

# Classes de testes de hipóteses

Rafael Izbicki

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

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CNPq/FAPESP

São Paulo, junho de 2010

## Classes de testes de hipóteses

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

Banca Examinadora:

- Prof. Dr. Luís Gustavo Esteves (orientador) - IME-USP.
- Prof. Dr. Sergio Wechsler - IME-USP.
- Prof. Dr. Verónica Andrea González-López - UNICAMP.

*"Há um tempo  
em que é preciso abandonar as roupas usadas  
que já tem a forma do nosso corpo  
e esquecer os nossos caminhos  
que nos levam sempre aos mesmos lugares.*

*É o tempo da travessia  
E se não ousarmos fazê-la  
Teremos ficado para sempre  
À margem de nós mesmos."*

— Fernando Pessoa



## Agradecimentos

Inicialmente agradeço ao verdadeiro responsável por este trabalho: meu orientador Professor Luís Gustavo Esteves. Desde o início dos meus estudos em estatística tive o prazer e a sorte de estar em contato com ele.

Agradeço aos meus pais, Deborah e Meyer, e aos meus avós, Ita, Leon e Rebecca, que me ensinaram que o conhecimento é a única aquisição impossível de nos ser tirada. Agradeço à minha irmã, Sarah, pela paciência que teve comigo durante todo este tempo.

Agradeço ao Professor Carlos Alberto de Bragança Pereira (Carlinhos) pela contribuição inestimável que teve neste trabalho, e que, assim como o Professor Sérgio Wechsler, sempre me estimulou a pensar.

Agradeço aos seguintes amigos que de alguma maneira contribuíram para esta dissertação, seja através de discussões concretas sobre os temas aqui abordados, seja através do apoio que me deram: André Venciguero Annunziato, Eduardo Yoshio Nakano, Fabiana Pansonato, Felipe Giersztajn, Fernando Vieira Bonassi, Gabriel Ribeiro da Cruz Peixoto, Luiz Fernando Oliveira Corte Real, Rafael Bassi Stern, Renata Trevisan Brunelli, Tiago Mendonça e Victor Fossaluzza.



## Resumo

Na Inferência Estatística, é comum, após a realização de um experimento, testar simultaneamente um conjunto de diferentes hipóteses de interesse acerca de um parâmetro desconhecido. Assim, para cada hipótese, realiza-se um teste de hipótese e, a partir disto, conclui-se algo sobre os parâmetros de interesse. O objetivo deste trabalho é avaliar a (falta de) concordância lógica entre as conclusões obtidas a partir dos testes realizados após a observação de um único experimento.

Neste estudo, é apresentada uma definição de classe de testes de hipóteses, uma função que para cada hipótese de interesse associa uma função de teste. São então avaliadas algumas propriedades que refletem como gostaríamos que testes para diferentes hipóteses se comportassem em termos de coerência lógica. Tais propriedades são exemplificadas através de classes de testes que as satisfazem. A seguir, consideram-se conjuntos de axiomas para classes. Estes axiomas são baseados nas propriedades mencionadas. Classes de testes usuais são investigadas com relação aos conjuntos de axiomas propostos. São também estudadas propriedades advindas de tais conjuntos de axiomas. Por fim, estuda-se um resultado que estabelece uma espécie de conexão entre testes de hipóteses e estimação pontual.

**Palavras-chave:** Testes simultâneos, Teoria da decisão, Regras de decisão, Monotonicidade, Propriedades lógicas.





## Abstract

In Statistical Inference, it is usual, after an experiment is performed, to test simultaneously a set of hypotheses of interest concerning an unknown parameter. Therefore, to each hypothesis, a statistical test is performed and a conclusion about the parameter is drawn based on it. The objective of this work is to evaluate the (lack of) logical coherence among conclusions obtained from tests conducted after the observation of a single experiment.

In this study, a definition of class of hypotheses tests, a function that associates a test function to each hypothesis of interest, is presented. Some properties that reflect what one could expect (in terms of logical coherence) from tests to different hypotheses are then evaluated. These properties are exemplified by classes of hypotheses tests that respect them. Then, sets of axioms based on the properties studied are proposed to classes of hypotheses tests. Usual classes of hypotheses tests are investigated with respect to these sets of axioms. Some properties related to these sets of axioms are then analyzed. At last, a result which seems to connect hypotheses testing and point estimation is stated.

**Keywords:** Simultaneous tests, Decision Theory, Decision rules, Monotonicity, Logical properties.



# Sumário

<b>Lista de Abreviaturas</b>	<b>xi</b>
<b>Lista de Símbolos</b>	<b>xiii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Considerações Preliminares e Objetivos . . . . .	1
1.2 Definições Básicas . . . . .	1
1.3 Organização do Trabalho . . . . .	4
<b>2 Propriedades de classes de testes de hipóteses</b>	<b>5</b>
2.1 Monotonicidade . . . . .	5
2.1.1 Monotonicidade e Admissibilidade . . . . .	13
2.2 Consonância da União . . . . .	15
2.3 Consonância da Intersecção . . . . .	19
2.4 Invertibilidade . . . . .	22
<b>3 Desideratas</b>	<b>25</b>
3.1 Desideratas . . . . .	25
3.2 Relações entre as Desideratas . . . . .	35
<b>4 Conclusões</b>	<b>37</b>
4.1 Considerações Finais . . . . .	37
4.2 Sugestões para Pesquisas Futuras . . . . .	37
<b>A Ultrafiltros</b>	<b>39</b>
<b>B Demonstração Adicional</b>	<b>43</b>
<b>Referências Bibliográficas</b>	<b>47</b>



## Lista de Abreviaturas

ANOVA	Análise de Variância.
FBST	Full Bayesian Significance Test.
TRVG	Teste da Razão de Verossimilhanças Generalizada.
UMP	Teste Uniformemente Mais Poderoso.
i.i.d.	Independentes e identicamente distribuídas.



## Lista de Símbolos

$\mathbb{I}(A)$	Função indicadora de $A$ ser uma afirmação verdadeira.
$I$	Conjunto de índices arbitrário.
$\Re$	Os reais.
$\mathbb{Z}$	Os inteiros.
$\mathbb{N}$	Os naturais.
$\Theta$	Espaço paramétrico.
$\chi$	Espaço amostral.
$\mathcal{L}$	Classe de testes de hipóteses.
$L_A$	Função de perda para a hipótese $A$ .
$V_x(\theta)$	Função de verossimilhança no ponto $\theta$ para a observação $x$ .
$\sigma(A)$	Uma $\sigma$ -álgebra do conjunto $A$ .
$\mathcal{P}(A)$	O conjunto das partes de $A$ .
$\mathcal{B}(A)$	O conjunto dos Borelianos de $A$ .





# Capítulo 1

## Introdução

### 1.1 Considerações Preliminares e Objetivos

Na Inferência Estatística, quando há interesse em se testar um determinado conjunto de hipóteses simultaneamente após a realização de um experimento, é comum buscar estabelecer critérios de otimalidade que os testes para cada uma das hipóteses (com duas alternativas cada: nula e alternativa), individualmente, devem satisfazer. Exemplos destas propriedades são a minimização de probabilidades de erros de tomada de decisão, sob a abordagem clássica, e a minimização de perda esperada, sob a abordagem bayesiana. Assim, são derivados os testes Uniformemente Mais Poderosos (UMP), Testes da Razão de Verossimilhanças Generalizada (TRVG) e os Testes de Bayes, entre outros. Contudo, raramente avaliam-se quais propriedades seriam razoáveis de se esperar que tais testes satisfizessem conjuntamente. Por exemplo, ao se realizar uma Análise de Variância (ANOVA), usualmente tem-se interesse em várias hipóteses relacionadas às igualdades das médias dos diferentes grupos. Ao testá-las das formas usuais, é comum incorrer em algumas contradições, sob o ponto de vista lógico. Por exemplo, é comum concluir-se, simultaneamente, que a média do primeiro grupo é igual à do segundo, que a média do segundo grupo é igual à do terceiro e que a média do primeiro grupo é diferente da do terceiro. Evidentemente, de um ponto de vista estritamente lógico, tal situação não poderia ocorrer. Neste trabalho, buscou-se especificar e estudar algumas propriedades que poderiam se esperar quando se tem interesse em realizar uma série de testes de hipóteses, cada um com duas alternativas. Exemplos de conjuntos (classes) de testes que satisfazem tais propriedades são mostrados. São também caracterizados conjuntos de propriedades desejáveis (*desideratas*).

### 1.2 Definições Básicas

Neste trabalho, denotamos por  $\Theta$  o espaço paramétrico, cujos elementos são denotados por  $\theta$ ,  $\chi$  é o espaço amostral, que tem como elementos  $x$ ,  $\sigma(A)$  é uma  $\sigma$ -álgebra de subconjuntos do conjunto não vazio  $A$  e  $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$  é uma família de distribuições de probabilidade em  $\sigma(\chi)$  indexadas por  $\theta$ . O símbolo  $V_x(\theta)$  é utilizado para designar a função de verossimilhança no ponto  $\theta$  gerada pela observação  $x$ . Ademais,  $\mathbb{I}(\cdot)$  é, como em Gerber et al. [8], a função indicadora da afirmação  $A$ : ela vale 1 quando  $A$  é verdadeira e 0 caso contrário. Por exemplo,  $\mathbb{I}(\theta < 0)$  vale 1 quando  $\theta < 0$  e

o caso contrário. Observa-se que esta definição é diferente da definição usual de função indicadora. O símbolo  $I$  é utilizado para designar um conjunto de índices arbitrário, podendo, em particular, ser não enumerável.

A seguir, definimos uma hipótese estatística. Tal definição encontra-se em Casella e Berger [2].

**Definição 1.1. (Hipótese estatística)** *Uma hipótese estatística é uma afirmação sobre um parâmetro populacional.*

*Nota:* Uma hipótese é usualmente formulada como  $H : \theta \in A$ , em que  $A$  é um subconjunto do espaço paramétrico. Neste trabalho, faremos o seguinte abuso de notação: utilizaremos  $A$  como sinônimo de  $H$ . Isto é, apesar da hipótese ser *uma afirmação* sobre os parâmetros, ela também será considerada *um subconjunto* do espaço paramétrico. Assim, serão feitas afirmações do tipo “Consideremos a hipótese  $A \subseteq \Theta$ ”, quando o formal seria “Consideremos a hipótese  $H : \theta \in A$ ”.

Em um problema de testes de hipóteses, a hipótese que se deseja testar em geral é chamada de *hipótese nula* ( $H_0$  ou  $H$ ). Definida esta hipótese, fica especificada a chamada *hipótese alternativa* ( $H_A$  ou  $H_1$ ), que nada mais é que o complementar de  $H_0$ . Neste trabalho, utilizaremos apenas as hipóteses nulas, de modo que as respectivas hipóteses alternativas ficam subentendidas.

**Definição 1.2. (Teste de hipótese)** *Seja  $(\chi, \sigma(\chi), \{\mathbb{P}_\theta : \theta \in \Theta\})$  um modelo estatístico. Seja  $A \subseteq \Theta$  uma hipótese. Um teste de hipótese para  $A$  é uma função  $\sigma(\chi)$ -mensurável  $\phi : \chi \rightarrow \{0, 1\}$  que, para cada possível observação  $x \in \chi$ , associa o número 0 ou 1. O valor 0 está associado à aceitação (ou não rejeição) de  $A$ , enquanto que o valor 1 está associado à rejeição (ou não aceitação) de  $A$ . O conjunto  $\phi^{-1}(\{1\})$  é chamado de região crítica (região de rejeição) do teste  $\phi$  para  $A$ .  $\phi^{-1}(\{0\})$  é chamado de região de aceitação do teste  $\phi$  para  $A$ .*

Um teste de hipótese se refere a um único problema de decisão com duas alternativas. Neste estudo, temos o objetivo de estudar a plausibilidade de várias hipóteses simultaneamente através de testes a serem conduzidos após a realização de um único experimento. Iremos supor que estamos interessados em um conjunto de hipóteses que formam uma  $\sigma$ -álgebra do espaço paramétrico. A ideia de uma *classe de testes de hipóteses*, que será definida a seguir, é, através de um único objeto,  $\mathcal{L}$ , associar a cada hipótese da  $\sigma$ -álgebra escolhida um teste de hipótese. Uma definição de classe de testes de hipóteses apropriada para nossa abordagem encontra-se em Silva [18], e a reproduzimos a seguir:

**Definição 1.3. (Classe de testes de hipóteses)** *Seja  $\sigma(\Theta)$  uma  $\sigma$ -álgebra de subconjuntos de  $\Theta$ . Seja  $\Psi = \{\phi : \chi \rightarrow \{0, 1\} \text{ tal que } \phi \text{ é } \sigma(\chi)\text{-mensurável}\}$  o conjunto de todas as funções de teste. Uma classe de testes de hipóteses é uma função  $\mathcal{L} : \sigma(\Theta) \rightarrow \Psi$  que, para cada hipótese  $A \in \sigma(\Theta)$ , associa o teste  $\mathcal{L}(A) \in \Psi$  para testar  $H_0 : \theta \in A$  contra  $H_A : \theta \notin A$ . Assim, para  $A \in \sigma(\Theta)$  e  $x \in \chi$ ,  $\mathcal{L}(A)(x)$  representa a decisão de aceitar, caso  $\mathcal{L}(A)(x) = 0$ , ou rejeitar, caso  $\mathcal{L}(A)(x) = 1$ , a hipótese  $A$  após observar o ponto  $x$ .*

A seguir, exemplificamos o conceito de classe de testes através de duas classes que são induzidas por testes usualmente utilizados.

**Exemplo 1.1. (Classe de Testes da Razão de Verossimilhanças Generalizada de nível  $\alpha$ )** Suponhamos que  $\Theta = \mathfrak{R}^n$  e seja  $\sigma(\Theta) = \mathcal{P}(\Theta)$  o conjunto das partes de  $\Theta$ . Para cada  $A \in \sigma(\Theta)$ , seja  $\mathcal{L}(A) : \chi \rightarrow \{0, 1\}$  definida como

$$\mathcal{L}(A)(x) = \mathbb{I} \left( \frac{\sup_{\theta \in A} V_x(\theta)}{\sup_{\theta \in \Theta} V_x(\theta)} \leq c_A \right), \quad \forall x \in \chi,$$

em que  $c_A \in [0, 1]$  é escolhido de tal forma que  $\sup_{\theta \in A} \mathbb{P}(\mathcal{L}(A)(X) = 1 | \theta) = \alpha$ ,  $\alpha \in (0, 1)$  fixado previamente. Em palavras, esta é a classe que associa a cada hipótese  $A \in \mathcal{P}(\Theta)$  o teste da Razão de Verossimilhanças Generalizada (TRVG) de nível  $\alpha$  para  $A$  (Casella e Berger [2]).

□

**Exemplo 1.2. (Testes baseados em probabilidades *a posteriori*)** Suponhamos o mesmo cenário do Exemplo 1.1, mas, desta vez, com  $\sigma(\Theta) = \mathcal{B}(\Theta)$ , os borelianos de  $\mathfrak{R}^n$ . Suponhamos também que está fixa uma medida de probabilidade *a priori*  $\mathbb{P}$  para  $\Theta$ . Para cada  $A \in \sigma(\Theta)$ , seja  $\mathcal{L}(A) : \chi \rightarrow \{0, 1\}$  definida como

$$\mathcal{L}(A)(x) = \mathbb{I} \left( \mathbb{P}(A|x) < \frac{1}{2} \right), \quad \forall x \in \chi,$$

onde  $\mathbb{P}(\cdot|x)$  é a distribuição *a posteriori* de  $\theta$ , dado  $x$ . Em palavras, esta é a classe que associa, a cada hipótese  $A \in \mathcal{B}(\mathfrak{R}^n)$ , o teste que aceita  $A$  quando sua probabilidade *a posteriori* é maior ou igual a  $1/2$ . Como, para cada  $A \in \sigma(\Theta)$ , o teste baseia-se apenas nas probabilidades *a posteriori* de  $A$ , doravante chamaremos tais testes de testes baseados em probabilidades *a posteriori*.

□

Notemos que a Definição (1.3) permite que o conjunto de hipóteses de interesse no Exemplo 1.2,  $\mathcal{B}(\mathfrak{R}^n)$ , seja tal que de fato é possível atribuir probabilidade a cada um de seus elementos. A classe apresentada neste exemplo não estaria bem definida se a  $\sigma$ -álgebra do Exemplo 1.1,  $\mathcal{P}(\Theta)$ , fosse utilizada (Royden [17]). Nota-se também que, alternativamente, uma classe de testes de hipóteses poderia ter sido definida em uma classe de subconjuntos de  $\Theta$  que não fosse uma  $\sigma$ -álgebra. Contudo, optou-se por defini-la desta forma para acomodar probabilidades sobre  $\Theta$ .

Do ponto de vista bayesiano, um teste de hipóteses é sempre derivado através de um procedimento de minimização da perdas esperadas (DeGroot [3]). Assim, fixadas uma medida de probabilidade para  $\theta$  e, para cada  $A \in \sigma(\Theta)$ , uma função de perda  $L_A : \{0, 1\} \times \Theta \rightarrow \mathfrak{R}$  (em que a decisão 0 representa a aceitação da hipótese nula, e a decisão 1, sua rejeição), podemos derivar um teste de Bayes para cada uma das hipóteses  $A \in \sigma(\Theta)$ . Como derivamos um teste de Bayes para cada uma das hipóteses em  $\sigma(\Theta)$ , podemos considerar a seguinte definição:

**Definição 1.4. (Classe de testes bayesianos gerada por uma família de funções de perda)** Seja  $(\chi \times \Theta, \sigma(\chi \times \Theta), \mathbb{P})$  um modelo estatístico bayesiano. Seja  $(L_A)_{A \in \sigma(\Theta)}$  uma família

de funções de perda, em que  $L_A : \{0, 1\} \times \Theta \rightarrow \mathfrak{R}$  é a função de perda associada ao teste para a hipótese  $A \in \sigma(\Theta)$ . Uma classe de testes bayesianos  $\mathcal{L}$  gerada por  $(L_A)_{A \in \sigma(\Theta)}$  é qualquer classe de testes de hipóteses definida sobre os elementos de  $\sigma(\Theta)$  tal que  $\mathcal{L}(A)$  é um teste de Bayes para a hipótese  $A$  contra  $\mathbb{P}, \forall A \in \sigma(\Theta)$ .

Notamos que com pequenas adaptações na formulação das funções de perda  $(L_A)_{A \in \sigma(\Theta)}$ , a Definição (1.4) contempla também testes bayesianos gerados por funções de perda que dependem também da amostra.

