mathbf{x})] \\
&\iff \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq \frac{a_1}{a_0 + a_1}.
\end{aligned}$$

Consequentemente,

$$d_x^* = 0 \iff \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) > \frac{a_1}{a_0 + a_1}.$$

□

Vale notar que há uma dificuldade prática de aplicar testes baseados em probabilidade a posteriori no contexto bayesiano quando o conjunto especificado por  $H_0$  possui medida de Lebesgue nula. Isso se deve ao fato de, no caso de espaços paramétricos contínuos, a probabilidade a priori de uma hipótese especificada por um conjunto de medida nula ser sempre zero (admitindo probabilidade absolutamente contínua para  $\theta$ ), resultando, sob certas condições, em uma probabilidade a posteriori também igual a zero. Consequentemente, nesse caso a hipótese  $H_0$  é sempre rejeitada em testes baseados em probabilidade a posteriori. A seguir, veremos uma alternativa para lidar com esse tipo de problema.

### 3.1.2 Full Bayesian Significance Test

Um problema de teste de hipóteses bastante controverso e que é tema de extensa discussão na literatura é o teste de hipóteses precisas, em que pelo menos uma das hipóteses tem medida zero. Geralmente, no caso unidimensional, as hipóteses precisas são da forma  $H_0 : \theta = \theta_0$ , em que  $\theta_0 \in \Theta$ . Além da dificuldade mencionada no final da seção anterior, há certa controvérsia sobre testes de tais hipóteses na investigação científica. Casella e Berger [8] e Berger [3] afirmam que testes de hipóteses precisas são comumente utilizados de forma equivocada, já que, em muitos cenários, testar se o parâmetro  $\theta$  é exatamente igual a um valor  $\theta_0$  não faz sentido em termos práticos. Entretanto, Robert [39] menciona algumas situações em que é concebível testar uma afirmação de tal forma, como no caso de testar se uma distribuição de mistura possui dois ou três componentes, e Stern [46] argumenta que algumas das hipóteses científicas mais relevantes são de fato precisas.

Uma solução para contornar as dificuldades práticas de lidar com hipóteses precisas consiste na utilização de uma distribuição a priori mista para  $\theta$ , em que  $\mathbb{P}(\theta = \theta_0) > 0$  (ver Schervish [43] e Robert [39], por exemplo). Outra solução típica envolve reformular a hipótese nula com a abordagem de regiões de equivalência prática (Kruschcke [24] e Spiegelhalter et al. [45]), em que a hipótese nula é definida, por exemplo, como  $H_0 : \theta \in [\theta_0 - \epsilon, \theta_0 + \epsilon]$ , onde  $\epsilon$  é normalmente um valor pequeno que depende do problema estudado. Uma discussão sobre testes de hipóteses precisas e algumas aplicações sob a perspectiva bayesiana podem ser encontradas em Berger e Delampady [5].

Outra forma de testar hipóteses precisas foi proposta por Pereira e Stern [14] com a criação do *Full Bayesian Significance Test* (FBST). Para a formulação do FBST é definida uma medida de evidência que tem o objetivo de quantificar a consistência entre os dados observados e uma hipótese de interesse, de forma análoga à utilização do p-valor nos testes de hipóteses frequentistas. A seguir, definiremos a medida de evidência e o FBST.

**Definição 3.2.** Considere o parâmetro  $\theta \in \Theta$ , os dados observados  $\mathbf{x} \in \mathcal{X}$  e a função de densidade a posteriori  $f(\theta|\mathbf{x})$ . Seja a hipótese  $H_0 : \theta \in \Theta_0$ . Seja  $T_x$  o seguinte subconjunto do espaço paramétrico

$$T_x = \left\{ \theta \in \Theta : f(\theta|\mathbf{x}) > \sup_{\theta \in \Theta_0} f(\theta|\mathbf{x}) \right\}, \quad (3.4)$$

ou seja, a região composta pelos pontos do espaço paramétrico que possuem densidade posterior maior que qualquer ponto pertencente à hipótese nula  $H_0$ . A medida de evidência de Pereira-Stern é definida como

$$Ev(\Theta_0, \mathbf{x}) = 1 - \mathbb{P}(\theta \in T_x | \mathbf{x}). \quad (3.5)$$

O procedimento do teste FBST consiste em rejeitar a hipótese nula  $H_0$  quando  $Ev(\Theta_0, \mathbf{x})$  é suficientemente “pequena” ou aceitar  $H_0$  quando  $Ev(\Theta_0, \mathbf{x})$  é suficientemente “grande”.

Algumas propriedades e vantagens do FBST, do ponto de vista teórico e prático, são descritas em Pereira, Stern e Wechsler [13] e Madruga, Pereira e Stern [29]. Um fato interessante, demonstrado por Madruga, Esteves e Wechsler [30], é a possibilidade de construir o FBST a partir da minimização do risco posterior (assim como o teste baseado em probabilidade a posteriori) quando uma função de perda específica, que depende de  $\mathbf{x}$ , é escolhida.

### 3.1.3 Testes baseados em Estimador por Região

Outra forma de formular um teste de hipótese consiste em construir uma região de confiança (terminologia frequentista) ou credibilidade (terminologia bayesiana) para  $\theta$  e comparar tal região com os subconjuntos do espaço paramétrico especificados por  $H_0$  e  $H_1$ . No caso frequentista temos que, em geral, dada uma região de confiança  $(1 - \alpha)$  para  $\theta$ ,  $R(1 - \alpha) \subseteq \Theta$ , um teste de hipótese de nível  $\alpha$  para  $H_0 : \theta \in \Theta_0$  fica determinado ao se decidir pela rejeição de  $H_0$  se  $\Theta_0 \not\subseteq R(1 - \alpha)$  e pela aceitação caso contrário (Casella e Berger [9]).

Analogamente, para o caso bayesiano, podemos propor um teste para a mesma hipótese precisa  $H_0 : \theta = \theta_0$ , no qual a hipótese é aceita se  $\theta_0$  pertence a uma região de credibilidade particular, e rejeitada caso contrário. Testes desse tipo são citados por Lindley [27], em que a região particular utilizada é a região denominada de *Highest Posterior Density* (HPD), cuja definição é dada a seguir, juntamente com as definições de região de credibilidade e do teste proposto acima.

**Definição 3.3.** Seja o parâmetro  $\theta \in \Theta$  com distribuição a posteriori  $f(\theta|\mathbf{x})$ . Uma região  $R \subseteq \Theta$  tal que

$$\int_R f(\theta|\mathbf{x}) d\theta \geq 1 - \alpha \quad (3.6)$$

é denominada região de credibilidade  $1 - \alpha$  para  $\theta$ , em relação a posteriori  $f(\theta|\mathbf{x})$ , em que  $\alpha \in (0, 1)$ . Em particular, a região  $R'$  de credibilidade  $1 - \alpha$  é chamada de *Highest Posterior Density* (HPD) se  $R'$  é da forma

$$R' = \{\theta \in \Theta : f(\theta|\mathbf{x}) \geq k(\alpha)\}, \quad (3.7)$$

em que  $k(\alpha) \geq 0$  é a maior constante  $k$  tal que

$$\int_{\{\theta: f(\theta|\mathbf{x}) \geq k\}} f(\theta|\mathbf{x}) \geq 1 - \alpha. \quad (3.8)$$

**Definição 3.4.** *Considere o parâmetro  $\theta \in \Theta \subseteq \mathbb{R}$  e as hipóteses  $H_0 : \theta = \theta_0$  e  $H_1 \neq \theta_0$ . Seja  $R_x(1 - \alpha) \subseteq \Theta$  uma região HPD de credibilidade  $1 - \alpha$  em relação a posteriori  $f(\theta|\mathbf{x})$ . O teste baseado em estimador por região é da forma*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \text{se } \theta_0 \in R_x(1 - \alpha) \\ 1, & \text{se } \theta_0 \notin R_x(1 - \alpha) \end{cases} . \quad (3.9)$$

Diferentemente de testes baseados em probabilidade a posteriori e em fatores de Bayes, testes baseados em estimadores por região no paradigma bayesiano não são muito discutidos nas principais referências de Teoria Bayesiana, possivelmente por serem criticados por não possuírem justificativa do ponto de vista da Teoria da Decisão (Thulin [48]). Entretanto, Thulin [48] fornece uma possível justificativa em termos de Teoria da Decisão e também destaca algumas vantagens de tais testes para hipóteses precisas como: desnecessidade de utilização de prioris mistas para garantir que  $H_0$  tenha probabilidade positiva a priori e ausência de discrepância entre soluções bayesianas e frequentistas conforme mais dados são coletados—fato que pode ocorrer em testes bayesianos mais comuns. Além disso, testes baseados em estimador por região são destacados aqui por possuírem vantagens práticas e também por sua importância teórica. Do ponto de vista teórico, veremos adiante que testes baseados em estimadores de região possuem propriedades interessantes, especificamente no caso agnóstico. Adicionalmente, pode-se demonstrar que existem algumas equivalências entre testes construídos dessa forma e testes de outros tipos, o que garante que esses outros testes tenham as mesmas propriedades dos testes baseados em estimador por região. Stern et al. [47], por exemplo, mostram que o teste FBST com  $H_0 : \theta_0 \in \Theta_0$  e  $H_1 : \theta_0 \in \Theta_1$ , no qual  $H_0$  é rejeitada se  $Ev(H_0, \mathbf{x}) \leq c$  e aceita caso contrário, é equivalente a um teste baseado em estimador por região da forma

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \text{se } R_x(1 - \alpha) \cap \Theta_0 \neq \emptyset \\ 1, & \text{se } R_x(1 - \alpha) \subseteq \Theta_1 \end{cases} , \quad (3.10)$$

em que  $R_x(1 - \alpha)$  é uma região HPD de credibilidade  $1 - \alpha$ .

Do ponto de vista prático, testes baseados em estimador por região são uma alternativa interessante, pois podem ser computacionalmente mais simples de implementar (ao menos para  $\Theta$  unidimensional) e mais fáceis de interpretar. Há casos em que encontrar a região de credibilidade específica é menos trabalhoso do que derivar e encontrar a região de rejeição de um teste (por exemplo, em um teste baseado em probabilidade a posteriori).

Vale notar também que é possível generalizar esse tipo de teste para hipóteses não precisas, de modo que uma hipótese da forma  $H_0 : \theta \in A$ ,  $A \subset \Theta$ , é aceita se a intersecção entre a região HPD e o conjunto  $A$  é não-vazia, por exemplo. A definição abaixo faz tal generalização.

**Definição 3.5.** *Sejam o espaço amostral  $\mathcal{X}$ , as hipóteses  $H_0 : \theta \in A$ ,  $H_1 : \theta \in A^c$  e  $R$  um estimador por região. Um teste de hipóteses baseado em estimador por região  $R$  para  $H_0$  contra  $H_1$  é dado por*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \text{se } R \cap A \neq \emptyset \\ 1, & \text{se } R \subseteq A^c \end{cases} . \quad (3.11)$$

Como se vê na definição acima, os testes baseados em estimador por região não se restringem à regiões de HPD ou regiões de credibilidade em geral, Outros estimadores por região (frequentistas ou bayesianos) podem ser construídos e testes análogos ao FBST, por exemplo, podem ser elaborados. Via de regra, métodos específicos de construção de regiões são escolhidos por critérios de otimalidade estatística. No contexto bayesiano, um possível critério é encontrar a região que possui o menor intervalo—ou volume, no caso de  $\Theta$  multidimensional—dentre todas as regiões possíveis de mesma credibilidade. Tal solução resulta justamente na região de HPD, para o caso particular de posterioris unimodais (Casella e Berger [9]). Outra possibilidade consiste em incorporar o “tamanho” da região em uma função de perda (de forma que regiões maiores são mais penalizadas) e encontrar a região



ótima que minimiza tal perda. Exemplos de funções de perda desse tipo podem ser encontrados em Robert [39] e Aitchison [1].

### 3.2 Testes de Hipóteses Agnósticos

Em um artigo de 1933, mesmo ano em que publicaram um dos seus principais trabalhos sobre testes de hipóteses, Neyman e Pearson [35] descrevem um teste estatístico, em que a hipótese  $H_0$  está sendo testada, como uma regra do tipo

- rejeitar  $H_0$  se  $T(X)$  pertencer à uma região  $r$ ;
- aceitar  $H_0$  se  $T(X)$  pertencer à uma região  $r'$ ;
- ficar em dúvida se  $T(X)$  pertencer à uma região  $r''$ ;

em que  $r$ ,  $r'$  e  $r''$  formam uma partição do espaço amostral  $\mathcal{X}$  e  $T(X)$  representa alguma função da amostra observada.

Testes de hipóteses em que é adicionada mais uma possibilidade de decisão também são mencionados em referências como Lehmann [26], Kaiser [22], Berger [3] e Robert [39]. Nas duas primeiras referências, a terceira possibilidade de decisão se refere a quando temos um teste em que queremos decidir não somente se duas médias são iguais, por exemplo, mas também qual das médias é maior—por exemplo, com três hipóteses  $H_0 : \theta_1 = \theta_2$ ,  $H_1 : \theta_1 < \theta_2$  e  $H_2 : \theta_1 > \theta_2$ . Já nas duas últimas referências, a terceira decisão adicional é equivalente à ideia de Neyman e Pearson acima, ou seja, se refere a ficar em dúvida em relação à hipótese de interesse  $H_0$ .

Apesar das menções a essa possibilidade, a comparação entre testes de duas e de três possíveis decisões—chamados aqui de tradicional e agnóstico, respectivamente—não é tão explorada nas referências citadas até aqui, tanto de teoria bayesiana quanto de testes de hipóteses de forma geral. No entanto, vale mencionar que problemas de decisão com três possíveis ações são uma das ideias centrais na área de Análise Sequencial, tema bastante explorado na Estatística e que tem como um dos trabalhos pioneiros o de Wald [49]. Inclusive, Berger [3] aborda de maneira detalhada a teoria e exemplos da Análise Sequencial sob a perspectiva bayesiana. Na Análise Sequencial, os problemas são formulados de forma que observações são coletadas uma a uma em um experimento e, a cada observação coletada, uma das três possíveis decisões (aceitação, rejeição ou coleta de mais dados) é tomada, sendo que o experimento é encerrado caso o teste resulte em aceitação ou rejeição da hipótese de interesse. Porém, apesar das similaridades, o foco do presente trabalho não é na construção e na problemática da Análise Sequencial, e sim na comparação entre os dois tipos de teste (tradicional e agnóstico), principalmente em termos de suas propriedades e aplicações em um contexto de experimentos não sequenciais, que consiste em um tema menos explorado de forma geral.

Mais recentemente, Berg [2] apresenta a alternativa de teste de três decisões para o caso em que são testadas hipóteses simples, citando desvantagens do teste tradicional e algumas vantagens do teste agnóstico, com destaque para o controle simultâneo dos erros de tipo I e II, fato que não ocorre no testes tradicionais sob a perspectiva frequentista. Essa vantagem de controle dos dois tipos de erros também é abordada de forma mais geral em Coscrato, Izbicki e Stern [10].

Izbicki e Esteves [19] e Silva et al. [12] apresentam e discutem propriedades desejáveis (monotonicidade, consonância da união, consonância da intersecção e invertibilidade) que garantem que testes aplicados de forma simultânea sejam consistentes em termos lógicos, demonstrando que os únicos testes tradicionais de duas decisões que satisfazem todas as propriedades simultaneamente são os testes baseados em estimadores pontuais que, em geral, não satisfazem critérios de otimalidade estatística.

De forma resumida e informal, a propriedade de monotonicidade garante que se uma hipótese  $A$  implica outra hipótese  $B$ , então uma classe de testes de hipóteses que rejeita  $B$ , deve rejeitar a hipótese  $A$  também. A consonância da união estabelece que se uma classe de testes de hipóteses rejeita cada uma das hipóteses  $A$  e  $B$  separadamente, a hipótese  $A \cup B$  deve ser rejeitada. Já

a consonância da intersecção estabelece que se uma classe de testes rejeita a hipótese  $A \cap B$ , então deveria rejeitar pelo menos uma das hipóteses  $A$  ou  $B$ , quando testadas separadamente. E finalmente, a propriedade de invertibilidade garante que se uma classe de testes rejeita a hipótese  $A$ , a hipótese complementar  $A^c$  deveria ser aceita.

A partir daí, Esteves et al. [17] mostram que testes de hipóteses agnósticos podem satisfazer todas essas propriedades lógicas e, adicionalmente, atender determinados critérios de otimalidade estatística, dadas algumas restrições. Além disso, Yao [50] também explora outras vantagens de problemas de decisão com três possíveis decisões em relação a problemas com duas possibilidades de decisão, sobretudo em termos de redução da perda esperada. Tal redução é uma das motivações do presente trabalho e será apresentada de forma mais generalizada que em Yao [50].

À primeira vista, incluir uma decisão em que necessariamente cometemos um erro e que não concluimos algo a respeito de  $H_0$  e  $H_1$  parece não ser muito desejável. Entretanto, como visto acima, essa alternativa pode trazer vantagens do ponto de vista lógico e estatístico. Por exemplo, para alguns tipos de problemas, tomar alguma decisão definitiva a respeito das hipóteses e cometer um erro de decisão pode ter um custo bem mais alto do que ficar em dúvida e decidir coletar mais dados para se chegar a uma análise mais conclusiva. Uma possível situação em que isso ocorre é no caso de uma decisão de aprovação de um novo medicamento de alto custo para uma doença não tão grave, de modo que os pacientes possam aguardar mais evidências sem muitos prejuízos em termos de saúde. Este cenário pode ser elaborado através de restrições em relação à função de perda, conforme será especificado abaixo.

A seguir, definimos formalmente testes de hipóteses agnósticos e perdas próprias, conceito definido em Esteves et al. [17] que possui relevância no contexto de testes agnósticos.

**Definição 3.6.** *Sejam o parâmetro  $\theta \in \Theta$  e as hipóteses  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$  em que  $\Theta_0$  e  $\Theta_1$  formam uma partição de  $\Theta$ . Um problema de decisão é chamado de teste de hipóteses agnóstico quando temos um espaço de decisão  $\mathcal{D} = \{0, \frac{1}{2}, 1\}$  e uma função de perda  $L : \mathcal{D} \times \Theta \rightarrow \mathbb{R}$  tal que:  $L(1, \theta) > L(0, \theta)$  e  $L(\frac{1}{2}, \theta) > L(0, \theta)$  para  $\theta \in \Theta_0$ ;  $L(1, \theta) < L(0, \theta)$  e  $L(\frac{1}{2}, \theta) < L(0, \theta)$  para  $\theta \in \Theta_1$ . A decisão  $d = 1$  é entendida como o ato de rejeitar a hipótese  $H_0$ , a decisão  $d = 0$  como o ato de aceitar a hipótese  $H_0$  e a decisão  $d = \frac{1}{2}$  como o ato de ficar agnóstico em relação às hipóteses  $H_0$  e  $H_1$ .*

**Definição 3.7** (Esteves et al. [17]). *Considere o parâmetro  $\theta \in \Theta$  e as hipóteses  $H_0 : \theta \in A$ ,  $H_1 : \theta \notin A$ ,  $A \subset \Theta$ . A função de perda  $L_A : \{0, \frac{1}{2}, 1\} \times \Theta \rightarrow \mathbb{R}$  é própria se*

$$\begin{cases} L_A(0, \theta) < L_A(\frac{1}{2}, \theta) < L_A(1, \theta), & \text{se } \theta \in A \\ L_A(0, \theta) > L_A(\frac{1}{2}, \theta) > L_A(1, \theta), & \text{se } \theta \notin A \\ L_A(\frac{1}{2}, \theta) < \frac{L_A(0, \theta) + L_A(1, \theta)}{2}, & \text{para } \theta \in \Theta \end{cases} \quad (3.12)$$

A definição de perda própria estabelece que o erro associado à decisão de ficar agnóstico é maior do que de tomar a decisão correta, mas é menor do que os erros de tipo I e II. Além disso, a última condição pode ser interpretada como afirmar que ficar agnóstico sobre a hipótese  $A$  é mais desejável do que sortear de forma aleatória simples entre as decisões de rejeitar e aceitar a hipótese  $A$ , segundo Esteves et al. [17].

Analogamente à função de perda  $0 - a_0 - a_1$  apresentada anteriormente para testes tradicionais, podemos construir uma função de perda simples, para testar  $H_0 : \theta \in \Theta_0$  contra  $H_1 : \theta \in \Theta_1$  no caso agnóstico, representada pela tabela abaixo.

$L(d, \theta)$	$\theta \in \Theta_0$	$\theta \in \Theta_1$
$d = 0$	0	$a_1$
$d = \frac{1}{2}$	$b_0$	$b_1$
$d = 1$	$a_0$	0

**Tabela 3.2:** Função de perda  $0 - a_0 - a_1 - b_0 - b_1$ .

Para que a perda da Tabela 3.2 seja própria, devemos ter que  $0 < b_0 < \frac{a_0}{2}$  e  $0 < b_1 < \frac{a_1}{2}$ . A seguir, calculamos o risco esperado do teste agnóstico  $\varphi$  com a perda acima.

**Exemplo 3.2.** *Sejam o parâmetro  $\theta \in \Theta$ , os dados observados  $\mathbf{x} \in \mathcal{X}$  e  $f(\theta, \mathbf{x})$  a função de densidade conjunta de  $(\theta, \mathbf{X})$  (associada a probabilidade  $\bar{\mathcal{P}}$ ). Deseja-se testar a hipótese  $H_0 : \theta \in \Theta_0$  contra  $H_1 : \theta \in \Theta_1$  através da função de teste agnóstico  $\varphi : \mathcal{X} \rightarrow \{0, \frac{1}{2}, 1\}$  com a perda  $0 - a_0 - a_1 - b_0 - b_1$ . A perda esperada de tal teste é dada por*

$$\begin{aligned}
\bar{\rho}(\varphi, \bar{\mathcal{P}}) &= \int_{\Theta} \int_{\mathcal{X}} L(\varphi(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\
&= 0 \int_{\Theta_0} \int_{\varphi^{-1}(\{0\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta + b_0 \int_{\Theta_0} \int_{\varphi^{-1}(\{\frac{1}{2}\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\
&\quad + a_0 \int_{\Theta_0} \int_{\varphi^{-1}(\{1\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta + a_1 \int_{\Theta_1} \int_{\varphi^{-1}(\{0\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\
&\quad + b_1 \int_{\Theta_1} \int_{\varphi^{-1}(\{\frac{1}{2}\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta + 0 \int_{\Theta_1} \int_{\varphi^{-1}(\{1\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\
&= a_0 \int_{\Theta_0} \int_{\varphi^{-1}(\{1\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta + a_1 \int_{\Theta_1} \int_{\varphi^{-1}(\{0\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\
&\quad + b_0 \int_{\Theta_0} \int_{\varphi^{-1}(\{\frac{1}{2}\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta + b_1 \int_{\Theta_1} \int_{\varphi^{-1}(\{\frac{1}{2}\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta. \tag{3.13}
\end{aligned}$$

A ideia de testes agnósticos é aplicável tanto na abordagem frequentista quanto na abordagem bayesiana. Porém, assim como nos testes de hipóteses tradicionais, trataremos apenas da abordagem bayesiana. No restante dessa seção definiremos alguns tipos de testes agnósticos que serão estudados no próximo capítulo. Tais definições são adaptadas de Esteves et al. [17].

De forma análoga ao caso tradicional, o teste bayesiano baseado em probabilidade a posteriori no caso agnóstico é derivado via Teoria da Decisão pela minimização do risco posterior da função de perda da Tabela 3.2. Veremos esses detalhes a seguir.

### 3.2.1 Testes Agnósticos baseados em Probabilidade a Posteriori

A definição de testes de hipóteses baseados em probabilidade a posteriori, adaptada de Esteves et al. [17], é dada abaixo.

**Definição 3.8** (Esteves et al. [17]). *Considere o espaço amostral  $\mathcal{X}$  e as hipóteses  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$ . Sejam  $c_1, c_2 \in (0, 1)$ , com  $c_1 \geq c_2$ . O teste de hipóteses agnóstico baseado em probabilidade a posteriori para  $H_0$  contra  $H_1$  é a função  $\varphi : \mathcal{X} \rightarrow \{0, \frac{1}{2}, 1\}$  dada por*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) > c_1 \\ \frac{1}{2}, & c_2 < \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq c_1 \\ 1, & \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq c_2 \end{cases} . \tag{3.14}$$

Logo, segundo a função de decisão  $\varphi$ , aceitamos a hipótese de interesse  $A$  se sua probabilidade posterior for maior que o limiar  $c_1$ , rejeitamos a hipótese caso a probabilidade posterior de  $A$  for

menor que o limiar  $c_2$  e ficamos agnósticos caso contrário.

A seguir, verificamos que testes agnósticos como em (3.14) são ótimos sob a perda da Tabela 3.2.

**Proposição 3.2.** *Sejam as hipóteses  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$ , em que  $\Theta_0$  e  $\Theta_1$  formam uma partição de  $\Theta$ . O teste agnóstico de Bayes considerando a perda própria  $0 - a_0 - a_1 - b_0 - b_1$  é o teste baseado em probabilidade a posteriori dado por*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ \frac{1}{2}, & \frac{b_1}{a_0 + b_1 - b_0} < \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ 1, & \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq \frac{b_1}{a_0 + b_1 - b_0} \end{cases} . \quad (3.15)$$

*Demonstração.* A decisão  $d_x^* = 1$  de rejeitar a hipótese  $H_0$  é tomada ao observar  $\mathbf{x}$  se, e somente se, o risco a posteriori de rejeitar  $H_0$  for menor ou igual ao risco a posteriori da decisão de aceitar  $H_0$  e menor ou igual ao risco a posteriori da decisão de ficar agnóstico. Logo, temos que

$$\begin{aligned} d_x^* = 1 &\iff \rho_x(1, \mathcal{P}_x) \leq \rho_x(0, \mathcal{P}_x) \quad \text{e} \quad \rho_x(1, \mathcal{P}_x) \leq \rho_x\left(\frac{1}{2}, \mathcal{P}_x\right) \\ &\iff \int_{\Theta} L(1, \theta) f(\theta | \mathbf{x}) d\theta \leq \int_{\Theta} L(0, \theta) f(\theta | \mathbf{x}) d\theta \quad \text{e} \\ &\quad \int_{\Theta} L(1, \theta) f(\theta | \mathbf{x}) d\theta \leq \int_{\Theta} L\left(\frac{1}{2}, \theta\right) f(\theta | \mathbf{x}) d\theta \\ &\iff \int_{\Theta_0} a_0 f(\theta | \mathbf{x}) d\theta \leq \int_{\Theta_1} a_1 f(\theta | \mathbf{x}) d\theta \quad \text{e} \\ &\quad \int_{\Theta_0} a_0 f(\theta | \mathbf{x}) d\theta \leq \int_{\Theta_0} b_0 f(\theta | \mathbf{x}) d\theta + \int_{\Theta_1} b_1 f(\theta | \mathbf{x}) d\theta \\ &\iff a_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq a_1 \mathbb{P}(\theta \in \Theta_1 | \mathbf{x}) \quad \text{e} \\ &\quad a_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq b_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) + b_1 \mathbb{P}(\theta \in \Theta_1 | \mathbf{x}) \\ &\iff a_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq a_1 [1 - \mathbb{P}(\theta \in \Theta_0 | \mathbf{x})] \quad \text{e} \\ &\quad a_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq b_0 \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) + b_1 [1 - \mathbb{P}(\theta \in \Theta_0 | \mathbf{x})] \\ &\iff \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq \frac{a_1}{a_0 + a_1} \quad \text{e} \quad \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq \frac{b_1}{a_0 + b_1 - b_0}. \end{aligned}$$

Dadas as restrições de perda própria, ou seja, que  $0 < b_0 < \frac{a_0}{2}$  e  $0 < b_1 < \frac{a_1}{2}$ , temos que  $\frac{b_1}{a_0 + b_1 - b_0} < \frac{a_1}{a_0 + a_1}$ . Assim,

$$d_x^* = 1 \iff \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) \leq \frac{b_1}{a_0 + b_1 - b_0}.$$

Similarmente, para a decisão  $d_x^* = 0$  temos

$$\begin{aligned}
d_x^* = 0 &\iff \rho_x(0, \mathcal{P}_x) < \rho_x(1, \mathcal{P}_x) \quad \text{e} \quad \rho_x(0, \mathcal{P}_x) < \rho_x\left(\frac{1}{2}, \mathcal{P}_x\right) \\
&\iff \int_{\Theta} L(0, \theta) f(\theta|\mathbf{x}) d\theta < \int_{\Theta} L(1, \theta) f(\theta|\mathbf{x}) d\theta \quad \text{e} \\
&\quad \int_{\Theta} L(0, \theta) f(\theta|\mathbf{x}) d\theta < \int_{\Theta} L\left(\frac{1}{2}, \theta\right) f(\theta|\mathbf{x}) d\theta \\
&\iff \int_{\theta \in \Theta_1} a_1 f(\theta|\mathbf{x}) d\theta < \int_{\theta \in \Theta_0} a_0 f(\theta|\mathbf{x}) d\theta \quad \text{e} \\
&\quad \int_{\theta \in \Theta_1} a_1 f(\theta|\mathbf{x}) d\theta < \int_{\theta \in \Theta_0} b_0 f(\theta|\mathbf{x}) d\theta + \int_{\theta \in \Theta_1} b_1 f(\theta|\mathbf{x}) d\theta \\
&\iff a_1 \mathbb{P}(\theta \in \Theta_1|\mathbf{x}) < a_0 \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) \quad \text{e} \\
&\quad a_1 \mathbb{P}(\theta \in \Theta_1|\mathbf{x}) < b_0 \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) + b_1 \mathbb{P}(\theta \in \Theta_1|\mathbf{x}) \\
&\iff a_1 [1 - \mathbb{P}(\theta \in \Theta_0|\mathbf{x})] < a_0 \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) \quad \text{e} \\
&\quad a_1 [1 - \mathbb{P}(\theta \in \Theta_0|\mathbf{x})] < b_0 \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) + b_1 [1 - \mathbb{P}(\theta \in \Theta_0|\mathbf{x})] \\
&\iff \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) > \frac{a_1}{a_0 + a_1} \quad \text{e} \quad \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\
&\iff \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) > \frac{a_1 - b_1}{a_1 + b_0 - b_1}.
\end{aligned}$$

E finalmente,

$$d_x^* = \frac{1}{2} \iff \frac{b_1}{a_0 + b_1 - b_0} < \mathbb{P}(\theta \in \Theta_0|\mathbf{x}) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1}.$$

□

### 3.2.2 Agnostic Full Bayesian Significance Test

A adaptação do FBST ao caso agnóstico é semelhante à adaptação feita nos testes baseados em probabilidade a posteriori, em que é adicionada a terceira ação de ficar agnóstico, com a diferença de que utilizamos a medida de evidência de Pereira-Stern em relação à hipótese nula  $H_0$  ao invés da probabilidade a posteriori da hipótese. Assim como no caso tradicional, o FBST agnóstico é uma opção para se testar hipóteses precisas para evitar o problema de medida nula da hipótese.