**Exemplo 1.3. (Testes baseados em probabilidades *a posteriori*)** Suponhamos o mesmo cenário do Exemplo 1.2 e que  $(L_A)_{A \in \sigma(\Theta)}$  é a família de funções de perda com  $L_A(0, \theta) = \mathbb{I}(\theta \notin A)$  e  $L_A(1, \theta) = \mathbb{I}(\theta \in A), \forall \theta \in \Theta, \forall A \in \sigma(\Theta)$ . Em palavras, perde-se uma unidade em caso de erro e não se perde nada em caso de acerto. Então, a classe  $\mathcal{L}$  definida no Exemplo 1.2 é uma classe de testes bayesianos gerada por esta família de funções de perda. A classe definida por

$$\mathcal{L}(A)(x) = \mathbb{I}\left(\mathbb{P}(A|x) \leq \frac{1}{2}\right), \forall A \in \sigma(\Theta) \text{ e } \forall x \in \chi,$$

também é uma classe de testes bayesianos gerada por esta mesma família de funções de perda.

□

### 1.3 Organização do Trabalho

No Capítulo 2, discutimos algumas propriedades que podem ser esperadas de classes de testes de hipóteses. Exemplos de classes que as satisfazem são mostrados. Além disso, algumas características relacionadas a estas propriedades são analisadas. No Capítulo 3, estudamos algumas *desideratas* (conjuntos de axiomas) propostas com base nas propriedades estudadas anteriormente. Finalmente, no Capítulo 4, concluímos a dissertação e discutimos alguns aspectos que não foram contemplados neste estudo. No Apêndice A, é apresentado o conceito de ultrafiltro, que é complementar ao Capítulo 2. O Apêndice B contém a demonstração do Exemplo 2.8.

## Capítulo 2

# Propriedades de classes de testes de hipóteses

A seguir, em cada seção, descreveremos uma propriedade possivelmente desejável para classes de testes de hipóteses. Cada seção sempre será introduzida por um exemplo que tem por finalidade motivar as considerações de tais propriedades. Estudamos também algumas características destas propriedades, bem como alguns exemplos de classes que as satisfazem.

### 2.1 Monotonicidade

**Exemplo 2.1.** Suponhamos que em um estudo caso-controle verifica-se, em cada indivíduo de uma amostra, seu genótipo em um determinado *locus*. Os resultados encontram-se na Tabela (2.1). Tais valores foram retirados de Lin et al. [12] em um estudo cujo objetivo era verificar a hipótese de que subunidades do gene  $GABA_A$  contribuem para uma condição conhecida como distúrbio do uso de metanfetamina.

Tabela 2.1: Frequências genotípicas amostrais

	AA	AB	BB	Total
Caso	55	83	50	188
Controle	24	42	39	105

Aqui, o conjunto dos possíveis genótipos é  $\mathbb{G} = \{AA, AB, BB\}$ . Seja  $\boldsymbol{\gamma} = (\gamma_{AA}, \gamma_{AB}, \gamma_{BB})$ , onde  $\gamma_i$  é a probabilidade de que um indivíduo do grupo caso tenha genótipo  $i$ ,  $i \in \mathbb{G}$ . Analogamente, seja  $\boldsymbol{\pi} = (\pi_{AA}, \pi_{AB}, \pi_{BB})$ , onde  $\pi_i$  é a probabilidade de um indivíduo do grupo controle tenha genótipo  $i$ ,  $i \in \mathbb{G}$ . O espaço paramétrico natural é

$\Theta = \{(\gamma_{AA}, \gamma_{AB}, \gamma_{BB}, \pi_{AA}, \pi_{AB}, \pi_{BB}) \in \mathbb{R}_+^6 : \sum_{i \in \mathbb{G}} \gamma_i = \sum_{i \in \mathbb{G}} \pi_i = 1\}$ , enquanto que o espaço amostral é  $\chi = \{(x_{AA}, x_{AB}, x_{BB}, y_{AA}, y_{AB}, y_{BB}) \in \mathbb{N}^6 : \sum_{i \in \mathbb{G}} x_i = 188 \text{ e } \sum_{i \in \mathbb{G}} y_i = 105\}$ . Considerando o modelo produto de multinomiais (Paulino e Singer [14]), temos que a verossimilhança é dada por

$$V_{\mathbf{x}, \mathbf{y}}(\boldsymbol{\theta}) \propto \prod_{i \in \mathbb{G}} \gamma_i^{x_i} \prod_{i \in \mathbb{G}} \pi_i^{y_i}, \quad \forall \boldsymbol{\theta} \in \Theta,$$

em que  $\mathbf{x}$  e  $\mathbf{y}$  são os vetores de observações do grupo caso e controle, respectivamente.

Neste contexto, duas hipóteses são de interesse em geral:  $H_0^G : \boldsymbol{\gamma} = \boldsymbol{\pi}$ , isto é, a hipótese de que

as proporções genóticas são iguais nos dois grupos e  $H_0^A : \gamma_{AA} + \frac{1}{2}\gamma_{AB} = \pi_{AA} + \frac{1}{2}\pi_{AB}$ , a hipótese de que as proporções alélicas são as mesmas nos dois grupos.

Os p-valores obtidos através dos testes Chi-quadrado usuais (Paulino e Singer [14]) para essas hipóteses são, respetivamente, 0,152 e 0,069 (estes valores foram calculados através do método de Monte Carlo; não foi utilizada a aproximação pela distribuição chi-quadrado). Deste modo, vemos que, ao nível de significância de  $\alpha = 10\%$ ,  $H_0^A$  é rejeitada, mas  $H_0^G$  não. Isto é, conclui-se que as proporções genóticas são consideradas as mesmas nos dois grupos, enquanto que as alélicas não. Contudo, se  $\theta$  é tal que vale  $H_0^G$ , então  $\gamma_i = \pi_i, \forall i \in \mathbb{G}$ . Desta forma,  $\theta$  é tal que  $\gamma_{AA} + \frac{1}{2}\gamma_{AB} = \pi_{AA} + \frac{1}{2}\pi_{AB}$  e, portanto, vale  $H_0^A$ . Em palavras, se as proporções genóticas são as mesmas nos dois grupos, então as proporções alélicas também devem ser as mesmas.

□

O exemplo anterior mostra uma situação em que aceitamos uma hipótese mas rejeitamos outra hipótese que é implicada pela primeira. Uma propriedade que poderia se esperar que testes de hipóteses satisfizessem é justamente a oposta: se uma hipótese é aceita, então outra hipótese que a contém também deve ser aceita. Tal propriedade foi estudada por alguns autores como Gabriel [7], Lehmann [10] e [11], Schervish [19], Lavine e Schervish [9], Rom e Holland [16] e Betz e Levin [1] (os dois últimos artigos a estudam especificamente no caso da ANOVA). Nestes trabalhos, esta propriedade é chamada de *coerência*. No presente estudo, preferimos designá-la por *monotonicidade*, como em Fossaluza [5] e Silva [18], uma vez que avaliamos outras propriedades que também poderiam ser chamadas de *coerência*.

**Definição 2.1. (Monotonicidade)** *Uma classe de testes  $\mathcal{L}$  é monótona (isto é, atende à monotonicidade) se, e somente se,  $\forall A, B \in \sigma(\Theta), A \subseteq B \Rightarrow \mathcal{L}(A) \geq \mathcal{L}(B)$ . Em palavras, se, após a observação  $x$ , uma hipótese é aceita (isto é,  $\mathcal{L}(A)(x) = 0$ ), qualquer hipótese que é implicada por ela também deve ser aceita (isto é,  $\mathcal{L}(B)(x) = 0$ ).*

**Exemplo 2.2. (Classe baseada em Testes da Razão de Verossimilhanças Generalizada)**

Seja  $c \in [0, 1]$  e consideremos a classe de testes definida por

$$\mathcal{L}(A)(x) = \mathbb{I} \left( \frac{\sup_{\theta \in A} V_x(\theta)}{\sup_{\theta \in \Theta} V_x(\theta)} \leq c \right), \forall A \in \sigma(\Theta) \text{ e } \forall x \in \mathcal{X}.$$

Temos que esta classe é monótona, uma vez que se  $A, B \in \sigma(\Theta)$  são tais que  $A \subseteq B$  e  $x \in \mathcal{X}$ , então

$$\mathcal{L}(A)(x) = 0 \Rightarrow \frac{\sup_{\theta \in A} V_x(\theta)}{\sup_{\theta \in \Theta} V_x(\theta)} > c \Rightarrow \frac{\sup_{\theta \in B} V_x(\theta)}{\sup_{\theta \in \Theta} V_x(\theta)} > c \Rightarrow \mathcal{L}(B)(x) = 0.$$

Nota-se que, ao contrário do Exemplo 1.1, aqui os testes designados para as diversas hipóteses de  $\sigma(\Theta)$  podem, em geral, ter níveis de significância diferentes, uma vez que o valor  $c$  é o mesmo para todos os testes. Gabriel [7] estuda algumas propriedades que testes desta forma possuem. Lá, contudo, o conjunto de hipóteses de interesse não é uma  $\sigma$ -álgebra do espaço paramétrico. Porém,

supõe-se que a hipótese formada pela intersecção de todas as hipóteses de interesse é não vazia. Uma propriedade estudada, por exemplo, refere-se à probabilidade de alguma falsa rejeição quando a hipótese formada pela intersecção de todas as outras é verdadeira.

□

**Exemplo 2.3. (Classe de testes FBST)** Suponhamos que  $\Theta = \mathfrak{R}^n$ ,  $\sigma(\Theta) = \mathcal{B}(\mathfrak{R}^n)$ , e que temos fixada uma distribuição *a priori* para  $\theta$ . Suponhamos que, para todo  $x \in \chi$ , existe  $f(\theta|x)$ , a função de densidade da distribuição *a posteriori* de  $\theta$ , dado  $x$ . Para cada hipótese  $A \in \sigma(\Theta)$ , seja

$$T_x^A = \{\theta \in \Theta : f(\theta|x) > \sup_A f(\theta|x)\}$$

o conjunto tangente à hipótese nula e seja

$$ev_x(A) = 1 - P(\theta \in T_x^A|x)$$

o valor de evidência de Pereira-Stern (Pereira e Stern [15]) para a hipótese  $A$ . Para  $A = \emptyset$ , definimos  $T_x^A = \Theta$ .

Podemos definir uma classe de testes  $\mathcal{L}$  através de

$$\mathcal{L}(A)(x) = \mathbb{I}(ev_x(A) \leq c), \forall A \in \sigma(\Theta) \text{ e } \forall x \in \chi,$$

em que  $c \in [0, 1]$  está fixo. Em palavras, aceitamos uma hipótese sempre que sua evidência for maior que  $c$ . Tal classe atende à monotonicidade. De fato, sejam  $A, B \in \sigma(\Theta)$  tais que  $A \subseteq B$  e  $x \in \chi$  tal que  $\mathcal{L}(A)(x) = 0$ . Temos que  $\sup_B f(\theta|x) \geq \sup_A f(\theta|x)$ . Assim,  $T_x^B \subseteq T_x^A$ , de modo que  $ev_x(A) \leq ev_x(B)$ . Assim,

$$\mathcal{L}(A)(x) = 0 \Rightarrow ev_x(A) > c \Rightarrow ev_x(B) > c \Rightarrow \mathcal{L}(B)(x) = 0.$$

Silva [18] faz um estudo mais aprofundado relacionando o FBST e a monotonicidade: lá é considerado não apenas um único valor  $c$  para todos os testes, mas sim diferentes valores, que são obtidos através de uma família de funções de perda que dão origem a esta classe.

□

Classes monótonas podem ser caracterizadas de diversas maneiras. O Teorema a seguir estabelece algumas destas caracterizações.

**Teorema 2.1.** *São equivalentes:*

1.  $\mathcal{L}$  é monótona
2.  $\forall A, B \in \sigma(\Theta)$ ,  $\mathcal{L}(A \cup B) \leq \mathcal{L}(A)\mathcal{L}(B) = \min\{\mathcal{L}(A), \mathcal{L}(B)\}$

3.  $\forall A, B \in \sigma(\Theta)$ ,  $\mathcal{L}(A \cap B) \geq 1 - (1 - \mathcal{L}(A))(1 - \mathcal{L}(B)) = \max\{\mathcal{L}(A), \mathcal{L}(B)\}$
4.  $\forall \{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ , tal que  $\cup_{i \in I} A_i \in \sigma(\Theta)$ ,  $\mathcal{L}(\cup_{i \in I} A_i) \leq \min\{\mathcal{L}(A_i)\}_{i \in I}$
5.  $\forall \{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ , tal que  $\cap_{i \in I} A_i \in \sigma(\Theta)$ ,  $\mathcal{L}(\cap_{i \in I} A_i) \geq \max\{\mathcal{L}(A_i)\}_{i \in I}$
6.  $\forall \{A_1, \dots, A_n, A'_1, \dots, A'_{n'}\} \subseteq \sigma(\Theta)$ , tal que  $A_i \cap A_j = A'_{i'} \cap A'_{j'} = \emptyset$ , para  $i \neq j$  e  $i' \neq j'$ ,  $\cup_{i=1}^n A_i = \cup_{i=1}^{n'} A'_{i'} \equiv A^*$  e  $\forall i \in \{1, \dots, n\}$ ,  $A_i = \bigcup_{j \in I} A'_j$  para algum  $I \subseteq \{1, \dots, n'\}$  (isto é,  $\{A'_1, \dots, A'_{n'}\}$  é uma partição de  $A^*$  mais fina que  $\{A_1, \dots, A_n\}$ ),

$$\sum_{i=1}^{n'} (1 - \mathcal{L}(A'_{i'})(x)) \geq 1 \Rightarrow \sum_{i=1}^n (1 - \mathcal{L}(A_i)(x)) \geq 1, \forall x \in \chi,$$

isto é, para toda partição finita mensurável de todo subconjunto  $A^* \in \sigma(\Theta)$ , aceitar ao menos um de seus elementos implica aceitar ao menos um elemento de qualquer partição menos fina de  $A^*$  que a considerada.

*Demonstração.* Provaremos aqui apenas algumas das implicações, uma vez que as outras tem prova análoga.

$1 \Rightarrow 4$ . Seja  $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$  tal que  $\cup_{i \in I} A_i \in \sigma(\Theta)$  e seja  $x \in \chi$ . Se  $\min_{i \in I} \mathcal{L}(A_i)(x) = 0$ , existe  $i \in I$  tal que  $\mathcal{L}(A_i)(x) = 0$ . Como  $A_i \subseteq \cup_{i \in I} A_i$ , temos, por 1., que  $\mathcal{L}(\cup_{i \in I} A_i)(x) = 0$ . Caso  $\min_{i \in I} \mathcal{L}(A_i)(x) = 1$ , evidentemente  $\mathcal{L}(\cup_{i \in I} A_i)(x) \leq \min_{i \in I} \mathcal{L}(A_i)(x)$ .

$4 \Rightarrow 1$ . Sejam  $A, B \in \sigma(\Theta)$ , com  $A \subseteq B$  e seja  $x \in \chi$ . Definimos  $A_1 = A$  e  $A_2 = B$ . Suponhamos que  $\mathcal{L}(A)(x) = 0$ . Seja  $I = \{1, 2\}$ . Temos, por 4., que  $\mathcal{L}(B)(x) = \mathcal{L}(A \cup B)(x) = \min\{\mathcal{L}(A)(x), \mathcal{L}(B)(x)\} = 0$ . Assim,  $\mathcal{L}(A)(x) \geq \mathcal{L}(B)(x)$ . Se  $x$  é tal que  $\mathcal{L}(A)(x) = 1$ , evidentemente  $\mathcal{L}(A)(x) \geq \mathcal{L}(B)(x)$ . Desta forma,  $\mathcal{L}(A) \geq \mathcal{L}(B)$ .

$1 \Rightarrow 5$ . Seja  $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$  tal que  $\cap_{i \in I} A_i \in \sigma(\Theta)$  e seja  $x \in \chi$ . Suponhamos que  $\max_{i \in I} \mathcal{L}(A_i)(x) = 1$ . Assim, para algum  $i \in I$ ,  $\mathcal{L}(A_i)(x) = 1$ . Como  $\cap_{i \in I} A_i \subseteq A_i$ , temos, por 1., que  $\mathcal{L}(\cap_{i \in I} A_i)(x) = 1$ . Se  $\max_{i \in I} \mathcal{L}(A_i)(x) = 0$ , evidentemente  $\mathcal{L}(\cap_{i \in I} A_i)(x) \geq \max_{i \in I} \mathcal{L}(A_i)(x)$ .

$5 \Rightarrow 1$ . Sejam  $A, B \in \sigma(\Theta)$ , com  $A \subseteq B$  e seja  $x \in \chi$ . Seja  $I = \{1, 2\}$ . Suponhamos que  $\mathcal{L}(B)(x) = 1$ . Definimos  $A_1 = A$  e  $A_2 = B$ . Temos por, 5., que  $\mathcal{L}(A)(x) = \mathcal{L}(A \cap B)(x) \geq \max\{\mathcal{L}(A)(x), \mathcal{L}(B)(x)\} = 1$ . Assim,  $\mathcal{L}(A)(x) \geq \mathcal{L}(B)(x)$ . Se  $x$  é tal que  $\mathcal{L}(B)(x) = 0$ , evidentemente  $\mathcal{L}(A)(x) \geq \mathcal{L}(B)(x)$ . Desta forma,  $\mathcal{L}(A) \geq \mathcal{L}(B)$ .

$2 \Rightarrow 6$ . Seja  $\{A_1, \dots, A_n, A'_1, \dots, A'_{n'}\} \subseteq \sigma(\Theta)$  como em 6 e seja  $x \in \chi$ . Suponhamos que  $\sum_{i=1}^{n'} (1 - \mathcal{L}(A'_{i'})(x)) \geq 1$ . Então, existe  $j$  tal que  $\mathcal{L}(A'_j)(x) = 0$ . Como  $\{A'_1, \dots, A'_{n'}\}$  é mais fina que  $\{A_1, \dots, A_n\}$ , então  $\exists i \in \{1, \dots, n\}$  e  $\exists I \subseteq \{1, \dots, n'\}$  tal que  $j \in I$  e  $A_i = \bigcup_{j \in I} A'_j$ . Assim, por 2., como  $\mathcal{L}(A'_j)(x) = 0$ , temos que  $\mathcal{L}(A_i)(x) = 0$ . Desta forma,  $\sum_{i=1}^n (1 - \mathcal{L}(A_i)(x)) \geq 1$ .