**Definição 3.9** (Esteves et al. [17]). *Considere o parâmetro  $\theta \in \Theta$ , os dados observados  $\mathbf{x} \in \mathcal{X}$  e a hipótese nula  $H_0 : \theta \in \Theta_0$ . Seja  $Ev(\Theta_0, \mathbf{x})$  a medida de Pereira-Stern referente à hipótese  $H_0$ . Sejam  $c_1, c_2 \in (0, 1)$ , com  $c_1 \geq c_2$ . O teste FBST agnóstico é a função  $\varphi : \mathcal{X} \rightarrow \{0, \frac{1}{2}, 1\}$  dada por*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & Ev(\Theta_0, \mathbf{x}) > c_1 \\ \frac{1}{2}, & c_2 < Ev(\Theta_0, \mathbf{x}) \leq c_1 \\ 1, & Ev(\Theta_0, \mathbf{x}) \leq c_2 \end{cases} . \quad (3.16)$$

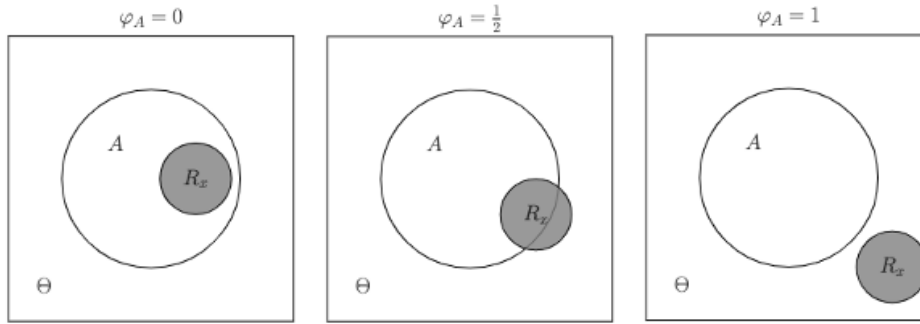
### 3.2.3 Testes Agnósticos baseados em Estimador por Região e GFBST

Conforme mencionado na seção anterior, uma das motivações para a construção de testes baseados em estimador por região é o fato de estes possuírem propriedades teóricas interessantes, particularmente no caso agnóstico. Esteves et al. [17] demonstram que, quando desejamos realizar testes simultâneos de várias hipóteses, testes agnósticos são logicamente consistentes se, e somente se, são baseados em estimador por região. Testes desse tipo são definidos abaixo.

**Definição 3.10** (Esteves et al. [17]). *Seja o espaço amostral  $\mathcal{X}$  e as hipóteses  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$ , em que  $\Theta_0$  e  $\Theta_1$  formam uma partição de  $\Theta$ . O teste de hipóteses agnóstico baseado no estimador por região  $R_x : \mathcal{X} \rightarrow \mathbb{P}(\Theta)$  é a função  $\varphi : \mathcal{X} \rightarrow \{0, \frac{1}{2}, 1\}$  dada por*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \text{se } R_x \subseteq \Theta_0 \\ 1, & \text{se } R_x \subseteq \Theta_1 \\ \frac{1}{2}, & \text{caso contrário} \end{cases} . \quad (3.17)$$

Tal definição é bastante intuitiva e é ilustrada na figura abaixo. Note que no caso de espaços paramétricos contínuos, não é possível aceitar uma hipótese  $\Theta_0$  que seja precisa, assim como nos testes baseados em probabilidade posterior. A região  $R$  pode ser baseada em um estimador qualquer, porém é interessante escolher uma região de confiança ou credibilidade com boas propriedades estatísticas, a fim de garantir concomitantemente consistência lógica, no caso de testes simultâneos de várias hipóteses, e otimalidade estatística para o teste.



**Figura 3.1:** Teste agnóstico baseado em estimador por região.

Outra motivação para a formulação dos testes baseados em estimador por região é a existência de equivalências entre esses testes e outros tipos de testes. Tal fato pode facilitar a interpretação e implementação prática de outros tipos de testes e também garantir a consistência lógica de testes que sejam equivalentes aos testes baseados em estimador por região. Coscrato, Izbicki e Stern [10] exploram algumas equivalências entre testes agnósticos unilaterais e bilaterais e testes agnósticos baseados em estimador por região de confiança, sob uma perspectiva frequentista.

Sob a perspectiva bayesiana, Stern et al. [47] constroem uma generalização do teste FBST para o caso agnóstico denominada de *Generalized Full Bayesian Significance Teste* (GFBST), definida a seguir.

**Definição 3.11** (Stern et al. [47]). *Considere as hipóteses  $H_0 : \theta_0 \in \Theta_0$  e  $H_1 : \theta_0 \in \Theta_1$ , em que  $\Theta_0$  e  $\Theta_1$  formam uma partição de  $\Theta$ . Seja  $R_x(1 - \alpha)$  a região HPD de probabilidade  $1 - \alpha$ . O teste GFBST para testar  $H_0$  contra  $H_1$  é a função  $\varphi : \mathcal{X} \rightarrow \{0, \frac{1}{2}, 1\}$  dada por*

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \text{se } R_x(1 - \alpha) \subseteq \Theta_0 \\ 1, & \text{se } R_x(1 - \alpha) \subseteq \Theta_1 \\ \frac{1}{2}, & \text{caso contrário} \end{cases} . \quad (3.18)$$

Note que o GFBST é um teste baseado em estimador por região que utiliza a região HPD, a mais conhecida no contexto bayesiano. Segundo Stern et al. [47], o teste GFBST também pode ser escrito da seguinte forma:

$$\varphi(\mathbf{x}) = \begin{cases} 0, & \text{se } Ev(\Theta_1, \mathbf{x}) \leq \alpha \\ 1, & \text{se } Ev(\Theta_0, \mathbf{x}) \leq \alpha \\ \frac{1}{2}, & \text{caso contrário} \end{cases}, \quad (3.19)$$

em que  $Ev(\Theta_0, \mathbf{x})$  e  $Ev(\Theta_1, \mathbf{x})$  são as medidas de evidência de Pereira-Stern referentes às hipóteses  $H_0$  e  $H_1$ , respectivamente.

Vale mencionar que tal teste utiliza as medidas de evidência de Pereira-Stern referentes às duas hipóteses  $H_0$  e  $H_1$ , ao contrário do teste definido pela função em (3.16), que só leva em consideração a medida de evidência em relação à  $H_0$ .

Alguns dos testes expostos nesse capítulo serão examinados quanto ao risco de Bayes no próximo capítulo.





## Capítulo 4

# Redução do Risco de Bayes e Determinação do Tamanho Amostral

Nesse capítulo, será apresentado o principal resultado deste trabalho, que mostra que o risco de um teste de Bayes agnóstico é menor que o risco de um teste de Bayes tradicional, dadas algumas restrições. Também será destacado um método de determinação de tamanho de amostra ótimo para testes de hipóteses e será demonstrado que o risco total do teste de hipótese incorporando a determinação do tamanho amostral ótimo em sua formulação também é menor no cenário agnóstico. Em seguida, serão apresentados exemplos de diferentes tipos de testes de hipóteses para alguns modelos estatísticos importantes com o objetivo de realizar um estudo comparativo entre os testes de hipóteses tradicionais e agnósticos. Os riscos de Bayes de tais testes, em função do tamanho amostral  $n$ , serão derivados, calculados e comparados através de exemplos numéricos. Além disso, os tamanhos amostrais ótimos serão obtidos pelo método escolhido para cada um dos exemplos, com o intuito de se avaliar se há diferenças expressivas entre os testes tradicionais e os testes agnósticos.

### 4.1 Redução do Risco de Bayes

Aqui apresentamos um teorema que estabelece que o risco de um teste de Bayes agnóstico é sempre menor que o risco de um teste de Bayes tradicional, sob algumas condições. Mais adiante, estudos comparativos serão elaborados de forma a ilustrar numericamente tal redução.

**Teorema 4.1.** *Considere o parâmetro  $\theta \in \Theta$  e a amostra  $\mathbf{x} \in \mathcal{X}$ . Seja  $\varphi_T : \mathcal{X} \rightarrow \{0, 1\}$  um teste de Bayes com relação a função de perda  $L_T : \{0, 1\} \times \Theta \rightarrow \mathbb{R}$ , e com risco de Bayes  $\bar{\rho}_T = \bar{\rho}(\varphi_T, \mathcal{P})$ . Seja  $\varphi_A : \mathcal{X} \rightarrow \{0, \frac{1}{2}, 1\} \times \Theta$  um teste de Bayes agnóstico com relação a função de perda  $L_A : \{0, \frac{1}{2}, 1\} \times \Theta \rightarrow \mathbb{R}$  tal que  $L_A(i, \theta) = L_T(i, \theta)$ , para  $i = 0, 1, \forall \theta \in \Theta$ . O risco de Bayes de tal teste é denotado por  $\bar{\rho}_A = \bar{\rho}(\varphi_A, \mathcal{P})$ . Se  $\mathbb{P}[\varphi_A^{-1}(\{\frac{1}{2}\})] > 0$ , ou seja, se existir probabilidade positiva de ficarmos agnósticos no teste  $\varphi_A$ , então  $\bar{\rho}_T > \bar{\rho}_A$ .*

*Demonstração.* Assumindo sem perda de generalidade que  $\Theta$  e  $\mathcal{X}$  são contínuos, temos que o risco de Bayes do teste tradicional é dado por

$$\begin{aligned}\bar{\rho}_T &= E[L_T(\varphi_T(\mathbf{X}, \theta))] = \int_{\mathcal{X}} \int_{\Theta} L_T(\varphi_T(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\theta d\mathbf{x} \\ &= \int_{\varphi_T^{-1}(\{0\})} \int_{\Theta} L_T(0, \theta) f(\theta|\mathbf{x}) f(\mathbf{x}) d\theta d\mathbf{x} + \int_{\varphi_T^{-1}(\{1\})} \int_{\Theta} L_T(1, \theta) f(\theta|\mathbf{x}) f(\mathbf{x}) d\theta d\mathbf{x} \\ &= \int_{\varphi_T^{-1}(\{0\})} \mathbb{E}[L_T(0, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} + \int_{\varphi_T^{-1}(\{1\})} \mathbb{E}[L_T(1, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x}.\end{aligned}$$

Considerando que o espaço amostral  $\mathcal{X}$  pode ser escrito como união de regiões de intersecção entre  $\varphi_T^{-1}(\{i\})$  e  $\varphi_A^{-1}(\{j\})$  para  $i = 0, 1$  e  $j = 0, \frac{1}{2}, 1$ , podemos escrever o risco de Bayes do teste tradicional

como

$$\bar{\rho}_T = \sum_{i \in \{0,1\}} \sum_{j \in \{0, \frac{1}{2}, 1\}} \left[ \int_{\varphi_T^{-1}(\{i\}) \cap \varphi_A^{-1}(\{j\})} \mathbb{E}[L_T(i, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \right]. \quad (4.1)$$

Note que, se tomarmos um  $\mathbf{x}$  qualquer tal que  $\mathbf{x} \in \varphi_A^{-1}(\{0\})$ , pelo fato dos testes  $\varphi_T$  e  $\varphi_A$  serem testes de Bayes e lembrando que  $L_A(i, \theta) = L_T(i, \theta)$ ,  $i = 0, 1, \forall \theta \in \Theta$ , temos que

$$\begin{aligned} \mathbf{x} \in \varphi_A^{-1}(\{0\}) &\iff \mathbb{E}[L_A(0, \theta) | \mathbf{X} = \mathbf{x}] < \mathbb{E}[L_A(1, \theta) | \mathbf{X} = \mathbf{x}] \\ &\quad \text{e} \quad \mathbb{E}[L_A(0, \theta) | \mathbf{X} = \mathbf{x}] < \mathbb{E}[L_A(1/2, \theta) | \mathbf{X} = \mathbf{x}] \\ &\implies \mathbb{E}[L_A(0, \theta) | \mathbf{X} = \mathbf{x}] < \mathbb{E}[L_A(1, \theta) | \mathbf{X} = \mathbf{x}] \\ &\implies \mathbb{E}[L_T(0, \theta) | \mathbf{X} = \mathbf{x}] < \mathbb{E}[L_T(1, \theta) | \mathbf{X} = \mathbf{x}] \\ &\implies \mathbf{x} \in \varphi_T^{-1}(\{0\}) \end{aligned}$$

Logo,

$$\varphi_A^{-1}(\{0\}) \subseteq \varphi_T^{-1}(\{0\}).$$

Usando o mesmo argumento acima, temos que  $\varphi_A^{-1}(\{1\}) \subseteq \varphi_T^{-1}(\{1\})$ .

Conseqüentemente,  $\varphi_T^{-1}(\{i\}) \cap \varphi_A^{-1}(\{i-1\}) = \emptyset$ ,  $i = 0, 1$ . Assim, o risco de Bayes  $\bar{\rho}_T$  em (4.1) se reduz a

$$\begin{aligned} \bar{\rho}_T &= \int_{\varphi_A^{-1}(\{0\})} \mathbb{E}[L_T(0, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\ &\quad + \int_{\varphi_T^{-1}(\{0\}) \cap \varphi_A^{-1}(\{\frac{1}{2}\})} \mathbb{E}[L_T(0, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\ &\quad + \int_{\varphi_T^{-1}(\{1\}) \cap \varphi_A^{-1}(\{\frac{1}{2}\})} \mathbb{E}[L_T(1, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\ &\quad + \int_{\varphi_A^{-1}(\{1\})} \mathbb{E}[L_T(1, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

E como  $L_A(i, \theta) = L_T(i, \theta)$ ,  $i = 0, 1$ , temos que

$$\begin{aligned} \bar{\rho}_T &= \int_{\varphi_A^{-1}(\{0\})} \mathbb{E}[L_A(0, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\ &\quad + \int_{\varphi_T^{-1}(\{0\}) \cap \varphi_A^{-1}(\{\frac{1}{2}\})} \mathbb{E}[L_A(0, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\ &\quad + \int_{\varphi_T^{-1}(\{1\}) \cap \varphi_A^{-1}(\{\frac{1}{2}\})} \mathbb{E}[L_A(1, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\ &\quad + \int_{\varphi_A^{-1}(\{1\})} \mathbb{E}[L_A(1, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x}. \end{aligned}$$

Adicionalmente, como  $\varphi_A$  é também teste de Bayes agnóstico, vale que

$$\mathbf{x} \in \varphi_A^{-1} \left( \left\{ \frac{1}{2} \right\} \right) \iff \mathbb{E} \left[ L_A \left( \frac{1}{2}, \theta \right) \middle| \mathbf{X} = \mathbf{x} \right] < \mathbb{E} \left[ L_A(i, \theta) \middle| \mathbf{X} = \mathbf{x} \right], \quad i = 0, 1.$$

E assim, supondo que  $\mathbb{P}[\varphi_A^{-1}(\{\frac{1}{2}\})] > 0$ , segue de resultado da Teoria da Medida (Cancelier [7], p.

45) que

$$\begin{aligned}
\bar{\rho}_T &> \int_{\varphi_A^{-1}(\{0\})} \mathbb{E}[L_A(0, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\
&+ \int_{\varphi_T^{-1}(\{0\}) \cap \varphi_A^{-1}(\{\frac{1}{2}\})} \mathbb{E}\left[L_A\left(\frac{1}{2}, \theta\right) \middle| \mathbf{X} = \mathbf{x}\right] f(\mathbf{x}) d\mathbf{x} \\
&+ \int_{\varphi_T^{-1}(\{1\}) \cap \varphi_A^{-1}(\{\frac{1}{2}\})} \mathbb{E}\left[L_A\left(\frac{1}{2}, \theta\right) \middle| \mathbf{X} = \mathbf{x}\right] f(\mathbf{x}) d\mathbf{x} \\
&+ \int_{\varphi_A^{-1}(\{1\})} \mathbb{E}[L_A(1, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} \\
&= \int_{\varphi_A^{-1}(\{0\})} \mathbb{E}[L_A(0, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} + \int_{\varphi_A^{-1}(\{\frac{1}{2}\})} \mathbb{E}\left[L_A\left(\frac{1}{2}, \theta\right) \middle| \mathbf{X} = \mathbf{x}\right] f(\mathbf{x}) d\mathbf{x} \\
&+ \int_{\varphi_A^{-1}(\{1\})} \mathbb{E}[L_A(1, \theta) | \mathbf{X} = \mathbf{x}] f(\mathbf{x}) d\mathbf{x} = \bar{\rho}_A.
\end{aligned}$$

Logo,

$$\bar{\rho}_T > \bar{\rho}_A.$$

□

O que o teorema nos garante é que, se em um teste de Bayes agnóstico temos probabilidade positiva de tomarmos a decisão de ficar agnóstico, o risco de tal teste é estritamente menor que o risco do teste de Bayes tradicional, para perdas que atendem as condições do enunciado. Vale notar que tal resultado é obtido através da suposição de que os testes são derivados a partir da minimização do risco a posteriori.

## 4.2 Determinação de tamanho amostral

Do ponto de vista da Teoria da Decisão Bayesiana, a forma mais direta de determinar o tamanho da amostra ideal para um experimento é através da minimização da perda esperada (ou risco). Porém, existe vasta literatura em que a adequabilidade de tal abordagem em cenários práticos é discutida e também em que outras abordagens bayesianas de cálculo amostral são propostas. Em um contexto de ensaios clínicos, Spiegelhalter et al. [44] menciona algumas das posições em relação ao debate de tratar um experimento clínico como um problema de decisão—especificando uma função de perda—ou de inferência, e também explora outros métodos de cálculo amostral que constituem uma abordagem híbrida, utilizando o ferramental tanto bayesiano quanto frequentista. Kunzmann et al. [25] também exploram e comparam três diferentes métodos híbridos de determinação de tamanho da amostra. Lindley [28] discute vantagens e desvantagens do método de maximização da utilidade esperada (ou minimização da perda esperada) em relação a outros métodos bayesianos propostos como, por exemplo, em Joseph et al. [21]. Apesar de algumas dificuldades em relação a aplicação em problemas reais, a abordagem de minimização da perda esperada será a proposta aqui, por sua simplicidade.

A ideia do método basicamente consiste em encontrar o tamanho da amostra  $n^*$  tal que o risco da função de perda específica é minimizado. Intuitivamente, quanto maior a amostra, mais informações são obtidas e menor o risco das decisões. Entretanto, normalmente temos um custo associado à coleta de uma amostra, de modo que muitas vezes não é factível coletar um número muito grande de observações e um *trade-off* entre perda esperada (em termos inferenciais) e custo de cada amostra se faz necessário. O método leva em consideração tal cenário e é proposto, por exemplo, em DeGroot [16], Raiffa e Schlaifer [37] e, mais recentemente, em Sahu e Smith [41]. Para sua formulação mais geral, considere o mesmo cenário proposto na seção 2.2.1, porém será

considerado adicionalmente o tamanho da amostra como componente da decisão a ser tomada e o termo  $c(n, \mathbf{x}, \theta)$ , que representa o custo associado à amostra  $\mathbf{X} = \mathbf{x}$  de tamanho  $n$  quando  $\Theta = \theta$ , de modo que a perda pode ser da forma

$$\bar{L}((n, \varphi), (\theta, \mathbf{x})) = L(\varphi(\mathbf{x}), \theta) + c(n, \mathbf{x}, \theta). \quad (4.2)$$

Nesse caso, o risco é dado por

$$\begin{aligned} \bar{\rho}((n, \varphi), \bar{\mathcal{P}}) &= E[L(\varphi(\mathbf{X}), \Theta)] + E[c(n, \mathbf{X}, \Theta)] \\ &= \int_{\Theta} \int_{\mathcal{X}} L(\varphi(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta + \int_{\Theta} \int_{\mathcal{X}} c(n, \mathbf{x}, \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta, \end{aligned} \quad (4.3)$$

onde assume-se que as duas esperanças podem ser somadas, ou seja, informalmente a função de perda  $L$  e o custo  $c$  podem ser comparados. Raiffa e Schlaifer [37] discutem em mais detalhes em quais situações tal premissa pode ser aplicada, como em situações em que somente valores monetários estão associados tanto à função de perda quanto ao custo de cada amostra, por exemplo. No geral, em situações reais, o custo não vai depender especificamente do  $\mathbf{x}$  observado e de  $\theta$ , de modo que usualmente define-se uma função simples  $c(n, \mathbf{x}, \theta) = c(n)$ , que depende exclusivamente do tamanho da amostra  $n$ . Possíveis escolhas incluem uma função linear da forma  $c(n) = nc$  ou uma função não-linear em que o custo marginal de cada observação diminui conforme o tamanho da amostra aumenta, dependendo do tipo de problema. A fim de simplificar, aqui adotaremos a função linear de modo que ao observarmos uma amostra  $\mathbf{x}$  de tamanho  $n$ , temos

$$\bar{\rho}((n, \varphi), \bar{\mathcal{P}}) = \int_{\Theta} \int_{\mathcal{X}} L(\varphi(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta + nc. \quad (4.4)$$

Note que o tamanho  $n$  também impacta o primeiro termo da função de perda e não somente o custo da observação, pois  $\mathbf{x}$  é a realização de uma amostra de tamanho  $n$ , como será visto nos exemplos a seguir. Uma forma de determinar o tamanho amostral ótimo  $n^*$  é de forma exaustiva, onde o risco em (4.4) é calculado para todo  $n \in \{1, 2, \dots, N\}$ ,  $N \in \mathbb{N}$ , e verifica-se qual o  $n$  específico que resulta no menor dos riscos calculados. Outra maneira consiste em assumir que  $n$  é contínuo, derivar a expressão do risco acima em função de  $n$  e igualá-la a zero, de forma a encontrar analiticamente o  $n$  que minimiza a função. Aqui, utilizaremos o primeiro método e Berger [3] elabora alguns exemplos em que é possível aplicar o segundo método.

Vale destacar que, sob as condições do Teorema 4.1 e considerando  $c(n, \mathbf{x}, \theta) = c(n)$  qualquer, o risco total do problema de teste de hipótese em conjunto com a determinação do tamanho amostral ótimo é menor sob a abordagem agnóstica relativamente à abordagem tradicional. O teorema abaixo estabelece tal relação.

**Teorema 4.2.** *Considere as mesmas condições do Teorema 4.1 em que  $\varphi_T$  e  $\varphi_A$  são testes de Bayes tradicional e agnóstico, respectivamente. Considere a função de perda  $\bar{L}_T((n, \varphi_T), (\theta, \mathbf{x})) = L_T(\varphi_T(\mathbf{x}), \theta) + c(n)$  para o teste de Bayes tradicional e suponha que  $(n_T^*, \varphi_T)$  é decisão de Bayes com risco  $\bar{\rho}_T^*$ . Considere também a função de perda  $\bar{L}_A((n, \varphi_A), (\theta, \mathbf{x})) = L_A(\varphi_A(\mathbf{x}), \theta) + c(n)$  para o teste de Bayes agnóstico e suponha que  $(n_A^*, \varphi_A)$  é decisão de Bayes com risco  $\bar{\rho}_A^*$ . Então,  $\bar{\rho}_A^* < \bar{\rho}_T^*$ .*

*Demonstração.* Como  $(n_A^*, \varphi_A)$  é decisão de Bayes para o caso agnóstico, temos que

$$\bar{\rho}_A^* = \bar{\rho}_A((n_A^*, \varphi_A), \bar{\mathcal{P}}) = \bar{\rho}_A(\varphi_A, \bar{\mathcal{P}}) + c(n_A^*) \leq \bar{\rho}_A(\varphi_A^{n_T^*}, \bar{\mathcal{P}}) + c(n_T^*),$$

onde  $\varphi_A^{n_T^*}$  é o teste agnóstico de Bayes para uma amostra de tamanho  $n_T^*$ .

Adicionalmente, do Teorema 4.1 segue que

$$\bar{\rho}_A(\varphi_A^{n_T^*}, \bar{\mathcal{P}}) + c(n_T^*) < \bar{\rho}_T(\varphi_T, \bar{\mathcal{P}}) + c(n_T^*) = \bar{\rho}_T^*,$$

Logo,

$$\bar{\rho}_A^* < \bar{\rho}_T^*.$$

□

No que segue, ilustramos os resultados desta seção através de alguns exemplos.

### 4.3 Estudo comparativo: Testes de Hipóteses Unilaterais

Para os estudos desta seção, considere que desejamos testar a hipótese unilateral  $H_0 : \theta \leq \theta_0$  contra  $H_1 : \theta > \theta_0$  a respeito do parâmetro  $\theta$ , em que  $\theta_0 \in \mathbb{R}$  é um valor de interesse fixado antes da realização do experimento.

#### 4.3.1 Teste baseado em Probabilidade a Posteriori: Modelo Normal com Variância conhecida

Considere um experimento em que será observada uma amostra representada pelo vetor aleatório  $\mathbf{X} = (X_1, \dots, X_n)$  onde assumimos que os  $X_1, \dots, X_n$ , dado  $\theta$ , são condicionalmente independentes e identicamente distribuídos (i.i.d.) segundo o modelo Normal de média  $\theta \in \mathbb{R}$  e variância  $v^2 > 0$  conhecida. Nas condições descritas, que denotamos por  $X_i | \theta \stackrel{iid}{\sim} \mathcal{N}(\theta, v^2)$ , temos que a média amostral  $\bar{X}$  possui, dado  $\theta$ , distribuição Normal tal que  $\bar{X} | \theta \sim \mathcal{N}(\theta, v^2/n)$ .

Assumindo que a priori de  $\theta$  segue distribuição Normal, digamos,  $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ , e considerando  $\mathbf{X}$  como descrito acima, temos por conjugação (DeGroot [16]) que a distribuição a posteriori do parâmetro  $\theta$ , dado  $\mathbf{x} = (x_1, \dots, x_n)$ , é

$$\theta | \mathbf{x} \sim \mathcal{N}(\mu_x, \sigma_x^2),$$

em que  $\mu_x = \frac{\mu_0 v^2 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2}$  e  $\sigma_x^2 = \frac{v^2 \sigma_0^2}{v^2 + n \sigma_0^2}$ .