$6 \Rightarrow 2$ . Sejam  $A, B \in \sigma(\Theta)$ . Seja  $x \in \chi$  tal que  $\mathcal{L}(A)(x) = 0$  (sem perda de generalidade; o



mesmo ocorre se  $\mathcal{L}(B)(x) = 0$ . Temos então

$$(1 - \mathcal{L}(A)(x)) + (1 - \mathcal{L}(B \setminus A)(x)) \geq 1,$$

de modo que, pela propriedade 6, considerando as partições de  $A \cup B$   $\{A, B \setminus A\}$  e  $\{A \cup B\}$ , temos

$$1 - \mathcal{L}(A \cup B)(x) \geq 1,$$

e, portanto,  $\mathcal{L}(A \cup B)(x) = 0$ . O caso  $\mathcal{L}(A)(x) = \mathcal{L}(B)(x) = 1$  é imediato. Assim, vale 2. □

*Observação:* Não é difícil ver que no Teorema 2.1, as propriedades 2. e 3. podem ser trocadas pelos respectivos casos finitos:  $\forall n \geq 1$  e  $\forall \{A_1, \dots, A_n\} \subseteq \sigma(\Theta)$ ,  $\mathcal{L}(\cup_{i=1}^n A_i) \leq \prod_{i=1}^n \mathcal{L}(A_i)$  e  $\forall \{A_1, \dots, A_n\} \subseteq \sigma(\Theta)$ ,  $\mathcal{L}(\cap_{i=1}^n A_i) \geq 1 - \prod_{i=1}^n (1 - \mathcal{L}(A_i))$ .

Com a finalidade de estudar monotonicidade sob um ponto de vista de Teoria da Decisão, iremos considerar algumas propriedades para famílias de funções de perda. Seja  $(L_A)_{A \in \sigma(\Theta)}$  uma família de funções de perda, onde  $L_A$  é a perda associada à hipótese  $A$ ,  $A \in \sigma(\Theta)$ . Neste trabalho, as duas propriedades a seguir são de interesse:

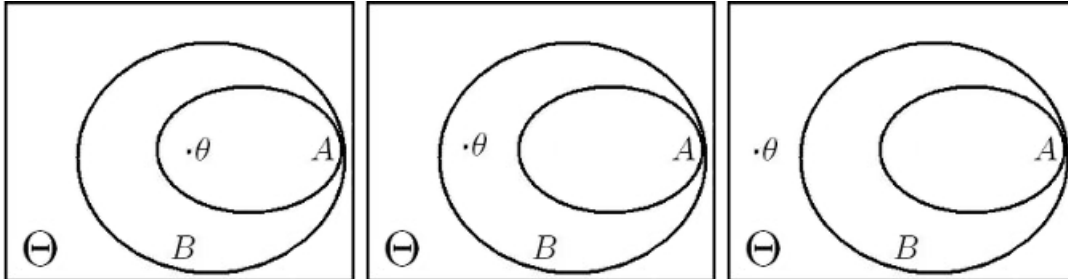
**L1**  $\forall A, B \in \sigma(\Theta)$  com  $A \subseteq B$  e  $\forall \theta \in \Theta$ ,  $L_A(0, \theta) - L_A(1, \theta) \geq L_B(0, \theta) - L_B(1, \theta)$

**L2**  $\forall A \in \sigma(\Theta)$ ,  $\theta \in A \Rightarrow L_A(0, \theta) \leq L_A(1, \theta)$  e  $\theta \in A^c \Rightarrow L_A(0, \theta) \geq L_A(1, \theta)$

### Interpretação:

**L1** Esta propriedade será analisada em três casos distintos, como mostrado na Figura ???. Chamando-se de *perda relativa* a diferença entre as perdas nas duas possíveis decisões (aceitar e rejeitar a hipótese), temos que:

Figura 2.1: Interpretação da propriedade L1



- Se  $\theta \in A$  (primeiro quadro da Figura (??)), acertamos quando decidimos por  $A$  e por  $B$ . A suposição diz que a perda relativa de aceitar  $A$  deve ser maior ou igual à perda relativa de aceitar  $B$ . Como  $B$  contém  $A$ , se  $\theta$  está em  $A$ ,  $\theta$  também está em  $B$ . Assim,

quando aceitamos  $B$ , parece razoável perder no máximo o mesmo que quando aceitamos  $A$  (há “mais elementos” em  $B$  que em  $A$ ).

- Se  $\theta \in B \setminus A$  (segundo quadro da Figura (??)), acertamos quando decidimos por  $A^c$  e por  $B$ . Assim, parece razoável que  $L_A(0, \theta) \geq L_A(1, \theta)$  e  $L_B(1, \theta) \geq L_B(0, \theta)$  (na realidade, estas desigualdades são consequência da propriedade **L2**). Mas isto implica  $L_A(0, \theta) - L_A(1, \theta) \geq 0 \geq L_B(0, \theta) - L_B(1, \theta)$ .
- Se  $\theta \in B^c$  (terceiro quadro da Figura (??)), acertamos quando decidimos por  $A^c$  e por  $B^c$ . A desigualdade da propriedade é equivalente a  $L_B(1, \theta) - L_B(0, \theta) \geq L_A(1, \theta) - L_A(0, \theta)$ . A suposição diz que a perda relativa de rejeitar  $B$  deve ser maior ou igual à perda relativa de rejeitar  $A$ . A interpretação é análoga ao primeiro caso.

**L2** Ao se tomar uma decisão correta, perde-se o mesmo ou menos que ao se tomar uma decisão incorreta.

**Exemplo 2.4.** As seguintes famílias de perdas satisfazem **L1** e **L2**:

- Perdas da forma da Tabela 2.2, com  $a_A, b_A \in \mathfrak{R}$ ,  $\forall A \in \sigma(\Theta)$ , e com as restrições  $a_A = b_{A^c}$ ,  $\forall A \in \sigma(\Theta)$  (isto é, aceitar  $A$  tem a mesma perda de rejeitar  $A^c$ ) e  $a_A \geq a_B \geq 0$ ,  $\forall A, B \in \sigma(\Theta)$  com  $A \subseteq B$ . Neste caso, vale que  $0 \leq b_A \leq b_B$ .

Tabela 2.2: Exemplo de função de perda

Decisão	Estado da natureza	
	$\theta \in A$	$\theta \in A^c$
0	0	$a_A$
1	$b_A$	0

Em particular, podemos tomar  $a_A = \lambda(A^c)$  e  $b_A = \lambda(A)$ , em que  $\lambda$  é uma medida finita em  $\sigma(\Theta)$ , com  $\lambda(\Theta) > 0$ .

- $L_A(0, \theta) = f(d(\theta, A))$  e  $L_A(1, \theta) = f(d(\theta, A^c))$ , em que  $d(\theta, A)$  é uma distância entre  $\theta$  e  $A$  e  $f$  é uma função não decrescente.

□

O Teorema a seguir estabelece algumas relações entre as propriedades **L1** e **L2** e classes monótonas.

**Teorema 2.2.** *Seja  $(L_A)_{A \in \sigma(\Theta)}$  uma família de funções de perda e  $\mathcal{L}$  uma classe bayesiana gerada por essa família. Suponhamos que  $\forall A \in \sigma(\Theta)$  e  $\forall x \in \mathcal{X}$ ,  $|\mathbb{E}[L_A(0, \theta)|x]| < \infty$  e  $|\mathbb{E}[L_A(1, \theta)|x]| < \infty$ . Então:*

1. *Se  $(L_A)_{A \in \sigma(\Theta)}$  atende a **L1**, então  $\mathcal{L}$  é monótona, qualquer que seja a distribuição a priori para  $\theta$ .*

2. Se  $(L_A)_{A \in \sigma(\Theta)}$  atende a **L2**, mas  $\exists A, B \in \sigma(\Theta)$ , com  $A \subset B$ , e  $\exists \theta_1 \in A$  e  $\theta_2 \in B^c$ , com  $\{\theta_1\}, \{\theta_2\} \in \sigma(\Theta)$ , tais que  $L_A(0, \theta) - L_A(1, \theta) < L_B(0, \theta) - L_B(1, \theta)$ ,  $\theta \in \{\theta_1, \theta_2\}$  (e, portanto, não vale **L1**), então existe uma distribuição a priori para a qual  $\mathcal{L}$  não é monótona. Isto é, se **L1** não vale, mas **L2** sim,  $\mathcal{L}$  pode não apresentar monotonicidade.

*Demonstração.* 1. Suponhamos que vale **L1**. Sejam  $A, B \in \sigma(\Theta)$  com  $A \subseteq B$  e  $x \in \chi$  tal que  $\mathcal{L}(A)(x) = 0$ . Seja  $\mu_x$  a medida de probabilidade a posteriori de  $\theta$ , dado  $x$ . Temos então

$$\begin{aligned} \mathcal{L}(A)(x) = 0 &\Rightarrow \int_{\Theta} L_A(0, \theta) d\mu_x(\theta) < \int_{\Theta} L_A(1, \theta) d\mu_x(\theta) \Rightarrow \int_{\Theta} L_A(0, \theta) - L_A(1, \theta) d\mu_x(\theta) < 0 \Rightarrow \\ &\Rightarrow \int_{\Theta} L_B(0, \theta) - L_B(1, \theta) d\mu_x(\theta) < 0 \Rightarrow \int_{\Theta} L_B(0, \theta) d\mu_x(\theta) < \int_{\Theta} L_B(1, \theta) d\mu_x(\theta) \Rightarrow \mathcal{L}(B)(x) = 0. \end{aligned}$$

2. Suponhamos que vale **L2** mas que  $\exists A, B \in \sigma(\Theta)$ , com  $A \subset B$ , tais que  $\exists \theta_1 \in A$  e  $\theta_2 \in B^c$  com  $L_A(0, \theta) - L_A(1, \theta) < L_B(0, \theta) - L_B(1, \theta)$ ,  $\theta \in \{\theta_1, \theta_2\}$ . Como vale **L2**, temos

$$L_A(0, \theta_1) - L_A(1, \theta_1) < L_B(0, \theta_1) - L_B(1, \theta_1) \leq 0 \leq L_A(0, \theta_2) - L_A(1, \theta_2) < L_B(0, \theta_2) - L_B(1, \theta_2).$$

Consideremos que  $x$  é observado e que, coincidentemente, a opinião do agente decisor sobre  $\theta$  é dada por

$$\mathbb{P}(\{\theta_1\}) = \frac{pV_x(\theta_2)}{pV_x(\theta_2) + (1-p)V_x(\theta_1)}; \quad \mathbb{P}(\{\theta_2\}) = 1 - \mathbb{P}(\{\theta_1\}),$$

de modo que a distribuição a posteriori de  $\theta$  dado  $x$ ,  $\mu_x$ , é dada por  $\mu_x(\{\theta_1\}) = p$  e  $\mu_x(\{\theta_2\}) = 1 - p$ ,  $0 < p < 1$ .

A prova será dividida em quatro casos:

- Se  $L_B(0, \theta_1) - L_B(1, \theta_1) < 0 < L_A(0, \theta_2) - L_A(1, \theta_2)$ , das hipóteses temos que vale:

$$L_A(0, \theta_1) - L_A(1, \theta_1) < L_B(0, \theta_1) - L_B(1, \theta_1) < 0 < L_A(0, \theta_2) - L_A(1, \theta_2) < L_B(0, \theta_2) - L_B(1, \theta_2).$$

Assim,

$$x_1 \equiv \frac{|L_A(0, \theta_2) - L_A(1, \theta_2)|}{|L_A(0, \theta_1) - L_A(1, \theta_1)|} < \frac{|L_B(0, \theta_2) - L_B(1, \theta_2)|}{|L_B(0, \theta_1) - L_B(1, \theta_1)|} \equiv x_2.$$

Se  $p \in (0, 1)$  é tal que  $x_1 < \frac{p}{1-p} < x_2$ , temos

$$\begin{aligned} \int_{\Theta} L_A(0, \theta) - L_A(1, \theta) d\mu_x(\theta) &= [L_A(0, \theta_1) - L_A(1, \theta_1)]p + [L_A(0, \theta_2) - L_A(1, \theta_2)](1-p) < 0 < \\ &< [L_B(0, \theta_1) - L_B(1, \theta_1)]p + [L_B(0, \theta_2) - L_B(1, \theta_2)](1-p) = \int_{\Theta} L_B(0, \theta) - L_B(1, \theta) d\mu_x(\theta). \end{aligned}$$

Assim, temos  $\mathcal{L}(A)(x) = 0$ , mas  $\mathcal{L}(B)(x) = 1$ , e, portanto, existe uma distribuição a priori para a qual  $\mathcal{L}$  não é monótona.

- Se  $L_B(0, \theta_1) - L_B(1, \theta_1) = 0 < L_A(0, \theta_2) - L_A(1, \theta_2)$ , das hipóteses temos que vale:

$$L_A(0, \theta_1) - L_A(1, \theta_1) < L_B(0, \theta_1) - L_B(1, \theta_1) = 0 < L_A(0, \theta_2) - L_A(1, \theta_2) < L_B(0, \theta_2) - L_B(1, \theta_2).$$

Se  $p < 1$  é tal que

$$\frac{|L_A(0, \theta_2) - L_A(1, \theta_2)|}{|L_A(0, \theta_1) - L_A(1, \theta_1)|} < \frac{p}{1-p},$$

temos

$$\begin{aligned} \int_{\Theta} L_A(0, \theta) - L_A(1, \theta) d\mu_x(\theta) &= [L_A(0, \theta_1) - L_A(1, \theta_1)]p + [L_A(0, \theta_2) - L_A(1, \theta_2)](1-p) < 0 < \\ &< [L_B(0, \theta_1) - L_B(1, \theta_1)]p + [L_B(0, \theta_2) - L_B(1, \theta_2)](1-p) = \int_{\Theta} L_B(0, \theta) - L_B(1, \theta) d\mu_x(\theta). \end{aligned}$$

Assim, temos  $\mathcal{L}(A)(x) = 0$ , mas  $\mathcal{L}(B)(x) = 1$ , e, novamente, existe uma distribuição *a priori* para a qual  $\mathcal{L}$  não é monótona.

- Se  $L_B(0, \theta_1) - L_B(1, \theta_1) < 0 = L_A(0, \theta_2) - L_A(1, \theta_2)$ , das hipóteses temos que vale:

$$L_A(0, \theta_1) - L_A(1, \theta_1) < L_B(0, \theta_1) - L_B(1, \theta_1) < 0 = L_A(0, \theta_2) - L_A(1, \theta_2) < L_B(0, \theta_2) - L_B(1, \theta_2).$$

Se  $p > 0$  é tal que

$$\frac{p}{1-p} < \frac{|L_B(0, \theta_2) - L_B(1, \theta_2)|}{|L_B(0, \theta_1) - L_B(1, \theta_1)|},$$

temos

$$\begin{aligned} \int_{\Theta} L_A(0, \theta) - L_A(1, \theta) d\mu_x(\theta) &= [L_A(0, \theta_1) - L_A(1, \theta_1)]p + [L_A(0, \theta_2) - L_A(1, \theta_2)](1-p) < 0 < \\ &< [L_B(0, \theta_1) - L_B(1, \theta_1)]p + [L_B(0, \theta_2) - L_B(1, \theta_2)](1-p) = \int_{\Theta} L_B(0, \theta) - L_B(1, \theta) d\mu_x(\theta). \end{aligned}$$

Assim, temos  $\mathcal{L}(A)(x) = 0$ , mas  $\mathcal{L}(B)(x) = 1$ , e, novamente, existe uma distribuição *a priori* para a qual  $\mathcal{L}$  não é monótona.

- Se  $L_B(0, \theta_1) - L_B(1, \theta_1) = 0 = L_A(0, \theta_2) - L_A(1, \theta_2)$ , das hipóteses temos que vale:

$$L_A(0, \theta_1) - L_A(1, \theta_1) < L_B(0, \theta_1) - L_B(1, \theta_1) = 0 = L_A(0, \theta_2) - L_A(1, \theta_2) < L_B(0, \theta_2) - L_B(1, \theta_2).$$

Para todo  $p \in (0, 1)$ , temos

$$\begin{aligned} \int_{\Theta} L_A(0, \theta) - L_A(1, \theta) d\mu_x(\theta) &= [L_A(0, \theta_1) - L_A(1, \theta_1)]p + [L_A(0, \theta_2) - L_A(1, \theta_2)](1-p) < 0 < \\ &< [L_B(0, \theta_1) - L_B(1, \theta_1)]p + [L_B(0, \theta_2) - L_B(1, \theta_2)](1-p) = \int_{\Theta} L_B(0, \theta) - L_B(1, \theta) d\mu_x(\theta). \end{aligned}$$

Assim, temos  $\mathcal{L}(A)(x) = 0$ , mas  $\mathcal{L}(B)(x) = 1$ , e, novamente, existe uma distribuição *a priori* para a qual  $\mathcal{L}$  não é monótona.

□

### 2.1.1 Monotonicidade e Admissibilidade

Sob o paradigma clássico, a noção de admissibilidade, definida a seguir, desempenha um papel fundamental.

**Definição 2.2. (Admissibilidade)** *Seja  $\phi$  uma regra de decisão e  $L$  uma função de perda. A regra  $\phi$  é inadmissível se existe uma regra  $\phi'$  tal que*

$$R(\theta, \phi') \leq R(\theta, \phi), \quad \forall \theta \in \Theta$$

e

$$\exists \theta_0 \in \Theta \text{ tal que } R(\theta_0, \phi') < R(\theta_0, \phi),$$

onde  $R(\theta, \phi) = \mathbb{E}[L(\theta, \phi(X))|\theta]$  é o risco da regra  $\phi$  no ponto  $\theta$ . Se  $\phi$  não é inadmissível, dizemos que  $\phi$  é admissível.

O seguinte exemplo mostra que a admissibilidade de dois testes não implica e nem é implicada pela monotonicidade.

**Exemplo 2.5.** Seja  $\Theta = \{\frac{1}{3}, \frac{1}{2}, \frac{2}{3}\}$  e  $X|\theta \sim \text{Bernoulli}(\theta)$ . Consideremos as hipóteses  $A = \{\frac{1}{3}\}$  e  $B = \{\frac{1}{3}, \frac{1}{2}\}$ , de modo que  $A \subseteq B$ . Considere que a função de perda utilizada para testar  $A$  é dada por  $L_A(0, \theta) = \mathbb{I}(\theta \notin A)$  e  $L_A(1, \theta) = \mathbb{I}(\theta \in A)$ , enquanto que a função de perda para testar  $B$  é  $L_B(0, \theta) = \mathbb{I}(\theta \notin B)$  e  $L_B(1, \theta) = \mathbb{I}(\theta \in B)$ . Em palavras, perde-se uma unidade em caso de erro e não se perde nada em caso de acerto. Para estas funções de perda, o risco nada mais é que a probabilidade de erro. As Tabelas (2.3) e (2.4) mostram, respectivamente, os possíveis testes para  $A$  e  $B$ , bem como seus riscos.

Tabela 2.3: Testes e riscos associados à hipótese  $A$

$R(\theta, \phi)$	Região de aceitação de $\phi$			
$\theta$	$\emptyset$	$\{0\}$	$\{1\}$	$\{0, 1\}$
1/3	1	1/3	2/3	0
1/2	0	1/2	1/2	1
2/3	0	1/3	2/3	1

Vemos, a partir destas tabelas, que o teste que aceita  $A$  em  $\{0, 1\}$  é admissível para testar  $A$ , bem como o teste que aceita  $B$  em  $\{0\}$  é admissível para testar  $B$ . Contudo, quando utilizamos

Tabela 2.4: Testes e riscos associados à hipótese  $B$ 

$R(\theta, \phi)$	Região de aceitação de $\phi$			
$\theta$	$\emptyset$	$\{0\}$	$\{1\}$	$\{0, 1\}$
1/3	1	1/3	2/3	0
1/2	1	1/2	1/2	0
2/3	0	1/3	2/3	1

estes testes e observamos  $X = 1$ , aceitamos  $A$  e rejeitamos  $B$ , de modo que não há monotonicidade. Assim, a admissibilidade de cada um dos testes não implica monotonicidade.

Além disso, suponhamos que usamos o teste que aceita  $A$  em  $\{1\}$  e o teste que aceita  $B$  em  $\{1\}$ . É fácil ver que estes testes são inadmissíveis (eles são dominados pelo teste que aceita  $A$  em  $\{0\}$  e pelo teste que aceita  $B$  em  $\{0\}$ , respectivamente), contudo sempre que  $A$  é aceita,  $B$  também é, de modo que há monotonicidade.

□

Embora, neste trabalho, o estudo de admissibilidade de cada um dos testes não está relacionado à monotonicidade, Lavine e Schervish [9] relacionam estes dois conceitos levando-se em conta uma única perda (conjunta) dos dois testes simultaneamente. A seguir, temos uma versão do resultado apresentado por Lavine e Schervish [9] para famílias de funções de perda atendendo **L1** e **L2**, que são generalizações das famílias apresentadas na versão original destes autores. Nela, as desigualdades de **L1** eram igualdades.

**Teorema 2.3.** *Sejam  $A$  e  $B$  duas hipóteses, com  $A \subset B$ . Suponhamos que as funções de perda para cada uma destas hipóteses,  $L_A$  e  $L_B$ , respectivamente, satisfaçam **L1** e **L2**. Suponhamos que, ao testar  $A$  e  $B$  simultaneamente, as perdas do teste simultâneo possam ser compostas da seguinte forma:*

$$L_{(A,B)} : \{0, 1\}^2 \times \Theta \rightarrow \mathfrak{R}+$$

$$((d_1, d_2), \theta) \rightarrow L_{(A,B)}((d_1, d_2), \theta) = L_A(d_1, \theta) + L_B(d_2, \theta)$$

*Isto é, a perda quando se testa  $A$  e  $B$  simultaneamente é a soma das perdas individuais de cada teste. Seja  $\phi$  a regra de decisão (simultânea) para  $(A, B)$  que aceita  $A$  ( $B$ ) em  $\chi_A$  ( $\chi_B$ ). Então, se existe  $\theta \in B \cap A^c$  tal que  $\mathbb{P}(\chi_A \cap \chi_B^c | \theta) > 0$ , e  $L_A(0, \theta) > L_A(1, \theta)$  ou  $L_B(1, \theta) > L_B(0, \theta)$  (isto é, para ao menos um destes casos, a desigualdade em **L2** é estrita), a regra  $\phi$  é inadmissível. Isto é, se não há monotonicidade, não há admissibilidade.*

*Demonstração.* Consideremos uma nova regra de decisão simultânea  $\psi$  para  $(A, B)$  que aceita  $A$

( $B$ ) em  $\chi_A \cap \chi_B$  ( $\chi_A \cup \chi_B$ ). Temos que  $\psi$  domina  $\phi$ . De fato, denotando por  $R(\theta, \phi)$  o risco em  $\theta$  da decisão  $\phi$ , temos que

$$\begin{aligned}
R(\theta, \phi) - R(\theta, \psi) &= \\
&= L_{(A,B)}((0, 1), \theta)\mathbb{P}(\chi_A \cap \chi_B^c | \theta) + L_{(A,B)}((0, 0), \theta)\mathbb{P}(\chi_A \cap \chi_B | \theta) \\
&+ L_{(A,B)}((1, 1), \theta)\mathbb{P}(\chi_A^c \cap \chi_B^c | \theta) + L_{(A,B)}((1, 0), \theta)\mathbb{P}(\chi_A^c \cap \chi_B | \theta) \\
&- L_{(A,B)}((0, 1), \theta)\mathbb{P}((\chi_A \cap \chi_B) \cap (\chi_A \cup \chi_B)^c | \theta) \\
&- L_{(A,B)}((0, 0), \theta)\mathbb{P}((\chi_A \cap \chi_B) \cap (\chi_A \cup \chi_B) | \theta) \\
&- L_{(A,B)}((1, 1), \theta)\mathbb{P}((\chi_A \cap \chi_B)^c \cap (\chi_A \cup \chi_B)^c | \theta) \\
&- L_{(A,B)}((1, 0), \theta)\mathbb{P}((\chi_A \cap \chi_B)^c \cap (\chi_A \cup \chi_B) | \theta) = \\
&= L_{(A,B)}((0, 1), \theta)\mathbb{P}(\chi_A \cap \chi_B^c | \theta) + L_{(A,B)}((0, 0), \theta)\mathbb{P}(\chi_A \cap \chi_B | \theta) \\
&+ L_{(A,B)}((1, 1), \theta)\mathbb{P}(\chi_A^c \cap \chi_B^c | \theta) + L_{(A,B)}((1, 0), \theta)\mathbb{P}(\chi_A^c \cap \chi_B | \theta) \\
&- L_{(A,B)}((0, 1), \theta)\mathbb{P}(\emptyset | \theta) - L_{(A,B)}((0, 0), \theta)\mathbb{P}(\chi_A \cap \chi_B | \theta) - L_{(A,B)}((1, 1), \theta)\mathbb{P}((\chi_A \cup \chi_B)^c | \theta) \\
&- L_{(A,B)}((1, 0), \theta)\mathbb{P}((\chi_A \cap \chi_B^c) \cup (\chi_A^c \cap \chi_B) | \theta) = \\
&= L_{(A,B)}((0, 1), \theta)\mathbb{P}(\chi_A \cap \chi_B^c | \theta) - L_{(A,B)}((1, 0), \theta)\mathbb{P}(\chi_A \cap \chi_B^c | \theta) = \\
&= [(L_A(0, \theta) - L_A(1, \theta)) + (L_B(1, \theta) - L_B(0, \theta))]\mathbb{P}(\chi_A \cap \chi_B^c | \theta) = (*)
\end{aligned}$$

Por **L1**, vemos que  $(*) \geq 0, \forall \theta \in \Theta$ . Além disso, seja  $\theta \in B \cap A^c$  tal que  $\mathbb{P}(\chi_A \cap \chi_B^c | \theta) > 0$ . Por hipótese, temos que vale **L2** e  $L_A(0, \theta) > L_A(1, \theta)$  ou  $L_B(1, \theta) > L_B(0, \theta)$ . Deste modo, vale que  $L_A(0, \theta) - L_B(0, \theta) + L_B(1, \theta) - L_A(1, \theta) > 0$  e, portanto,  $(*) > 0$ . Assim,  $\psi$  domina  $\phi$  e esta é, portanto, inadmissível. □

## 2.2 Consonância da União

**Exemplo 2.6.** Suponhamos que, em uma eleição majoritária em dois turnos (Nicolau [13]), há três candidatos, cujas respectivas proporções de eleitores são dadas por  $\theta_1, \theta_2$  e  $\theta_3$ . Simplificadamente, desconsideraremos votos nulos e em branco. Estamos interessados nas quatro hipóteses a seguir:

$$\begin{aligned}
H_0^{2^\circ} &: \bigcup_{i=1}^3 \{\theta_i > \frac{1}{2}\}, \\
H_0^1 &: \{\theta_1 > \frac{1}{2}\}, \\
H_0^2 &: \{\theta_2 > \frac{1}{2}\} \text{ e} \\
H_0^3 &: \{\theta_3 > \frac{1}{2}\}.
\end{aligned}$$

Deste modo,  $H_0^{2^\circ}$  é a hipótese de que não haverá segundo turno na eleição, enquanto que  $H_0^i$  é a hipótese de que o  $i$ -ésimo candidato irá ganhar no primeiro turno,  $i = 1, 2, 3$ .

Suponhamos que, ao entrevistar um conjunto de 410 eleitores, observamos um vetor  $X = (X_1, X_2, X_3)$ , em que  $X_i$  é o número de eleitores na amostra que votam no candidato  $i$ ,  $i = 1, 2, 3$ . Suponhamos  $X|\theta \sim \text{Multinomial}(410, \theta)$ . Suponhamos também que, *a priori*, temos  $(\theta_1, \theta_2, \theta_3) \sim \text{Dirichlet}(1, 1, 1)$ . Se a amostra observada é  $x = (200, 200, 10)$ , teremos, *a posteriori*,  $(\theta_1, \theta_2, \theta_3)|x \sim \text{Dirichlet}(201, 201, 11)$ , de modo que teremos as seguintes probabilidades *a posteriori*:

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i=1}^3 \{\theta_i > \frac{1}{2}\} | x\right) &= 0,588; \quad \mathbb{P}\left(\{\theta_1 > \frac{1}{2}\} | x\right) = 0,294; \\
\mathbb{P}\left(\{\theta_2 > \frac{1}{2}\} | x\right) &= 0,294; \quad \mathbb{P}\left(\{\theta_3 > \frac{1}{2}\} | x\right) = 0,000;
\end{aligned}$$

Se utilizamos a classe de testes descrita no Exemplo (1.2), isto é, aceitamos as hipóteses que tem probabilidade *a posteriori* maior ou igual a 0,5, então aceitamos  $H_0^{2^\circ}$  mas rejeitamos  $H_0^i$ ,  $i = 1, 2, 3$ . Do ponto de vista lógico, temos uma contradição: se não haverá segundo turno, algum dos candidatos deve ter mais votos que os demais somados. Contudo, concluímos que não haverá segundo turno, mesmo concluindo que todos os candidatos tem no máximo 50% dos votos.

□

**Definição 2.3. (Consonância da união finita)** Uma classe de testes  $\mathcal{L}$  atende à consonância da união finita se, e somente se,  $\forall A, B \in \sigma(\Theta)$ ,  $\mathcal{L}(A \cup B) \geq \mathcal{L}(A)\mathcal{L}(B) = \min\{\mathcal{L}(A), \mathcal{L}(B)\}$ . Em palavras, se aceitamos a união de dois conjuntos, devemos aceitar ao menos um dos componentes.

*Nota 1:* A palavra *consonância* aparece em Gabriel [7] para designar uma propriedade parecida, mas em um contexto diferente do atual. Lá o conjunto de hipóteses de interesse não forma, necessariamente, uma  $\sigma$ -álgebra do espaço paramétrico.



*Nota 2:* Não é difícil verificar, por indução, que  $\mathcal{L}$  atende à consonância da união finita se, e somente se,  $\forall n \geq 1$  e  $\forall A_1, \dots, A_n \in \sigma(\Theta)$ ,  $\mathcal{L}(\bigcup_{i=1}^n A_i) \geq \prod_{i=1}^n \mathcal{L}(A_i)$ .

Podemos considerar ainda outras formas de consonância da união:

**Definição 2.4. (Consonância da união enumerável)** Uma classe de testes  $\mathcal{L}$  atende à consonância da união enumerável se, e somente se,  $\forall \{A_n\}_{n \in \mathbb{N}} \subseteq \sigma(\Theta)$ ,  $\mathcal{L}(\bigcup_{n \geq 1} A_n) \geq \prod_{n \geq 1} \mathcal{L}(A_n) \equiv \lim_{n \rightarrow \infty} \prod_{i=1}^n \mathcal{L}(A_i) = \min\{\mathcal{L}(A_n)\}_{n \in \mathbb{N}}$ .

**Definição 2.5. (Consonância da união não enumerável)** Uma classe de testes  $\mathcal{L}$  atende à consonância da união não enumerável se, e somente se,  $\forall \{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ , tal que  $\bigcup_{i \in I} A_i \in \sigma(\Theta)$ ,  $\mathcal{L}(\bigcup_{i \in I} A_i) \geq \min\{\mathcal{L}(A_i)\}_{i \in I}$ , em que  $I$  é um conjunto de índices arbitrário.

**Teorema 2.4.** *Seja  $\mathcal{L}$  uma classe de testes de hipóteses. Então, se  $\mathcal{L}$  satisfaz a consonância da união não enumerável,  $\mathcal{L}$  satisfaz a consonância da união enumerável. Se esta vale, então  $\mathcal{L}$  satisfaz a consonância da união finita. Ademais, não vale a volta de nenhuma destas duas implicações.*

*Demonstração.* As implicações do Teorema são triviais, pois um conjunto finito é um conjunto enumerável e o conjunto  $I$  da Definição 2.5 pode, em particular, ser enumerável.

Para mostrar que  $\mathcal{L}$  satisfazer a consonância da união finita não implica que  $\mathcal{L}$  satisfaz consonância da união enumerável, seja  $\Theta = \mathbb{N}$  e  $\sigma(\Theta) = \mathcal{P}(\Theta)$ . Consideremos uma classe de testes  $\mathcal{L}$  definida por  $\mathcal{L}(A) = I(A \text{ é finito})$ . Não é difícil verificar que  $\mathcal{L}$  respeita a consonância da união finita, mas não a consonância da união enumerável.

De fato, sejam  $A, B \in \sigma(\Theta)$ . Então

$$\begin{aligned} \mathcal{L}(A \cup B) = 0 &\Rightarrow A \cup B \text{ é infinito} \Rightarrow A \text{ é infinito ou } B \text{ é infinito} \Rightarrow \\ \mathcal{L}(A) = 0 \text{ ou } \mathcal{L}(B) &= 0. \end{aligned}$$

Assim, vale a consonância da união finita. Para verificar que não vale a consonância da união enumerável, basta notar que  $\mathcal{L}(\{2, 3, \dots\}) = 0 < 1 = \min_{\theta \geq 2} \mathcal{L}(\{\theta\})$ .

Um contraexemplo análogo mostra que existem classes de testes que satisfazem a consonância da união enumerável, mas não a consonância da união não enumerável.

□

O Teorema a seguir tem a finalidade de caracterizar classes de testes que satisfazem monotonicidade e consonância da união (finita, enumerável ou não enumerável, dependendo da estrutura de  $\Theta$ ). Ele também mostra que, se desejamos criar uma classe de testes que satisfaz essas duas propriedades simultaneamente, basta criar os testes para cada hipótese simples. Isto é, definidos estes testes, há uma forma única de estendê-los para as outras hipóteses da  $\sigma$ -álgebra de modo a obter uma classe de testes que satisfaça as propriedades de interesse.

**Teorema 2.5.** *Seja  $\mathcal{L}$  uma classe de testes construída da seguinte forma: para cada  $\theta \in \Theta$ , está fixado um teste  $\mathcal{L}(\{\theta\})$  (estamos supondo que  $\{\theta\} \in \sigma(\Theta)$ ,  $\forall \theta \in \Theta$ ). Para cada  $A \in \sigma(\Theta)$ , definimos  $\mathcal{L}(A) \equiv \min_{\theta \in A} \mathcal{L}(\{\theta\})$ . Em palavras, o teste para  $A$  é o teste da intersecção-união (Casella e Berger [2]) para  $A$ , com base nos testes para as hipóteses  $\{\theta\}$ ,  $\theta \in A$ . Então*

1.  $\mathcal{L}$  atende à consonância da união não enumerável, bem como à monotonicidade.
2. Seja  $\mathcal{L}'$  uma classe que atende à monotonicidade, com  $\mathcal{L}'(\{\theta\}) = \mathcal{L}(\{\theta\})$ ,  $\forall \theta \in \Theta$ . Temos que:
  - (a) Se  $\Theta$  é finito e  $\mathcal{L}'$  atende à consonância da união finita, então  $\mathcal{L}' = \mathcal{L}$ .
  - (b) Se  $\Theta$  é enumerável e  $\mathcal{L}'$  atende à consonância da união enumerável, então  $\mathcal{L}' = \mathcal{L}$ .
  - (c) Se  $\Theta$  é não enumerável e  $\mathcal{L}'$  atende à consonância da união não enumerável, então  $\mathcal{L}' = \mathcal{L}$ .

*Demonstração.*

1. Seja  $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ , tal que  $\cup_{i \in I} A_i \in \sigma(\Theta)$ . Temos que

$$\mathcal{L}(\cup_{i \in I} A_i) = (\text{def}) = \min_{\theta \in \cup_{i \in I} A_i} \mathcal{L}(\{\theta\}) = \min_{i \in I} \min_{\theta \in A_i} \mathcal{L}(\{\theta\}) = (\text{def}) = \min_{i \in I} \mathcal{L}(A_i).$$

Assim,  $\mathcal{L}(\cup_{i \in I} A_i) \leq \min_{i \in I} \mathcal{L}(A_i)$  e, pelo Teorema 2.1, vale a monotonicidade. Além disso,  $\mathcal{L}(\cup_{i \in I} A_i) \geq \min_{i \in I} \mathcal{L}(A_i)$  e, pela definição, vale a consonância da união não enumerável.

2. Mostraremos apenas o caso (a), pois os outros são análogos. Seja  $A \in \sigma(\Theta)$ . Como, por hipótese, vale a monotonicidade na classe  $\mathcal{L}'$ , temos, pelo Teorema 2.1, que  $\mathcal{L}'(A) \leq \min_{\theta \in A} \mathcal{L}'(\{\theta\})$ . Analogamente, pela consonância da união finita, temos que  $\mathcal{L}'(A) \geq \min_{\theta \in A} \mathcal{L}'(\{\theta\})$ . Assim,  $\mathcal{L}'(A) = \min_{\theta \in A} \mathcal{L}'(\{\theta\}) = \min_{\theta \in A} \mathcal{L}(\{\theta\}) = \mathcal{L}(A)$ .

□

**Exemplo 2.7.** Seja  $\mathcal{L}$  a classe de testes de hipóteses construída da seguinte forma: suponhamos que, para cada  $\theta_0 \in \Theta$ ,

$$\mathcal{L}(\{\theta_0\})(x) = \mathbb{I} \left( \frac{V_x(\theta_0)}{\sup_{\theta \in \Theta} V_x(\theta)} \leq c \right), \forall x \in \mathcal{X},$$

como em 2.2. Então, a extensão dos testes acima para  $\sigma(\Theta)$  descrita no Teorema 2.5 é justamente a classe do Exemplo 2.2. Isto é, esta é a única extensão dos testes para as hipóteses simples  $\{\theta\}$ ,  $\theta \in \Theta$ , para uma classe de testes que respeita tanto a monotonicidade quanto a consonância da união (finita, enumerável ou não enumerável, dependendo da estrutura de  $\Theta$ ).