Seja  $\varphi_T$  o teste de Bayes tradicional construído a partir da função de perda  $0 - a_0 - a_1$  representada pela Tabela 3.1. Utilizando a expressão (3.3) derivada no Capítulo 3, temos que  $\varphi_T(\mathbf{x}) = 1 \iff \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) = \mathbb{P}(\theta \leq \theta_0 | \mathbf{x}) \leq \frac{a_1}{a_0 + a_1}$ . Logo, o teste de Bayes  $\varphi_T : \mathcal{X} \rightarrow \{0, 1\}$ , baseado em probabilidade a posteriori das hipóteses para o modelo Normal pode ser escrito como

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & \Phi\left(\frac{\theta_0 - \mu_x}{\sigma_x}\right) > \frac{a_1}{a_0 + a_1} \\ 1, & \Phi\left(\frac{\theta_0 - \mu_x}{\sigma_x}\right) \leq \frac{a_1}{a_0 + a_1} \end{cases}, \quad (4.5)$$

em que  $\Phi(\cdot)$  é a função de distribuição (acumulada) de uma variável aleatória Normal padrão. Tal resultado é obtido através da padronização da distribuição Normal a posteriori de  $\theta$ , conforme visto acima. Substituindo as expressões da média  $\mu_x$  e variância  $\sigma_x^2$  a posteriori, podemos expressar a

função de decisão do teste em termos da estatística suficiente  $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ . Assim,

$$\begin{aligned}\varphi_T(\mathbf{x}) = 0 &\iff \Phi\left(\frac{\theta_0 - \mu_x}{\sigma_x}\right) > \frac{a_1}{a_0 + a_1} \\ &\iff \theta_0 - \mu_x > \Phi^{-1}\left(\frac{a_1}{a_0 + a_1}\right) \sigma_x \\ &\iff \frac{\mu_0 v^2 + n\sigma_0^2 \bar{x}}{v^2 + n\sigma_0^2} < \theta_0 - \Phi^{-1}\left(\frac{a_1}{a_0 + a_1}\right) \sqrt{\frac{v^2 \sigma_0^2}{v^2 + n\sigma_0^2}} \\ &\iff \bar{x} < \theta_0 + \frac{v^2(\theta_0 - \mu_0)}{n\sigma_0^2} - \Phi^{-1}\left(\frac{a_1}{a_0 + a_1}\right) \frac{v}{n\sigma_0} \sqrt{v^2 + n\sigma_0^2}.\end{aligned}$$

Na última desigualdade acima, definindo

$$k_n = \theta_0 + \frac{v^2(\theta_0 - \mu_0)}{n\sigma_0^2} - \Phi^{-1}\left(\frac{a_1}{a_0 + a_1}\right) \frac{v}{n\sigma_0} \sqrt{v^2 + n\sigma_0^2},$$

segue que

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & \bar{x} < k_n \\ 1, & \bar{x} \geq k_n \end{cases}. \quad (4.6)$$

Avaliamos agora o risco de Bayes (ou risco de  $\varphi_T$ ), denotado por  $\bar{\rho}(\varphi_T, \bar{\mathcal{P}})$ , que pode ser calculado através da expressão (3.2), obtendo

$$\bar{\rho}(\varphi_T, \bar{\mathcal{P}}) = a_0 \int_{\varphi_T^{-1}(\{1\})} \int_{\theta \leq \theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} + a_1 \int_{\varphi_T^{-1}(\{0\})} \int_{\theta > \theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x}. \quad (4.7)$$

Para cada uma das integrais duplas acima, que denotamos por  $I_1$  e  $I_2$ , podemos inverter a ordem de integração e decompor a função de densidade conjunta nas funções de densidade condicional e marginal. Logo,

$$\begin{aligned}I_1 &= a_0 \int_{\varphi_T^{-1}(\{1\})} \int_{\theta \leq \theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} = a_0 \int_{\theta \leq \theta_0} \int_{\varphi_T^{-1}(\{1\})} f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\ &= a_0 \int_{\theta \leq \theta_0} \left[ \int_{\varphi_T^{-1}(\{1\})} f(\mathbf{x}|\theta) d\mathbf{x} \right] f(\theta) d\theta.\end{aligned}$$

A integral  $\int_{\varphi_T^{-1}(\{1\})} f(\mathbf{x}|\theta) d\mathbf{x}$  é a probabilidade da amostra  $\mathbf{X}$ , condicional a  $\theta$ , pertencer à região de rejeição do teste ( $\bar{X} \geq k_n$ ). Assim, temos que

$$\begin{aligned}I_1 &= a_0 \int_{\theta \leq \theta_0} \mathbb{P}(\bar{X} \geq k_n | \theta) f(\theta) d\theta = a_0 \int_{\theta \leq \theta_0} \mathbb{P}\left(\frac{\bar{X} - \theta}{\sqrt{(v^2/n)}} \geq \frac{k_n - \theta}{\sqrt{(v^2/n)}} \middle| \theta\right) f(\theta) d\theta \\ &= a_0 \int_{\theta \leq \theta_0} \left[ 1 - \Phi\left(\frac{k_n - \theta}{\sqrt{(v^2/n)}}\right) \right] f(\theta) d\theta = a_0 \int_{\mathbb{R}} \left[ 1 - \Phi\left(\frac{k_n - \theta}{\sqrt{(v^2/n)}}\right) \right] \mathbb{1}_{(-\infty, \theta_0]}(\theta) f(\theta) d\theta.\end{aligned}$$

Analogamente, para a segunda integral de (4.7), obtemos

$$\begin{aligned}I_2 &= a_1 \int_{\theta > \theta_0} \mathbb{P}(\bar{X} < k_n | \theta) f(\theta) d\theta = a_1 \int_{\theta > \theta_0} \mathbb{P}\left(\frac{\bar{X} - \theta}{\sqrt{(v^2/n)}} < \frac{k_n - \theta}{\sqrt{(v^2/n)}} \middle| \theta\right) f(\theta) d\theta \\ &= a_1 \int_{\theta > \theta_0} \left[ \Phi\left(\frac{k_n - \theta}{\sqrt{(v^2/n)}}\right) \right] f(\theta) d\theta = a_1 \int_{\mathbb{R}} \left[ \Phi\left(\frac{k_n - \theta}{\sqrt{(v^2/n)}}\right) \right] \mathbb{1}_{(\theta_0, +\infty)}(\theta) f(\theta) d\theta.\end{aligned}$$

Uma forma de aproximar as integrais é através do método de Monte Carlo, conforme descrito em

Robert e Casella [40]. O método consiste em aproximar o valor de uma integral da forma

$$\mathbb{E}_f[h(\theta)] = \int_{\Theta} h(\theta)f(\theta)d\theta$$

através da média empírica

$$\bar{h} = \frac{1}{M} \sum_{j=1}^M h(\theta_j), \quad j = 1, 2, \dots, M,$$

calculada a partir de uma amostra  $(\theta_1, \dots, \theta_M)$ , onde cada  $\theta_j$  é gerado i.i.d. pela função de densidade  $f(\theta)$ , que neste caso é a priori  $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ . Como os  $\theta_j$ 's são gerados no espaço paramétrico  $\Theta = \mathbb{R}$ , as funções  $h_1(\theta_j)$  e  $h_2(\theta_j)$  correspondentes às integrais  $I_1$  e  $I_2$  são, respectivamente,

$$h_1(\theta_j) = a_0 \left[ 1 - \Phi \left( \frac{k_n - \theta_j}{\sqrt{(v^2/n)}} \right) \right] \mathbb{1}_{(-\infty, \theta_0]}(\theta_j)$$

e

$$h_2(\theta_j) = a_1 \left[ \Phi \left( \frac{k_n - \theta_j}{\sqrt{(v^2/n)}} \right) \right] \mathbb{1}_{(\theta_0, +\infty)}(\theta_j).$$

Vamos construir agora o teste de Bayes agnóstico  $\varphi_A$ , sob a função de perda  $a_0 - a_1 - b_0 - b_1$ , representada na Tabela 3.2. Conforme a expressão (3.15), o teste de Bayes agnóstico para as hipóteses unilaterais pode ser escrito como

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \Phi \left( \frac{\theta_0 - \mu_x}{\sigma_x} \right) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ \frac{1}{2}, & \frac{b_1}{a_0 + b_1 - b_0} < \Phi \left( \frac{\theta_0 - \mu_x}{\sigma_x} \right) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ 1, & \Phi \left( \frac{\theta_0 - \mu_x}{\sigma_x} \right) \leq \frac{b_1}{a_0 + b_1 - b_0} \end{cases}. \quad (4.8)$$

Escrevendo a função de decisão do teste em termos da estatística suficiente  $\bar{x}$ , aceitamos  $H_0$  se

$$\begin{aligned} \varphi_A(\mathbf{x}) = 0 &\iff \Phi \left( \frac{\theta_0 - \mu_x}{\sigma_x} \right) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ &\iff \theta_0 - \mu_x > \Phi^{-1} \left( \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right) \sigma_x \\ &\iff \bar{x} < \theta_0 + \frac{v^2(\theta_0 - \mu_0)}{n\sigma_0^2} - \Phi^{-1} \left( \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right) \frac{v}{n\sigma_0} \sqrt{v^2 + n\sigma_0^2}. \end{aligned}$$

Seja  $k_n^{(1)} = \theta_0 + \frac{v^2(\theta_0 - \mu_0)}{n\sigma_0^2} - \Phi^{-1} \left( \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right) \frac{v}{n\sigma_0} \sqrt{v^2 + n\sigma_0^2}$ . Analogamente, rejeitamos  $H_0$  se

$$\begin{aligned} \varphi_A(\mathbf{x}) = 1 &\iff \Phi \left( \frac{\theta_0 - \mu_x}{\sigma_x} \right) \leq \frac{b_1}{a_0 + b_1 - b_0} \\ &\iff \theta_0 - \mu_x \leq \Phi^{-1} \left( \frac{b_1}{a_0 + b_1 - b_0} \right) \sigma_x \\ &\iff \bar{x} \geq \theta_0 + \frac{v^2(\theta_0 - \mu_0)}{n\sigma_0^2} - \Phi^{-1} \left( \frac{b_1}{a_0 + b_1 - b_0} \right) \frac{v}{n\sigma_0} \sqrt{v^2 + n\sigma_0^2}. \end{aligned}$$

Seja  $k_n^{(2)} = \theta_0 + \frac{v^2(\theta_0 - \mu_0)}{n\sigma_0^2} - \Phi^{-1} \left( \frac{b_1}{a_0 + b_1 - b_0} \right) \frac{v}{n\sigma_0} \sqrt{v^2 + n\sigma_0^2}$ .

Conseqüentemente, temos

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \bar{x} < k_n^{(1)} \\ \frac{1}{2}, & k_n^{(1)} \leq \bar{x} < k_n^{(2)} \\ 1, & \bar{x} \geq k_n^{(2)} \end{cases} . \quad (4.9)$$

Note que, no caso de perda própria, temos  $k_n^{(1)} < k_n^{(2)}$  e, portanto,  $\mathbb{P}(\varphi_A(\mathbf{x}) = \frac{1}{2}) > 0$ .

A seguir, calculamos o risco do teste  $\varphi_A$ , denotado por  $\bar{\rho}(\varphi_A, \bar{\mathcal{P}})$ , que é calculado pela expressão (3.13). Temos que

$$\begin{aligned} \bar{\rho}(\varphi_A, \bar{\mathcal{P}}) &= a_0 \int_{\varphi_A^{-1}(\{1\})} \int_{\theta \leq \theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} + a_1 \int_{\varphi_A^{-1}(\{0\})} \int_{\theta > \theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} \\ &+ b_0 \int_{\varphi_A^{-1}(\{\frac{1}{2}\})} \int_{\theta \leq \theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} + b_1 \int_{\varphi_A^{-1}(\{\frac{1}{2}\})} \int_{\theta > \theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x}. \end{aligned} \quad (4.10)$$

Denotamos as parcelas em (4.10) por  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$  e utilizamos o mesmo procedimento para o cálculo do risco do teste de Bayes tradicional  $\varphi_T$ :

$$\begin{aligned} I_1 &= a_0 \int_{\mathbb{R}} \left[ 1 - \Phi \left( \frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}} \right) \right] \mathbb{1}_{(-\infty, \theta_0]}(\theta) f(\theta) d\theta; \\ I_2 &= a_1 \int_{\mathbb{R}} \left[ \Phi \left( \frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}} \right) \right] \mathbb{1}_{(\theta_0, +\infty)}(\theta) f(\theta) d\theta; \\ I_3 &= b_0 \int_{\mathbb{R}} \left[ \Phi \left( \frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}} \right) - \Phi \left( \frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}} \right) \right] \mathbb{1}_{(-\infty, \theta_0]}(\theta) f(\theta) d\theta; \\ I_4 &= b_1 \int_{\mathbb{R}} \left[ \Phi \left( \frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}} \right) - \Phi \left( \frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}} \right) \right] \mathbb{1}_{(\theta_0, +\infty)}(\theta) f(\theta) d\theta. \end{aligned}$$

De forma análoga, as integrais podem ser aproximadas pelo método de Monte Carlo descrito anteriormente. A seguir, elaboramos um exemplo numérico desse cenário.

**Exemplo 4.1.** Sejam  $X_1, \dots, X_n$  em que  $X_i | \theta \stackrel{iid}{\sim} \mathcal{N}(\theta, 1)$ , para  $i = 1, \dots, n$ . Suponha que iremos testar as hipóteses  $H_0 : \theta \leq 3$  e  $H_1 : \theta > 3$  e que desejamos comparar os riscos de Bayes dos testes tradicional e agnóstico baseados em probabilidade a posteriori, em função do tamanho amostral  $n$ . As perdas escolhidas para o teste tradicional são  $a_0 = 1$ ,  $a_1 = 1$  e  $a_0 = 1$ ,  $a_1 = 1$ ,  $b_0 = \frac{1}{4}$  e  $b_1 = \frac{1}{4}$  para o teste agnóstico.

Assumindo inicialmente uma priori para  $\theta$  de distribuição Normal qualquer,  $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ , e substituindo os valores fornecidos nas expressões (4.6) e (4.9), os testes obtidos são

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & \bar{x} < 3 + \frac{3-\mu_0}{n\sigma_0^2} \\ 1, & \bar{x} \geq 3 + \frac{3-\mu_0}{n\sigma_0^2} \end{cases}$$

e

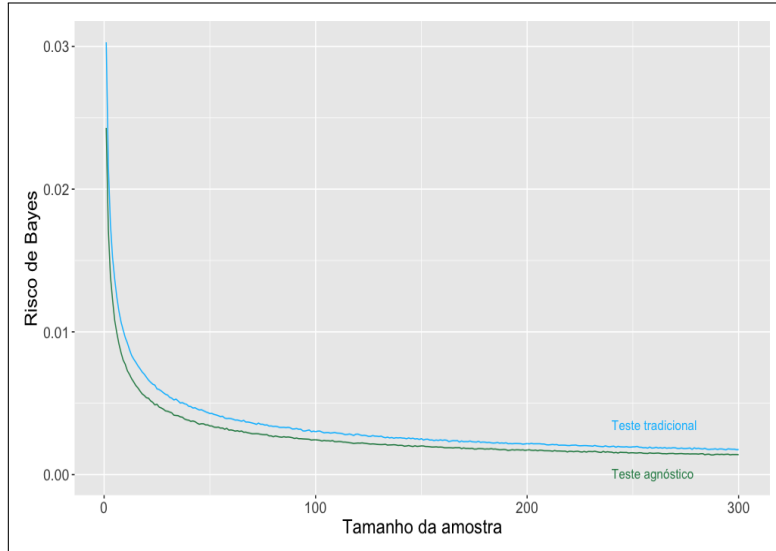
$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \bar{x} < 3 + \frac{3-\mu_0}{n\sigma_0^2} - \frac{\Phi^{-1}(3/4)\sqrt{1+n\sigma_0^2}}{n\sigma_0} \\ \frac{1}{2}, & 3 + \frac{3-\mu_0}{n\sigma_0^2} - \frac{\Phi^{-1}(3/4)\sqrt{1+n\sigma_0^2}}{n\sigma_0} \leq \bar{x} < 3 + \frac{3-\mu_0}{n\sigma_0^2} - \frac{\Phi^{-1}(1/4)\sqrt{1+n\sigma_0^2}}{n\sigma_0} \\ 1, & \bar{x} \geq 3 + \frac{3-\mu_0}{n\sigma_0^2} - \frac{\Phi^{-1}(1/4)\sqrt{1+n\sigma_0^2}}{n\sigma_0} \end{cases} .$$

Vale notar que no primeiro teste quando a média da priori é igual a 3, a quantidade  $k_n$  independe de  $n$ , diferentemente do teste agnóstico em que  $k_n^{(1)}$  e  $k_n^{(2)}$  dependem de  $n$ . Tal fato ocorre pois as



perdas  $a_0$  e  $a_1$  foram definidas iguais, resultando em  $\Phi^{-1}(1/2) = 0$  na expressão de  $k_n$ .

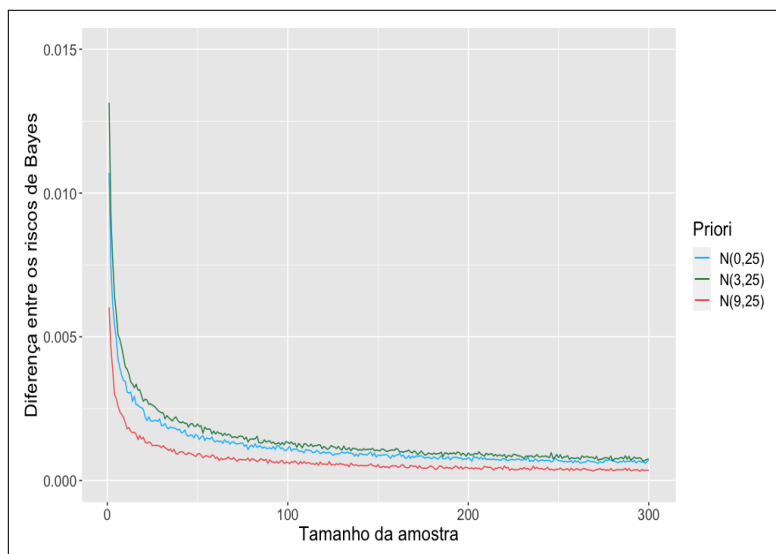
Suponha agora que  $\mu_0 = 0$  e  $\sigma_0^2 = 100$  de forma a calcular numericamente o risco de Bayes de cada um dos testes para diferentes tamanhos amostrais, com  $n$  variando de 1 a 300. Para o cálculo, as integrais foram aproximadas pelo método de Monte Carlo exposto anteriormente com  $M = 10^6$ . A Figura 4.1 ilustra os resultados obtidos.



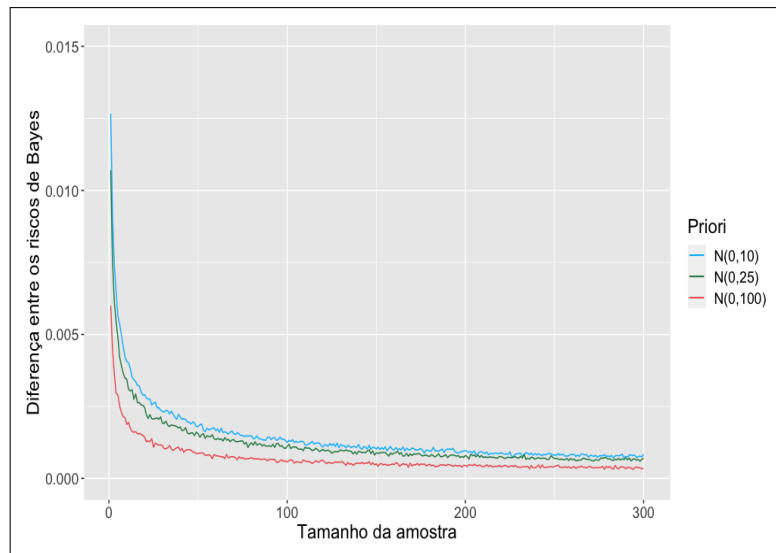
**Figura 4.1:** *Riscos de Bayes dos testes unilaterais para modelo Normal - Priori  $\mathcal{N}(0, 100)$ .*

Como esperado, o risco de Bayes decresce conforme  $n$  aumenta—teoricamente as funções são estritamente decrescentes, porém há pequenas oscilações no gráfico devido ao fato de os valores das integrais serem aproximados e não exatos—e o risco do teste agnóstico é menor que o risco do teste não agnóstico para todo  $n$ , fato que foi demonstrado no teorema 4.1 (note que nesse caso  $L_A(i, \theta) = L_T(i, \theta)$ , para  $i = 0, 1, \forall \theta \in \Theta$ ).

Para avaliar o efeito de prioris distintas para  $\theta$  nos riscos de Bayes de cada teste, foram escolhidas prioris de distribuição Normal com parâmetros distintos e, para cada uma delas, foi calculada a diferença absoluta entre o risco do teste não agnóstico e o risco do teste agnóstico para todo  $n$ . Algumas comparações são apresentadas nas figuras a seguir.



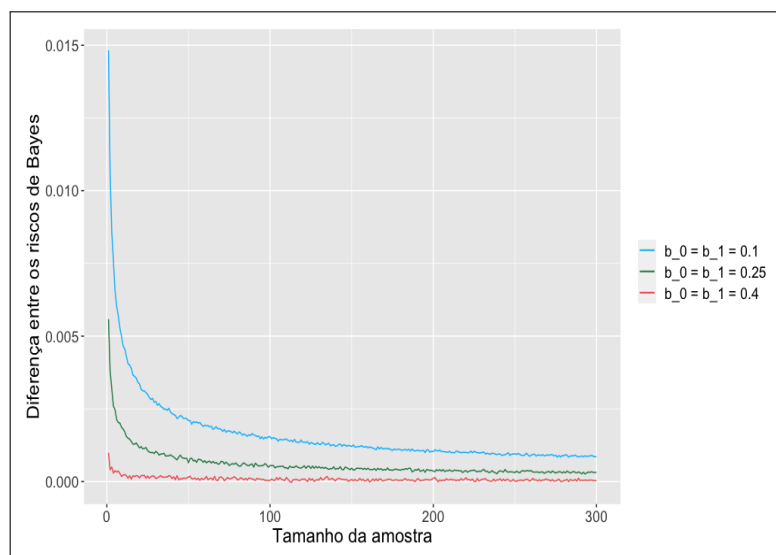
**Figura 4.2:** *Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes unilaterais para modelo Normal.*



**Figura 4.3:** Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes unilaterais para modelo Normal.

Na Figura 4.2, pode-se observar que a diferença entre os riscos dos testes é maior quando a média da priori é igual a 3, ou seja, quando a priori está centrada no ponto que separa as regiões de  $H_0$  e  $H_1$ . Conforme a média da priori se distancia de 3, menor a diferença absoluta entre os riscos dos testes. Já no caso da figura 4.3, observa-se que, para uma mesma média, quanto maior a variância da priori, menor a diferença entre os riscos dos testes.

Outra possibilidade de análise é comparar os riscos dos dois testes conforme variam-se as perdas  $b_0$  e  $b_1$  associadas aos erros cometidos quando a decisão do teste é ficar agnóstico. Se considerarmos perdas em que  $a_0 = a_1 = 1$  para os dois testes e  $b_0 = b_1$  para o teste agnóstico, o intervalo correspondente à decisão do teste de ficar agnóstico se simplifica a  $b_0 \leq \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) < 1 - b_0$ . Considere três cenários em que  $b_0 (= b_1)$  vale 0,1, 0,25 e 0,4. Seja  $\theta \sim \mathcal{N}(0, 100)$  a priori. A Figura 4.4 ilustra a diferença entre os riscos de Bayes dos dois testes para esses exemplos.



**Figura 4.4:** Diferença absoluta entre os riscos de Bayes para diferentes perdas - Testes unilaterais para modelo Normal.

Quanto mais próximos de 0,5 os valores de  $b_0$  e  $b_1$ , menores os intervalos em que a decisão é de ficar agnóstico ( $b_0 \leq \mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) < 1 - b_0$ ) e, conseqüentemente, mais próximos os valores dos riscos dos dois testes. Quando  $b_0 = b_1 = 0,5$ , não há mais probabilidade positiva de ficar agnóstico e

os dois testes se tornam equivalentes no caso particular do teste de Bayes baseado em probabilidade a posteriori das hipóteses, com  $a_0 = a_1 = 1$ .

Por fim, considere o método de cálculo de tamanho amostral proposto na seção 4.2. Supondo um custo  $c = 0,0001$  na expressão de custo linear  $c(n) = cn$ , foram encontrados os tamanhos amostrais ótimos  $n_T^*$  e  $n_A^*$  que minimizam as respectivas perdas esperadas:

$$\begin{aligned}\bar{\rho}_T^{(n)} &= \bar{\rho}(\varphi_T, \bar{\mathcal{P}}) + nc \\ \bar{\rho}_A^{(n)} &= \bar{\rho}(\varphi_A, \bar{\mathcal{P}}) + nc,\end{aligned}$$

em que  $n \in \mathbb{N}$  e  $\bar{\rho}(\varphi_T, \bar{\mathcal{P}})$  e  $\bar{\rho}(\varphi_A, \bar{\mathcal{P}})$  são dados pelas expressões (4.7) e (4.10), respectivamente. A Tabela 4.1 apresenta os resultados obtidos para diferentes escolhas de prioris, em que  $\delta(\%) = 100 * \frac{n_T^* - n_A^*}{n_T^*}$ .

Priori $\theta$	$n_T^*$	$n_A^*$	$\delta(\%)$	$\bar{\rho}_T^*$	$\bar{\rho}_A^*$
$\mathcal{N}(0, 10)$	46	41	10,9	0,0094	0,0079
$\mathcal{N}(0, 25)$	41	40	2,4	0,0083	0,0066
$\mathcal{N}(0, 100)$	25	24	4,0	0,0060	0,0049
$\mathcal{N}(3, 10)$	64	57	10,9	0,0125	0,0105
$\mathcal{N}(3, 25)$	45	40	11,1	0,0094	0,0079
$\mathcal{N}(3, 100)$	32	24	25,0	0,0056	0,0051
$\mathcal{N}(6, 10)$	48	35	27,1	0,0092	0,0085
$\mathcal{N}(6, 25)$	41	35	14,6	0,0082	0,0071
$\mathcal{N}(6, 100)$	32	25	21,9	0,0053	0,0048
$\mathcal{N}(9, 10)$	19	16	15,8	0,0038	0,0032
$\mathcal{N}(9, 25)$	26	24	7,7	0,0060	0,0050
$\mathcal{N}(9, 100)$	28	25	10,7	0,0049	0,0042

**Tabela 4.1:** Tamanhos amostrais e riscos para os testes de Bayes unilaterais - Modelo Normal.

Notamos que os tamanhos amostrais obtidos são menores para o teste agnóstico. Vale destacar que dependendo do número  $M$  de iterações do método de Monte Carlo e da semente aleatória utilizada na rotina computacional, os resultados podem variar devido ao fato de o cálculo das integrais ser aproximado e não exato. Observa-se também que os maiores tamanhos amostrais e riscos de Bayes são obtidos no caso de priori  $\mathcal{N}(3, 10)$ , em que temos, a priori, maior crença que  $\theta$  está próximo do ponto que divide as regiões de  $H_0$  e  $H_1$ . O caso de menor tamanho amostral e risco é obtido com a priori  $\mathcal{N}(9, 10)$ , que corresponde a quando temos maior crença por  $H_1$  a priori. Ao comparar os tamanhos amostrais ótimos para a mesma priori, constatamos casos em que  $n_T^*$  são de 10 a 20% maiores que  $n_A^*$ , o que pode constituir uma redução de custo expressiva no caso agnóstico.

Além disso, com exceção das prioris de média  $\mu_0 = 9$ , observa-se que, para prioris de mesma média, quanto maior a variância (resultando em uma priori menos informativa), menores os tamanhos amostrais e os riscos de Bayes. Enquanto que, para prioris de mesma variância, quanto mais distante a média da priori de  $\theta_0$ , menor o tamanho amostral necessário, já que temos mais certeza a priori a respeito de qual das hipóteses é verdadeira. Tais resultados da relação dos parâmetros da priori com os tamanhos amostrais obtidos são condizentes com os observados em Sahu e Smith [41], apesar de o método de cálculo amostral ser um pouco diferente do proposto aqui.

### 4.3.2 Teste de Bayes com Perda baseada em Norma $L_1$ : Modelo Normal com Variância conhecida

Em seguida, vamos explorar outra função de perda, que depende da distância do parâmetro  $\theta$  à hipótese escolhida em caso de erro de tomada de decisão. Para esse exemplo, considere novamente as hipóteses  $H_0 : \theta \leq 0$  e  $H_1 : \theta > 0$  para o mesmo modelo Normal com variância conhecida. A função de perda escolhida, denotada aqui de  $0 - L_1$ , é a representada pela tabela abaixo.

$L_T(d, \theta)$	$\theta \leq 0$	$\theta > 0$
$d = 0$	0	$ \theta  = \theta$
$d = 1$	$-\theta$	0

**Tabela 4.2:** Função de perda  $0 - L_1$ .

A seguir, derivamos os testes de Bayes tradicional e agnóstico e seus respectivos riscos.

Para derivar o teste de Bayes tradicional, note que tomamos a decisão de rejeitar  $H_0$  se o risco de tal decisão é menor que o risco de aceitar  $H_0$ . Assim,

$$\begin{aligned}
 d_x^* = 1 &\iff \rho_x(1, \mathcal{P}_x) < \rho_x(0, \mathcal{P}_x) \\
 &\iff \int_{\Theta} L_T(1, \theta) f(\theta | \mathbf{x}) d\theta < \int_{\Theta} L_T(0, \theta) f(\theta | \mathbf{x}) d\theta \\
 &\iff \int_{\theta \leq 0} -\theta f(\theta | \mathbf{x}) d\theta < \int_{\theta > 0} \theta f(\theta | \mathbf{x}) d\theta \\
 &\iff \int_{\theta \leq 0} \theta f(\theta | \mathbf{x}) d\theta + \int_{\theta > 0} \theta f(\theta | \mathbf{x}) d\theta > 0 \\
 &\iff \mathbb{E}[\theta | \mathbf{X} = \mathbf{x}] > \iff \mu_x > 0 \\
 &\iff \frac{\mu_0 v^2 + n\sigma_0^2 \bar{x}}{v^2 + n\sigma_0^2} > 0 \iff \bar{x} > -\frac{\mu_0 v^2}{n\sigma_0^2}.
 \end{aligned}$$

Consequentemente, a função do teste é escrita como

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & \bar{x} < -\frac{\mu_0 v^2}{n\sigma_0^2} \\ 1, & \bar{x} > -\frac{\mu_0 v^2}{n\sigma_0^2} \end{cases}. \quad (4.11)$$

O cálculo do risco é dado pela definição apresentada na equação (2.6), resultando em

$$\begin{aligned}
 \bar{\rho}(\varphi_T, \bar{\mathcal{P}}) &= \int_{\Theta} \int_{\mathcal{X}} L_T(\varphi_T(\mathbf{x}), \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\
 &= \int_{\theta \leq 0} \int_{\varphi_T^{-1}(\{1\})} -\theta f(\theta, \mathbf{x}) d\mathbf{x} d\theta + \int_{\theta > 0} \int_{\varphi_T^{-1}(\{0\})} \theta f(\theta, \mathbf{x}) d\mathbf{x} d\theta \\
 &= - \int_{\theta \leq 0} \left[ \int_{\varphi_T^{-1}(\{1\})} f(\mathbf{x} | \theta) d\mathbf{x} \right] \theta f(\theta) d\theta + \int_{\theta > 0} \left[ \int_{\varphi_T^{-1}(\{0\})} f(\mathbf{x} | \theta) d\mathbf{x} \right] \theta f(\theta) d\theta \\
 &= - \int_{\theta \leq 0} \mathbb{P} \left( \bar{X} > -\frac{\mu_0 v^2}{n\sigma_0^2} \middle| \theta \right) \theta f(\theta) d\theta + \int_{\theta > 0} \mathbb{P} \left( \bar{X} < -\frac{\mu_0 v^2}{n\sigma_0^2} \middle| \theta \right) \theta f(\theta) d\theta \\
 &= - \int_{\theta \leq \theta_0} \left[ 1 - \Phi \left( \frac{-\frac{\mu_0 v^2}{n\sigma_0^2} - \theta}{\sqrt{(v^2/n)}} \right) \right] \theta f(\theta) d\theta + \int_{\theta > \theta_0} \left[ \Phi \left( \frac{-\frac{\mu_0 v^2}{n\sigma_0^2} - \theta}{\sqrt{(v^2/n)}} \right) \right] \theta f(\theta) d\theta.
 \end{aligned} \quad (4.12)$$

No caso agnóstico, vamos considerar a função de perda própria representada pela tabela abaixo,

a qual denotamos de  $0 - L_1^{(p)}$ , em que  $p \in (0, \frac{1}{2})$ , a fim de garantir que a função de perda seja própria.