De fato, seja  $A \in \sigma(\Theta)$ . Para provar o que foi afirmado, basta notar que temos

$$\mathcal{L}(A)(x) \equiv \min_{\theta_0 \in A} \mathcal{L}(\{\theta_0\})(x) = \min_{\theta_0 \in A} \mathbb{I} \left( \frac{V_x(\theta_0)}{\sup_{\theta \in \Theta} V_x(\theta)} \leq c \right) = \mathbb{I} \left( \frac{\sup_{\theta \in A} V_x(\theta)}{\sup_{\theta \in \Theta} V_x(\theta)} \leq c \right)$$

□

**Exemplo 2.8.** Seja  $\mathcal{L}$  a classe de testes de hipóteses construída da seguinte forma: suponhamos que, para cada  $\theta \in \Theta$ ,

$$\mathcal{L}(\{\theta\})(x) = \mathbb{I}(ev_x(\{\theta\}) \leq c), \forall x \in \chi,$$

como em 2.3. Então, se  $\Theta$  é finito ou enumerável, a extensão dos testes acima para  $\sigma(\Theta)$  descrita no Teorema 2.5 é justamente a classe do Exemplo 2.3. Se  $\Theta$  é não enumerável, este resultado vale contanto que  $\forall x \in \chi$  e  $\forall a \in \mathfrak{R}^+$ ,  $\mathbb{P}(\{\theta : f(\theta|x) = a\}|x) = 0$ , isto é, se as possíveis distribuições *a posteriori* não tenham nenhum “patamar”. A demonstração deste fato encontra-se no Apêndice B

□

### 2.3 Consonância da Intersecção

**Exemplo 2.9. (ANOVA)** Suponhamos que  $X_1, \dots, X_{20} | (\mu_1, \mu_2, \mu_3, \sigma^2)$  são condicionalmente independentes e identicamente distribuídas (i.i.d.’s)  $N(\mu_1, \sigma^2)$ ;  $X_{21}, \dots, X_{40} | (\mu_1, \mu_2, \mu_3, \sigma^2)$  são condicionalmente i.i.d.’s  $N(\mu_2, \sigma^2)$  e  $X_{41}, \dots, X_{60} | (\mu_1, \mu_2, \mu_3, \sigma^2)$  são condicionalmente i.i.d.’s  $N(\mu_3, \sigma^2)$ . Consideremos as seguintes hipóteses:

$$H_0^{(1,2,3)} : \mu_1 = \mu_2 = \mu_3$$

$$H_0^{(1,2)} : \mu_1 = \mu_2$$

$$H_0^{(1,3)} : \mu_1 = \mu_3$$

e suponhamos que observamos as seguintes médias e desvios-padrão:

$$\bar{X}_1 = 0,15; S_1 = 1,09$$

$$\bar{X}_2 = -0,13; S_2 = 0,5$$

$$\bar{X}_3 = -0,38; S_3 = 0,79.$$

Utilizando-se o TRVG para testar cada um destas hipóteses, temos os seguintes p-valores:

$$p_{H_0^{(1,2,3)}} = 0,0498$$

$$p_{H_0^{(1,2)}} = 0,2564$$

$$p_{H_0^{(1,3)}} = 0,0920.$$

Assim, ao nível de significância de  $\alpha = 5\%$ , rejeitamos  $H_0^{(1,2,3)}$ , mas não rejeitamos nem  $H_0^{(1,2)}$  nem  $H_0^{(1,3)}$ . Desta forma, concluímos que ao menos dois dos três grupos (tratamentos) tem médias diferentes. Contudo, quando comparamos o primeiro com o segundo, concluímos que estes tem médias iguais, assim como quando comparamos o primeiro com o terceiro. Evidentemente, este conjunto de afirmações é contraditório, sob o ponto de vista lógico, uma vez que a partir da não rejeição de  $H_0^{(1,2)}$  e  $H_0^{(1,3)}$  deveríamos esperar a não rejeição de  $H_0^{(1,2,3)}$ . Esta propriedade é definida a seguir.

□

**Definição 2.6. (Consonância da intersecção finita)** Uma classe de testes  $\mathcal{L}$  atende à consonância da intersecção finita se, e somente se,  $\forall A, B \in \sigma(\Theta)$ ,  $\mathcal{L}(A \cap B) \leq 1 - (1 - \mathcal{L}(A))(1 - \mathcal{L}(B)) = \max\{\mathcal{L}(A), \mathcal{L}(B)\}$ . Em palavras, se aceitamos dois conjuntos, devemos aceitar sua intersecção.

*Nota:* Não é difícil verificar, por indução, que  $\mathcal{L}$  atende à consonância da intersecção finita se, e somente se,  $\forall n \geq 1$  e  $\forall A_1, \dots, A_n \in \sigma(\Theta)$ ,  $\mathcal{L}(\bigcap_{i=1}^n A_i) \leq 1 - \prod_{i=1}^n (1 - \mathcal{L}(A_i))$ .

Podemos considerar outras formas de consonância da intersecção:

**Definição 2.7. (Consonância da intersecção enumerável)** Uma classe de testes  $\mathcal{L}$  atende à consonância da intersecção enumerável se, e somente se,  $\forall \{A_n\}_{n \in \mathbb{N}} \subseteq \sigma(\Theta)$ ,  $\mathcal{L}(\bigcap_{n \geq 1} A_n) \leq 1 - \prod_{n \geq 1} (1 - \mathcal{L}(A_n)) = \max\{\mathcal{L}(A_n)\}_{n \in \mathbb{N}}$

**Definição 2.8. (Consonância da intersecção não enumerável)** Uma classe de testes  $\mathcal{L}$  atende à consonância da intersecção não enumerável se, e somente se,  $\forall \{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ , tal que  $\bigcap_{i \in I} A_i \in \sigma(\Theta)$ ,  $\mathcal{L}(\bigcap_{i \in I} A_i) \leq \max\{\mathcal{L}(A_i)\}_{i \in I}$

**Teorema 2.6.** *Seja  $\mathcal{L}$  uma classe de testes de hipóteses. Então, se  $\mathcal{L}$  satisfaz a consonância da intersecção não enumerável,  $\mathcal{L}$  satisfaz a consonância da intersecção enumerável. Se esta vale, então  $\mathcal{L}$  satisfaz a consonância da intersecção finita. Ademais, não vale a volta de nenhuma destas duas implicações.*

*Demonstração.* As implicações do Teorema são triviais, pois um conjunto finito é um conjunto enumerável e o conjunto  $I$  da Definição 2.8 pode, em particular, ser enumerável.

Para mostrar que  $\mathcal{L}$  satisfazer a consonância da intersecção finita não implica que  $\mathcal{L}$  satisfaz consonância da intersecção enumerável, seja  $\Theta = \mathbb{N}$  e  $\sigma(\Theta) = \mathcal{P}(\Theta)$ . Consideremos uma classe de testes  $\mathcal{L}$  definida por  $\mathcal{L}(A) = 1$  ( $A^c$  é infinito).  $\mathcal{L}$  respeita a consonância da intersecção finita, mas não a consonância da intersecção enumerável. De fato, sejam  $A, B \in \sigma(\Theta)$ . Então

$$\begin{aligned} \mathcal{L}(A \cap B) = 1 &\Rightarrow (A \cap B)^c = A^c \cup B^c \text{ é infinito} \Rightarrow A^c \text{ é infinito ou } B^c \text{ é infinito} \Rightarrow \\ \mathcal{L}(A) = 1 \text{ ou } \mathcal{L}(B) = 1. \end{aligned}$$

Assim, vale a consonância da intersecção finita. Para verificar que não vale a consonância da intersecção enumerável, consideremos  $A_i = \{i\}^c$ ,  $\forall i \geq 2$ . Temos que  $\mathcal{L}(A_i) = 0$ ,  $\forall i \geq 2$ , uma vez que  $A_i^c$  é finito. Contudo,  $\mathcal{L}(\bigcap_{i \geq 2} A_i) = \mathcal{L}(\{1\}) = 1$ , pois  $\{1\}^c$  é infinito.

Um contraexemplo análogo mostra que existem classes de testes que satisfazem a consonância da intersecção enumerável, mas não a consonância da intersecção não enumerável.  $\square$

**Exemplo 2.10.** Seja  $\mathcal{L}$  a classe de testes de hipóteses construída da seguinte forma: suponhamos que, para cada  $A \in \sigma(\Theta)$ ,

$$\mathcal{L}(A)(x) = \mathbb{I}(R(x) \not\subseteq A), \forall x \in \chi,$$

em que  $R(x)$  é um estimador por região de  $\theta$ , isto é,  $R(x) \in \sigma(\Theta)$ .  $\mathcal{L}$  satisfaz a consonância da intersecção não enumerável assim como a monotonicidade. De fato, se  $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$  é tal que  $\bigcap_{i \in I} A_i \in \sigma(\Theta)$  e  $x \in \chi$ , temos

$$\mathcal{L}(\bigcap_{i \in I} A_i)(x) = 0 \iff R(x) \subseteq \bigcap_{i \in I} A_i \iff R(x) \subseteq A_i, \forall i \in I \iff \max\{\mathcal{L}(A_i)(x)\}_{i \in I} = 0.$$

O resultado segue da definição de consonância da intersecção não enumerável e do item 5. do Teorema 2.1.  $\square$

O Teorema a seguir é análogo ao Teorema 2.5. Ele tem a finalidade de caracterizar classes de testes que satisfazem monotonicidade e consonância da intersecção (finita, enumerável ou não enumerável, dependendo da estrutura de  $\Theta$ ). Ele também mostra que, se desejamos criar uma classe de testes que satisfaz essas duas propriedades simultaneamente, basta criar os testes para cada hipótese  $\Theta \setminus \{\theta\}$ ,  $\theta \in \Theta$ . Isto é, definidos estes testes, há uma forma única de estendê-los para as outras hipóteses da  $\sigma$ -álgebra, de modo a obter uma classe de testes que satisfaça as propriedades de interesse.

**Teorema 2.7.** *Seja  $\mathcal{L}$  uma classe de testes construída da seguinte forma: para cada  $\theta \in \Theta$ , está fixado um teste  $\mathcal{L}(\Theta \setminus \{\theta\})$  (estamos supondo que  $\{\theta\} \in \sigma(\Theta)$ ,  $\forall \theta \in \Theta$ ). Para cada  $A \in \sigma(\Theta)$ , seja  $\mathcal{L}(A) \equiv \max_{\theta \in A^c} \mathcal{L}(\Theta \setminus \{\theta\})$ . Em palavras, o teste para  $A$  é o teste da união-intersecção (Casella e Berger [2]) para  $A$ , com base nos testes para as hipóteses  $\Theta \setminus \{\theta\}$ ,  $\theta \in A$ . Então*

1.  $\mathcal{L}$  atende à consonância da intersecção não enumerável, bem como à monotonicidade.
2. Seja  $\mathcal{L}'$  é uma classe que atende à monotonicidade, com  $\mathcal{L}'(\Theta \setminus \{\theta\}) = \mathcal{L}(\Theta \setminus \{\theta\})$ ,  $\forall \theta \in \Theta$ . Temos que:
  - (a) Se  $\Theta$  é finito e  $\mathcal{L}'$  atende à consonância da intersecção finita, então  $\mathcal{L}' = \mathcal{L}$ .
  - (b) Se  $\Theta$  é enumerável e  $\mathcal{L}'$  atende à consonância da intersecção enumerável, então  $\mathcal{L}' = \mathcal{L}$ .

(c) Se  $\Theta$  é não enumerável e  $\mathcal{L}'$  atende à consonância da intersecção não enumerável, então  $\mathcal{L}' = \mathcal{L}$ .

*Demonstração.*

1. Seja  $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$ , tal que  $\bigcap_{i \in I} A_i \in \sigma(\Theta)$ . Temos que

$$\mathcal{L}(\bigcap_{i \in I} A_i) = (\text{def}) = \max_{\theta \in (\bigcap_{i \in I} A_i)^c} \mathcal{L}(\Theta \setminus \{\theta\}) = \max_{i \in I} \max_{\theta \in A_i^c} \mathcal{L}(\Theta \setminus \{\theta\}) = (\text{def}) = \max_{i \in I} \mathcal{L}(A_i).$$

Assim,  $\mathcal{L}(\bigcap_{i \in I} A_i) \geq \max_{i \in I} \mathcal{L}(A_i)$  e, pelo Teorema 2.1, vale a monotonicidade. Além disso,  $\mathcal{L}(\bigcap_{i \in I} A_i) \leq \max_{i \in I} \mathcal{L}(A_i)$  e, pela definição, vale a consonância da intersecção não enumerável.

2. Mostraremos apenas o caso (a), pois os outros são análogos. Seja  $A \in \sigma(\Theta)$ . Como, por hipótese, vale a monotonicidade na classe  $\mathcal{L}'$ , temos, pelo Teorema 2.1, que  $\mathcal{L}'(A) \geq \max_{\theta \in A^c} \mathcal{L}'(\Theta \setminus \{\theta\})$ . Analogamente, pela consonância da intersecção finita,  $\mathcal{L}'(A) \leq \max_{\theta \in A^c} \mathcal{L}'(\Theta \setminus \{\theta\})$ . Assim, temos  $\mathcal{L}'(A) = \max_{\theta \in A^c} \mathcal{L}'(\Theta \setminus \{\theta\}) = \max_{\theta \in A^c} \mathcal{L}(\Theta \setminus \{\theta\}) = \mathcal{L}(A)$ .  $\square$

## 2.4 Invertibilidade

O seguinte exemplo é tradicional nos cursos de introdução à estatística e mostra o cuidado que se deve ter, sob a abordagem clássica, ao escolher o “rótulo” de uma determinada hipótese: hipótese nula ou hipótese alternativa.

**Exemplo 2.11.** Suponhamos que  $X|\theta \sim Normal(\mu, 1)$  e que desejamos testar as duas seguintes hipóteses (nulas):

$$\begin{aligned} H_0^{\leq} &: \mu \leq 0 \\ H_0^{\geq} &: \mu > 0 \end{aligned}$$

Os Testes Uniformemente Mais Poderosos (UMP) para estas hipóteses (que também são os TRVG) tem as respectivas regiões críticas, ao nível 5%:

$$\{x \in \mathfrak{R} : x > 1,64\} \text{ e } \{x \in \mathfrak{R} : x < -1,64\}.$$

Assim, se observamos  $x = 1,0$ , não rejeitamos  $H_0^{\leq}$  nem  $H_0^{\geq}$ . Em palavras, concluímos que a média é menor ou igual a 0, mas também concluímos que ela é maior que 0. Desta forma, dependendo de qual a hipótese que é definida como hipótese nula, essa não será rejeitada.

$\square$

**Definição 2.9. (Invertibilidade)** Uma classe de testes  $\mathcal{L}$  atende à invertibilidade se, e somente se,  $\forall A \in \sigma(\Theta)$ ,  $\mathcal{L}(A) = 1 - \mathcal{L}(A^c)$ . Em palavras, é irrelevante qual hipótese designamos por hipótese nula e qual designamos por hipótese alternativa, a rejeição de  $A$  deve ser equivalente a não rejeição de  $A^c$ .

**Exemplo 2.12.** Suponhamos que  $(L_A)_{A \in \sigma(\Theta)}$  é uma família de funções de perda com

$$L_A(0, \theta) = a_A \mathbb{I}(\theta \notin A) \text{ e } L_A(1, \theta) = b_A \mathbb{I}(\theta \in A), \quad \forall \theta \in \Theta,$$

com  $a_A = b_{A^c} > 0$ ,  $\forall A \in \sigma(\Theta)$ . Seja  $\theta_0 = \theta_0(x) \in \Theta$  um ponto qualquer e seja  $\mathcal{L}$  definida por

$$\mathcal{L}(A)(x) = \mathbb{I} \left( \mathbb{P}(A|x) < \frac{a_A}{a_A + b_A} \right) + \mathbb{I} \left( \mathbb{P}(A|x) = \frac{a_A}{a_A + b_A} \text{ e } \theta_0 \notin A \right), \quad \forall A \in \sigma(\Theta) \text{ e } \forall x \in \mathcal{X}.$$

Em palavras, aceitamos  $A$  sempre que sua probabilidade *a posteriori* é maior que  $\frac{a_A}{a_A + b_A}$ , ou então quando sua probabilidade *a posteriori* é exatamente  $\frac{a_A}{a_A + b_A}$  e um ponto  $\theta_0$ , fixado previamente, está em  $A$ . Observa-se que o ponto  $\theta_0$  pode depender de  $x$ , podendo ser, portanto, algum estimador pontual de  $\theta$ .

$\mathcal{L}$  satisfaz a invertibilidade uma vez que

$$\begin{aligned} \mathcal{L}(A)(x) = 0 &\iff \left( \mathbb{P}(A|x) > \frac{a_A}{a_A + b_A} \right) \text{ ou } \left( \mathbb{P}(A|x) = \frac{a_A}{a_A + b_A} \text{ e } \theta_0 \in A \right) \iff \\ &\left( \mathbb{P}(A^c|x) < \frac{b_A}{a_A + b_A} \right) \text{ ou } \left( \mathbb{P}(A^c|x) = \frac{b_A}{a_A + b_A} \text{ e } \theta_0 \notin A^c \right) \iff \\ &\left( \mathbb{P}(A^c|x) < \frac{a_{A^c}}{b_{A^c} + a_{A^c}} \right) \text{ ou } \left( \mathbb{P}(A^c|x) = \frac{a_{A^c}}{b_{A^c} + a_{A^c}} \text{ e } \theta_0 \notin A^c \right) \iff \mathcal{L}(A^c)(x) = 1. \end{aligned}$$

Notamos que  $\mathcal{L}$  é uma classe de testes bayesianos gerada pela família  $(L_A)_{A \in \sigma(\Theta)}$ . Do ponto de vista bayesiano, quando  $\mathbb{P}(A|x) = \frac{a_A}{a_A + b_A}$ , a decisão de aceitar  $A$  tem mesmo valor da decisão de rejeitar  $A$ . A classe apresentada foi escolhida dentre todas as classes de testes geradas pela família  $(L_A)_{A \in \sigma(\Theta)}$  justamente por respeitar a invertibilidade (esta escolha não é única). Contudo, do ponto de vista da Teoria da Decisão, todas as classes de testes geradas por  $(L_A)_{A \in \sigma(\Theta)}$  são equivalentes.

□





## Capítulo 3

# Desideratas

Neste capítulo, estudaremos três conjuntos de axiomas (*desideratas*) para classes de testes que podemos considerar a partir das propriedades descritas no Capítulo (2). Inicialmente, descreveremos cada uma delas, mostrando algumas de suas propriedades, bem como as exemplificando. Por fim, analisaremos as relações entre as desideratas propostas.

### 3.1 Desideratas

**D1**  $\mathcal{L}$  deve satisfazer

1. Invertibilidade ( $\forall A \in \sigma(\Theta)$ ,  $\mathcal{L}(A) = 1 - \mathcal{L}(A^c)$ )
2. Monotonicidade ( $\forall A, B \in \sigma(\Theta)$ ,  $A \subseteq B \Rightarrow \mathcal{L}(A) \geq \mathcal{L}(B)$ )

□

**Exemplo 3.1.** Suponhamos que  $(L_A)_{A \in \sigma(\Theta)}$  é uma família de funções de perda com

$$L_A(0, \theta) = a_A \mathbb{I}(\theta \notin A) \text{ e } L_A(1, \theta) = b_A \mathbb{I}(\theta \in A), \forall \theta \in \Theta,$$

com  $a_A = b_{A^c}$  e  $a_A \geq a_B > 0 \forall A, B \in \sigma(\Theta)$  com  $A \subseteq B$  (e, portanto,  $0 < b_A \leq b_B$ ). Seja  $\theta_0 = \theta_0(x) \in \Theta$  um ponto qualquer e seja  $\mathcal{L}$  definida por

$$\mathcal{L}(A)(x) = \mathbb{I}\left(\mathbb{P}(A|x) < \frac{a_A}{a_A + b_A}\right) + \mathbb{I}\left(\mathbb{P}(A|x) = \frac{a_A}{a_A + b_A} \text{ e } \theta_0 \notin A\right), \forall A \in \sigma(\Theta) \text{ e } \forall x \in \mathcal{X}.$$

Como  $(L_A)_{A \in \sigma(\Theta)}$  satisfaz **L1**, definida no Capítulo 2, pelo Teorema 2.2, vale a monotonicidade. Além disso,  $\mathcal{L}$  satisfaz a invertibilidade, pois trata-se de um caso particular do Exemplo (2.12). Assim,  $\mathcal{L}$  satisfaz **D1**.

Notemos que  $\mathcal{L}$  é uma classe de testes bayesianos gerada pela família  $(L_A)_{A \in \sigma(\Theta)}$ . Este exemplo é análogo ao Exemplo (2.12), no qual uma particular classe foi escolhida dentre todas as classes geradas pela mesma família de funções de perda justamente por respeitar a invertibilidade.

□

**Teorema 3.1.** *Se  $\mathcal{L}$  é uma classe de testes que satisfaz a desiderata **D1**, então*

1.  $\mathcal{L}(\Theta) = 0$ ,
2.  $\mathcal{L}(\emptyset) = 1$ ,
3. se  $x \in \chi$  e se  $A, B \in \sigma(\Theta)$  são disjuntos, então  $\mathcal{L}(A)(x) = 0 \Rightarrow \mathcal{L}(B)(x) = 1$ ,
4.  $\forall \{A_1, \dots, A_n\}$  partição finita mensurável de  $\Theta$ ,  $\sum_{i=1}^n (1 - \mathcal{L}(A_i)) \leq 1$  (isto é, aceita-se no máximo um dos elementos da partição),
5.  $\forall \{A_1, A_2, \dots\}$  partição enumerável mensurável de  $\Theta$ ,  $\sum_{i \geq 1} (1 - \mathcal{L}(A_i)) \leq 1$ .

*Demonstração.*

1. Temos, pela invertibilidade, que  $\mathcal{L}(\emptyset) = 1 - \mathcal{L}(\Theta)$ . Como vale a monotonicidade, temos, pelo Teorema 2.1, que  $\mathcal{L}(\Theta) = \mathcal{L}(\emptyset \cup \Theta) \leq \mathcal{L}(\emptyset)\mathcal{L}(\Theta) = (1 - \mathcal{L}(\Theta))\mathcal{L}(\Theta) = \mathcal{L}(\Theta) - \mathcal{L}^2(\Theta) = 0$ .

2. Segue de 1., do fato que  $\emptyset^c = \Theta$  e da invertibilidade.

3. Sejam  $x \in \chi$  e  $A, B \in \sigma(\Theta)$  disjuntos, com  $\mathcal{L}(A)(x) = 0$ . Pela invertibilidade,  $\mathcal{L}(A^c)(x) = 1$ . Como  $B \subseteq A^c$ , pela monotonicidade vale que  $\mathcal{L}(B)(x) = 1$ .

4. Seja  $\{A_1, \dots, A_n\}$  uma partição finita mensurável de  $\Theta$  e seja  $x \in \chi$ . Suponhamos que  $\exists i$  tal que  $\mathcal{L}(A_i)(x) = 0$ . Como  $A_i \cap A_j = \emptyset \forall j \neq i$ , por 3. temos que  $\mathcal{L}(A_j)(x) = 1, \forall j \neq i$ , de modo que  $\sum_{i=1}^n (1 - \mathcal{L}(A_i)(x)) = 1$ . Se não existe  $i$  tal que  $\mathcal{L}(A_i)(x) = 1$ , então  $\sum_{i=1}^n (1 - \mathcal{L}(A_i)(x)) = 0$ .

5. Demonstração análoga à 4. □

**D2**  $\mathcal{L}$  deve satisfazer

1. Invertibilidade ( $\forall A \in \sigma(\Theta), \mathcal{L}(A) = 1 - \mathcal{L}(A^c)$ )
2. Monotonicidade e consonância da união finita ( $\forall A, B \in \sigma(\Theta), \mathcal{L}(A \cup B) = \mathcal{L}(A)\mathcal{L}(B)$ )

□

*Nota:*  $\mathcal{L}$  satisfaz monotonicidade e consonância da união finita simultaneamente se, e somente se,  $\forall A, B \in \sigma(\Theta), \mathcal{L}(A \cup B) = \mathcal{L}(A)\mathcal{L}(B)$ . Isto decorre da definição de consonância da união finita e do Teorema (2.1).

**Exemplo 3.2.** Uma classe que satisfaz a desiderata **D2** é a definida por:

$$\mathcal{L}(A)(x) = \mathbb{I}(\arg \sup_{\theta \in \Theta} V_x(\theta) \notin A), \quad \forall x \in \chi \text{ e } \forall A \in \sigma(\Theta).$$

Em palavras, aceitamos  $A$  se, e somente se, o estimador de máxima verossimilhança de  $\theta$  pertence a  $A$ . De fato, se  $A, B \in \sigma(\Theta)$  e  $x \in \chi$  é tal que  $\mathcal{L}(A)(x) = 0$ , temos

$$\mathcal{L}(A)(x) = 0 \iff \arg \sup_{\theta \in \Theta} V_x(\theta) \in A \iff \arg \sup_{\theta \in \Theta} V_x(\theta) \notin A^c \iff \mathcal{L}(A^c)(x) = 1,$$

de modo que vale a invertibilidade. Além disso,

$$\begin{aligned} \mathcal{L}(A \cup B)(x) = 0 &\iff \arg \sup_{\theta \in \Theta} V_x(\theta) \in A \cup B \iff \arg \sup_{\theta \in \Theta} V_x(\theta) \in A \text{ ou} \\ &\arg \sup_{\theta \in \Theta} V_x(\theta) \in B \iff \mathcal{L}(A)(x) = 0 \text{ ou } \mathcal{L}(B)(x) = 0, \end{aligned}$$

de modo que vale a monotonicidade (pelo Teorema (2.1)), assim como a consonância da união finita.

□

**Exemplo 3.3.** Fixada uma distribuição *a priori* para  $\theta$  e admitindo-se que exista  $f(\theta|x)$ , a densidade *a posteriori* de  $\theta$ . Seja

$$\mathcal{L}(A)(x) = \mathbb{I}(\arg \sup_{\theta \in \Theta} f(\theta|x) \notin A), \forall x \in \chi \text{ e } \forall A \in \sigma(\Theta).$$

Esta classe satisfaz a desiderata **D2**. A demonstração deste fato é análoga à demonstração do Exemplo 3.2.

□

**Teorema 3.2.** *Se  $\mathcal{L}$  é uma classe de testes que satisfaz a desiderata **D2**, então:*

1.  $\mathcal{L}(\emptyset) = 1$ ,
2.  $\mathcal{L}(\Theta) = 0$ ,
3. se  $x \in \chi$  e se  $A, B \in \sigma(\Theta)$  são disjuntos, então  $\mathcal{L}(A)(x) = 0 \Rightarrow \mathcal{L}(B)(x) = 1$ ,
4. vale a consonância da intersecção finita  
 $(\forall A, B \in \sigma(\Theta), \mathcal{L}(A \cap B) \leq 1 - (1 - \mathcal{L}(A))(1 - \mathcal{L}(B)) = \max\{\mathcal{L}(A), \mathcal{L}(B)\})$ ,
5.  $\forall \{A_1, \dots, A_n\}$  partição finita mensurável de  $\Theta$ ,  $\sum_{i=1}^n (1 - \mathcal{L}(A_i)) = 1$ , isto é, aceita-se um, e somente um, elemento da partição.

*Demonstração.*

- 1., 2. e 3. tem provas análogas à do Teorema 3.1.

4. Sejam  $A, B \in \sigma(\Theta)$ . Usando a consonância da união finita e a invertibilidade, temos que

$$\mathcal{L}(A \cap B) = 1 - \mathcal{L}((A \cap B)^c) = 1 - \mathcal{L}(A^c \cup B^c) = 1 - \mathcal{L}(A^c)\mathcal{L}(B^c) = 1 - (1 - \mathcal{L}(A))(1 - \mathcal{L}(B)).$$

5. Seja  $\{A_1, \dots, A_n\}$  partição finita mensurável de  $\Theta$ , e  $x \in \chi$ . Temos, por 2. e pela consonância da união finita, que  $\prod_{i=1}^n \mathcal{L}(A_i) \leq \mathcal{L}(\bigcup_{i=1}^n A_i) = \mathcal{L}(\Theta) = 0$ , de modo que  $\exists i_0 \in \{1, \dots, n\}$  tal que  $\mathcal{L}(A_{i_0})(x) = 0$ . Mas, para  $j \neq i_0$ , temos, por 3., que  $\mathcal{L}(A_j)(x) = 1, \forall j \neq i_0$ , uma vez que  $A_j \cap A_{i_0}, \forall j \neq i_0$ . Assim,  $\sum_{i=1}^n (1 - \mathcal{L}(A_i)(x)) = 1$ . □

A desiderata **D2** pode ser caracterizada de diversas formas. O Teorema a seguir contém algumas delas.

**Teorema 3.3.** *São equivalentes:*

1.  $\mathcal{L}$  satisfaz a desiderata **D2**

2.  $\forall \{A_1, \dots, A_n\}$  partição finita mensurável de  $\Theta$ ,  $\sum_{i=1}^n (1 - \mathcal{L}(A_i)) = 1$

3.  $\mathcal{L}$  satisfaz:

(a) *Invertibilidade* ( $\forall A \in \sigma(\Theta), \mathcal{L}(A) = 1 - \mathcal{L}(A^c)$ )

(b) *Monotonicidade e consonância da intersecção finita*

( $\forall A, B \in \sigma(\Theta), \mathcal{L}(A \cap B) = 1 - (1 - \mathcal{L}(A))(1 - \mathcal{L}(B))$ )

4.  $\mathcal{L}$  satisfaz:

(a)  $\mathcal{L}(\emptyset) = 1$

(b)  $\mathcal{L}(\Theta) = 0$

(c) *Consonância da intersecção finita* ( $\forall A, B \in \sigma(\Theta), \mathcal{L}(A \cap B) \leq 1 - (1 - \mathcal{L}(A))(1 - \mathcal{L}(B))$ )

(d) *Consonância da união finita* ( $\forall A, B \in \sigma(\Theta), \mathcal{L}(A \cup B) \geq \mathcal{L}(A)\mathcal{L}(B)$ )

*Demonstração.*

$1 \Rightarrow 2$ . Já está provado no Teorema 3.2.

$2 \Rightarrow 1$ . Sejam  $A, B \in \sigma(\Theta)$  e  $x \in \chi$ . Consideremos a partição  $A_1 = A$  e  $A_2 = A^c$ . Temos que  $(1 - \mathcal{L}(A)) + (1 - \mathcal{L}(A^c)) = 1$ , de modo que  $\mathcal{L}(A) = 1 - \mathcal{L}(A^c)$  e vale, portanto, a invertibilidade.

Seja  $x \in \chi$ . Sem perda de generalidade, consideremos  $\mathcal{L}(A)(x) = 0$ . Consideremos agora a partição  $A_1 = A, A_2 = B \setminus A$  e  $A_3 = (A \cup B)^c$ . Temos que  $(1 - \mathcal{L}(A)(x)) + (1 - \mathcal{L}(B \setminus A))(x) + (1 - \mathcal{L}((A \cup B)^c)(x)) = 1$ . Assim,  $\mathcal{L}(A)(x) = 0 \Rightarrow \mathcal{L}((A \cup B)^c)(x) = 1$ , e da invertibilidade, que já foi mostrada,  $\mathcal{L}(A \cup B)(x) = 0$ . Desta forma, pelo item 2. do Teorema 2.1, vale a monotonicidade.

Considerando a mesma partição, mas supondo  $\mathcal{L}(A \cup B)(x) = 0$ , temos pela invertibilidade que  $\mathcal{L}((A \cup B)^c)(x) = 1$ . Assim, ou  $\mathcal{L}(A)(x) = 0$  ou  $\mathcal{L}(B \setminus A)(x) = 0$ . No segundo caso, pela

monotonicidade, que já foi mostrada,  $\mathcal{L}(B)(x) = 0$ . Assim,  $\mathcal{L}(A \cup B)(x) = 0 \Rightarrow \mathcal{L}(A)(x)\mathcal{L}(B)(x) = 0$ , e, portanto, vale a consonância da união finita.

$1 \Rightarrow 3$ . Já está provado no Teorema 3.2.

$3 \Rightarrow 1$ . Basta mostrar que  $3 \Rightarrow$  consonância da união finita. De fato, sejam  $A, B \in \sigma(\Theta)$ . Usando a consonância da intersecção finita e a invertibilidade, temos que

$$\mathcal{L}(A \cup B) = 1 - \mathcal{L}((A \cup B)^c) = 1 - \mathcal{L}(A^c \cap B^c) = 1 - [1 - (1 - \mathcal{L}(A^c))(1 - \mathcal{L}(B^c))] = \mathcal{L}(A)\mathcal{L}(B).$$

$1 \Rightarrow 4$ . Já está provado no Teorema 3.2.

$4 \Rightarrow 1$ . Basta mostrar que  $4 \Rightarrow$  invertibilidade e monotonicidade. Para verificar que vale a invertibilidade, seja  $A \in \sigma(\Theta)$ . Temos, pela consonância da intersecção finita e por  $\mathcal{L}(\emptyset) = 1$ , que  $1 - (1 - \mathcal{L}(A))(1 - \mathcal{L}(A^c)) = \mathcal{L}(A \cap A^c) = \mathcal{L}(\emptyset) = 1$ . Desta forma,  $(1 - \mathcal{L}(A))(1 - \mathcal{L}(A^c)) = 0$ , de modo que  $\mathcal{L}(A)\mathcal{L}(A^c) = \mathcal{L}(A) + \mathcal{L}(A^c) - 1$ . Mas, pela consonância da união finita e por  $\mathcal{L}(\Theta) = 0$ ,  $\mathcal{L}(A)\mathcal{L}(A^c) = \mathcal{L}(A \cup A^c) = \mathcal{L}(\Theta) = 0$ . Assim, temos que  $\mathcal{L}(A) + \mathcal{L}(A^c) - 1 = 0$  e, portanto, vale a invertibilidade.

Para verificar a monotonicidade, sejam  $A, B \in \sigma(\Theta)$ , com  $A \subseteq B$ , e  $x \in \chi$ . Se  $\mathcal{L}(A)(x) = 0$ , temos, pela invertibilidade (que já foi mostrada),  $\mathcal{L}(A^c)(x) = 1 - \mathcal{L}(A)(x) = 1$ . Como  $\mathcal{L}(A \cap B^c)(x) = \mathcal{L}(\emptyset)(x) = 1$  e vale a consonância da intersecção finita,  $1 = \mathcal{L}(A \cap B^c)(x) \leq 1 - (1 - \mathcal{L}(A)(x))(1 - \mathcal{L}(B^c)(x)) = 1 - 1(1 - \mathcal{L}(B^c)(x))$ , de modo que  $\mathcal{L}(B^c)(x) = 1$ . Novamente usando a invertibilidade,  $\mathcal{L}(B)(x) = 0$  e, assim, vale a monotonicidade.

□

**D3**  $\mathcal{L}$  deve satisfazer

1. Invertibilidade ( $\forall A \in \sigma(\Theta)$ ,  $\mathcal{L}(A) = 1 - \mathcal{L}(A^c)$ )
2. Monotonicidade e consonância da união enumerável

$$(\forall \{A_n\}_{n \in \mathbb{N}} \subseteq \sigma(\Theta), \mathcal{L}(\cup_{n \geq 1} A_n) = \prod_{n \geq 1} \mathcal{L}(A_n))$$

□

**Exemplo 3.4.** As classes definidas nos Exemplos 3.2 e 3.3 satisfazem também à desiderata **D3**. A demonstração é análoga às demonstrações apresentadas nestes exemplos.

□

**Teorema 3.4.** Se  $\mathcal{L}$  é uma classe de testes que satisfaz a desiderata **D3**, então:

1.  $\mathcal{L}(\emptyset) = 1$ ,

2.  $\mathcal{L}(\Theta) = 0$ ,
3. se  $x \in \chi$  e se  $A, B \in \sigma(\Theta)$  são disjuntos, então  $\mathcal{L}(A)(x) = 0 \Rightarrow \mathcal{L}(B)(x) = 1$ ,
4. vale a consonância da intersecção enumerável  
 $(\forall n \geq 1 \text{ e } \forall A_1, \dots, A_n \in \sigma(\Theta), \mathcal{L}(\bigcap_{i=1}^n A_i) \leq 1 - \prod_{i=1}^n (1 - \mathcal{L}(A_i)))$ ,
5.  $\forall \{A_1, A_2, \dots\}$  partição enumerável mensurável de  $\Theta$ ,  $\sum_{i \geq 1} (1 - \mathcal{L}(A_i)) = 1$ .

*Demonstração.* As demonstrações são análogas às do Teorema 3.2. □

Analogamente à desiderata **D2**, a desiderata **D3** também pode ser caracterizada de diversas formas, como mostra o Teorema a seguir.