$L_A(d, \theta)$	$\theta \leq 0$	$\theta > 0$
$d = 0$	0	$ \theta  = \theta$
$d = \frac{1}{2}$	$-p\theta$	$p\theta$
$d = 1$	$-\theta$	0

**Tabela 4.3:** Função de perda  $0 - L_1^{(p)}$ .

Para derivar o teste agnóstico de Bayes, procedemos da mesma forma que na seção 3.2.1. Assim,

$$\begin{aligned}
d_x^* = 1 &\iff \rho_x(1, \mathcal{P}_x) < \rho_x(0, \mathcal{P}_x) \quad \text{e} \quad \rho_x(1, \mathcal{P}_x) < \rho_x\left(\frac{1}{2}, \mathcal{P}_x\right) \\
&\iff \int_{\Theta} L_A(1, \theta) f(\theta|\mathbf{x}) d\theta < \int_{\Theta} L_A(0, \theta) f(\theta|\mathbf{x}) d\theta \quad \text{e} \\
&\quad \int_{\Theta} L_A(1, \theta) f(\theta|\mathbf{x}) d\theta < \int_{\Theta} L_A\left(\frac{1}{2}, \theta\right) f(\theta|\mathbf{x}) d\theta \\
&\iff \int_{\theta \leq 0} -\theta f(\theta|\mathbf{x}) d\theta < \int_{\theta > 0} \theta f(\theta|\mathbf{x}) d\theta \quad \text{e} \\
&\quad \int_{\theta \leq 0} -\theta f(\theta|\mathbf{x}) d\theta < \int_{\theta \leq 0} -p\theta f(\theta|\mathbf{x}) d\theta + \int_{\theta > 0} p\theta f(\theta|\mathbf{x}) d\theta \\
&\iff \mu_x > 0 \quad \text{e} \quad \int_{\theta \leq 0} -\theta f(\theta|\mathbf{x}) d\theta < p \left( \int_{\theta > 0} \theta f(\theta|\mathbf{x}) d\theta - \int_{\theta \leq 0} \theta f(\theta|\mathbf{x}) d\theta \right).
\end{aligned}$$

Desenvolvendo as integrais acima, temos

$$\begin{aligned}
\int_{\theta \leq 0} \theta f(\theta|\mathbf{x}) d\theta &= \int_{-\infty}^0 \theta \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(\theta-\mu_x)^2}{2\sigma_x^2}} d\theta \quad (t = \theta - \mu_x) \\
&= \int_{-\infty}^{-\mu_x} (t + \mu_x) \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{t^2}{2\sigma_x^2}} dt \\
&= \mu_x \int_{-\infty}^{-\mu_x} \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{t^2}{2\sigma_x^2}} dt + \int_{-\infty}^{-\mu_x} \frac{t}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{t^2}{2\sigma_x^2}} dt \quad \left( \times \frac{\sqrt{\frac{\sigma_x^2}{2\pi}}}{-\sqrt{\frac{\sigma_x^2}{2\pi}}} \right) \\
&= \mu_x \mathbb{P}\left(Z < \frac{-\mu_x}{\sigma_x}\right) - \sqrt{\frac{\sigma_x^2}{2\pi}} \int_{-\infty}^{-\mu_x} \frac{-t}{\sigma_x^2} e^{-\frac{t^2}{2\sigma_x^2}} dt \quad (Z \sim \mathcal{N}(0, 1)) \\
&= \mu_x \Phi\left(\frac{-\mu_x}{\sigma_x}\right) - \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}}.
\end{aligned}$$

$$\begin{aligned}
\int_{\theta>0} \theta f(\theta|\mathbf{x})d\theta &= \int_0^{+\infty} \theta \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(\theta-\mu_x)^2}{2\sigma_x^2}} d\theta \quad (t = \theta - \mu_x) \\
&= \int_{-\mu_x}^{+\infty} (t + \mu_x) \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{t^2}{2\sigma_x^2}} dt \\
&= \mu_x \int_{-\mu_x}^{+\infty} \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{t^2}{2\sigma_x^2}} dt + \int_{-\mu_x}^{+\infty} \frac{t}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{t^2}{2\sigma_x^2}} dt \quad \left( \times \frac{-\sqrt{\frac{\sigma_x^2}{2\pi}}}{-\sqrt{\frac{\sigma_x^2}{2\pi}}} \right) \\
&= \mu_x \mathbb{P} \left( Z > \frac{-\mu_x}{\sigma_x} \right) - \sqrt{\frac{\sigma_x^2}{2\pi}} \int_{-\mu_x}^{+\infty} \frac{-t}{\sigma_x^2} e^{-\frac{t^2}{2\sigma_x^2}} dt \quad (Z \sim \mathcal{N}(0, 1)) \\
&= \mu_x \left[ 1 - \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}}.
\end{aligned}$$

Logo,

$$\begin{aligned}
d_x^* = 1 &\iff \mu_x > 0 \quad \text{e} \quad \int_{\theta \leq 0} -\theta f(\theta|\mathbf{x})d\theta < p \left( \int_{\theta>0} \theta f(\theta|\mathbf{x})d\theta - \int_{\theta \leq 0} \theta f(\theta|\mathbf{x})d\theta \right) \\
&\iff \mu_x > 0 \quad \text{e} \quad (1-2p) \left[ \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} - \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] - p\mu_x < 0.
\end{aligned}$$

Utilizando a mesma lógica para a decisão de aceitar  $H_0$ , obtemos

$$\begin{aligned}
d_x^* = 0 &\iff \rho_x(0, \mathcal{P}_x) < \rho_x(1, \mathcal{P}_x) \quad \text{e} \quad \rho_x(0, \mathcal{P}_x) < \rho_x \left( \frac{1}{2}, \mathcal{P}_x \right) \\
&\iff \int_{\Theta} L(0, \theta) f(\theta|\mathbf{x})d\theta < \int_{\Theta} L(1, \theta) f(\theta|\mathbf{x})d\theta \quad \text{e} \\
&\quad \int_{\Theta} L(0, \theta) f(\theta|\mathbf{x})d\theta < \int_{\Theta} L \left( \frac{1}{2}, \theta \right) f(\theta|\mathbf{x})d\theta \\
&\iff \mu_x < 0 \quad \text{e} \quad \int_{\theta>0} \theta f(\theta|\mathbf{x})d\theta < p \left( \int_{\theta>0} \theta f(\theta|\mathbf{x})d\theta - \int_{\theta \leq 0} \theta f(\theta|\mathbf{x})d\theta \right) \\
&\iff \mu_x < 0 \quad \text{e} \quad (1-2p) \left[ \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} - \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + (1-p)\mu_x < 0.
\end{aligned}$$

Portanto, o teste agnóstico de Bayes para a perda  $0 - L_1^{(p)}$  é dado pela seguinte expressão

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \mu_x < 0 \quad \text{e} \quad (1-2p) \left[ \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} - \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + (1-p)\mu_x < 0 \\ 1, & \mu_x > 0 \quad \text{e} \quad (1-2p) \left[ \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} - \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] - p\mu_x < 0 \\ \frac{1}{2}, & \text{caso contrário} \end{cases} \quad (4.13)$$

Finalmente, o risco de Bayes de tal teste é dado por

$$\begin{aligned}
\bar{\rho}(\varphi_A, \bar{\mathcal{P}}) &= \int_{\Theta} \int_{\mathcal{X}} L_A(\varphi_A, \theta) f(\theta, \mathbf{x}) d\mathbf{x} d\theta = \int_{\mathcal{X}} \int_{\Theta} L_A(\varphi_A, \theta) f(\theta, \mathbf{x}) d\theta d\mathbf{x} \\
&= \int_{\varphi_T^{-1}(\{1\})} \int_{\theta \leq 0} -\theta f(\theta, \mathbf{x}) d\theta d\mathbf{x} + \int_{\varphi_T^{-1}(\{0\})} \int_{\theta > 0} \theta f(\theta, \mathbf{x}) d\theta d\mathbf{x} \\
&\quad + \int_{\varphi_T^{-1}(\{\frac{1}{2}\})} \int_{\theta \leq 0} -p\theta f(\theta, \mathbf{x}) d\theta d\mathbf{x} + \int_{\varphi_T^{-1}(\{\frac{1}{2}\})} \int_{\theta > 0} p\theta f(\theta, \mathbf{x}) d\theta d\mathbf{x}. \quad (4.14)
\end{aligned}$$

Diferentemente do que foi feito no teste tradicional, nesse caso realizamos a decomposição da função de densidade conjunta de  $(\theta, \mathbf{X})$  utilizando as funções de densidade condicional  $f(\theta|\mathbf{x})$  e marginal  $f(\mathbf{x})$ . Denotando as quatro parcelas acima de  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , respectivamente, e lembrando que  $\mu_x = \frac{\mu_0 v^2 + n\sigma_0^2 \bar{x}}{v^2 + n\sigma_0^2}$ , segue que

$$\begin{aligned} I_1 &= \int_R \int_{\theta \leq 0} -\theta f(\theta, \mathbf{x}) d\theta d\mathbf{x} = - \int_R \int_{\theta \leq 0} \theta f(\theta|\mathbf{x}) f(\mathbf{x}) d\theta d\mathbf{x} \\ &= - \int_R \left[ \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) - \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right] f(\mathbf{x}) d\mathbf{x}; \\ I_2 &= \int_A \int_{\theta > 0} \theta f(\theta, \mathbf{x}) d\theta d\mathbf{x} = \int_A \int_{\theta > 0} \theta f(\theta|\mathbf{x}) f(\mathbf{x}) d\theta d\mathbf{x} \\ &= \int_A \left\{ \mu_x \left[ 1 - \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right\} f(\mathbf{x}) d\mathbf{x}; \\ I_3 &= \int_G \int_{\theta \leq 0} -p\theta f(\theta, \mathbf{x}) d\theta d\mathbf{x} = - \int_G \int_{\theta \leq 0} p\theta f(\theta|\mathbf{x}) f(\mathbf{x}) d\theta d\mathbf{x} \\ &= -p \int_G \left[ \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) - \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right] f(\mathbf{x}) d\mathbf{x}; \\ I_4 &= \int_G \int_{\theta > 0} p\theta f(\theta, \mathbf{x}) d\theta d\mathbf{x} = \int_G \int_{\theta > 0} p\theta f(\theta|\mathbf{x}) f(\mathbf{x}) d\theta d\mathbf{x} \\ &= p \int_G \left\{ \mu_x \left[ 1 - \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right\} f(\mathbf{x}) d\mathbf{x}; \end{aligned}$$

onde

$$\begin{aligned} R &= \left\{ \bar{x} \in \mathbb{R} : \mu_x > 0 \quad \text{e} \quad (1-2p) \left[ \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} - \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] - p\mu_x < 0 \right\}; \\ A &= \left\{ \bar{x} \in \mathbb{R} : \mu_x < 0 \quad \text{e} \quad (1-2p) \left[ \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} - \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + (1-p)\mu_x < 0 \right\}; \\ G &= \left\{ \bar{x} \in \mathbb{R} : \bar{x} \in (R \cup A)^c \right\}. \end{aligned}$$

Utilizando o método de Monte Carlo, podemos aproximar as integrais  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , que são da forma  $\int_{\mathcal{X}} h(\mathbf{x}) f(\mathbf{x}) d\mathbf{x}$ , simulando da distribuição marginal de  $\mathbf{X}$  (ou de  $\bar{X}$ ). É possível mostrar que (Paulino et al. [36], p.167), se  $X_i|\theta \stackrel{iid}{\sim} \mathcal{N}(\theta, v^2)$  e  $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ , então a distribuição marginal de  $\bar{X}$  é dada por

$$\bar{X} \sim \mathcal{N} \left( \mu_0, \sigma_0^2 + \frac{v^2}{n} \right).$$

A partir daí, para cada valor simulado de  $\bar{x}$ , conseguimos calcular  $\mu_x$  e  $\sigma_x^2$  e checar para tais valores, qual seria a decisão do teste de acordo com as regiões acima, de forma a aproximar as integrais através das médias empíricas

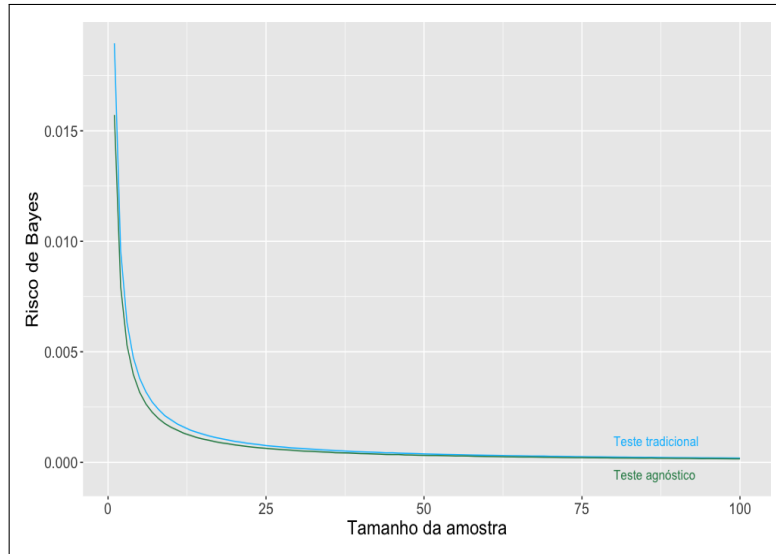
$$\bar{h}_i = \frac{1}{M} \sum_{j=1}^M h_i(\bar{x}_j), \quad i = 1, 2, 3, 4; \quad j = 1, 2, \dots, M,$$

onde  $\bar{x}_1, \dots, \bar{x}_M$  é amostra gerada da distribuição de  $\bar{X}$  e

$$\begin{aligned} h_1(\bar{x}_j) &= - \left[ \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) - \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right] \mathbb{1}_R(\bar{x}_j); \\ h_2(\bar{x}_j) &= \left\{ \mu_x \left[ 1 - \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right\} \mathbb{1}_A(\bar{x}_j); \\ h_3(\bar{x}_j) &= -p \left[ \mu_x \Phi \left( \frac{-\mu_x}{\sigma_x} \right) - \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right] \mathbb{1}_G(\bar{x}_j); \\ h_4(\bar{x}_j) &= p \left\{ \mu_x \left[ 1 - \Phi \left( \frac{-\mu_x}{\sigma_x} \right) \right] + \sqrt{\frac{\sigma_x^2}{2\pi}} e^{-\frac{\mu_x^2}{2\sigma_x^2}} \right\} \mathbb{1}_G(\bar{x}_j). \end{aligned}$$

**Exemplo 4.2.** Sejam  $X_1, \dots, X_n$  em que  $X_i | \theta \stackrel{iid}{\sim} \mathcal{N}(\theta, 1)$ , para  $i = 1, \dots, n$ . Desejamos testar a hipótese  $H_0 : \theta \leq 0$  contra a hipótese  $H_1 : \theta > 0$  e comparar os riscos de Bayes dos testes tradicional e agnóstico, em função do tamanho amostral  $n$ , considerando as funções de perda  $0 - L_1$  e  $0 - L_1^{(p)}$ , em que  $p = \frac{1}{4}$ .

Assumindo uma priori Normal para  $\theta$  com parâmetros  $\mu_0 = 3$  e  $\sigma_0^2 = 100$ , aproximamos numericamente o risco de Bayes para cada um dos testes para tamanhos amostrais variando de 1 a 100 e aproximando as integrais do cálculo do risco pelo método de Monte Carlo utilizando  $M = 10^6$  iterações. A figura abaixo ilustra os resultados obtidos.

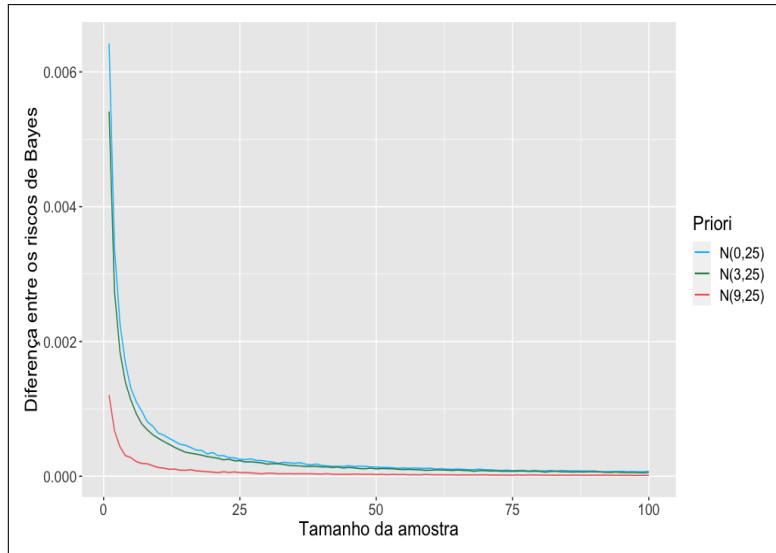


**Figura 4.5:** Riscos de Bayes dos testes unilaterais para modelo Normal com função de perda  $L_1$  - Priori  $\mathcal{N}(3, 100)$ .

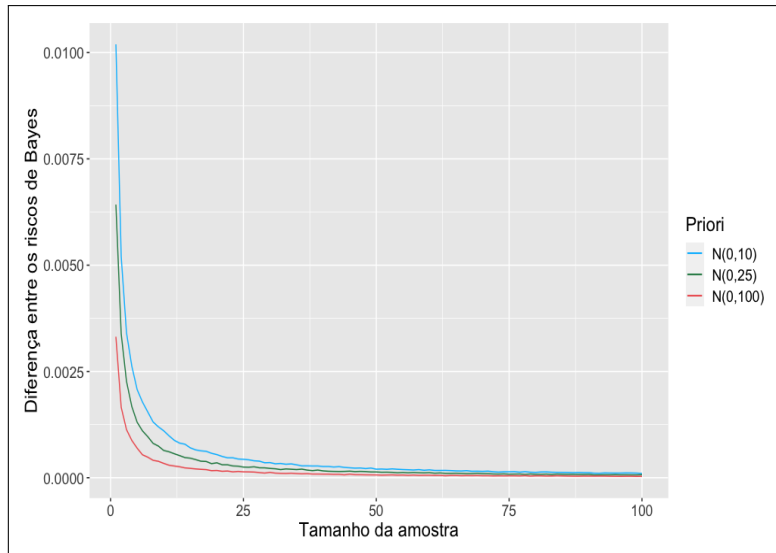
Nesse caso, notamos que os riscos dos dois testes ficam muito próximos se compararmos com o exemplo anterior que utiliza as perdas  $0 - a_0 - a_1$  e  $0 - a_0 - a_1 - b_0 - b_1$ , mas a curva do risco do caso agnóstico sempre fica abaixo do caso tradicional, como esperado. Observa-se também que o risco de Bayes decai de forma mais acentuada conforme o tamanho da amostra  $n$  aumenta.

Assim como no exemplo anterior, considere agora diferentes escolhas de prioris para  $\theta$ , a fim de comparar a diferença absoluta entre os riscos de Bayes das duas abordagens para cada uma das prioris. Tais comparações são ilustradas pelas duas figuras a seguir.





**Figura 4.6:** Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes unilaterais com perda  $0 - L_1$  para modelo Normal.



**Figura 4.7:** Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes unilaterais com perda  $0 - L_1^{(p)}$  para modelo Normal.

Observa-se que, para prioris de mesma variância, a diferença entre os riscos dos testes diminui conforme a média da priori se distancia de 0 (ponto que divide  $H_0$  e  $H_1$ ). Já para prioris de mesma média, quanto maior a variância da priori, menor a diferença entre os riscos dos testes. Tais comportamentos também foram observados no exemplo anterior.

Por fim, através do método de cálculo de tamanho amostral utilizado nos exemplos anteriores e assumindo um custo  $c = 0,0001$ , foram encontrados os tamanhos amostrais  $n_T^*$  e  $n_A^*$ , para diversas escolhas de priori, que são exibidos na tabela abaixo.

Priori $\theta$	$n_T^*$	$n_A^*$	$\delta(\%)$	$\bar{\rho}_T^*$	$\bar{\rho}_A^*$
$\mathcal{N}(-3, 10)$	20	18	10,0	0,0020	0,0018
$\mathcal{N}(-3, 25)$	18	17	5,6	0,0018	0,0016
$\mathcal{N}(-3, 100)$	14	13	7,1	0,0014	0,0012
$\mathcal{N}(0, 10)$	24	24	0,0	0,0026	0,0022
$\mathcal{N}(0, 25)$	19	18	5,3	0,0021	0,0018
$\mathcal{N}(0, 100)$	14	13	7,1	0,0014	0,0013
$\mathcal{N}(3, 10)$	19	18	5,3	0,0021	0,0018
$\mathcal{N}(3, 25)$	18	18	0,0	0,0018	0,0015
$\mathcal{N}(3, 100)$	13	12	7,7	0,0015	0,0013
$\mathcal{N}(6, 10)$	10	9	10,0	0,0010	0,0009
$\mathcal{N}(6, 25)$	15	13	13,3	0,0013	0,0012
$\mathcal{N}(6, 100)$	12	11	8,3	0,0014	0,0013
$\mathcal{N}(9, 10)$	3	3	0,0	0,0003	0,0002
$\mathcal{N}(9, 25)$	9	8	11,1	0,0009	0,0008
$\mathcal{N}(9, 100)$	11	10	9,1	0,0012	0,0011

**Tabela 4.4:** Tamanhos amostrais e riscos para os testes de Bayes unilaterais com perda  $L_1$  - Modelo Normal.

Observa-se que, de forma geral, os tamanhos amostrais obtidos são menores quando comparados com o exemplo anterior, o que era esperado, dado o decaimento mais rápido do risco de Bayes observado na Figura 4.5. Ao comparar o teste tradicional com o teste agnóstico, notamos que os tamanhos amostrais no cenário agnóstico são menores ou iguais que os obtidos no cenário tradicional, e nos casos de  $n$  igual, o risco de Bayes no cenário agnóstico é menor. Vale ressaltar, que nesse caso, se fixarmos a mesma priori, as diferenças entre os tamanhos de amostra ótimos dos dois cenários são menores, se comparadas com o exemplo anterior. Note também que, assim como no caso anterior, os maiores tamanhos amostrais são observados quando, a priori, temos mais incerteza em relação a  $H_0$  ou  $H_1$  (ou seja, quando  $\mu_0 = 0$ ) e os menores tamanhos são obtidos quando temos mais certeza a respeito de algumas das hipóteses (ou seja, quando a média da priori está mais distante de zero).

### 4.3.3 Teste baseado em Probabilidade a Posteriori: Modelo Bernoulli

Considere um experimento em que será observada uma amostra  $\mathbf{X} = (X_1, \dots, X_n)$  em que assumimos que  $X_1, \dots, X_n$ , dado  $\theta$ , são condicionalmente i.i.d com distribuição Bernoulli de parâmetro  $\theta$ , condições essas que denotamos por  $X_i | \theta \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$ , para  $i = 1, \dots, n$  e  $\theta \in (0, 1)$ . Vamos testar as hipóteses  $H_0 : \theta \leq \theta_0$  e  $H_1 : \theta > \theta_0$ ,  $\theta_0 \in (0, 1)$ .

Para o modelo Bernoulli, temos, por conjugação, que se tomarmos a priori para  $\theta$  uma distribuição Beta, digamos,  $\theta \sim \text{Beta}(\alpha, \beta)$ ,  $\alpha > 0$ ,  $\beta > 0$ , a distribuição a posteriori de  $\theta$  é Beta tal que

$$\theta | \mathbf{x} \sim \text{Beta} \left( \alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i \right).$$

O teste de Bayes construído a partir da função de perda  $0 - a_0 - a_1$  para  $H_0$  contra  $H_1$ , baseado em probabilidade a posteriori das hipóteses, pode ser escrito como

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & F_{\theta|\mathbf{x}}(\theta_0) > \frac{a_1}{a_0 + a_1} \\ 1, & F_{\theta|\mathbf{x}}(\theta_0) \leq \frac{a_1}{a_0 + a_1} \end{cases}, \quad (4.15)$$

em que  $F_{\theta|\mathbf{x}}(\cdot)$  é a função de distribuição de  $\theta | \mathbf{x}$ .

Como neste modelo o espaço amostral  $\mathcal{X}$  é discreto ( $X_i \in \{0, 1\}$ ), para o cálculo do risco de Bayes substituímos uma das integrais de Riemann por uma soma múltipla (em várias dimensões)

na expressão (3.2), de modo que

$$\begin{aligned} \bar{\rho}(\varphi_T, \bar{\mathcal{P}}) &= E[L(\varphi_T(\mathbf{X}), \theta)] = \sum_{\mathbf{x} \in \mathcal{X}} \int_{\Theta} L(\varphi_T(\mathbf{x}), \theta) f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &= a_0 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{1\})} \int_{\theta \leq \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) + a_1 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{0\})} \int_{\theta > \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}). \end{aligned} \quad (4.16)$$

Acima, decompos a distribuição conjunta de  $(\theta, \mathbf{X})$  na função de densidade condicional de  $\theta$  dado  $\mathbf{x}$ ,  $f(\theta|\mathbf{x})$ , e na função de probabilidade marginal  $\mathbb{P}(\mathbf{X} = \mathbf{x})$ . Adicionalmente, temos que  $S(\mathbf{X}) = \sum_{i=1}^n X_i$  é estatística suficiente para o modelo Bernoulli e, conforme demonstrado no apêndice, sua distribuição no modelo proposto é *Beta – Binomial*( $n, \alpha, \beta$ ),  $\alpha > 0$ ,  $\beta > 0$ , com função de probabilidade

$$\mathbb{P}(S(\mathbf{X}) = s) = \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \mathbb{1}_{\{0, \dots, n\}}(s),$$

onde  $B(.,.)$  é a função beta e  $s \in \{0, 1, \dots, n\}$ .

Para o primeiro termo da equação (4.16), temos

$$\begin{aligned} a_0 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{1\})} \int_{\theta \leq \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) &= a_0 \sum_{s \in R} \mathbb{P}(\theta \leq \theta_0 | s) \mathbb{P}(S(\mathbf{X}) = s) \\ &= a_0 \sum_{s \in R} \left[ F_{\theta|s}(\theta_0) \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \right], \end{aligned}$$

onde  $R = \left\{ s \in \{0, 1, \dots, n\} : F_{\theta|s}(\theta_0) \leq \frac{a_1}{a_0 + a_1} \right\}$ .

De forma análoga para o segundo termo, obtemos

$$\begin{aligned} a_1 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{0\})} \int_{\theta > \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) &= a_1 \sum_{s \in A} \mathbb{P}(\theta > \theta_0 | s) \mathbb{P}(S(\mathbf{X}) = s) \\ &= a_1 \sum_{s \in A} \left[ (1 - F_{\theta|s}(\theta_0)) \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \right], \end{aligned}$$

onde  $A = \left\{ s \in \{0, 1, \dots, n\} : F_{\theta|s}(\theta_0) > \frac{a_1}{a_0 + a_1} \right\}$ .

Logo, o risco de Bayes pode ser escrito como

$$\begin{aligned} \bar{\rho}(\varphi_T, \bar{\mathcal{P}}) &= \sum_{s=0}^n \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} [a_0 F_{\theta|s}(\theta_0) \mathbb{1}_R(s) + a_1 (1 - F_{\theta|s}(\theta_0)) \mathbb{1}_A(s)] \\ &= \sum_{s=0}^n \rho'_T(s). \end{aligned} \quad (4.17)$$

Para o cálculo numérico do risco, podemos encontrar os valores da função de distribuição da posteriori Beta e da função de probabilidade da Beta-Binomial de forma computacional. Já para encontrar os valores de  $s$  que levam à rejeição e aceitação da hipótese nula, basta avaliar para cada valor possível de  $s = 0, 1, \dots, n$ , se  $F_{\theta|s}(\theta_0) \leq \frac{a_1}{a_0 + a_1}$  ou  $F_{\theta|s}(\theta_0) > \frac{a_1}{a_0 + a_1}$ . Nesse caso, diferentemente do modelo Normal, obtemos um valor exato para o risco de Bayes.

Agora, considere um teste de Bayes agnóstico para a mesma situação, construído a partir da função de perda  $a_0 - a_1 - b_0 - b_1$ . O teste unilateral para o modelo proposto pode ser escrito

como

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & F_{\theta|\mathbf{x}}(\theta_0) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ \frac{1}{2}, & \frac{b_1}{a_0 + b_1 - b_0} < F_{\theta|\mathbf{x}}(\theta_0) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ 1, & F_{\theta|\mathbf{x}}(\theta_0) \leq \frac{b_1}{a_0 + b_1 - b_0} \end{cases} . \quad (4.18)$$

Para o caso agnóstico, o risco de Bayes é dado por

$$\begin{aligned} \bar{\rho}(\varphi_A, \bar{\mathcal{P}}) &= a_0 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{1\})} \int_{\theta \leq \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &+ a_1 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{0\})} \int_{\theta > \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &+ b_0 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{\frac{1}{2}\})} \int_{\theta \leq \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &+ b_1 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{\frac{1}{2}\})} \int_{\theta > \theta_0} f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}). \end{aligned} \quad (4.19)$$

Analogamente ao caso anterior e denotando as parcelas da equação acima por  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , respectivamente, temos

$$\begin{aligned} I_1 &= a_0 \sum_{s \in R} \left[ F_{\theta|s}(\theta_0) \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \right]; \\ I_2 &= a_1 \sum_{s \in A} \left[ (1 - F_{\theta|s}(\theta_0)) \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \right]; \\ I_3 &= b_0 \sum_{s \in G} \left[ F_{\theta|s}(\theta_0) \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \right]; \\ I_4 &= b_1 \sum_{s \in G} \left[ (1 - F_{\theta|s}(\theta_0)) \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \right], \end{aligned}$$

onde

$$\begin{aligned} R &= \left\{ s \in \{0, 1, \dots, n\} : F_{\theta|s}(\theta_0) \leq \frac{b_1}{a_0 + b_1 - b_0} \right\} \\ A &= \left\{ s \in \{0, 1, \dots, n\} : F_{\theta|s}(\theta_0) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right\} \\ G &= \left\{ s \in \{0, 1, \dots, n\} : \frac{b_1}{a_0 + b_1 - b_0} < F_{\theta|s}(\theta_0) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right\}. \end{aligned}$$

Simplificando, temos que o risco é dado por

$$\begin{aligned} \bar{\rho}(\varphi_A, \bar{\mathcal{P}}) &= \sum_{s=0}^n \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \left[ a_0 F_{\theta|s}(\theta_0) \mathbb{1}_R(s) + a_1 (1 - F_{\theta|s}(\theta_0)) \mathbb{1}_A(s) \right. \\ &\quad \left. + b_0 F_{\theta|s}(\theta_0) \mathbb{1}_G(s) + b_1 (1 - F_{\theta|s}(\theta_0)) \mathbb{1}_G(s) \right] \\ &= \sum_{s=0}^n \rho'_A(s). \end{aligned} \quad (4.20)$$

**Exemplo 4.3.** Sejam  $X_1, \dots, X_n$  em que  $X_i|\theta \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$ , para  $i = 1, \dots, n$  e  $\theta \in (0, 1)$ . Suponha que iremos testar as hipóteses  $H_0 : \theta \leq 0,5$  e  $H_1 : \theta > 0,5$ . Para comparar os riscos de Bayes dos dois testes expostos acima, os valores escolhidos para a função de perda para o teste tradicional são  $a_0 = 1, a_1 = 1$  e  $a_0 = 1, a_1 = 1, b_0 = \frac{1}{4}, b_1 = \frac{1}{4}$  para a função de perda do teste agnóstico.

Substituindo os valores das perdas, temos os seguintes testes ótimos

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & F_{\theta|s}(0,5) > 0,5 \\ 1, & F_{\theta|s}(0,5) \leq 0,5 \end{cases}$$

e

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & F_{\theta|s}(0,5) > 0,75 \\ \frac{1}{2}, & 0,25 < F_{\theta|s}(0,5) \leq 0,75 \\ 1, & F_{\theta|s}(0,5) \leq 0,25 \end{cases} ,$$

em que  $s = \sum_{i=1}^n x_i$ .

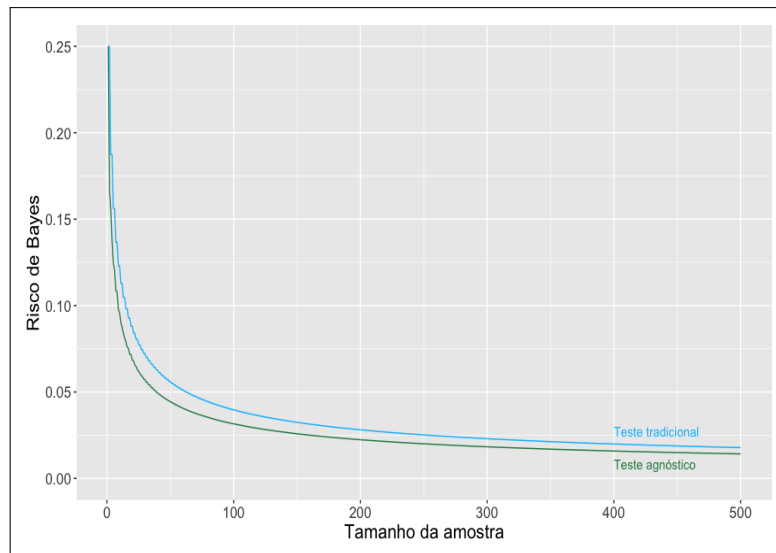
De forma a ilustrar o cálculo dos riscos de Bayes pelas expressões (4.16) e (4.20), considere a priori  $\theta \sim \text{Beta}(1, 1)$  e que  $n = 10$ . A tabela abaixo apresenta, para cada valor de  $s$ , qual o valor de  $F_{\theta|s}(0,5)$ , as decisões de cada teste e os valores das parcelas  $\rho'_T(s)$  e  $\rho'_A(s)$ .

s	$F_{\theta s}(0,5)$	$\varphi_T$	$\rho'_T(s)$	$\varphi_A$	$\rho'_A(s)$
0	0,9995	0	0,00004	0	0,00004
1	0,9941	0	0,00053	0	0,00053
2	0,9673	0	0,00297	0	0,00297
3	0,8868	0	0,01030	0	0,01030
4	0,7256	0	0,02495	$\frac{1}{2}$	0,22727
5	0,5000	1	0,04545	$\frac{1}{2}$	0,22727
6	0,2744	1	0,02495	$\frac{1}{2}$	0,22727
7	0,1133	1	0,01030	1	0,01030
8	0,0327	1	0,00297	1	0,00297
9	0,0059	1	0,00053	1	0,00053
10	0,0005	1	0,00004	1	0,00004

**Tabela 4.5:** Exemplo de cálculo dos riscos de Bayes de cada teste para  $n = 10$ .

Logo, os riscos do teste tradicional e agnóstico valem, respectivamente,  $\sum_{s=0}^n \rho'_T(s) = 0,12305$  e  $\sum_{s=0}^n \rho'_A(s) = 0,09588$ .

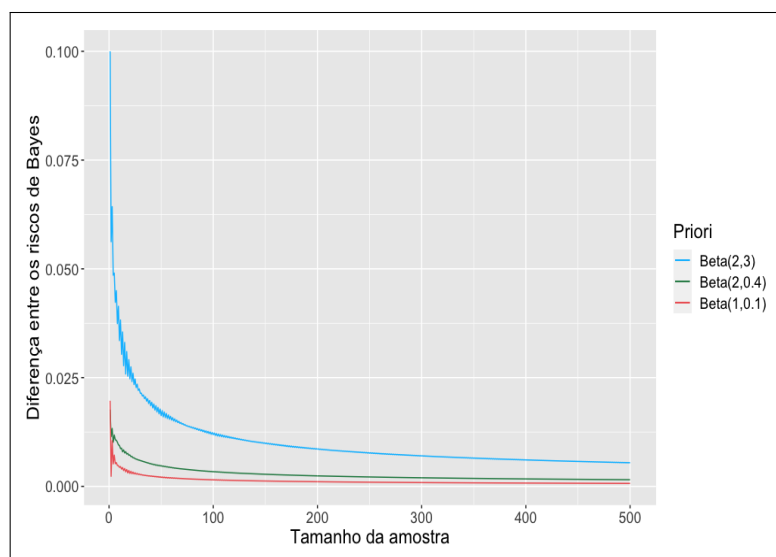
Repetindo o procedimento acima, considerando a mesma priori e variando  $n$  de 1 a 500, temos os seguintes resultados representados pela figura abaixo.



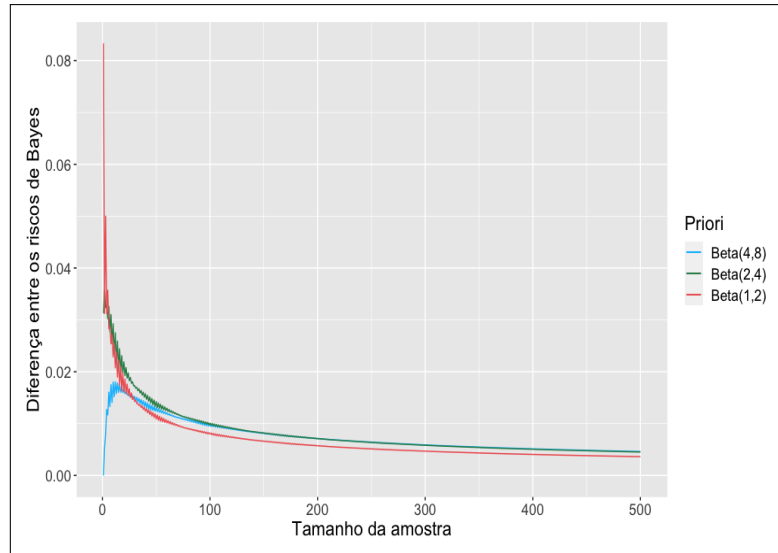
**Figura 4.8:** *Riscos de Bayes dos testes unilaterais para modelo Bernoulli - Priori  $Beta(1,1)$ .*

Assim como nos exemplos anteriores, os riscos diminuem conforme  $n$  aumenta e o risco do teste agnóstico é menor que o risco do teste tradicional para todo  $n$  avaliado.

Consideramos agora diferentes escolhas de priori a fim de comparar a diferença absoluta entre os riscos de Bayes da cada teste e determinar os tamanhos amostrais ótimos assumindo um custo  $c = 0,001$ . As figuras 4.9 e 4.10 e a tabela 4.6 apresentam os resultados obtidos. A esperança e a variância de cada priori de  $\theta$  são apresentadas na tabela de forma a facilitar a comparação.



**Figura 4.9:** *Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias semelhantes - Testes unilaterais para modelo Bernoulli.*



**Figura 4.10:** Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes unilaterais para modelo Bernoulli.

Priori $\theta$	$\mathbb{E}(\theta)$	$Var(\theta)$	$n_T^*$	$n_A^*$	$\delta(\%)$	$\bar{\rho}_T^*$	$\bar{\rho}_A^*$
$Beta(1, 1)$	0,50	0,083	33	29	12,1	0,0679	0,0576
$Beta(2, 2)$	0,50	0,050	41	35	14,6	0,0904	0,0771
$Beta(4, 4)$	0,50	0,028	53	44	17,0	0,1142	0,0979
$Beta(1, 2)$	0,33	0,056	32	28	12,5	0,0679	0,0576
$Beta(2, 4)$	0,33	0,031	35	29	17,1	0,0783	0,0675
$Beta(4, 8)$	0,33	0,017	25	23	8,0	0,0820	0,0685
$Beta(2, 3)$	0,40	0,040	40	34	15,0	0,0904	0,0771
$Beta(2, 0,4)$	0,83	0,041	17	14	17,6	0,0384	0,0335
$Beta(1, 0,1)$	0,91	0,039	10	8	20,0	0,0384	0,0335

**Tabela 4.6:** Tamanhos amostrais e riscos para os testes de Bayes unilaterais - Modelo Bernoulli.

Ao analisar a Figura 4.9, notamos que, assim como nos exemplos do modelo Normal, para prioris de variância similar, quanto mais o valor da esperança de  $\theta$  se distancia de  $\theta_0 (= 0,5)$ , menor a diferença entre os riscos dos testes, e menores os tamanhos amostrais ótimos. Já para prioris de mesma média, observamos na Figura 4.10 que a diferença entre os riscos é menor para a priori de maior variância. Na Tabela 4.6, vale destacar que para as prioris de esperança igual a 0,5, o menor tamanho amostral é obtido na priori de maior variância. Já quando a esperança é igual a 0,33, o menor tamanho amostral é obtido na priori de menor variância. Uma forma de se interpretar tais resultados é da seguinte forma: quanto maior nossa crença de que o parâmetro está próximo de  $\theta_0$  (ponto que separa as regiões de  $H_0$  e  $H_1$ ), mais observações são necessárias para decidirmos a respeito de  $H_0$  e  $H_1$ ; ao passo que, para as prioris de esperança de  $\theta$  igual a 0,33, quanto menor a variância (ou seja, temos maior crença, a priori, em  $H_0$ ), menos observações são necessárias para se chegar a alguma conclusão sobre as hipóteses. Finalmente, os tamanhos amostrais ótimos do cenário tradicional são, em geral, de 10 a 20% maiores que os tamanhos amostrais do cenário agnóstico, quando a mesma priori é considerada para os dois cenários. Tais resultados foram similares aos obtidos no exemplo 4.1, em que foi utilizado o modelo Normal também sob perdas  $0 - a_0 - a_1$  e  $0 - a_0 - a_1 - b_0 - b_1$ .

## 4.4 Estudo comparativo: Testes de Hipóteses Intervalares

Para os estudos desta seção, considere que desejamos testar a hipótese  $H_0 : \theta \in \Theta_0$  contra  $H_1 : \theta \in \Theta_1$ , em que  $\Theta \subseteq \mathbb{R}$ ,  $\Theta_0 = \{\theta \in \Theta : \theta_1 \leq \theta \leq \theta_2\}$ ,  $\Theta_1 = \{\theta \in \Theta : \theta < \theta_1 \text{ ou } \theta > \theta_2\}$ ,  $\theta_1, \theta_2 \in \Theta$ , com  $\theta_1 < \theta_2$ .

### 4.4.1 Teste baseado em Probabilidade a Posteriori: Modelo Normal com Variância conhecida

Considerando que  $H_0$  é um intervalo limitado de números reais e que  $(\theta, \mathbf{X})$  é contínuo, temos que

$$\mathbb{P}(\theta \in \Theta_0 | \mathbf{x}) = \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | \mathbf{x}) = F_{\theta|\mathbf{x}}(\theta_2) - F_{\theta|\mathbf{x}}(\theta_1).$$

Assim o teste de Bayes sob função de perda  $0 - a_0 - a_1$ , para o modelo Normal proposto em 4.3.1, é dado por

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & \Phi\left(\frac{\theta_2 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{\theta_1 - \mu_x}{\sigma_x}\right) > \frac{a_1}{a_0 + a_1} \\ 1, & \Phi\left(\frac{\theta_2 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{\theta_1 - \mu_x}{\sigma_x}\right) \leq \frac{a_1}{a_0 + a_1} \end{cases}, \quad (4.21)$$

em que  $\Phi(\cdot)$  é a função de distribuição da variável aleatória normal  $Z \sim \mathcal{N}(0, 1)$  e  $\mu_x = \frac{\mu_0 v^2 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2}$  e  $\sigma_x^2 = \frac{v^2 \sigma_0^2}{v^2 + n \sigma_0^2}$ .

Para o cálculo do risco de Bayes decomposmos a função de densidade conjunta nas funções condicional  $f(\theta|\mathbf{x})$  e marginal  $f(\mathbf{x})$ , o que resulta em

$$\begin{aligned} \bar{\rho}(\varphi_T, \bar{\mathcal{P}}) &= a_0 \int_{\varphi_T^{-1}(\{1\})} \int_{\Theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} + a_1 \int_{\varphi_T^{-1}(\{0\})} \int_{\Theta_1} f(\theta, \mathbf{x}) d\theta d\mathbf{x} \\ &= a_0 \int_{\varphi_T^{-1}(\{1\})} \left[ \int_{\Theta_0} f(\theta|\mathbf{x}) d\theta \right] f(\mathbf{x}) d\mathbf{x} + a_1 \int_{\varphi_T^{-1}(\{0\})} \left[ \int_{\Theta_1} f(\theta|\mathbf{x}) d\theta \right] f(\mathbf{x}) d\mathbf{x} \\ &= a_0 \int_{\varphi_T^{-1}(\{1\})} \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \\ &\quad + a_1 \int_{\varphi_T^{-1}(\{0\})} \mathbb{P}(\theta < \theta_1 \cup \theta > \theta_2 | \mathbf{x}) f(\mathbf{x}) d\mathbf{x} \\ &= a_0 \int_{\varphi_T^{-1}(\{1\})} \left[ \Phi\left(\frac{\theta_2 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{\theta_1 - \mu_x}{\sigma_x}\right) \right] f(\mathbf{x}) d\mathbf{x} \\ &\quad + a_1 \int_{\varphi_T^{-1}(\{0\})} \left[ 1 - \Phi\left(\frac{\theta_2 - \mu_x}{\sigma_x}\right) + \Phi\left(\frac{\theta_1 - \mu_x}{\sigma_x}\right) \right] f(\mathbf{x}) d\mathbf{x}. \end{aligned} \quad (4.22)$$

As integrais são aproximadas novamente pelo método de Monte Carlo calculando as médias empíricas

$$\bar{h}_i = \frac{1}{M} \sum_{j=1}^M h_i(\bar{x}_j), \quad i = 1, 2; \quad j = 1, 2, \dots, M,$$



em que

$$\begin{aligned} h_1(\bar{x}_j) &= a_0 \left[ \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] \mathbb{1}_R(\bar{x}_j); \\ h_2(\bar{x}_j) &= a_1 \left[ 1 - \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) + \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] \mathbb{1}_A(\bar{x}_j); \\ R &= \left\{ \bar{x} \in \mathbb{R} : \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \leq \frac{a_1}{a_0 + a_1} \right\}; \\ A &= \left\{ \bar{x} \in \mathbb{R} : \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) > \frac{a_1}{a_0 + a_1} \right\}. \end{aligned}$$

Já o teste de Bayes agnóstico, construído a partir da função de perda  $a_0 - a_1 - b_0 - b_1$ , é dado por

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ \frac{1}{2}, & \frac{b_1}{a_0 + b_1 - b_0} < \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ 1, & \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \leq \frac{b_1}{a_0 + b_1 - b_0} \end{cases}. \quad (4.23)$$

Procedendo de maneira análoga ao caso anterior, a expressão do risco de Bayes á dada por

$$\begin{aligned} \bar{\rho}(\varphi_A, \bar{\mathcal{P}}) &= a_0 \int_{\varphi_A^{-1}(\{1\})} \int_{\Theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} + a_1 \int_{\varphi_A^{-1}(\{0\})} \int_{\Theta_1} f(\theta, \mathbf{x}) d\theta d\mathbf{x} \\ &\quad + b_0 \int_{\varphi_A^{-1}(\{\frac{1}{2}\})} \int_{\Theta_0} f(\theta, \mathbf{x}) d\theta d\mathbf{x} + b_1 \int_{\varphi_A^{-1}(\{\frac{1}{2}\})} \int_{\Theta_1} f(\theta, \mathbf{x}) d\theta d\mathbf{x}. \end{aligned} \quad (4.24)$$

Denotando as integrais em (4.24) por  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , respectivamente, obtemos

$$\begin{aligned} I_1 &= a_0 \int_{\varphi_A^{-1}(\{1\})} \left[ \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] f(\mathbf{x}) d\mathbf{x}; \\ I_2 &= a_1 \int_{\varphi_A^{-1}(\{0\})} \left[ 1 - \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) + \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] f(\mathbf{x}) d\mathbf{x}; \\ I_3 &= b_0 \int_{\varphi_A^{-1}(\{\frac{1}{2}\})} \left[ \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] f(\mathbf{x}) d\mathbf{x}; \\ I_4 &= b_1 \int_{\varphi_A^{-1}(\{\frac{1}{2}\})} \left[ 1 - \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) + \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] f(\mathbf{x}) d\mathbf{x}; \end{aligned}$$

que são aproximadas pelo método de Monte Carlo, respectivamente, através da simulação de  $\bar{x}_1, \dots, \bar{x}_M$ , e calculando as seguintes expressões:

$$\begin{aligned} h_1(\bar{x}_j) &= a_0 \left[ \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] \mathbb{1}_R(\bar{x}_j); \\ h_2(\bar{x}_j) &= a_1 \left[ 1 - \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) + \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] \mathbb{1}_A(\bar{x}_j); \\ h_3(\bar{x}_j) &= b_0 \left[ \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) - \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] \mathbb{1}_G(\bar{x}_j); \\ h_4(\bar{x}_j) &= b_1 \left[ 1 - \Phi \left( \frac{\theta_2 - \mu_x}{\sigma_x} \right) + \Phi \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right) \right] \mathbb{1}_G(\bar{x}_j); \end{aligned}$$

onde

$$R = \left\{ \bar{x} \in \mathbb{R} : \Phi\left(\frac{\theta_2 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{\theta_1 - \mu_x}{\sigma_x}\right) \leq \frac{b_1}{a_0 + b_1 - b_0} \right\};$$

$$A = \left\{ \bar{x} \in \mathbb{R} : \Phi\left(\frac{\theta_2 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{\theta_1 - \mu_x}{\sigma_x}\right) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right\};$$

$$G = \left\{ \bar{x} \in \mathbb{R} : \frac{b_1}{a_0 + b_1 - b_0} < \Phi\left(\frac{\theta_2 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{\theta_1 - \mu_x}{\sigma_x}\right) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right\}.$$

**Exemplo 4.4.** Sejam  $X_1, \dots, X_n$  em que  $X_i | \theta \stackrel{iid}{\sim} \mathcal{N}(\theta, 1)$ , para  $i = 1, \dots, n$ . As hipóteses a serem testadas são  $H_0 : 2 \leq \theta \leq 4$  e  $H_1 : \theta < 2$  ou  $\theta > 4$ . Com o intuito de comparar os riscos de Bayes de cada tipo de teste em função do tamanho amostral  $n$ , consideremos  $a_0 = 1$ ,  $a_1 = 1$ ,  $b_0 = \frac{1}{4}$  e  $b_1 = \frac{1}{4}$  para as funções de perda.

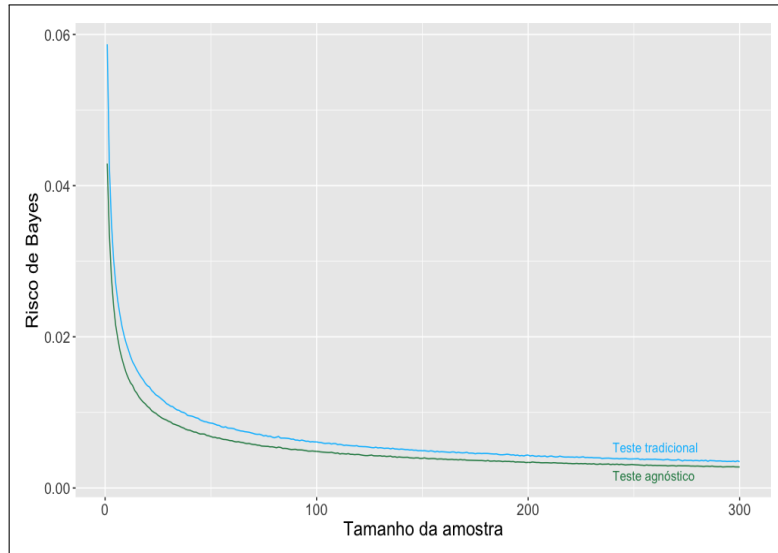
Substituindo os valores acima em (4.21) e (4.23), os dois testes de Bayes são dados por

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & \Phi\left(\frac{4 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{2 - \mu_x}{\sigma_x}\right) > 0,5 \\ 1, & \Phi\left(\frac{4 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{2 - \mu_x}{\sigma_x}\right) \leq 0,5 \end{cases}$$

e

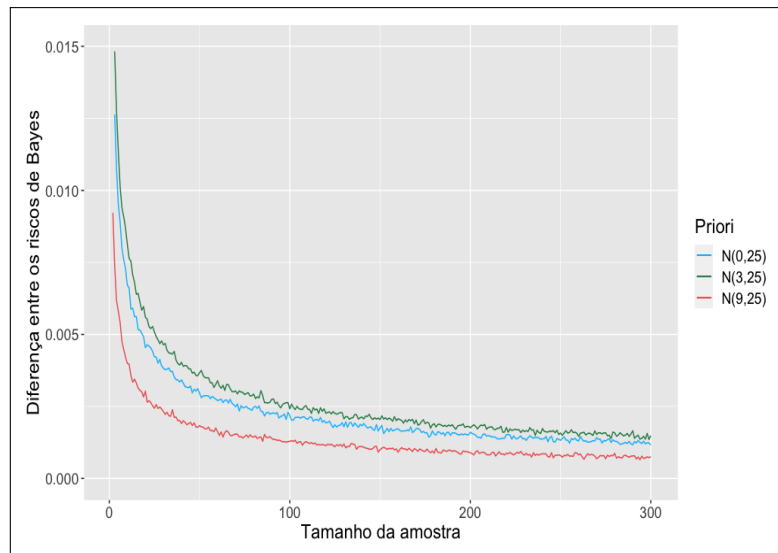
$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \Phi\left(\frac{4 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{2 - \mu_x}{\sigma_x}\right) > 0,75 \\ \frac{1}{2}, & 0,25 < \Phi\left(\frac{4 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{2 - \mu_x}{\sigma_x}\right) \leq 0,75 \\ 1, & \Phi\left(\frac{4 - \mu_x}{\sigma_x}\right) - \Phi\left(\frac{2 - \mu_x}{\sigma_x}\right) \leq 0,25 \end{cases}.$$

Supondo uma priori com parâmetros  $\mu_0 = 0$  e  $\sigma_0^2 = 100$ , o risco de Bayes é calculado para cada um dos testes para tamanhos amostrais variando de 1 a 300. As integrais para o cálculo do risco foram aproximadas pelo método de Monte Carlo com  $M = 10^6$  iterações. A figura abaixo ilustra os resultados obtidos.

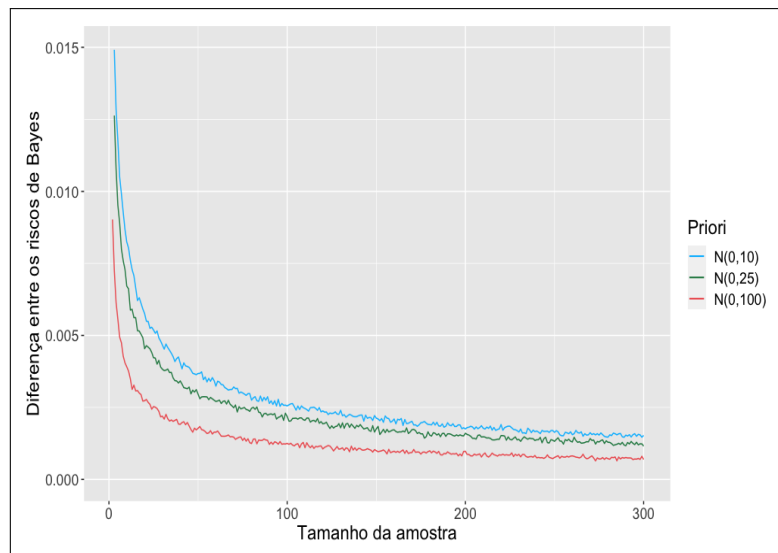


**Figura 4.11:** Riscos de Bayes dos testes intervalares para modelo Normal - Priori  $\mathcal{N}(0, 100)$ .

A seguir, são consideradas diferentes escolhas de prioris para  $\theta$ . As Figuras 4.12 e 4.13 comparam a diferença absoluta entre os dois riscos de Bayes para cada priori, sendo que a primeira compara prioris de mesma variância e médias diferentes e a segunda compara prioris de mesma média e variâncias diferentes.



**Figura 4.12:** *Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes intervalares para modelo Normal.*



**Figura 4.13:** *Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes intervalares para modelo Normal.*

Por fim, assumindo um custo  $c = 0,0001$ , foram encontrados os tamanhos amostrais ótimos  $n_T^*$  e  $n_A^*$ , que são exibidos na tabela a seguir.

Priori $\theta$	$n_T^*$	$n_A^*$	$\delta(\%)$	$\bar{\rho}_T^*$	$\bar{\rho}_A^*$
$\mathcal{N}(0, 10)$	79	66	16,5	0,0143	0,0125
$\mathcal{N}(0, 25)$	75	57	24,0	0,0120	0,0110
$\mathcal{N}(0, 100)$	50	38	24,0	0,0084	0,0077
$\mathcal{N}(3, 10)$	101	84	16,8	0,0189	0,0165
$\mathcal{N}(3, 25)$	77	64	16,9	0,0141	0,0123
$\mathcal{N}(3, 100)$	45	36	20,0	0,0093	0,0083
$\mathcal{N}(6, 10)$	78	63	19,2	0,0143	0,0127
$\mathcal{N}(6, 25)$	63	54	14,3	0,0131	0,0113
$\mathcal{N}(6, 100)$	47	39	17,0	0,0088	0,0077
$\mathcal{N}(9, 10)$	31	27	12,9	0,0067	0,0057
$\mathcal{N}(9, 25)$	46	41	10,9	0,0091	0,0077
$\mathcal{N}(9, 100)$	46	38	17,4	0,0077	0,0068

**Tabela 4.7:** Tamanhos amostrais e riscos para os testes de Bayes intervalares - Modelo Normal.

Os resultados são bastante similares aos obtidos no exemplo 4.1 da seção 4.3.1, tanto na comparação entre  $n_T^*$  e  $n_A^*$  para a mesma priori, quanto no impacto da média e variância da priori na diferença entre os riscos e nos tamanhos amostrais ótimos.

#### 4.4.2 Teste baseado em Probabilidade a Posteriori: Modelo Poisson

Considere um experimento em que será observada uma amostra  $\mathbf{X} = (X_1, \dots, X_n)$  na qual assumimos que  $X_1, \dots, X_n$ , dado  $\theta$ , são condicionalmente i.i.d com distribuição Poisson de parâmetro  $\theta$ , denotado por  $X_i | \theta \stackrel{iid}{\sim} \text{Poisson}(\theta)$ , para  $i = 1, \dots, n$  e  $\theta > 0$ . Utilizando outro resultado de conjugação (DeGroot [16]), sabe-se que se tomarmos como priori para  $\theta$  uma distribuição Gama, digamos,  $\theta \sim \mathcal{G}(\alpha, \beta)$ ,  $\alpha > 0$ ,  $\beta > 0$ , a distribuição de  $\theta | \mathbf{x}$  é também Gama tal que

$$\theta | \mathbf{x} \sim \mathcal{G} \left( \alpha + \sum_{i=1}^n x_i, \beta + n \right).$$

Assim, o teste de Bayes, para testar  $H_0 : \theta_1 \leq \theta \leq \theta_2$  contra  $H_1 : \theta < \theta_1$  ou  $\theta > \theta_2$ ,  $0 < \theta_1 < \theta_2$ , sob perda  $0 - a_0 - a_1$ , fica dado por

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & F_{\theta|\mathbf{x}}(\theta_2) - F_{\theta|\mathbf{x}}(\theta_1) > \frac{a_1}{a_0 + a_1} \\ 1, & F_{\theta|\mathbf{x}}(\theta_2) - F_{\theta|\mathbf{x}}(\theta_1) \leq \frac{a_1}{a_0 + a_1} \end{cases}, \quad (4.25)$$

em que  $F_{\theta|\mathbf{x}}(\cdot)$  é a função de distribuição da posteriori dada acima.