**Teorema 3.5.** *São equivalentes:*

1.  $\mathcal{L}$  satisfaz a desiderata **D3**
2.  $\forall \{A_1, A_2, \dots\}$  partição enumerável mensurável de  $\Theta$ ,  $\sum_{i \geq 1} (1 - \mathcal{L}(A_i)) = 1$
3.  $\mathcal{L}$  satisfaz:
  - (a) *Invertibilidade* ( $\forall A \in \sigma(\Theta), \mathcal{L}(A) = 1 - \mathcal{L}(A^c)$ )
  - (b) *Monotonicidade e consonância da intersecção enumerável*  
 $(\forall \{A_n\}_{n \in \mathbb{N}} \subseteq \sigma(\Theta), \mathcal{L}(\bigcap_{n \geq 1} A_n) = 1 - \prod_{n \geq 1} (1 - \mathcal{L}(A_n)))$
4.  $\mathcal{L}$  satisfaz:
  - (a)  $\mathcal{L}(\emptyset) = 1$
  - (b)  $\mathcal{L}(\Theta) = 0$
  - (c) *Consonância da intersecção enumerável*  
 $(\forall \{A_n\}_{n \in \mathbb{N}} \subseteq \sigma(\Theta), \mathcal{L}(\bigcap_{n \geq 1} A_n) \leq 1 - \prod_{n \geq 1} (1 - \mathcal{L}(A_n)))$
  - (d) *Consonância da união enumerável* ( $\forall \{A_n\}_{n \in \mathbb{N}} \subseteq \sigma(\Theta), \mathcal{L}(\bigcup_{n \geq 1} A_n) \geq \prod_{n \geq 1} \mathcal{L}(A_n)$ )

*Demonstração.* As demonstrações são análogas às do Teorema 3.3. □

O Lema a seguir será útil na demonstração do Teorema 3.6, que fornece uma caracterização das classes de testes que satisfazem a desiderata **D3**.

**Lema 3.1.** *Seja  $\mathcal{L}$  uma classe de testes. Para cada  $x \in \chi$ , seja:*

$$\begin{aligned} f_x &: \sigma(\Theta) \longrightarrow \{0, 1\} \\ A \in \sigma(\Theta) &\mapsto f_x(A) = \mathcal{L}(A)(x) \end{aligned}$$

*Isto é,  $f_x$  associa, a cada  $A \in \sigma(\Theta)$ ,  $f_x(A) = \mathcal{L}(A)(x)$ . Então, se vale a desiderata **D3**, para cada  $x \in \chi$  a função  $f_x$  é contínua no sentido de que se  $(A_n)_{n \geq 1}$ ,  $A_n \in \sigma(\Theta)$ ,  $\forall n \geq 1$ , é uma seqüência tal que  $\liminf A_n = \limsup A_n \equiv \lim A_n$ , então  $f_x(\lim A_n) = \lim f_x(A_n)$ .*

*Demonstração.* Seja  $x \in \chi$ .

(1) Seja  $(A_n)_{n \geq 1}$  um seqüência crescente (isto é,  $A_i \subseteq A_{i+1} \forall i \geq 1$ ), com  $A_n \in \sigma(\Theta)$ ,  $\forall n \geq 1$ . Temos que  $\mathcal{L}(\lim A_n)(x) = \mathcal{L}\left(\bigcup_{n \geq 1} A_n\right)(x) = \lim \mathcal{L}(A_n)(x)$  (este limite está bem definido, pois, da monotonicidade, segue que  $(\mathcal{L}(A_n)(x))_{n \geq 1}$  é uma seqüência não crescente, limitada inferiormente por 0). De fato, pela monotonicidade e consonância da união enumerável:

$$\mathcal{L}\left(\bigcup_{n \geq 1} A_n\right)(x) = \prod_{n \geq 1} \mathcal{L}(A_n)(x) = (*)$$

Se existe  $i$  tal que  $\mathcal{L}(A_i)(x) = 0$ , então  $\mathcal{L}(A_j)(x) = 0 \forall j \geq i$  (pela monotonicidade), de modo que  $(*) = 0 = \lim \mathcal{L}(A_n)(x)$ .

Se não existe  $i$  tal que  $\mathcal{L}(A_i)(x) = 0$ , então  $(*) = 1 = \lim \mathcal{L}(A_n)(x)$ .

(2) Seja  $(A_n)_{n \geq 1}$  um seqüência decrescente (isto é,  $A_i \supseteq A_{i+1} \forall i \geq 1$ ), com  $A_n \in \sigma(\Theta)$ ,  $\forall n \geq 1$ . Temos que  $\mathcal{L}(\lim A_n)(x) = \mathcal{L}\left(\bigcap_{n \geq 1} A_n\right)(x) = \lim \mathcal{L}(A_n)(x)$  (este limite está bem definido, pois, da monotonicidade, segue que  $(\mathcal{L}(A_n)(x))_{n \geq 1}$  é uma seqüência não decrescente, limitada superiormente por 1). De fato, pela monotonicidade e consonância da intersecção enumerável (que vale, como mostrado no Teorema 3.4):

$$\mathcal{L}\left(\bigcap_{n \geq 1} A_n\right)(x) = 1 - \prod_{n \geq 1} (1 - \mathcal{L}(A_n)(x)) = (*)$$

Se existe  $i$  tal que  $\mathcal{L}(A_i)(x) = 1$ , então  $\mathcal{L}(A_j)(x) = 1 \forall j \geq i$  (pela monotonicidade), de modo que  $(*) = 1 = \lim \mathcal{L}(A_n)(x)$ .

Se não existe  $i$  tal que  $\mathcal{L}(A_i)(x) = 1$ , então  $(*) = 0 = \lim \mathcal{L}(A_n)(x)$ .

(3) Seja  $(A_n)_{n \geq 1}$  uma seqüência, com  $A_n \in \sigma(\Theta)$ ,  $\forall n \geq 1$ . Suponhamos que  $\liminf A_n = \limsup A_n \equiv \lim A_n \equiv A$ . De (1) e (2), temos que:

$$\begin{aligned}
\mathcal{L}(A)(x) &= \mathcal{L}\left(\bigcup_{n \geq 1} \bigcap_{k \geq n} A_k\right)(x) = \lim_n \mathcal{L}\left(\bigcap_{k \geq n} A_k\right)(x) = \limsup_n \mathcal{L}\left(\bigcap_{k \geq n} A_k\right)(x) \geq \\
&\limsup_n \mathcal{L}(A_n)(x) \geq \liminf_n \mathcal{L}(A_n)(x) \geq \liminf_n \mathcal{L}\left(\bigcup_{k \geq n} A_k\right)(x) = \lim_n \mathcal{L}\left(\bigcup_{k \geq n} A_k\right)(x) = \\
&\mathcal{L}\left(\bigcap_{n \geq 1} \bigcup_{k \geq n} A_k\right)(x) = \mathcal{L}(A)(x),
\end{aligned}$$

em que a primeira e a última desigualdades decorrem da monotonicidade.

Desta forma,  $\mathcal{L}(A)(x) = \lim \mathcal{L}(A_n)$ .

□

**Teorema 3.6.** *Suponha que  $\Theta = \mathfrak{R}^n$  e  $\mathcal{B}(\Theta) \subseteq \sigma(\Theta)$ . Então, uma classe  $\mathcal{L}$  satisfaz a desiderata **D3** se, e somente se, para cada  $x \in \chi$ ,  $\exists! \theta_0 = \theta_0(x) \in \Theta$  tal que  $\mathcal{L}(A)(x) = 0 \iff \theta_0 \in A$ ,  $\forall A \in \sigma(\Theta)$ . Isto é, vale **D3** se, e somente se,  $\forall A \in \sigma(\Theta)$ ,  $\mathcal{L}(A)(x) = \mathbb{I}(\theta_0(x) \notin A)$  para algum  $\theta_0(x) \in \Theta$ .*

*Demonstração.* ( $\Rightarrow$ )

Seja  $x \in \chi$ . Para cada  $(y_1, \dots, y_n) \in \mathbb{Z}^n$ , definimos

$$A_{y_1, \dots, y_n}^0 = \{(x_1, \dots, x_n) \in \mathfrak{R}^n : y_j \leq x_j < y_j + 1, \forall j = 1, \dots, n\}.$$

A Figura (??) mostra um exemplo típico de conjunto desta forma, para o caso de  $\mathfrak{R}^2$ .

Notemos que  $A_{y_1, \dots, y_n}^0 \cap A_{z_1, \dots, z_n}^0 = \emptyset$  se  $\exists k \in \{1, \dots, n\}$  tal que  $y_k \neq z_k$ . Além disso, temos que

$$\bigcup_{(y_1, \dots, y_n) \in \mathbb{Z}^n} A_{y_1, \dots, y_n}^0 = \mathfrak{R}^n (*).$$

Da união enumerável em (\*) e do item 5. do Teorema 3.4, temos que  $\exists! (y_1^0, \dots, y_n^0) \in \mathbb{Z}^n$  tal que

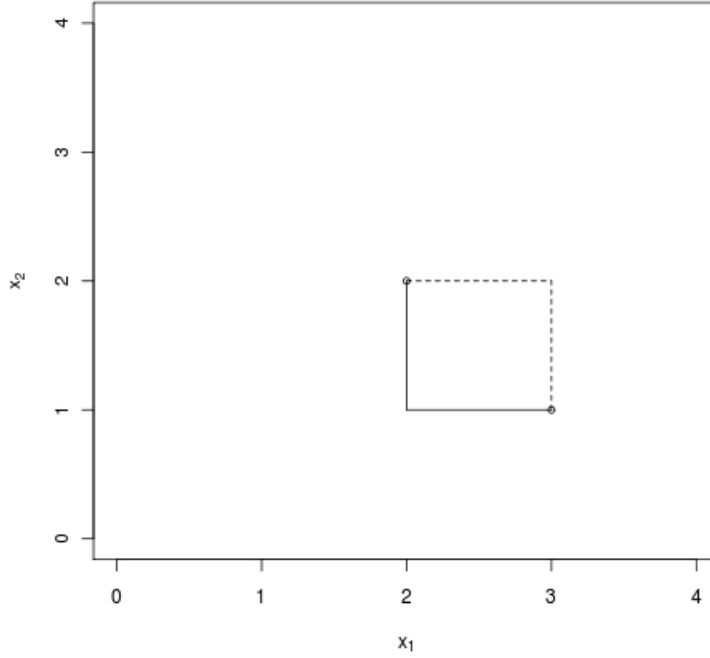
$$\mathcal{L}\left(A_{y_1^0, \dots, y_n^0}^0\right)(x) = 0.$$

Fixado o vetor  $(y_1^0, \dots, y_n^0)$ , consideremos agora os seguintes  $2^n$  conjuntos. Para cada  $(y_1, \dots, y_n) \in \{0, \frac{1}{2}\}^n$ , seja

$$A_{y_1, \dots, y_n}^1 = \{(x_1, \dots, x_n) \in \mathfrak{R}^n : y_j^0 + y_j \leq x_j < y_j^0 + \frac{1}{2} + y_j, \forall j = 1, \dots, n\}.$$

Temos que  $A_{y_1, \dots, y_n}^1 \cap A_{z_1, \dots, z_n}^1 = \emptyset$  se  $\exists k \in \{1, \dots, n\}$  tal que  $y_k \neq z_k$ . Além disso, temos que



Figura 3.1: Exemplo: o conjunto  $A_{2,1}^0$  em  $\mathbb{R}^2$ 

$$\bigcup_{(y_1, \dots, y_n) \in \{0, \frac{1}{2}\}^n} A_{y_1, \dots, y_n}^1 = A_{y_1^0, \dots, y_n^0}^0.$$

Assim, como  $\mathcal{L}\left(\left(A_{y_1^0, \dots, y_n^0}^0\right)^c\right)(x) = 1$ , temos, pelo item 5. do Teorema 3.4, que  $\exists! (y_1^1, \dots, y_n^1) \in \{0, \frac{1}{2}\}^n$  tal que

$$\mathcal{L}\left(A_{y_1^1, \dots, y_n^1}^1\right)(x) = 0.$$

Procedemos desta forma, sucessivamente. No passo  $m$ ,  $m \geq 1$ , definimos os  $2^n$  conjuntos: para cada  $(y_1, \dots, y_n) \in \{0, \frac{1}{2^m}\}^n$ , seja

$$A_{y_1, \dots, y_n}^m = \{(x_1, \dots, x_n) \in \mathbb{R}^n : \sum_{l=0}^{m-1} y_j^l + y_j \leq x_j < \sum_{l=0}^{m-1} y_j^l + \frac{1}{2^m} + y_j, \forall j = 1, \dots, n\}.$$

Novamente, temos que  $A_{y_1, \dots, y_n}^m \cap A_{z_1, \dots, z_n}^m = \emptyset$  se  $\exists k$  tal que  $y_k \neq z_k$ . Além disso, temos que

$$\bigcup_{(y_1, \dots, y_n) \in \{0, \frac{1}{2^m}\}^n} A_{y_1, \dots, y_n}^m = A_{y_1^{m-1}, \dots, y_n^{m-1}}^{m-1}.$$

Assim, como  $\mathcal{L}\left(\left(A_{y_1^{m-1}, \dots, y_n^{m-1}}^{m-1}\right)^c\right)(x) = 1$ ,  $\exists! (y_1^m, \dots, y_n^m) \in \{0, \frac{1}{2^m}\}^n$  tal que

$$\mathcal{L}\left(A_{y_1^m, \dots, y_n^m}^m\right)(x) = 0.$$

No final deste processo, obtemos a sequência

$$\overline{A}_{y_1^1, \dots, y_n^1}^1, \overline{A}_{y_1^2, \dots, y_n^2}^2, \dots, \overline{A}_{y_1^m, \dots, y_n^m}^m, \dots,$$

em que  $\overline{A}$  representa o fecho do conjunto  $A$ . Esta sequência é composta por conjuntos fechados encaixados cujo diâmetro vai a zero (Engelking [4]). Assim, pelo Teorema de Cantor (Engelking [4]),  $\exists! \theta_0 \in \Theta$  tal que  $\lim_m \overline{A}_{y_1^m, \dots, y_n^m}^m = \{\theta_0\}$ . Como  $A_{y_1^m, \dots, y_n^m}^m \subseteq \overline{A}_{y_1^m, \dots, y_n^m}^m$ , e  $\mathcal{L}\left(A_{y_1^m, \dots, y_n^m}^m\right)(x) = 0$ , temos  $\mathcal{L}\left(\overline{A}_{y_1^m, \dots, y_n^m}^m\right)(x) = 0$ ,  $\forall m \geq 1$ . Logo, pela continuidade de  $\mathcal{L}(\cdot)(x)$  (Lema 3.1),

$$\mathcal{L}(\{\theta_0\})(x) = \mathcal{L}\left(\bigcap_{m \geq 1} \overline{A}_{y_1^m, \dots, y_n^m}^m\right)(x) = \mathcal{L}\left(\lim_m \overline{A}_{y_1^m, \dots, y_n^m}^m\right)(x) = \lim_m \mathcal{L}\left(\overline{A}_{y_1^m, \dots, y_n^m}^m\right)(x) = 0$$

Assim, pela monotonicidade, se  $A$  é tal que  $\theta_0 \in A$ , temos  $\mathcal{L}(A)(x) = 0$ . Se  $A$  é tal que  $\theta_0 \notin A$ , temos, pelo item 3. do Teorema 3.4 que  $\mathcal{L}(A)(x) = 1$ , pois  $\mathcal{L}(A^c)(x) = 0$ .

Desta forma, se  $\mathcal{L}$  satisfaz a desiderata **D3**, para cada  $x \in \chi$ ,  $\exists! \theta_0 = \theta_0(x) \in \Theta$  tal que  $\mathcal{L}(A)(x) = 0 \iff \theta_0 \in A$ ,

( $\Leftarrow$ )

Seja  $x \in \chi$ . Seja  $\theta_0 = \theta_0(x) \in \Theta$  tal que  $\mathcal{L}(A)(x) = 0 \iff \theta_0 \in A$ . Temos que

$$\mathcal{L}(A)(x) = 0 \iff \theta_0 \in A \iff \theta_0 \notin A^c \iff \mathcal{L}(A^c)(x) = 1,$$

de modo que  $\mathcal{L}(A)(x) = 1 - \mathcal{L}(A^c)(x)$ . Assim, vale a invertibilidade.

Seja  $\{A_n\}_{n \in \mathbb{N}} \subseteq \sigma(\Theta)$ . Temos que

$$\begin{aligned} \mathcal{L}\left(\bigcup_{n \geq 1} A_n\right)(x) = 0 &\iff \theta_0 \in \bigcup_{n \geq 1} A_n \iff \exists i \in \mathbb{N} \text{ t.q. } \theta_0 \in A_i \iff \\ &\iff \exists i \text{ t.q. } \mathcal{L}(A_i)(x) = 0 \iff \prod_{n \geq 1} \mathcal{L}(A_n)(x) = 0, \end{aligned}$$

de modo que  $\mathcal{L}\left(\bigcup_{n \geq 1} A_n\right)(x) = \prod_{n \geq 1} \mathcal{L}(A_n)(x)$ ,  $\forall x \in \chi$ . Assim, vale a monotonicidade, bem como a consonância da união enumerável. □

Observamos que no Teorema anterior,  $\theta_0$  pode depender do ponto amostral observado,  $x$ . Isto é,  $\theta_0 = \theta_0(X)$  é um estimador pontual. Na realidade, este Teorema estabelece uma espécie de

conexão entre testes de hipóteses e estimação pontual: se acreditamos que as propriedades da desiderata **D3** devem ser satisfeitas e, portanto, escolhemos uma classe de testes que as satisfazem, necessariamente aceitaremos uma hipótese  $A$  se, e somente se, um determinado ponto (função da amostra) pertencer a ela. Notamos que os exemplos apresentados que obedecem a desiderata **D2** também são testes deste tipo. Podemos nos perguntar se há exemplos de classes de testes que satisfazem à desiderata **D2** mas que não satisfazem **D3**. A próxima seção tem como finalidade mostrar que todas as desideratas propostas são diferentes.

### 3.2 Relações entre as Desideratas

**Teorema 3.7.** *Seja  $\mathcal{L}$  uma classe de testes de hipóteses. Então*

$$\mathcal{L} \text{ satisfaz } \mathbf{D3} \begin{matrix} \Rightarrow \\ \not\Leftarrow \end{matrix} \mathcal{L} \text{ satisfaz } \mathbf{D2} \begin{matrix} \Rightarrow \\ \not\Leftarrow \end{matrix} \mathcal{L} \text{ satisfaz } \mathbf{D1}$$

*Demonstração.* As implicações seguem imediatamente das definições. Assim, exibiremos apenas contraexemplos para mostrar que **D1** não implica **D2**, que por sua vez não implica **D3**.