Assim como no modelo Bernoulli da seção 4.3.3, o espaço amostral  $\mathcal{X}$  é discreto ( $X_i \in \mathbb{N}$ ). Consequentemente, o risco de Bayes é dado pela expressão

$$\begin{aligned} \bar{\rho}(\varphi_T, \bar{\mathcal{P}}) &= E[L(\varphi_T(\mathbf{X}, \Theta))] = \sum_{\mathbf{x} \in \mathcal{X}} \int_{\Theta} L(\varphi_T(\mathbf{x}, \theta)) f(\theta|\mathbf{x}) d\theta \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &= a_0 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{1\})} \left[ \int_{\Theta_0} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &\quad + a_1 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{0\})} \left[ \int_{\Theta_1} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}), \end{aligned} \quad (4.26)$$

em que  $\Theta_0 = \{\theta > 0 : \theta_1 \leq \theta \leq \theta_2\}$  e  $\Theta_1 = \{\theta > 0 : \theta < \theta_1 \text{ ou } \theta > \theta_2\}$ .

Nesse caso, a estatística suficiente para o modelo é  $S(\mathbf{X}) = \sum_{i=1}^n X_i$ , cuja distribuição marginal

é *Binomial – Negativa*  $\left(\alpha, \frac{\beta}{\beta+n}\right)$  quando  $\alpha \in \mathbb{N}^*$ , conforme demonstrado no apêndice, com função de probabilidade

$$\mathbb{P}(S(\mathbf{X}) = s) = \frac{\Gamma(\alpha + s)}{\Gamma(\alpha)s!} \left(\frac{\beta}{\beta+n}\right)^\alpha \left(\frac{n}{\beta+n}\right)^s \mathbb{1}_{\mathbb{N}}(s),$$

onde  $\Gamma(\cdot)$  representa a função gama.

Para o primeiro termo da expressão (4.26), temos:

$$\begin{aligned} a_0 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{1\})} \left[ \int_{\Theta_0} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}) &= a_0 \sum_{\mathbf{x} \in \varphi_T^{-1}(\{1\})} \mathbb{P}(\theta_1 \leq \theta \leq \theta_2|\mathbf{x}) \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &= a_0 \sum_{s \in R} \mathbb{P}(\theta_1 \leq \theta \leq \theta_2|s) \mathbb{P}(S(\mathbf{X}) = s), \end{aligned}$$

em que  $R = \left\{s \in \mathbb{N} : \mathbb{P}(\theta_1 \leq \theta \leq \theta_2|s) \leq \frac{a_1}{a_0+a_1}\right\}$ .

De forma análoga, para o segundo termo, obtemos:

$$\begin{aligned} a_1 \sum_{\mathbf{x} \in \varphi^{-1}(\{0\})} \left[ \int_{\Theta_1} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}) &= a_1 \sum_{\mathbf{x} \in \varphi^{-1}(\{0\})} \mathbb{P}(\theta < \theta_1 \cup \theta > \theta_2|\mathbf{x}) \mathbb{P}(\mathbf{X} = \mathbf{x}) \\ &= a_1 \sum_{s \in A} [1 - \mathbb{P}(\theta_1 \leq \theta \leq \theta_2|s)] \mathbb{P}(S(\mathbf{X}) = s), \end{aligned}$$

em que  $A = \left\{s \in \mathbb{N} : \mathbb{P}(\theta_1 \leq \theta \leq \theta_2|s) > \frac{a_1}{a_0+a_1}\right\}$ .

O valor aproximado do risco de Bayes pode ser obtido através de uma simulação de Monte Carlo que, nesse caso, visa aproximar uma série ao invés de uma integral. Simulam-se  $M$  vezes da distribuição de  $S(\mathbf{X}) = \sum_{i=1}^n X_i$ , que é conhecida conforme visto acima, gerando observações  $s_1, \dots, s_M$  e, a partir daí, calcula-se a seguinte expressão:

$$\frac{1}{M} \sum_{j=1}^M \left\{ \left[ a_0 \mathbb{P}(\theta_1 \leq \theta \leq \theta_2|s_j) \mathbb{1}_R(s_j) \right] + \left[ a_1 (1 - \mathbb{P}(\theta_1 \leq \theta \leq \theta_2|s_j)) \mathbb{1}_A(s_j) \right] \right\}.$$

Agora vamos examinar o teste de Bayes agnóstico construído a partir da função de perda  $a_0 - a_1 - b_0 - b_1$ , baseado em probabilidade a posteriori das hipóteses. Tal teste, no caso intervalar, é dado por

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & F_{\theta|\mathbf{x}}(\theta_2) - F_{\theta|\mathbf{x}}(\theta_1) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ \frac{1}{2}, & \frac{b_1}{a_0 + b_1 - b_0} < F_{\theta|\mathbf{x}}(\theta_2) - F_{\theta|\mathbf{x}}(\theta_1) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \\ 1, & F_{\theta|\mathbf{x}}(\theta_2) - F_{\theta|\mathbf{x}}(\theta_1) \leq \frac{b_1}{a_0 + b_1 - b_0} \end{cases} \quad (4.27)$$

O risco de Bayes no caso agnóstico é dado por

$$\begin{aligned}
\bar{\rho}(\varphi_A, \bar{P}) &= a_0 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{1\})} \left[ \int_{\Theta_0} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}) \\
&+ a_1 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{0\})} \left[ \int_{\Theta_1} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}) \\
&+ b_0 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{\frac{1}{2}\})} \left[ \int_{\Theta_0} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}) \\
&+ b_1 \sum_{\mathbf{x} \in \varphi_A^{-1}(\{\frac{1}{2}\})} \left[ \int_{\Theta_1} f(\theta|\mathbf{x}) d\theta \right] \mathbb{P}(\mathbf{X} = \mathbf{x}).
\end{aligned} \tag{4.28}$$

Seguindo a mesma lógica do caso anterior, denotamos os termos da equação acima por  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , respectivamente. Assim,

$$\begin{aligned}
I_1 &= a_0 \sum_{s \in R} \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s) \mathbb{P}(S(\mathbf{X}) = s); \\
I_2 &= a_1 \sum_{s \in A} [1 - \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s)] \mathbb{P}(S(\mathbf{X}) = s); \\
I_3 &= b_0 \sum_{s \in G} \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s) \mathbb{P}(S(\mathbf{X}) = s); \\
I_4 &= b_1 \sum_{s \in G} [1 - \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s)] \mathbb{P}(S(\mathbf{X}) = s);
\end{aligned}$$

em que

$$\begin{aligned}
R &= \left\{ s \in \mathbb{N} : \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s) \leq \frac{b_1}{a_0 + b_1 - b_0} \right\}; \\
A &= \left\{ s \in \mathbb{N} : \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s) > \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right\}; \\
G &= \left\{ s \in \mathbb{N} : \frac{b_1}{a_0 + b_1 - b_0} < \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s) \leq \frac{a_1 - b_1}{a_1 + b_0 - b_1} \right\}.
\end{aligned}$$

Simulando  $M$  observações  $s_1, \dots, s_M$  da distribuição marginal de  $S(\mathbf{X}) = \sum_{i=1}^n X_i$ , aproximamos o risco de Bayes por

$$\begin{aligned}
\frac{1}{M} \sum_{j=1}^M \left\{ \left[ a_0 \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s_j) \mathbb{1}_R(s_j) \right] + \left[ a_1 (1 - \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s_j)) \mathbb{1}_A(s_j) \right] \right. \\
\left. + \left[ b_0 \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s_j) \mathbb{1}_G(s_j) \right] + \left[ b_1 (1 - \mathbb{P}(\theta_1 \leq \theta \leq \theta_2 | s_j)) \mathbb{1}_G(s_j) \right] \right\}.
\end{aligned}$$

**Exemplo 4.5.** Seja a amostra  $\mathbf{X} = X_1, \dots, X_n$  em que  $X_i | \theta \stackrel{iid}{\sim} \text{Poisson}(\theta)$ , para  $i = 1, \dots, n$  e  $\theta > 0$ . Considere as hipóteses  $H_0 : 3 \leq \theta \leq 5$  e  $H_1 : \theta < 3$  ou  $\theta > 5$  e os valores  $a_0 = 1$ ,  $a_1 = 1$ ,  $b_0 = \frac{1}{4}$  e  $b_1 = \frac{1}{4}$  para as funções de perda. Deseja-se calcular o risco de Bayes em função de  $n$  para os testes tradicional e agnóstico, baseados em probabilidade a posteriori.

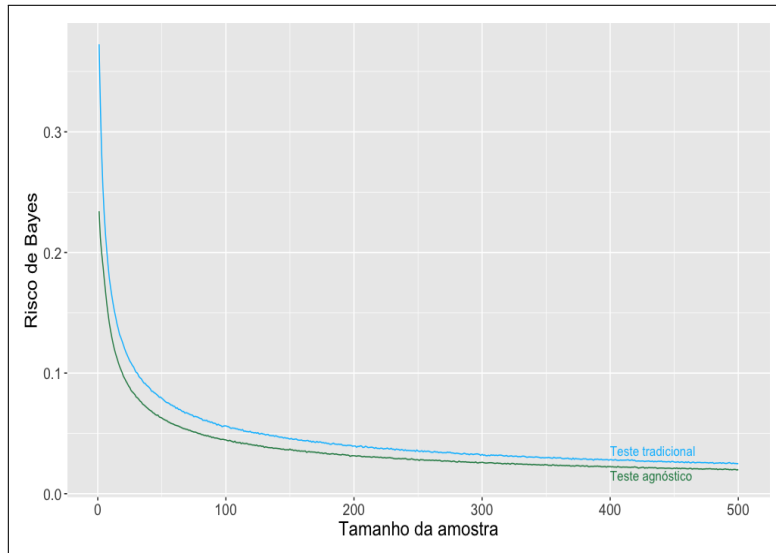
Substituindo os valores em (4.25) e (4.27), os testes de Bayes são dados por

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & F_{\theta|\mathbf{x}}(5) - F_{\theta|\mathbf{x}}(3) > 0,5 \\ 1, & F_{\theta|\mathbf{x}}(5) - F_{\theta|\mathbf{x}}(3) \leq 0,5 \end{cases}$$

e

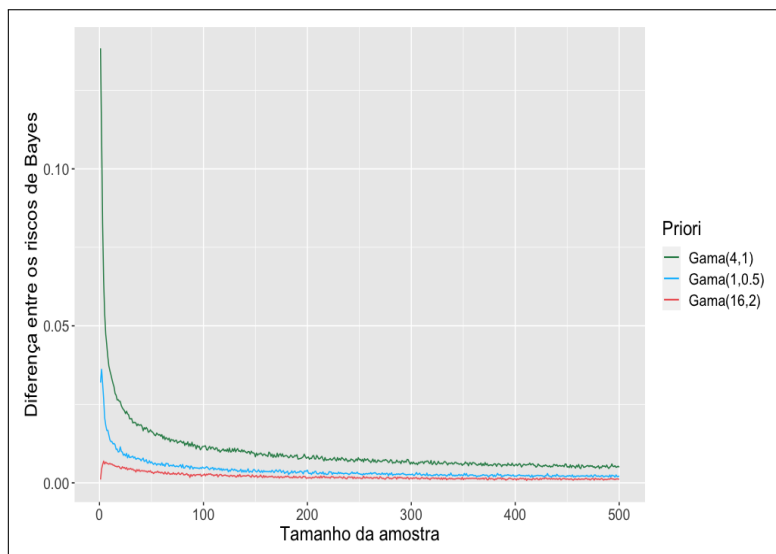
$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & F_{\theta|\mathbf{x}}(5) - F_{\theta|\mathbf{x}}(3) > 0,75 \\ \frac{1}{2}, & 0,25 < F_{\theta|\mathbf{x}}(5) - F_{\theta|\mathbf{x}}(3) \leq 0,75 \\ 1, & F_{\theta|\mathbf{x}}(5) - F_{\theta|\mathbf{x}}(3) \leq 0,25 \end{cases} .$$

Supondo como priori para  $\theta$  uma distribuição Gama com parâmetros  $\alpha = 4$  e  $\beta = 1$  e simulando com  $M = 10^5$  iterações pelo método de Monte Carlo descrito acima, obtemos os resultados representados na figura abaixo, para  $n$  variando de 1 a 500.

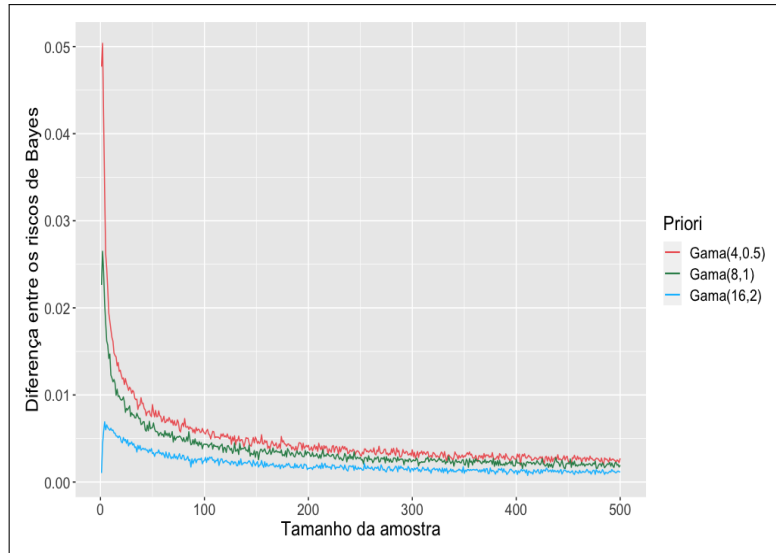


**Figura 4.14:** *Riscos de Bayes dos testes intervalares para modelo Poisson - Priori  $\mathcal{G}(4,1)$ .*

Considerando prioris distintas e um custo por unidade amostral  $c = 0,001$ , as figuras e a tabela a seguir apresentam os resultados das análises. A esperança e variância de  $\theta$  para cada priori são apresentadas na Tabela 4.8 a fim de facilitar a comparação.



**Figura 4.15:** *Diferença absoluta entre os riscos de Bayes para prioris com médias diferentes e variâncias iguais - Testes intervalares para modelo Poisson.*



**Figura 4.16:** Diferença absoluta entre os riscos de Bayes para prioris com mesma média e variâncias diferentes - Testes intervalares para modelo Poisson.

Priori $\theta$	$E(\theta)$	$Var(\theta)$	$n_T^*$	$n_A^*$	$\delta(\%)$	$\bar{\rho}_T^*$	$\bar{\rho}_A^*$
$\mathcal{G}(4, 2)$	2	1	26	19	26,9	0,0510	0,0472
$\mathcal{G}(2, 1)$	2	2	26	23	11,5	0,0507	0,0432
$\mathcal{G}(8, 2)$	4	2	48	40	16,7	0,0991	0,0862
$\mathcal{G}(2, 0,5)$	4	8	35	30	14,3	0,0689	0,0595
$\mathcal{G}(1, 0,5)$	2	4	25	20	20,0	0,0455	0,0394
$\mathcal{G}(4, 1)$	4	4	42	35	16,7	0,0850	0,0743
$\mathcal{G}(16, 2)$	8	4	13	10	23,1	0,0306	0,0278
$\mathcal{G}(8, 1)$	8	8	24	19	20,8	0,0426	0,0382
$\mathcal{G}(4, 0,5)$	8	16	27	22	18,5	0,0530	0,0465

**Tabela 4.8:** Tamanhos amostrais e riscos para os testes de Bayes intervalares - Modelo Poisson.

Ao analisar a Figura 4.15, observa-se que a menor diferença absoluta entre os riscos ocorre na priori de maior média ( $\mathcal{G}(16, 2)$ ), assim como nos exemplos anteriores. Já na Figura 4.16, nota-se um comportamento distinto aos exemplos anteriores em que, para prioris de mesma média, quanto menor a variância da priori, menor a diferença entre os riscos dos testes. Os maiores tamanhos amostrais ótimos são obtidos quando a esperança de  $\theta$  da priori é igual a 4. O menor tamanho amostral é observado na priori cuja esperança é igual a 8 (valor mais distante dos pontos que delimitam  $H_0$  e  $H_1$ ) e que possui a menor variância dentre as prioris de mesma média (ou seja, dentre as prioris analisadas, é a priori em que temos maior crença em  $H_1$ , a priori). Quanto à comparação entre  $n_T^*$  e  $n_A^*$ , os tamanhos amostrais no cenário tradicional são na faixa de 10 a 20% maiores que no cenário agnóstico, valores similares aos observados nos outros modelos com perdas  $0 - a_0 - a_1$  e  $0 - a_0 - a_1 - b_0 - b_1$ .

## 4.5 Estudo comparativo: Testes de Hipóteses Precisas

Nessa seção, estamos interessados em testar a hipótese precisa  $H_0 : \theta = \theta_0$ ,  $\theta_0 \in \Theta \subseteq \mathbb{R}$ . Diferentemente dos casos anteriores, estudaremos testes baseados em estimador por região utilizando novamente as funções de perda  $0 - a_0 - a_1$  e  $0 - a_0 - a_1 - b_0 - b_1$ . Vale destacar que sob essas funções de perda, os testes baseados em estimador por região não são necessariamente testes de Bayes, ou seja, não são derivados pela minimização do risco a posteriori. Conseqüentemente, o



Teorema 4.1 não é aplicável em tal contexto e não há a garantia de obtenção de um risco de Bayes menor no cenário agnóstico. Tais testes serão estudados aqui em um caráter mais exploratório para testar hipóteses precisas e avaliar se obtemos resultados similares aos estudos anteriores, mesmo não havendo a garantia formal de redução de risco. Uma alternativa seria encontrar uma função de perda que faça com que o GFBST, por exemplo, seja um teste de Bayes, de forma similar ao que foi feito em Madruga, Pereira e Stern [30] para o FBST. Entretanto, tais formulações não serão contempladas aqui.

Como foi visto na seção 3.2.3, no caso de hipóteses precisas não há possibilidade de aceitar uma hipótese precisa  $H_0$  em testes baseados em estimador por região. Tendo em vista tal fato, aqui consideramos a abordagem de região de equivalência prática, adaptando as hipóteses originais de interesse para  $H_0 : \theta \in \Theta_0$  e  $H_1 : \theta \in \Theta_1$ , em que  $\Theta_0 = \{\theta \in \Theta : \theta \in [\theta_0 - \epsilon, \theta_0 + \epsilon]\}$  e  $\Theta_1 = \{\theta \in \Theta : \theta \notin [\theta_0 - \epsilon, \theta_0 + \epsilon]\}$ , com  $\epsilon > 0$ . A seguir, estudaremos os testes sob essa abordagem para o modelo de distribuição Normal univariado e, na seção 4.5.2, para o caso bivariado.

#### 4.5.1 Teste baseado em estimador por região: Modelo Normal Univariado com Variância conhecida

Considere uma amostra  $\mathbf{X} = (X_1, \dots, X_n)$  em que  $X_i | \theta \stackrel{iid}{\sim} \mathcal{N}(\theta, v^2)$ ,  $v^2 > 0$  conhecida, para  $i = 1, \dots, n$ . Vamos testar a hipótese  $H_0 : \theta \in [\theta_0 - \epsilon, \theta_0 + \epsilon]$  contra  $H_1 : \theta \notin [\theta_0 - \epsilon, \theta_0 + \epsilon]$  através de testes baseados em estimador por região, para os casos tradicional e agnóstico.

Assumindo que a priori de  $\theta$  segue distribuição Normal, digamos,  $\theta \sim \mathcal{N}(\mu_0, \sigma_0^2)$ , temos que a região de HPD de credibilidade  $1 - \alpha$  é dada pelo intervalo

$$IC_{HPD} = \left[ \mu_x - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x, \mu_x + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \right],$$

em que  $\mu_x = \frac{\mu_0 v^2 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2}$  e  $\sigma_x^2 = \frac{v^2 \sigma_0^2}{v^2 + n \sigma_0^2}$ .

Seja o teste tradicional  $\varphi_T$  baseado em estimador por região, dado pela equação (3.11). Considerando a região HPD acima, temos

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & IC_{HPD} \cap H_0 \neq \emptyset \\ 1, & IC_{HPD} \subseteq H_1 \end{cases},$$

que podemos reescrever como

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & \theta_0 - \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \leq \mu_x \leq \theta_0 + \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \\ 1, & \mu_x < \theta_0 - \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \quad \text{ou} \quad \mu_x > \theta_0 + \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \end{cases}. \quad (4.29)$$

Deixando o teste em função da estatística suficiente  $\bar{x}$ , obtemos

$$\begin{aligned} \varphi_T(\mathbf{x}) = 1 &\iff \mu_x < \theta_0 - \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \quad \text{ou} \quad \mu_x > \theta_0 + \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \\ &\iff \frac{\sigma_0^2 \mu_0 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2} < \theta_0 - \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \quad \text{ou} \\ &\quad \frac{\sigma_0^2 \mu_0 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2} > \theta_0 + \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \\ &\iff \bar{x} < \frac{(\theta_0 - \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2} \quad \text{ou} \\ &\quad \bar{x} > \frac{(\theta_0 + \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2}. \end{aligned}$$

Sejam  $k_n^{(1)} = \frac{(\theta_0 - \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2}$  e  $k_n^{(2)} = \frac{(\theta_0 + \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2}$ .

Logo,

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & k_n^{(1)} < \bar{x} < k_n^{(2)} \\ 1, & \bar{x} \leq k_n^{(1)} \quad \text{ou} \quad \bar{x} \geq k_n^{(2)} \end{cases} . \quad (4.30)$$

Similarmente aos casos anteriores do modelo Normal, o risco do teste tradicional considerando a perda  $0 - a_0 - a_1$  é dado por

$$\begin{aligned} \bar{\rho}(\varphi_T, \bar{P}) &= a_0 \int_{\Theta_0} \left[ \Phi \left( \frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}} \right) + 1 - \Phi \left( \frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}} \right) \right] f(\theta) d\theta \\ &+ a_1 \int_{\Theta_1} \left[ \Phi \left( \frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}} \right) - \Phi \left( \frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}} \right) \right] f(\theta) d\theta, \end{aligned} \quad (4.31)$$

onde as integrais podem ser aproximadas pelo método de Monte Carlo.

Agora, considere o teste agnóstico baseado em estimador por região, dado pela equação (3.17). Temos que

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \text{se } IC_{HPD} \subseteq H_0 \\ 1, & \text{se } IC_{HPD} \subseteq H_1 \\ \frac{1}{2}, & \text{caso contrário} \end{cases} ,$$

que equivale a

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & \text{se } \theta_0 - \epsilon + \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \leq \mu_x \leq \theta_0 + \epsilon - \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \\ 1, & \text{se } \mu_x < \theta_0 - \epsilon - \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \quad \text{ou} \quad \mu_x > \theta_0 + \epsilon + \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \\ \frac{1}{2}, & \text{caso contrário} \end{cases} . \quad (4.32)$$

Note que, para que a região HPD possa estar contida no intervalo especificado por  $H_0$ , temos uma restrição quanto ao tamanho amostral. Assim, para que seja possível aceitar a hipótese  $H_0$ , devemos ter que

$$\Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \leq \epsilon \iff \Phi \left( \frac{1-\alpha}{2} \right) \sqrt{\frac{v^2 \sigma_0^2}{v^2 + n \sigma_0^2}} \leq \epsilon \iff n \geq \frac{v^2 \left( \Phi \left( \frac{1-\alpha}{2} \right)^2 \sigma_0^2 - \epsilon^2 \right)}{\sigma_0^2 \epsilon^2} .$$

Escrevendo o teste em função da estatística suficiente  $\bar{x}$ , rejeitamos  $H_0$  se

$$\begin{aligned} \varphi_A(\mathbf{x}) = 1 &\iff \mu_x < \theta_0 - \epsilon - \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \quad \text{ou} \quad \mu_x > \theta_0 + \epsilon + \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \\ &\iff \frac{\sigma_0^2 \mu_0 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2} < \theta_0 - \epsilon - \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \quad \text{ou} \\ &\quad \frac{v^2 \mu_0 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2} > \theta_0 + \epsilon + \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x \\ &\iff \bar{x} < \frac{(\theta_0 - \epsilon - \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2} \quad \text{ou} \\ &\quad \bar{x} > \frac{(\theta_0 + \epsilon + \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2} . \end{aligned}$$

Defina  $k_n^{(1)} = \frac{(\theta_0 - \epsilon - \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2}$  e  $k_n^{(4)} = \frac{(\theta_0 + \epsilon + \Phi \left( \frac{1-\alpha}{2} \right) \sigma_x)(v^2 + n \sigma_0^2) - \sigma_0^2 \mu_0}{n \sigma_0^2}$ .

Analogamente, aceitamos  $H_0$  se

$$\begin{aligned}\varphi_A(\mathbf{x}) = 0 &\iff \theta_0 - \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \leq \mu_x \leq \theta_0 + \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \\ &\iff \theta_0 - \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \leq \frac{\sigma_0^2 \mu_0 + n\sigma_0^2 \bar{x}}{v^2 + n\sigma_0^2} \leq \theta_0 + \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x \\ &\iff \bar{x} \leq \frac{(\theta_0 + \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n\sigma_0^2) - \sigma_0^2 \mu_0}{n\sigma_0^2} \quad \text{e} \\ &\quad \bar{x} \geq \frac{(\theta_0 - \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n\sigma_0^2) - \sigma_0^2 \mu_0}{n\sigma_0^2}.\end{aligned}$$

Defina  $k_n^{(2)} = \frac{(\theta_0 - \epsilon + \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n\sigma_0^2) - \sigma_0^2 \mu_0}{n\sigma_0^2}$  e  $k_n^{(3)} = \frac{(\theta_0 + \epsilon - \Phi\left(\frac{1-\alpha}{2}\right) \sigma_x)(v^2 + n\sigma_0^2) - \sigma_0^2 \mu_0}{n\sigma_0^2}$ .

Assim, o teste agnóstico baseado em estimador por região fica dado por

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & k_n^{(2)} \leq \bar{x} \leq k_n^{(3)} \\ \frac{1}{2}, & k_n^{(1)} \leq \bar{x} < k_n^{(2)} \quad \text{ou} \quad k_n^{(3)} \leq \bar{x} < k_n^{(4)} \\ 1, & \bar{x} < k_n^{(1)} \quad \text{ou} \quad \bar{x} > k_n^{(4)} \end{cases}. \quad (4.33)$$

Denotando as parcelas em (3.13) de  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , respectivamente, o risco do teste agnóstico é calculado por

$$\bar{\rho}(\varphi_A, \bar{\mathcal{P}}) = I_1 + I_2 + I_3 + I_4, \quad (4.34)$$

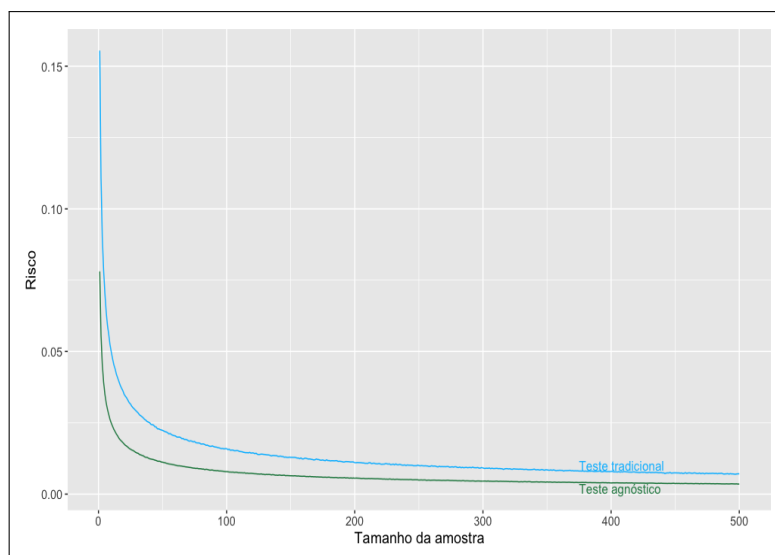
onde

$$\begin{aligned}I_1 &= a_0 \int_{\Theta_0} \left[ \Phi\left(\frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}}\right) + 1 - \Phi\left(\frac{k_n^{(4)} - \theta}{\sqrt{(v^2/n)}}\right) \right] f(\theta) d\theta; \\ I_2 &= a_1 \int_{\Theta_1} \left[ \Phi\left(\frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}}\right) - \Phi\left(\frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}}\right) \right] f(\theta) d\theta; \\ I_3 &= b_0 \int_{\Theta_0} \left[ \Phi\left(\frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}}\right) - \Phi\left(\frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}}\right) + \Phi\left(\frac{k_n^{(4)} - \theta}{\sqrt{(v^2/n)}}\right) - \Phi\left(\frac{k_n^{(3)} - \theta}{\sqrt{(v^2/n)}}\right) \right] f(\theta) d\theta; \\ I_4 &= b_1 \int_{\Theta_1} \left[ \Phi\left(\frac{k_n^{(2)} - \theta}{\sqrt{(v^2/n)}}\right) - \Phi\left(\frac{k_n^{(1)} - \theta}{\sqrt{(v^2/n)}}\right) + \Phi\left(\frac{k_n^{(4)} - \theta}{\sqrt{(v^2/n)}}\right) - \Phi\left(\frac{k_n^{(3)} - \theta}{\sqrt{(v^2/n)}}\right) \right] f(\theta) d\theta.\end{aligned}$$

Novamente, as integrais podem ser aproximadas pelo método de Monte Carlo simulando da distribuição de  $\theta$ .

**Exemplo 4.6.** Seja uma amostra  $\mathbf{X} = (X_1, \dots, X_n)$  em que  $X_i | \theta \stackrel{iid}{\sim} \mathcal{N}(\theta, 1)$ , para  $i = 1, \dots, n$ . Desejamos testar a hipótese  $H_0 : \theta \in [0 - \epsilon, 0 + \epsilon]$  contra  $H_1 : \theta \notin [0 - \epsilon, 0 + \epsilon]$ , com  $\epsilon = 0,01$ . Vamos comparar os riscos dos testes tradicional e agnóstico baseados em estimador por região em função do tamanho amostral  $n$ , considerando as funções de perda  $0 - a_0 - a_1$  e  $0 - a_0 - a_1 - b_0 - b_1$ , com  $a_0 = 1$ ,  $a_1 = 1$ ,  $b_0 = \frac{1}{4}$ ,  $b_1 = \frac{1}{4}$ . Seja  $\alpha = 0,05$  de modo que a região HPD é de 0,95 de credibilidade.

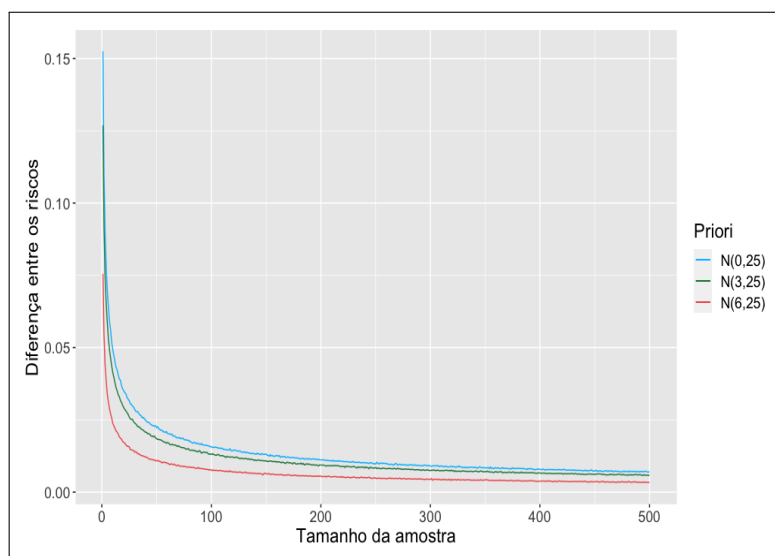
Assumindo como priori para  $\theta$  uma distribuição Normal com parâmetros  $\mu_0 = 0$  e  $\sigma_0^2 = 100$ , calculamos numericamente o risco dos testes para tamanhos amostrais variando de 1 a 500. As integrais para o cálculo do risco foram aproximadas pelo método de Monte Carlo com  $M = 10^6$  iterações. A figura abaixo ilustra os resultados obtidos.



**Figura 4.17:** Riscos dos testes baseados em estimador por região para modelo Normal univariado - Priori  $\mathcal{N}(0, 100)$ .

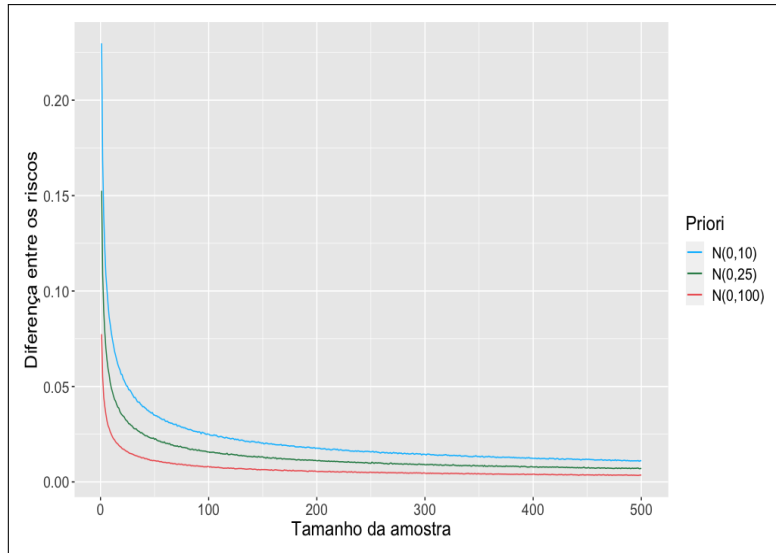
Note que o risco do teste agnóstico é menor que o risco do teste tradicional para todo  $n$  considerado, apesar de não termos formalmente provado tal redução. Vale destacar que, para essa priori, esse  $\alpha$  e esse  $\epsilon$ , temos a restrição de que  $n \geq 385$  para que seja possível aceitar  $H_0$  no teste agnóstico.

Em seguida, os gráficos abaixo comparam a diferença absoluta entre os dois riscos para diferentes escolhas de prioris para  $\theta$ , variando a média e a variância.



**Figura 4.18:** Diferença absoluta entre os riscos para prioris com médias diferentes e variâncias iguais - Testes baseados em estimador por região para modelo Normal univariado.

Nas Figuras 4.18 e 4.19 observamos resultados similares aos obtidos nas seções 4.3.1 e 4.3.2, em que as menores diferenças entre os riscos ocorrem nas prioris de maior variância (quando fixamos a média) e nas prioris de maior média (quando fixamos a variância).



**Figura 4.19:** Diferença absoluta entre os riscos para prioris com mesma média e variâncias diferentes - Testes baseados em estimador por região para modelo Normal univariado.

Finalmente, a tabela abaixo exhibe os resultados do método de cálculo amostral utilizado nos exemplos anteriores, com custo  $c = 0,0001$ .

Priori $\theta$	$n_T^*$	$n_A^*$	$\delta(\%)$	$\bar{\rho}_T^*$	$\bar{\rho}_A^*$
$\mathcal{N}(-3, 10)$	138	81	41,3	0,0268	0,0175
$\mathcal{N}(-3, 25)$	119	67	43,7	0,0239	0,0159
$\mathcal{N}(-3, 100)$	72	56	22,2	0,0175	0,0100
$\mathcal{N}(0, 10)$	189	105	44,4	0,0361	0,0240
$\mathcal{N}(0, 25)$	153	85	44,4	0,0252	0,0171
$\mathcal{N}(0, 100)$	87	55	36,8	0,0168	0,0106
$\mathcal{N}(3, 10)$	140	86	38,6	0,0265	0,0170
$\mathcal{N}(3, 25)$	126	70	44,4	0,0232	0,0156
$\mathcal{N}(3, 100)$	90	48	46,7	0,0158	0,0108
$\mathcal{N}(6, 10)$	61	38	37,7	0,0104	0,0067
$\mathcal{N}(6, 25)$	87	51	41,4	0,0164	0,0107
$\mathcal{N}(6, 100)$	76	47	38,2	0,0149	0,0095
$\mathcal{N}(9, 10)$	13	8	38,5	0,0024	0,0015
$\mathcal{N}(9, 25)$	43	33	23,3	0,0094	0,0054
$\mathcal{N}(9, 100)$	66	43	34,8	0,0128	0,0080

**Tabela 4.9:** Tamanhos amostrais e riscos para os testes baseados em estimador por região - Modelo Normal univariado.

Nesse caso, observamos reduções maiores (em torno de 30 a 40%) dos tamanhos amostrais ótimos do teste agnóstico em relação ao teste tradicional, se compararmos com os outros testes já estudados. Em relação ao impacto da priori, os resultados são bastante similares aos obtidos nos exemplos dos testes unilaterais para o modelo Normal.

### 4.5.2 Teste baseado em estimador por região: Modelo Normal Bivariado com Variância conhecida

Sejam  $\mathbf{X} = (X_1, \dots, X_n)$  e  $\mathbf{Y} = (Y_1, \dots, Y_n)$  vetores aleatórios que, dado  $\boldsymbol{\theta} = (\theta_1, \theta_2)$ , são independentes tais que  $X_i | \boldsymbol{\theta} \stackrel{iid}{\sim} N(\theta_1, v^2)$  e  $Y_i | \boldsymbol{\theta} \stackrel{iid}{\sim} N(\theta_2, v^2)$ , com variância  $v^2 > 0$  conhecida. Desejamos testar se as médias das duas amostras são iguais ( $\theta_1 = \theta_2$ ). Porém, para que haja possibilidade de aceitar  $H_0$ , adaptamos as hipóteses de interesse para  $H_0 : (\theta_1, \theta_2) \in \Theta_0$  e  $H_1 : (\theta_1, \theta_2) \in \Theta_1$ ,  $\Theta_0, \Theta_1 \in \Theta \subseteq \mathbb{R}^2$ , em que  $\Theta_0 = \{(\theta_1, \theta_2) \in \mathbb{R}^2 : \theta_1 - \epsilon < \theta_2 < \theta_1 + \epsilon\}$  e  $\Theta_1 = \{(\theta_1, \theta_2) \in \mathbb{R}^2 : \theta_2 \leq \theta_1 - \epsilon \text{ ou } \theta_2 \geq \theta_1 + \epsilon\}$ . A figura a seguir ilustra a região especificada por  $H_0$ , destacada em azul.

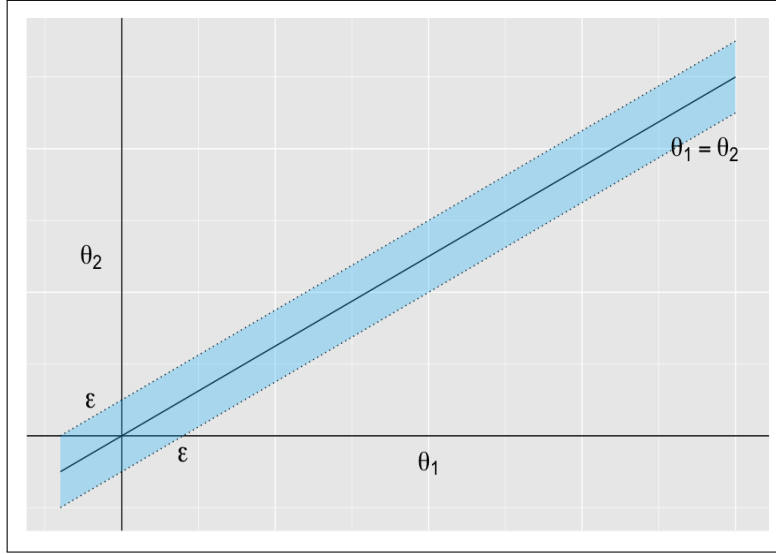


Figura 4.20: Figura ilustrativa da hipótese  $H_0$  no caso bivariado.

Assumindo a priori que  $\theta_1$  e  $\theta_2$  são independentes, com  $\theta_1 \sim \theta_2 \sim \mathcal{N}(\mu_0, \sigma_0^2)$ , temos, por conjugação, que a distribuição posteriori conjunta (dado  $\mathbf{X} = \mathbf{x} = (x_1, \dots, x_n)$ ,  $\mathbf{Y} = \mathbf{y} = (y_1, \dots, y_n)$ ) dos parâmetros segue distribuição Normal bivariada dada por

$$\theta_1, \theta_2 | \mathbf{X} = \mathbf{x}, \mathbf{Y} = \mathbf{y} \sim \mathcal{N}_2 \left( \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \right),$$

em que

$$\mu_x = \frac{v^2 \mu_0 + n \sigma_0^2 \bar{x}}{v^2 + n \sigma_0^2}, \quad \mu_y = \frac{v^2 \mu_0 + n \sigma_0^2 \bar{y}}{v^2 + n \sigma_0^2}, \quad \sigma_x^2 = \sigma_y^2 = \frac{v^2 \sigma_0^2}{v^2 + n \sigma_0^2}.$$

Note que a posteriori é proporcional à

$$f(\theta_1, \theta_2 | \mathbf{x}, \mathbf{y}) \propto e^{-\left\{ \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right)^2 + \left( \frac{\theta_2 - \mu_y}{\sigma_y} \right)^2 \right\}}.$$

Como a região HPD para  $\boldsymbol{\theta} = (\theta_1, \theta_2)$  é a região de pontos com densidade maior que um certo valor  $d$ , temos que

$$R_{HPD} = \left\{ (\theta_1, \theta_2) \in \mathbb{R}^2 : \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right)^2 + \left( \frac{\theta_2 - \mu_y}{\sigma_y} \right)^2 \leq d \right\},$$

que corresponde a uma circunferência centrada em  $(\mu_x, \mu_y)$  e raio  $r = \sigma_x \sqrt{d}$ , uma vez que  $\sigma_x = \sigma_y$ .

Conseqüentemente, para avaliar

$$\mathbb{P} \left\{ (\theta_1, \theta_2) \in \mathbb{R}^2 : \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right)^2 + \left( \frac{\theta_2 - \mu_y}{\sigma_y} \right)^2 \leq d \mid \mathbf{x}, \mathbf{y} \right\}$$

note que as variáveis aleatórias  $Z_1 = \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right)$  e  $Z_2 = \left( \frac{\theta_2 - \mu_y}{\sigma_y} \right)$ , dado  $\mathbf{X} = \mathbf{x}$ ,  $\mathbf{Y} = \mathbf{y}$ , seguem ambas a distribuição normal padrão e, portanto,  $Z_1^2$  e  $Z_2^2$  possuem distribuição qui-quadrado com um grau de liberdade,  $Z_1^2 \sim Z_2^2 \sim \chi_1^2$ . Além disso, pela independência de  $\theta_1$  e  $\theta_2$  a posteriori, temos que  $Z_1^2 + Z_2^2 \sim \chi_2^2$ . Logo,

$$\mathbb{P} \left\{ (\theta_1, \theta_2) \in \mathbb{R}^2 : \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right)^2 + \left( \frac{\theta_2 - \mu_y}{\sigma_y} \right)^2 \leq d \mid \mathbf{x}, \mathbf{y} \right\} = \mathbb{P} \{ \chi_2^2 \leq d \}.$$

Assim, para obter a região HPD de  $\boldsymbol{\theta}$  de credibilidade  $1 - \alpha$ , basta tomar  $d = q_{1-\alpha}$ , em que  $q_{1-\alpha}$  é o percentil  $1 - \alpha$  da distribuição qui-quadrado com dois graus de liberdade. Nesse caso, a região HPD é dada por

$$\begin{aligned} R_{HPD} &= \left\{ (\theta_1, \theta_2) \in \mathbb{R}^2 : \left( \frac{\theta_1 - \mu_x}{\sigma_x} \right)^2 + \left( \frac{\theta_2 - \mu_y}{\sigma_x} \right)^2 \leq q_{1-\alpha} \right\} \\ &= \{ (\theta_1, \theta_2) \in \mathbb{R}^2 : (\theta_1 - \mu_x)^2 + (\theta_2 - \mu_y)^2 \leq q_{1-\alpha} \sigma_x^2 \}. \end{aligned}$$

Para o teste baseado em estimador por região, rejeitamos a hipótese  $H_0$  quando a região de HPD está contida na região complementar à  $H_0$ . Dessa forma, ao supor inicialmente que  $\mu_y > \mu_x$ , a rejeição ocorre quando a distância euclidiana entre o centro da circunferência da região de HPD e a reta  $\theta_2 = \theta_1 + \epsilon$  é maior que o raio  $r$ . Assim,

$$\begin{aligned} \varphi_T(\mathbf{x}, \mathbf{y}) = 1 &\iff \sqrt{\left( \mu_x - \frac{\mu_x + \mu_y - \epsilon}{2} \right)^2 + \left( \mu_y - \frac{\mu_x + \mu_y + \epsilon}{2} \right)^2} > r \\ &\iff 2(\mu_y - \mu_x - \epsilon)^2 > 4r^2 \\ &\iff \mu_y - \mu_x - \epsilon > r\sqrt{2} \\ &\iff \mu_y - \mu_x > \epsilon + r\sqrt{2}. \end{aligned}$$

Analogamente, se  $\mu_x > \mu_y$ , temos que

$$\varphi_T(\mathbf{x}, \mathbf{y}) = 1 \iff \mu_x - \mu_y > \epsilon + r\sqrt{2}.$$

Logo, podemos escrever o teste tradicional  $\varphi_T$  como

$$\varphi_T(\mathbf{x}, \mathbf{y}) = \begin{cases} 0, & |\mu_x - \mu_y| < \epsilon + r\sqrt{2} \\ 1, & |\mu_x - \mu_y| > \epsilon + r\sqrt{2} \end{cases}. \quad (4.35)$$

Reescrevendo o teste em função das estatísticas suficientes  $\bar{x}$  e  $\bar{y}$ , temos que

$$\begin{aligned} \varphi_T(\mathbf{x}, \mathbf{y}) = 0 &\iff |\mu_x - \mu_y| < \epsilon + r\sqrt{2} \\ &\iff \left| \frac{v^2 \mu_0 + n\sigma_0^2 \bar{x}}{v^2 + n\sigma_0^2} - \frac{v^2 \mu_0 + n\sigma_0^2 \bar{y}}{v^2 + n\sigma_0^2} \right| < \epsilon + \sigma_x \sqrt{2q_{1-\alpha}} \\ &\iff |\bar{x} - \bar{y}| < \left( \frac{v^2 + n\sigma_0^2}{n\sigma_0^2} \right) (\epsilon + \sigma_x \sqrt{2q_{1-\alpha}}) \end{aligned}$$

e

$$\begin{aligned}
\varphi_T(\mathbf{x}, \mathbf{y}) = 1 &\iff |\mu_x - \mu_y| > \epsilon + r\sqrt{2} \\
&\iff \left| \frac{v^2\mu_0 + n\sigma_0^2\bar{x}}{v^2 + n\sigma_0^2} - \frac{v^2\mu_0 + n\sigma_0^2\bar{y}}{v^2 + n\sigma_0^2} \right| > \epsilon + \sigma_x\sqrt{2q_{1-\alpha}} \\
&\iff |\bar{x} - \bar{y}| > \left( \frac{v^2 + n\sigma_0^2}{n\sigma_0^2} \right) (\epsilon + \sigma_x\sqrt{2q_{1-\alpha}}).
\end{aligned}$$

Seja  $k_n = \left( \frac{v^2 + n\sigma_0^2}{n\sigma_0^2} \right) (\epsilon + \sigma_x\sqrt{2q_{1-\alpha}})$ . Escrevemos o teste como

$$\varphi_T(\mathbf{x}) = \begin{cases} 0, & |\bar{x} - \bar{y}| < k_n \\ 1, & |\bar{x} - \bar{y}| > k_n \end{cases}. \quad (4.36)$$

Para o cálculo do risco do teste tradicional, define-se a variável aleatória  $W = \bar{X} - \bar{Y}$ , que possui distribuição  $W|\boldsymbol{\theta} \sim \mathcal{N}\left(\theta_1 - \theta_2, \sqrt{\frac{2v^2}{n}}\right)$ . Considerando a função de perda  $0 - a_0 - a_1$ , segue que

$$\begin{aligned}
\bar{\rho}(\varphi_T, \bar{\mathcal{P}}) &= \int_{\mathbb{R}} \int_{\Theta} L(\varphi_T(w, \boldsymbol{\theta})) f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = \\
&= a_0 \int_R \int_{\Theta_0} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw + a_1 \int_A \int_{\Theta_1} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw,
\end{aligned} \quad (4.37)$$

onde  $R = \{w \in \mathbb{R} : |w| > k_n\}$  e  $A = \{w \in \mathbb{R} : |w| < k_n\}$ .

Desenvolvendo a primeira integral, que denotamos por  $I_1$ , temos

$$\begin{aligned}
I_1 &= a_0 \int_{\Theta_0} \int_R f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = a_0 \int_{\Theta_0} \left[ \int_R f(w|\boldsymbol{\theta}) dw \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\Theta_0} \mathbb{P}(|W| > k_n | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\Theta_0} \mathbb{P}(W < -k_n \text{ ou } W > k_n | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\Theta_0} \left[ \Phi\left(\frac{-k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) + 1 - \Phi\left(\frac{k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\mathbb{R}^2} \left[ \Phi\left(\frac{-k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) + 1 - \Phi\left(\frac{k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) \right] \mathbb{1}_{\Theta_0}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}.
\end{aligned}$$

Já para a segunda integral, denotada por  $I_2$ , segue que

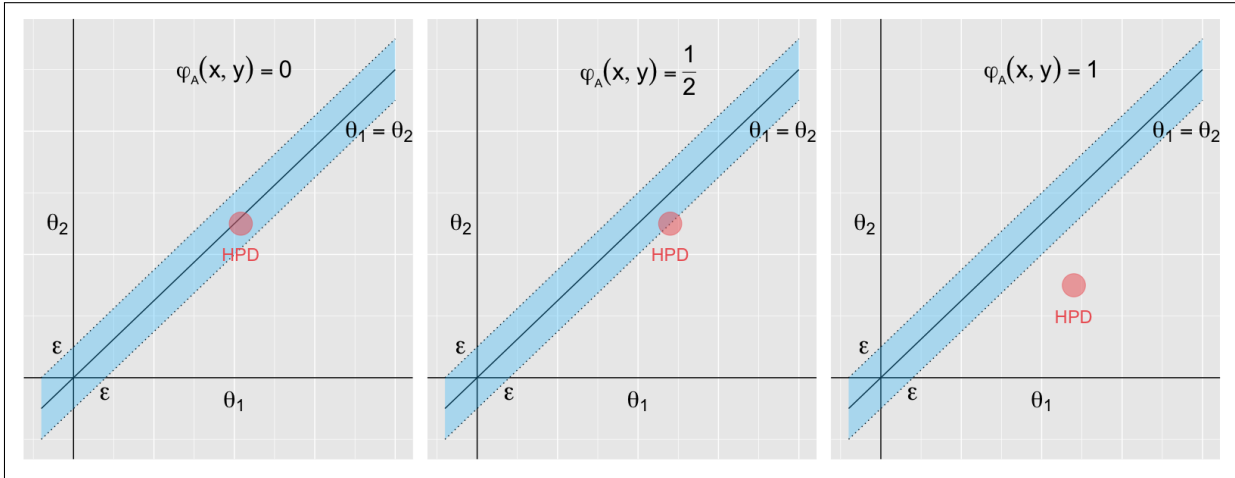
$$\begin{aligned}
I_2 &= a_1 \int_A \int_{\Theta_1} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = a_1 \int_{\Theta_1} \left[ \int_A f(w|\boldsymbol{\theta}) dw \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\Theta_1} \mathbb{P}(|W| < k_n | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\Theta_1} \mathbb{P}(-k_n < W < k_n | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\Theta_1} \left[ \Phi\left(\frac{k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) - \Phi\left(\frac{-k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\mathbb{R}^2} \left[ \Phi\left(\frac{k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) - \Phi\left(\frac{-k_n - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}}\right) \right] \mathbb{1}_{\Theta_1}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}.
\end{aligned}$$

Para aproximar as integrais pelo método de Monte Carlo, podemos simular as prioris de  $\theta_1$  e  $\theta_2$



de forma independente, utilizando a distribuição Normal  $\mathcal{N}(\mu_0, \sigma_0^2)$ .

Para o caso agnóstico, considere a figura abaixo que ilustra geometricamente as situações em que aceitamos  $H_0$ , rejeitamos  $H_0$  ou ficamos agnósticos.



**Figura 4.21:** Figura ilustrativa do teste agnóstico baseado em estimador por região no caso bivariado.

Supondo que  $\mu_y > \mu_x$ , a partir da equação (3.17) e utilizando raciocínio similar ao teste tradicional, aceitamos  $H_0$  se

$$\begin{aligned} \varphi_A(\mathbf{x}, \mathbf{y}) = 0 &\iff \sqrt{\left(\mu_x - \frac{\mu_x + \mu_y}{2}\right)^2 + \left(\mu_y - \frac{\mu_x + \mu_y}{2}\right)^2} < \frac{\epsilon\sqrt{2}}{2} - r \\ &\iff (\mu_x - \mu_y)^2 + (\mu_y - \mu_x)^2 < 4\left(\frac{\epsilon\sqrt{2}}{2} - r\right)^2 \\ &\iff (\mu_y - \mu_x)^2 < 2\left(\frac{\epsilon\sqrt{2}}{2} - r\right)^2 \\ &\iff \mu_y - \mu_x > \epsilon - r\sqrt{2}. \end{aligned}$$

Já a rejeição de  $H_0$ , temos o mesmo cenário que no teste tradicional. Logo, podemos escrever o teste agnóstico  $\varphi_A$  como

$$\varphi_A(\mathbf{x}, \mathbf{y}) = \begin{cases} 0, & |\mu_x - \mu_y| < \epsilon - r\sqrt{2} \\ \frac{1}{2}, & \epsilon - r\sqrt{2} < |\mu_x - \mu_y| < \epsilon + r\sqrt{2} \\ 1, & |\mu_x - \mu_y| > \epsilon + r\sqrt{2} \end{cases}, \quad (4.38)$$

em que  $r$  é o raio da circunferência dada pela região HPD. Note que, para que seja possível aceitar  $H_0$ , há a restrição de que  $r < \frac{\sqrt{2}\epsilon}{2}$ . Assim, temos uma restrição em relação ao tamanho amostral de modo que  $n > \frac{v^2(2\sigma_0^2q_{1-\alpha} - \epsilon^2)}{\sigma_0^2\epsilon^2}$ .

Reescrevendo o teste em função das estatísticas suficientes  $\bar{x}$  e  $\bar{y}$ , obtemos

$$\begin{aligned} \varphi_A(\mathbf{x}, \mathbf{y}) = 0 &\iff |\mu_x - \mu_y| < \epsilon - r\sqrt{2} \\ &\iff \left| \frac{v^2\mu_0 + n\sigma_0^2\bar{x}}{v^2 + n\sigma_0^2} - \frac{v^2\mu_0 + n\sigma_0^2\bar{y}}{v^2 + n\sigma_0^2} \right| < \epsilon - \sigma_x\sqrt{2q_{1-\alpha}} \\ &\iff |\bar{x} - \bar{y}| < \left(\frac{v^2 + n\sigma_0^2}{n\sigma_0^2}\right) (\epsilon - \sigma_x\sqrt{2q_{1-\alpha}}) \end{aligned}$$

e

$$\begin{aligned}
\varphi_A(\mathbf{x}) = 1 &\iff |\mu_x - \mu_y| > \epsilon + r\sqrt{2} \\
&\iff \left| \frac{v^2\mu_0 + n\sigma_0^2\bar{x}}{v^2 + n\sigma_0^2} - \frac{v^2\mu_0 + n\sigma_0^2\bar{y}}{v^2 + n\sigma_0^2} \right| > \epsilon + \sigma_x\sqrt{2q_{1-\alpha}} \\
&\iff |\bar{x} - \bar{y}| > \left( \frac{v^2 + n\sigma_0^2}{n\sigma_0^2} \right) (\epsilon + \sigma_x\sqrt{2q_{1-\alpha}}).
\end{aligned}$$

Sejam  $k_n^{(1)} = \left( \frac{v^2 + n\sigma_0^2}{n\sigma_0^2} \right) (\epsilon - \sigma_x\sqrt{2q_{1-\alpha}})$  e  $k_n^{(2)} = \left( \frac{v^2 + n\sigma_0^2}{n\sigma_0^2} \right) (\epsilon + \sigma_x\sqrt{2q_{1-\alpha}})$ . Então,

$$\varphi_A(\mathbf{x}) = \begin{cases} 0, & |\bar{x} - \bar{y}| < k_n^{(1)} \\ \frac{1}{2}, & k_n^{(1)} < |\bar{x} - \bar{y}| < k_n^{(2)} \\ 1, & |\bar{x} - \bar{y}| > k_n^{(2)} \end{cases} . \quad (4.39)$$

Analogamente ao teste tradicional, se definirmos  $W = \bar{X} - \bar{Y}$ , o risco para o teste agnóstico fica dado por

$$\begin{aligned}
\bar{\rho}(\varphi_A, \bar{\mathcal{P}}) &= \int_{\mathbb{R}} \int_{\Theta} L(\varphi_A(w, \boldsymbol{\theta})) f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = \\
&= a_0 \int_R \int_{\Theta_0} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw + a_1 \int_A \int_{\Theta_1} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw \\
&\quad + b_0 \int_G \int_{\Theta_0} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw + b_1 \int_G \int_{\Theta_1} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw,
\end{aligned} \quad (4.40)$$

onde  $R = \{w \in \mathbb{R} : |w| > k_n^{(2)}\}$ ,  $A = \{w \in \mathbb{R} : |w| < k_n^{(1)}\}$  e  $G = \{w \in \mathbb{R} : k_n^{(1)} < |w| < k_n^{(2)}\}$ .

Denotando as integrais de  $I_1$ ,  $I_2$ ,  $I_3$  e  $I_4$ , respectivamente, temos

$$\begin{aligned}
I_1 &= a_0 \int_R \int_{\Theta_0} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = a_0 \int_{\Theta_0} \left[ \int_R f(w|\boldsymbol{\theta}) dw \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\Theta_0} \mathbb{P}(|W| > k_n^{(2)} | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\Theta_0} \mathbb{P}(W < -k_n^{(2)} \text{ ou } W > k_n^{(2)} | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\Theta_0} \left[ \Phi \left( \frac{-k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) + 1 - \Phi \left( \frac{k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_0 \int_{\mathbb{R}^2} \left[ \Phi \left( \frac{-k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) + 1 - \Phi \left( \frac{k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right] \mathbb{1}_{\Theta_0}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}.
\end{aligned}$$

$$\begin{aligned}
I_2 &= a_1 \int_A \int_{\Theta_1} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = a_1 \int_{\Theta_1} \left[ \int_A f(w|\boldsymbol{\theta}) dw \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\Theta_1} \mathbb{P}(|W| < k_n^{(1)} | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\Theta_1} \mathbb{P}(-k_n^{(1)} < W < k_n^{(1)} | \boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\Theta_1} \left[ \Phi \left( \frac{k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{-k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= a_1 \int_{\mathbb{R}^2} \left[ \Phi \left( \frac{k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{-k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right] \mathbb{1}_{\Theta_1}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}.
\end{aligned}$$

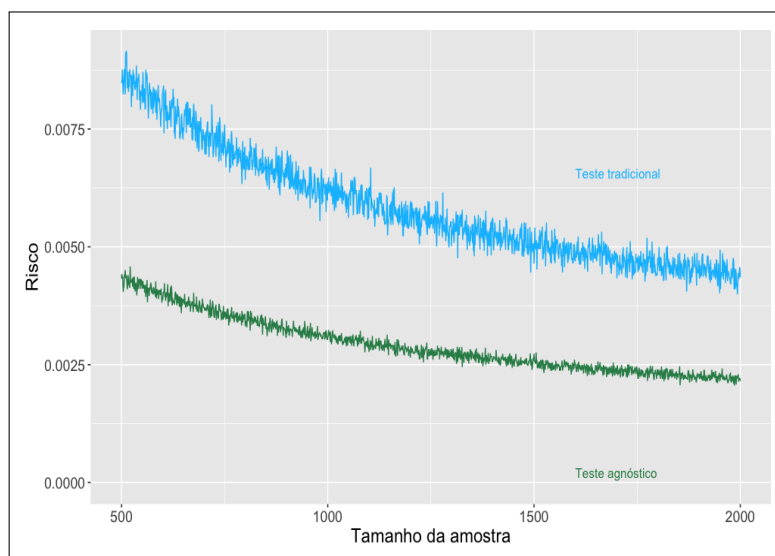
$$\begin{aligned}
I_3 &= b_0 \int_G \int_{\Theta_0} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = b_0 \int_{\Theta_0} \left[ \int_G f(w|\boldsymbol{\theta}) dw \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= b_0 \int_{\Theta_0} \left[ \mathbb{P} \left( k_n^{(1)} < |W| < k_n^{(2)} \mid \boldsymbol{\theta} \right) \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= b_0 \int_{\Theta_0} \left[ \mathbb{P} \left( -k_n^{(2)} < W < -k_n^{(1)} \text{ ou } k_n^{(1)} < W < k_n^{(2)} \mid \boldsymbol{\theta} \right) \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} = \\
&= b_0 \int_{\Theta_0} \left[ \Phi \left( \frac{-k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{-k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right. \\
&\quad \left. + \Phi \left( \frac{k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= b_0 \int_{\mathbb{R}^2} \left[ \Phi \left( \frac{-k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{-k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right. \\
&\quad \left. + \Phi \left( \frac{k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right] \mathbb{1}_{\Theta_0}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}.
\end{aligned}$$

$$\begin{aligned}
I_4 &= b_1 \int_G \int_{\Theta_1} f(\boldsymbol{\theta}, w) d\boldsymbol{\theta} dw = b_1 \int_{\Theta_1} \left[ \int_G f(w|\boldsymbol{\theta}) dw \right] f(\boldsymbol{\theta}) d\boldsymbol{\theta} \\
&= b_1 \int_{\mathbb{R}^2} \left[ \Phi \left( \frac{-k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{-k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right. \\
&\quad \left. + \Phi \left( \frac{k_n^{(2)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) - \Phi \left( \frac{k_n^{(1)} - (\theta_1 - \theta_2)}{\sqrt{2v^2/n}} \right) \right] \mathbb{1}_{\Theta_1}(\boldsymbol{\theta}) f(\boldsymbol{\theta}) d\boldsymbol{\theta}.
\end{aligned}$$

**Exemplo 4.7.** Sejam as amostras  $\mathbf{X} = (X_1, \dots, X_n)$  e  $\mathbf{Y} = (Y_1, \dots, Y_n)$  que, dado  $\boldsymbol{\theta}$  são independentes entre si e tais que  $X_i|\boldsymbol{\theta} \stackrel{iid}{\sim} \mathcal{N}(\theta_1, 1)$  e  $Y_i|\boldsymbol{\theta} \stackrel{iid}{\sim} \mathcal{N}(\theta_2, 1)$ , para  $i = 1, \dots, n$ . As hipóteses a serem testadas são  $H_0 : (\theta_1, \theta_2) \in \Theta_0$  e  $H_1 : (\theta_1, \theta_2) \in \Theta_1$ ,  $\Theta_0 = \{(\theta_1, \theta_2) \in \mathbb{R}^2 : \theta_1 - \epsilon < \theta_2 < \theta_1 + \epsilon\}$  e  $\Theta_1 = \{(\theta_1, \theta_2) \in \mathbb{R}^2 : \theta_2 \leq \theta_1 - \epsilon \text{ ou } \theta_2 \geq \theta_1 + \epsilon\}$ , com  $\epsilon = 0,1$ . Para o cálculo dos riscos dos testes baseados em estimador por região, considere as funções de perda  $0 - a_0 - a_1$  e  $0 - a_0 - a_1 - b_0 - b_1$ , com  $a_0 = 1$ ,  $a_1 = 1$ ,  $b_0 = \frac{1}{4}$ ,  $b_1 = \frac{1}{4}$ . Seja  $\alpha = 0,05$  de modo que a HPD é uma região de 0,95 de credibilidade.

Supondo como prioris para  $\theta_1$  e  $\theta_2$  a distribuição Normal com parâmetros  $\mu_0 = 0$  e  $\sigma_0^2 = 100$ , calculamos o risco para cada um dos testes para  $n$  variando de 500 a 2000, pois nesse caso temos restrições de valores mais altas para  $n$  para que seja possível aceitar  $H_0$ . As integrais para o cálculo do risco foram aproximadas pelo método de Monte Carlo com  $M = 10^5$  iterações. O gráfico abaixo

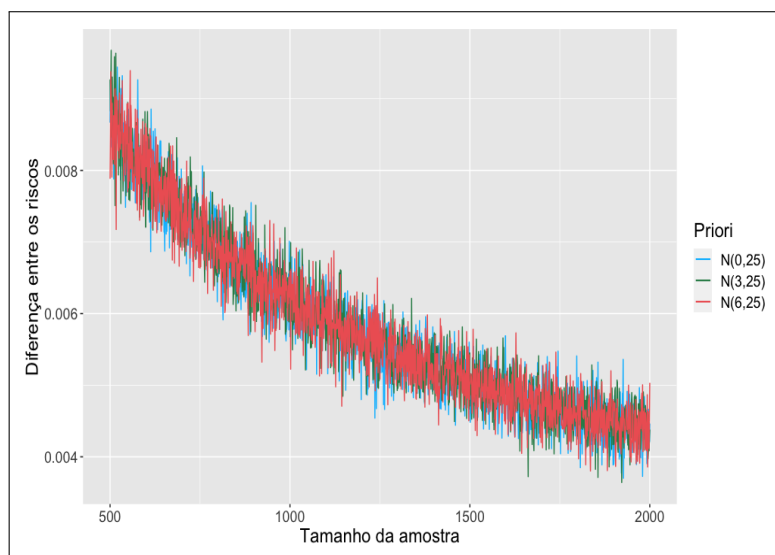
ilustra os riscos dos testes tradicional e agnóstico em função de  $n$ .



**Figura 4.22:** Riscos dos testes baseados em estimador por região para modelo Normal bivariado - Prioris  $\mathcal{N}(0, 100)$ .

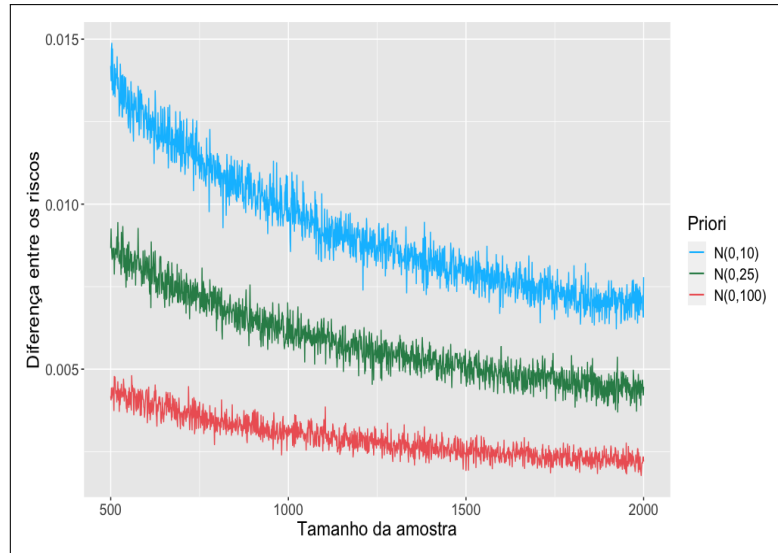
Assim como no caso univariado, o risco do teste agnóstico é menor que o risco do teste tradicional para todo  $n$  considerado. Para essa priori, temos a restrição de que  $n \geq 1199$  para que possamos aceitar  $H_0$  no teste agnóstico.

As figuras a seguir retratam a diferença absoluta entre os riscos dos dois testes para distintas escolhas de prioris.



**Figura 4.23:** Diferença absoluta entre os riscos para prioris com médias diferentes e variâncias iguais - Testes baseados em estimador por região para modelo Normal bivariado.

Diferentemente de todos os exemplos anteriores, observa-se na Figura 4.23 que não há impacto na mudança da média da priori em termos da diferença entre os riscos dos dois testes. Tal fato ocorre pois os termos  $k_n$ ,  $k_n^{(1)}$  e  $k_n^{(2)}$  não dependem da média da priori  $\mu_0$  nas expressões (4.36) e (4.39). Porém, como os mesmos termos dependem da variância da priori  $\sigma_0^2$ , notamos na Figura 4.24 que a variância de fato impacta na diferença entre os riscos. Nesse caso, quanto maior a variância da priori, menor a diferença de risco entre os testes, mesmo comportamento observado em todos os exemplos já estudados, com exceção do modelo Poisson. Vale destacar que não investigamos o caso em que as distribuições marginais a priori de  $\theta_1$  e  $\theta_2$  são distintas (com médias distintas).



**Figura 4.24:** Diferença absoluta entre os riscos para prioris com mesma média e variâncias diferentes - Testes baseados em estimador por região para modelo Normal bivariado.

Por fim, a tabela a seguir apresenta os resultados dos tamanhos amostrais ótimos  $n_T^*$  e  $n_A^*$  para cada teste, assumindo um custo  $c = 0,0000025$ . Tal custo foi escolhido de modo que os tamanhos ótimos obtidos fossem grandes o suficiente para garantir que  $H_0$  possa ser aceita.

Priori $\theta$	$n_T^*$	$n_A^*$	$\delta(\%)$	$\bar{\rho}_T^*$	$\bar{\rho}_A^*$
$\mathcal{N}(-3, 10)$	1863	1295	30,5	0,0134	0,0072
$\mathcal{N}(-3, 25)$	1767	1107	37,4	0,0086	0,0046
$\mathcal{N}(-3, 100)$	1138	665	41,6	0,0052	0,0031
$\mathcal{N}(0, 10)$	1923	1331	30,8	0,0132	0,0070
$\mathcal{N}(0, 25)$	1840	1227	33,3	0,0084	0,0045
$\mathcal{N}(0, 100)$	981	645	34,3	0,0056	0,0032
$\mathcal{N}(3, 10)$	1948	1504	22,8	0,0131	0,0070
$\mathcal{N}(3, 25)$	1856	1082	41,7	0,0083	0,0046
$\mathcal{N}(3, 100)$	1196	714	40,3	0,0050	0,0029
$\mathcal{N}(6, 10)$	1941	1369	29,5	0,0131	0,0070
$\mathcal{N}(6, 25)$	1702	1067	37,3	0,0088	0,0047
$\mathcal{N}(6, 100)$	787	716	9,0	0,0061	0,0033
$\mathcal{N}(9, 10)$	1988	1516	23,7	0,0128	0,0071
$\mathcal{N}(9, 25)$	1706	856	49,8	0,0087	0,0047
$\mathcal{N}(9, 100)$	1329	680	48,8	0,0047	0,0028

**Tabela 4.10:** Tamanhos amostrais e riscos para os testes baseados em estimador por região - Modelo Normal bivariado.

Podemos perceber que, quando fixamos a média da priori, quanto maior a variância, menor o tamanho amostral ótimo dos dois testes. Já quando fixamos a variância e variamos a média, não há um padrão claro em relação aos tamanhos amostrais. Na comparação de  $n_T^*$  e  $n_A^*$ , fixada a priori, observamos variações percentuais do caso agnóstico em relação ao caso tradicional em todos os cenários e com valores similares ao exemplo anterior do modelo univariado, com reduções expressivas em torno de 30 a 40% para a maior parte das prioris.



# Capítulo 5

## Conclusões

### 5.1 Considerações Finais

Neste trabalho, foram exploradas algumas potenciais vantagens adicionais da utilização de testes agnósticos como alternativa ao uso de testes tradicionais, além das vantagens lógicas em testes simultâneos exploradas em trabalhos anteriores. Foi demonstrado que, para testes de Bayes agnósticos e tradicionais sob perdas iguais para as decisões de rejeição e aceitação da hipótese nula, o risco de Bayes do teste agnóstico é estritamente menor que o risco de Bayes do teste tradicional, desde que haja probabilidade positiva de o teste resultar na decisão de ficar agnóstico. Além disso, testes de hipóteses bayesianos para diversos modelos estatísticos foram derivados e implementados em rotinas computacionais, tanto no cenário tradicional quanto no agnóstico. Os testes de duas e três decisões foram comparados em termos de risco de Bayes e quanto ao problema de determinação de tamanho amostral ótimo, para várias escolhas de distribuição a priori para os parâmetros de interesse.

Através de exemplos numéricos, notamos que a escolha dos parâmetros da priori pode impactar de forma expressiva os tamanhos amostrais ótimos obtidos pelo método proposto em DeGroot [16] e Sahu e Smith [41]. Adicionalmente, vimos que testes agnósticos podem levar a uma redução de custo substancial no desenho de experimentos, já que os tamanhos amostrais ótimos obtidos nos testes tradicionais foram, em geral, 10 a 20% maiores que os obtidos nos testes agnósticos. Tais resultados foram observados para a maior parte dos testes sob perda  $0 - a_0 - a_1 - b_0 - b_1$ , com  $b_0 (= b_1) = \frac{1}{4}$ . Contudo, nos testes sob perda  $0 - L_1$  e  $0 - L_1^{(p)}$ , as diferenças entre os testes em termos de risco de Bayes e tamanhos amostrais ótimos foram menos expressivas, o que nos faz concluir que as vantagens em termos de redução de risco e tamanho da amostra dependerão da função de perda escolhida.

Nos testes baseados em probabilidade a posteriori, conforme os valores de  $b_0$  e  $b_1$  aumentam e os testes de Bayes tradicionais e agnósticos se aproximam, a diferença entre os riscos de Bayes dos dois tipos de testes tende a reduzir. O comportamento de redução do risco também foi observado inclusive para outros testes bayesianos que não são necessariamente testes ótimos sob as perdas escolhidas, ou seja, que não são derivados pela minimização do risco a posteriori. Especificamente, os testes agnósticos baseados em estimador por região, estudados em caráter exploratório, também apresentaram uma redução no risco e nos tamanhos amostrais ótimos em relação aos testes baseados em estimador por região no cenário tradicional.

Portanto, dadas as vantagens já estudadas em outros trabalhos e as exploradas aqui, a utilização de testes agnósticos pode ser uma alternativa interessante aos testes de duas decisões do ponto de vista prático e teórico, principalmente em situações de testes simultâneos para o mesmo conjunto de dados observados e em problemas de decisão nos quais tomar alguma decisão definitiva a respeito das hipóteses (aceitação ou rejeição) pode ser mais custoso do que se abster de uma decisão de forma temporária. Já em situações em que se faz necessária alguma decisão definitiva—ou seja, não é de interesse do agente decisor ter a possibilidade de ficar agnóstico—uma alternativa seria construir o teste de forma que a probabilidade de ficar agnóstico é baixa a partir de penalidades

maiores para a decisão de ficar agnóstico relativamente às penalidades de aceitação ou rejeição da hipótese nula. Em tal cenário, apesar de as vantagens dos testes agnósticos em termos de redução do risco de Bayes e do tamanho amostral possivelmente serem menos expressivas, há o benefício de o teste não resultar em conclusões logicamente inconsistentes em problemas de testes simultâneos.



## 5.2 Sugestões para Pesquisas Futuras

- Explorar se existem outras vantagens dos testes agnósticos em relação aos testes tradicionais.
- Estudar, implementar e comparar testes FBST e GFBST, utilizando funções de perda que garantam que estes sejam testes de Bayes.
- Aplicar outros métodos de determinação de tamanho de amostra de forma a avaliar se reduções do tamanho amostral ótimo são obtidas nos testes agnósticos.
- Estudar condições mais gerais em relação às funções de perda sob as quais possamos garantir a redução do risco de Bayes em testes agnósticos. Especificamente, determinar quais condições sobre  $\bar{L}(0, \theta)$  e  $\bar{L}(1/2, \theta)$ , comparativamente a  $L(0, \theta)$ , asseguram a redução do risco de Bayes, mantendo-se iguais as penalidades associadas à rejeição da hipótese nula.
- Estudar comparativamente testes tradicionais e agnósticos quanto ao risco de Bayes em outros modelos estatísticos de interesse como modelos lineares, modelos para dados categorizados e modelos com parâmetros de incômodo (*nuisance*), por exemplo.
- Estudar possíveis generalizações dos resultados obtidos para testes com mais de duas ações possíveis (além de rejeitar ou não rejeitar  $H_0$ ), particionando o espaço paramétrico em mais de dois subconjuntos.



# Apêndice A

## Demonstrações adicionais

### A.1 Distribuição marginal do modelo Beta-Bernoulli

Seja uma amostra  $\mathbf{X} = (X_1, \dots, X_n)$  em que  $X_i|\theta \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$ , para  $i = 1, \dots, n$  e  $\theta \in (0, 1)$ . Assumindo que  $\theta$  possui distribuição priori  $\theta \sim \text{Beta}(\alpha, \beta)$ , a distribuição marginal da estatística suficiente  $S(\mathbf{X}) = \sum_{i=1}^n X_i$  é *Beta - Binomial*( $n, \alpha, \beta$ ) com função de probabilidade

$$\mathbb{P}(S(\mathbf{X}) = s) = \binom{n}{s} \frac{B(s + \alpha, n - s + \beta)}{B(\alpha, \beta)} \mathbb{1}_{\{0, \dots, n\}}(s),$$

em que  $s = \sum_{i=1}^n x_i$ .

*Demonstração.* Como  $X_i|\theta \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$ , temos que  $\sum_{i=1}^n X_i|\theta \stackrel{iid}{\sim} \text{Binomial}(n, \theta)$ . Segue que

$$\begin{aligned} \mathbb{P}(S = s) &= \int_0^1 \mathbb{P}(S = s|\theta) f(\theta) d\theta \\ &= \int_0^1 \binom{n}{s} \theta^s (1 - \theta)^{n-s} \mathbb{1}_{\{0, \dots, n\}}(s) \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1} d\theta \\ &= \binom{n}{s} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \mathbb{1}_{\{0, \dots, n\}}(s) \int_0^1 \theta^{\alpha+s-1} (1 - \theta)^{\beta+n-s-1} d\theta \\ &= \binom{n}{s} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha + s)\Gamma(\beta + n - s)}{\Gamma(\alpha + \beta + n)} \mathbb{1}_{\{0, \dots, n\}}(s) \\ &= \binom{n}{s} \frac{\beta(\alpha + s, \beta + n - s)}{\beta(\alpha, \beta)} \mathbb{1}_{\{0, \dots, n\}}(s). \end{aligned}$$

Logo,

$$S(\mathbf{X}) = \sum_{i=1}^n X_i \sim \text{Beta - Binomial}(n, \alpha, \beta).$$

□

## A.2 Distribuição marginal do modelo Poisson-Gama

Seja uma amostra  $\mathbf{X} = (X_1, \dots, X_n)$  em que  $X_i|\theta \stackrel{iid}{\sim} \text{Poisson}(\theta)$ , para  $i = 1, \dots, n$  e  $\theta > 0$ . Assumindo como priori para  $\theta$  uma distribuição  $\theta \sim \text{Gama}(\alpha, \beta)$ ,  $\alpha \in \mathbb{N}^*$ ,  $\beta > 0$ , a distribuição marginal da estatística suficiente  $S(\mathbf{X}) = \sum_{i=1}^n X_i$  é *Binomial – Negativa*  $\left(\alpha, \frac{\beta}{\beta+n}\right)$ , com função de probabilidade

$$\mathbb{P}(S(\mathbf{X}) = s) = \frac{\Gamma(\alpha + s)}{\Gamma(\alpha)s!} \left(\frac{\beta}{\beta + n}\right)^\alpha \left(\frac{n}{\beta + n}\right)^s \mathbb{1}_{\mathbb{N}}(s),$$

em que  $s = \sum_{i=1}^n x_i$ .

*Demonstração.* Se  $X_i|\theta \stackrel{iid}{\sim} \text{Poisson}(\theta)$ , então  $\sum_{i=1}^n X_i|\theta \stackrel{iid}{\sim} \text{Poisson}(n\theta)$ . Assim,

$$\begin{aligned} \mathbb{P}(S = s) &= \int_0^{+\infty} \mathbb{P}(S = s|\theta) f(\theta) d\theta = \int_0^{+\infty} \frac{(n\theta)^s e^{-n\theta}}{s!} \mathbb{1}_{\mathbb{N}}(s) \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} d\theta \\ &= \frac{\beta^\alpha n^s}{\Gamma(\alpha)s!} \mathbb{1}_{\mathbb{N}}(s) \int_0^{+\infty} \theta^{\alpha+s-1} e^{-(\beta+n)\theta} d\theta = \frac{\beta^\alpha n^s}{\Gamma(\alpha)s!} \mathbb{1}_{\mathbb{N}}(s) \frac{\Gamma(\alpha + s)}{(\beta + n)^{s+\alpha}} \\ &= \frac{\Gamma(\alpha + s)}{\Gamma(\alpha)s!} \left(\frac{\beta}{\beta + n}\right)^\alpha \left(\frac{n}{\beta + n}\right)^s \mathbb{1}_{\mathbb{N}}(s). \end{aligned}$$

Logo,

$$S(\mathbf{X}) = \sum_{i=1}^n X_i \sim \text{Binomial – Negativa} \left(\alpha, \frac{\beta}{\beta + n}\right).$$

□

# Referências Bibliográficas

- [1] J. Aitchison. Two papers on the comparison of Bayesian and frequentist approaches to statistical problems of prediction: Bayesian tolerance regions. *Journal of the Royal Statistical Society: Series B (Methodological)*. 15
- [2] N. Berg. No-decision classification: an alternative to testing for statistical significance. *The Journal of Socio-Economics*, 33(5):631–650, 2004. 15
- [3] J. O. Berger. *Statistical Decision Theory and Bayesian Analysis*. Springer, New York, 2nd edition, 1985. 4, 12, 15, 26
- [4] J. O. Berger. Could Fisher, Jeffreys and Neyman have agreed on testing? *Statistical Science*, 18(1):1–32, 2003. 9
- [5] J. O. Berger e M. Delampady. Testing precise hypotheses. *Statistical Science*, 2(3):317–335, 1987. 12
- [6] J. M. Bernardo e A. F. M. Smith. *Bayesian Theory*. John Wiley Sons, Chichester, 2000. 5
- [7] G. Cancelier. *Uma Introdução à Teoria da Medida e Integração*. Trabalho de Conclusão de Curso, Universidade Federal de Santa Catarina, 2019. 24
- [8] G. Casella e R. L. Berger. Reconciling Bayesian and frequentist evidence in the one-sided testing problem. *Journal of the American Statistical Association*, 82(398):106–111, 1987. 12
- [9] G. Casella e R. L. Berger. *Statistical Inference*. Duxbury Advanced Series, Pacific Grove, 2nd edition, 2002. 13, 14
- [10] V. Coscrato, R. Izbicki, e R. B. Stern. Agnostic tests can control the type I and type II errors simultaneously. *Brazilian Journal of Probability and Statistics*, 34(2):230–250, 2020. 15, 20
- [11] D. R. Cox, E. Spjøtvoll, S. Johansen, W. R. van Zwet, J. F. Bithell, O. Barndorff-Nielsen, e M. Keuls. The role of significance tests [with discussion and reply]. *Scandinavian Journal of Statistics*, 4(2):49–70, 1977. 9
- [12] G. M. da Silva, L. G. Esteves, V. Fossaluzza, R. Izbicki, e S. Wechsler. A Bayesian decision-theoretic approach to logically-consistent hypothesis testing. *Entropy*, 17(10):6534–6559, 2015. 15
- [13] C. de Bragança Pereira, Alberto, J. M. Stern, e S. Wechsler. Can a significance test be genuinely Bayesian? *Bayesian Analysis*, 3(1):79–100, 2008. 13
- [14] C. A. de Bragança Pereira e J. M. Stern. Evidence and credibility: Full Bayesian significance test for precise hypotheses. *Entropy*, 1(4):99–110, 1999. 12
- [15] B. de Finetti. *Probability: Induction and Statistics*. Wiley, New York, 1972. 4
- [16] M. H. DeGroot. *Optimal Statistical Decisions*. MacGraw-Hill, New York, 1970. 4, 5, 6, 25, 27, 50, 69

- [17] L. G. Esteves, R. Izbicki, J. M. Stern, e R. B. Stern. The logical consistency of simultaneous agnostic hypothesis tests. *Entropy*, 18(7), 2016. 1, 16, 17, 19, 20
- [18] S. Greenland, S. J. Senn, K. J. Rothman, J. Carlin, C. Poole, S. Goodman, e D. Altman. Statistical tests, p values, confidence intervals, and power: a guide to misinterpretations. *Eur J Epidemiol.*, 31(10):337–350, 2016. 1
- [19] R. Izbicki e L. G. Esteves. Logical consistency in simultaneous statistical test procedures. *Logic Journal of the IGP*, 23(5):732–758, 2015. 1, 15
- [20] H. Jeffreys. *The Theory of Probability*. Oxford University Press, 3rd edition, 1961. 9
- [21] L. Joseph, D. B. Wolfson, e R. D. Berger. Sample size calculations for binomial proportions via highest posterior density intervals. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 44(2):143–154, 1995. 25
- [22] H. F. Kaiser. Directional statistical decisions. *Psychological Review*, 67(3):160–167, 1960. 15
- [23] R. Kohavi, D. Tang, e Y. Xu. *Trustworthy Online Controlled Experiments*. Cambridge University Press, 2020. 1
- [24] J. K. Kruschke. Rejecting or accepting parameter values in Bayesian estimation. *Advances in Methods and Practices in Psychological Science*, 1(2):270–280, 2018. 12
- [25] K. Kunzmann, M. J. Grayling, K. M. Lee, D. S. Robertson, K. Rufibach, e J. M. S. Wason. A review of Bayesian perspectives on sample size derivation for confirmatory trials. *The American Statistician*, 75(4):424–432, 2021. 25
- [26] E. L. Lehmann. Some principles of the theory of testing hypotheses. *The Annals of Mathematical Statistics*, 21(1):1–26, 1950. 15
- [27] D. V. Lindley. *Introduction to Probability and Statistics from a Bayesian Viewpoint: Part 2*. Cambridge University Press, New York, 1965. 13
- [28] D. V. Lindley. The choice of sample size. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 46(2):129–138, 1997. 25
- [29] R. M. Madruga, C. de Bragança Pereira, e J. M. Stern. Bayesian evidence test for precise hypotheses. *Journal of Statistical Planning and Inference*, 117(2):185–198, 2003. 13
- [30] R. M. Madruga, L. G. Esteves, e S. Wechsler. On the bayesianity of pereira-stern tests. *Test*, 10(2):291–299, 2001. 13, 55
- [31] B. B. McShane, D. Gal, A. Gelman, C. Robert, e J. L. Tackett. Abandon statistical significance. *The American Statistician*, 73(sup1):235–245, 2019. 1
- [32] P. E. Meehl. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 46(4):806–834, 1978. 1
- [33] J. V. Neumann e O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 3rd edition, 1953. 3, 5
- [34] J. Neyman e E. S. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 231:289–337, 1933. 1
- [35] J. Neyman e E. S. Pearson. The testing of statistical hypotheses in relation to probabilities a priori. *Mathematical Proceedings of the Cambridge Philosophical Society*, 29(4):492–510, 1933. 15

- [36] C. D. Paulino, M. A. A. Turkman, B. Murteira, e G. L. Silva. *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa, 2nd edition, 2018. 37
- [37] H. Raiffa e R. Schlaifer. *Applied Statistical Decision Theory*. John Wiley & Sons, New York, wiley classics library edition, 2000. 3, 6, 25, 26
- [38] F. P. Ramsey. Truth and probability. In H. Arló-Costa, V. F. Hendricks, e J. van Benthem, editors, *Readings in Formal Epistemology: Sourcebook*, pages 21–45. Springer International Publishing, Cham, 2016. 3, 5
- [39] C. P. Robert. *The Bayesian Choice*. Springer, New York, 2nd edition, 2007. 12, 15
- [40] C. P. Robert e G. Casella. *Monte Carlo Statistical Methods*. Springer, New York, 2nd edition, 2004. 29
- [41] S. K. Sahu e T. M. F. Smith. A Bayesian method of sample size determination with practical applications. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 169(2):235–253, 2006. 25, 33, 69
- [42] L. J. Savage. *The Foundations of Statistics*. Dover Publications, New York, 2nd edition, 1972. 4, 5
- [43] M. J. Schervish. *Theory of Statistics*. Springer, New York, 1st edition, 1995. 9, 10, 12
- [44] D. J. Spiegelhalter, K. R. Abrams, e J. P. Myles. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. John Wiley & Sons, Chichester, 2004. 25
- [45] D. J. Spiegelhalter, L. S. Freedman, e M. K. B. Parmar. Bayesian approaches to randomized trials. *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 157(3):357–416, 1994. 12
- [46] J. M. Stern. Cognitive constructivism, eigen-solutions, and sharp statistical hypotheses. *Cybernetics & Human Knowing*, 14(1):9–36, 2007a. 12
- [47] J. M. Stern, R. Izbicki, L. G. Esteves, e R. B. Stern. Logically-consistent hypothesis testing and the hexagon of oppositions. *Logic Journal of the IGPL*, 25(5):741–757, 2017. 14, 20
- [48] M. Thulin. Decision-theoretic justifications for Bayesian hypothesis testing using credible sets. *Journal of Statistical Planning and Inference*, 146:133–138, 2014. 14
- [49] A. Wald. *Sequential Analysis*. John Wiley & Sons, New York, 1947. 15
- [50] Y. Yao. The superiority of three-way decisions in probabilistic rough set models. *Information Sciences*, 181(6):1080–1096, 2011. 16