Um exemplo de classe de testes que atende a **D1** mas não a **D2** pode ser obtido se consideramos  $\Theta = [0, 1]$ ,  $\sigma(\Theta) = \mathcal{B}(\Theta)$ ,

$\theta \sim Unif(0, 1)$  e  $X|\theta \sim Ber(\theta)$ . Suponhamos que a classe  $\mathcal{L}$  é definida por

$$\mathcal{L}(A)(x) = \mathbb{I}\left(\mathbb{P}(A|x) < \frac{1}{2}\right) + \mathbb{I}\left(\mathbb{P}(A|x) = \frac{1}{2} \text{ e } \mathbb{E}[\theta|x] \notin A\right) \quad \forall A \in \sigma(\Theta) \text{ e } \forall x \in \chi.$$

Tal classe é um caso particular da estudada no Exemplo 3.1 (com  $\theta_0(x) = \mathbb{E}[\theta|x]$ ), de modo que  $\mathcal{L}$  satisfaz **D1**. Contudo, sejam  $A = [0; 0, 25]$  e  $B = [0, 25; 0, 5]$ . Temos que  $\mathbb{P}(A \cup B|x = 0) = 0,75$ ;  $\mathbb{P}(A|x = 0) = 0,4375$  e  $\mathbb{P}(B|x = 0) = 0,3125$ , de modo que  $\mathcal{L}(A \cup B)(0) = 0$ , enquanto que  $\mathcal{L}(A)(0) = \mathcal{L}(B)(0) = 1$ . Assim,  $\mathcal{L}$  não satisfaz a consonância da união finita e, conseqüentemente, **D2**.

Para mostrar que **D3**  $\not\Rightarrow$  **D2**, seja  $\Theta = \mathfrak{R}^n$  e  $\sigma(\Theta) = \mathcal{P}(\Theta)$ . Seja  $\mathbb{U}$  um ultrafiltro não trivial em  $\Theta$  (que existe, como mostrado no Teorema A.1 do Apêndice A). Podemos definir uma classe  $\mathcal{L}$  através de

$$\mathcal{L}(A)(x) = \mathbb{I}(A \notin \mathbb{U}), \quad \forall A \in \sigma(\Theta) \text{ e } \forall x \in \chi.$$

Temos que  $\mathcal{L}$  satisfaz a desiderata **D2**. De fato, sejam  $A, B \in \sigma(\Theta)$  e  $x \in \chi$ . Temos que:

$$\mathcal{L}(A)(x) = 0 \iff A \in \mathbb{U} \iff A^c \notin \mathbb{U} \iff \mathcal{L}(A^c)(x) = 1.$$

Assim, vale a invertibilidade. Além disso, se  $A, B \in \sigma(\Theta)$  são tais que  $A \subseteq B$  e  $x \in \chi$ , temos

$$\mathcal{L}(A)(x) = 0 \Rightarrow A \in \mathbb{U} \Rightarrow B \in \mathbb{U} \Rightarrow \mathcal{L}(B)(x) = 0,$$

de modo que vale a monotonicidade. Vale também a consonância da intersecção finita. De fato, se  $A, B \in \sigma(\Theta)$  e  $x \in \chi$ , temos

$$1 - (1 - \mathcal{L}(A)(x))(1 - \mathcal{L}(B)(x)) = 0 \Rightarrow \mathcal{L}(A)(x) = 0 \text{ e } \mathcal{L}(B)(x) = 0 \Rightarrow A \in \mathbb{U} \text{ e } B \in \mathbb{U} \Rightarrow \\ A \cap B \in \mathbb{U} \Rightarrow \mathcal{L}(A \cap B)(x) = 0.$$

Deste modo,  $\mathcal{L}$  satisfaz a desiderata **D2** (item 3. do Teorema 3.3). Como  $\mathbb{U}$  é não trivial, não existe  $\theta_0 \in \Theta$  tal que  $A \in \mathbb{U} \iff \theta_0 \in A$ . Isto é,  $\nexists \theta_0 \in \Theta$  tal que  $\mathcal{L}(A)(x) = 0 \iff \theta_0 \in A$ . Assim, pelo Teorema 3.6, temos que  $\mathcal{L}$  não satisfaz **D3**.

□

## Capítulo 4

# Conclusões

### 4.1 Considerações Finais

Neste trabalho, analisamos algumas propriedades que podem ser esperadas de testes de hipóteses simultâneos, após a realização de um único experimento. Vimos que muitos dos testes usualmente utilizados não respeitam as propriedades estudadas, de modo que, frequentemente, há falta de concordância lógica entre as conclusões obtidas através de testes simultâneos. Vimos também que, sob o ponto de vista da Teoria da Decisão, é possível criar classes de testes que satisfazem tais propriedades. Por fim, vimos um teorema que mostra que, impondo certas restrições a uma classe de testes de hipóteses, pode-se estabelecer uma espécie de concordância entre testes de hipóteses e estimação pontual.

### 4.2 Sugestões para Pesquisas Futuras

- Relacionar testes com múltiplas decisões com testes de duas decisões. Isto é, verificar o que se espera que ocorra, em termos de concordância, entre um teste com  $n$  possíveis decisões, que tem o objetivo de escolher um elemento dentre uma partição do espaço paramétrico com  $n$  elementos, e  $n$  testes com duas decisões, cujos objetivos são testar cada um dos elementos da partição contra seus complementares.
- Considerar perdas conjuntas para testes simultâneos, como feito neste trabalho apenas para o caso de duas hipóteses sob a perspectiva frequentista.
- Avaliar mais a fundo o problema de concordância de testes de hipóteses com estimação pontual e também com estimação intervalar.
- Avaliar a relação entre classes monótonas e classes admissíveis, considerando-se admissibilidade tanto de um ponto de vista clássico quanto bayesiano.



## Apêndice A

### Ultrafiltros

Neste apêndice, discutimos o conceito de ultrafiltro, que é utilizado na demonstração do Teorema (3.7). Mais informações sobre ultrafiltros podem ser encontradas em Fraïssé [6].

**Definição A.1. (*Filtro*)** Um conjunto  $\mathbb{F} \subseteq \mathcal{P}(X)$  é um filtro sobre um conjunto  $X$  se, e somente se:

$$\mathbf{F1} \quad \emptyset \notin \mathbb{F}$$

$$\mathbf{F2} \quad A \in \mathbb{F} \text{ e } A \subseteq B \Rightarrow B \in \mathbb{F}$$

$$\mathbf{F3} \quad A, B \in \mathbb{F} \Rightarrow A \cap B \in \mathbb{F}$$

**Definição A.2. (*Ultrafiltro*)** Um conjunto  $\mathbb{U} \subseteq \mathcal{P}(X)$  é um ultrafiltro sobre um conjunto  $X$  se, e somente se,  $\mathbb{U}$  é um filtro e  $\forall A \subseteq X$ , ou  $A \in \mathbb{U}$  ou  $A^c \in \mathbb{U}$ .

*Observação:* Neste estudo, um ultrafiltro pode ser interpretado como os conjuntos (hipóteses nulas) que são aceitos após observar a amostra  $x$ .

**Definição A.3. (*Ultrafiltro trivial*)** Um conjunto  $\mathbb{U} \subseteq \mathcal{P}(X)$  é um ultrafiltro trivial se, e somente se,  $\exists x \in X$  tal que, para  $A \subseteq X$ ,  $A \in \mathbb{U} \iff x \in A$ .

É fácil verificar que um ultrafiltro trivial é, em particular, um ultrafiltro.

**Definição A.4. (*Ultrafiltro não trivial*)** Um conjunto  $\mathbb{U} \subseteq \mathcal{P}(X)$  é um ultrafiltro não trivial se, e somente se,  $\mathbb{U}$  é ultrafiltro e  $\mathbb{U}$  não é trivial.

**Teorema A.1.** *Seja  $X$  um conjunto infinito. Existe um ultrafiltro não trivial sobre  $X$ .*

*Demonstração.*

1. Primeiramente, mostraremos que se  $\mathbb{F}$  é um filtro maximal sobre  $X$ , então  $\mathbb{F}$  é um ultrafiltro.

**Observações:** Um elemento  $s$  é maximal em um conjunto parcialmente ordenado  $(S, \leq)$  quando não existe elemento de  $S$  que é maior que  $s$ , segundo a ordem parcial  $\leq$  (Engelking [4]).  $S' \subseteq S$  é um subconjunto totalmente ordenado segundo a ordem  $\leq$  quando  $\forall A, B \in S'$ ,  $A \leq B$  ou  $B \leq A$ . Além disso,  $m \in S$  é um majorante de  $S'$  em  $S$  quando  $p \leq m$ ,  $\forall p \in S'$ . No presente caso, a ordem parcial considerada é a ordem da inclusão ( $\subseteq$ ) (Engelking [4]).

Suponhamos, pelo absurdo, que  $\mathbb{F}$  é filtro maximal, mas que não é ultrafiltro. Como  $\mathbb{F}$  não é ultrafiltro,  $\exists A \subseteq X$  tal que  $A \notin \mathbb{F}$  e  $A^c \notin \mathbb{F}$ . Podemos adicionar  $A$  a  $\mathbb{F}$  de modo a tentar estender  $\mathbb{F}$  para um novo filtro  $\mathbb{F}'$  que o contém. Para que o resultado da extensão seja de fato um filtro, temos também que adicionar  $A \cap Y$ ,  $\forall Y \in \mathbb{F}$ , assim como adicionar todos os conjuntos que contém  $A \cap Y$ ,  $\forall Y \in \mathbb{F}$ . Se  $A \cap Y \neq \emptyset$ ,  $\forall Y \in \mathbb{F}$ ,  $\mathbb{F}'$  será, de fato, um filtro. Contudo, por hipótese,  $\mathbb{F}$  é maximal, de modo que  $\mathbb{F}'$  não poderia ser um filtro. Assim,  $\exists Y \in \mathbb{F}$  tal que  $A \cap Y = \emptyset$ . Utilizando-se um raciocínio análogo para  $A^c$ , vemos que  $\exists Z \in \mathbb{F}$  tal que  $A^c \cap Z = \emptyset$ . Mas,  $A \cap Y = A^c \cap Z = \emptyset \Rightarrow Y \cap Z = \emptyset$ , o que é uma contradição, pois  $\mathbb{F}$  é filtro e, portanto,  $Y \cap Z \in \mathbb{F}$  mas  $\emptyset \notin \mathbb{F}$ . Assim, temos que se  $\mathbb{F}$  é filtro maximal, então  $\mathbb{F}$  é ultrafiltro.

2. Mostraremos agora que todo filtro está contido em um ultrafiltro.

Seja  $\mathcal{F}$  o conjunto de todos os filtros sobre  $X$ . Se  $\mathcal{F}' \subseteq \mathcal{F}$  é um subconjunto totalmente ordenado de  $\mathcal{F}$ , temos que  $\mathcal{F}'$  possui um majorante em  $\mathcal{F}$ . De fato, basta tomar  $\mathbb{F} = \bigcup \mathcal{F}'$ . Não é difícil verificar que  $\mathbb{F}$  é, de fato, um filtro. Deste modo, pelo Lema de Zorn (Engelking [4]), temos que existe um filtro maximal. Além disso, todo filtro está contido em um filtro maximal. De fato, se  $\mathbb{F}_1$  é um filtro que não está contido em nenhum filtro maximal, então não existe nenhum filtro  $\mathbb{F}'_1$  tal que  $\mathbb{F}_1 \subset \mathbb{F}'_1$  e, portanto,  $\mathbb{F}_1$  é filtro maximal. Desta forma, todo filtro está contido em um ultrafiltro (por 1.).

3. Finalmente, mostraremos que existe um ultrafiltro não trivial.

Consideremos o filtro dado por

$$\mathbb{F}_0 = \{A \subseteq X : A^c \text{ é finito}\}.$$

Nenhum filtro  $\mathbb{F}'$  tal que  $\mathbb{F}_0 \subseteq \mathbb{F}'$  contém conjuntos finitos. De fato, se  $A \subseteq X$  é finito, de modo que  $A^c \in \mathbb{F}_0$ , e  $\mathbb{F}'$  é tal que  $\mathbb{F}_0 \subseteq \mathbb{F}'$  e  $A \in \mathbb{F}'$  (conforme construção em 1.), então  $A \in \mathbb{F}'$  e  $A^c \in \mathbb{F}'$ , de modo que  $A \cap A^c = \emptyset \in \mathbb{F}'$ , contradição. Em particular, um ultrafiltro que contém  $\mathbb{F}_0$  (que existe, por 2.) não possui conjuntos finitos. Assim, existe um ultrafiltro não trivial  $\mathbb{F}^*$ , caso contrário existiria  $x_0 \in X$  tal que  $\{x_0\} \in \mathbb{F}^*$ , absurdo.

□



A demonstraco acima depende diretamente do Axioma da Escolha, uma vez que o Lema de Zorn   equivalente a este Axioma (Engelking [4]). Na realidade, a exist ncia de ultrafiltros no triviais   mais fraca que o Axioma da Escolha, mas no pode ser demonstrada utilizando-se apenas os Axiomas de Zermelo Fraenkel de Teoria do Conjuntos (Fraiss  [6]). Desta forma, a construico expl cita de exemplos de ultrafiltros no triviais no   simples.



## Apêndice B

### Demonstração Adicional

Neste Apêndice, demonstramos o que foi afirmado no Exemplo 2.8:

Seja  $\mathcal{L}$  a classe de testes de hipóteses construída da seguinte forma: suponhamos que, para cada  $\theta \in \Theta$ ,

$$\mathcal{L}(\{\theta\})(x) = \mathbb{I}(ev_x(\{\theta\}) \leq c), \forall x \in \chi,$$

como em 2.3. Então, se  $\Theta$  é finito ou enumerável, a extensão dos testes acima para  $\sigma(\Theta)$  descrita no Teorema 2.5 é justamente a classe do Exemplo 2.3. Se  $\Theta$  é não enumerável, este resultado vale contanto que  $\forall x \in \chi$  e  $\forall a \in \mathfrak{R}^+$ ,  $\mathbb{P}(\{\theta : f(\theta|x) = a\}|x) = 0$ . Isto é, esta é a única extensão dos testes para as hipóteses simples  $\{\theta\}$ ,  $\theta \in \Theta$ , para uma classe de testes que respeita tanto a monotonicidade quanto a consonância da união (finita, enumerável ou não enumerável, dependendo da estrutura de  $\Theta$ ).

A prova será dividida em três casos.

- Suponhamos que  $\Theta$  é finito.

Seja  $\{A_1, \dots, A_n\} \subseteq \sigma(\Theta)$ . Queremos mostrar, inicialmente, que

$$ev_x(\cup_{i=1}^n A_i) = \max_{1 \leq i \leq n} ev_x(A_i). \quad (1)$$

De fato, temos

$$\begin{aligned} ev_x(\cup_{i=1}^n A_i) &= \mathbb{P}(\{\theta : f(\theta|x) \leq \sup_{\theta \in \cup_{i=1}^n A_i} \mathbb{P}(\{\theta\}|x)\}|x) = \mathbb{P}(\{\theta : f(\theta|x) \leq \max_{\theta \in \cup_{i=1}^n A_i} \mathbb{P}(\{\theta\}|x)\}|x) \\ &= \mathbb{P}(\{\theta : f(\theta|x) \leq \mathbb{P}(\{\theta^*\}|x)\}|x) = ev_x(\{\theta^*\}), \end{aligned}$$

em que  $\theta^* = \arg \max_{\theta \in \cup_{i=1}^n A_i} \mathbb{P}(\{\theta\}|x)$ .

Seja  $i_0 \in \{1, \dots, n\}$  tal que  $\theta^* \in A_{i_0}$ . Um argumento análogo mostra que  $ev_x(\{\theta^*\}) = ev_x(A_{i_0})$  e, portanto,  $ev_x(\cup_{i=1}^n A_i) = ev_x(A_{i_0}) \leq \max_{1 \leq i \leq n} ev_x(A_i)$ . Como  $ev_x(\cup_{i=1}^n A_i) \geq ev_x(A_i)$ ,  $\forall i \in \{1, \dots, n\}$ , segue (1).

Para verificar que a classe  $\mathcal{L}$  definida no Teorema 2.5 é como a descrita no Exemplo 2.8, sejam

$A \in \sigma(\Theta)$  e  $x \in \chi$ . Temos que

$$\begin{aligned} \mathcal{L}(A)(x) &\equiv \min_{\theta_0 \in A} \mathcal{L}(\{\theta_0\})(x) = \min_{\theta_0 \in A} \mathbb{I}(ev_x(\{\theta_0\}) \leq c) = \mathbb{I}(ev_x(\{\theta_0\}) \leq c, \forall \theta_0 \in A) = \\ &= \mathbb{I}(\sup_{\theta_0 \in A} ev_x(\{\theta_0\}) \leq c) = \mathbb{I}(ev_x(A) \leq c), \end{aligned}$$

em que a última igualdade segue de (1).

- Suponhamos que  $\Theta$  é enumerável.

Seja  $\{A_i\}_{i \in \mathbb{N}} \subseteq \sigma(\Theta)$ . Queremos mostrar, inicialmente, que

$$ev_x(\cup_{i \in \mathbb{N}} A_i) = \sup_{i \in \mathbb{N}} ev_x(A_i). \quad (2)$$

Para tanto, mostraremos que  $\exists \theta^* \in \cup_{i \in \mathbb{N}} A_i$  tal que  $\sup_{\theta \in \cup_{i \in \mathbb{N}} A_i} \mathbb{P}(\{\theta\}|x) = \mathbb{P}(\{\theta^*\}|x)$ . De fato, suponhamos, pelo absurdo, que

$$\sup_{\theta \in \cup_{i \in \mathbb{N}} A_i} \mathbb{P}(\{\theta\}|x) > \mathbb{P}(\{\theta'\}|x), \forall \theta' \in \cup_{i \in \mathbb{N}} A_i. \quad (*)$$

Se vale  $\sup_{\theta \in \cup_{i \in \mathbb{N}} A_i} \mathbb{P}(\{\theta\}|x) = 0$ , o resultado é imediato. Seja então  $\theta_1 \in \cup_{i \in \mathbb{N}} A_i$  tal que  $\mathbb{P}(\{\theta_1\}|x) > 0$ . Como, por hipótese, vale (\*), temos que  $\exists \theta_2 \in \cup_{i \in \mathbb{N}} A_i$  tal que  $\mathbb{P}(\{\theta_2\}|x) > \mathbb{P}(\{\theta_1\}|x)$ . Procedendo desta maneira, sucessivamente, obtemos uma sequência  $(\theta_i)_{i \geq 1}$  tal que  $\mathbb{P}(\{\theta_{i+1}\}|x) > \mathbb{P}(\{\theta_i\}|x)$ ,  $\forall i \geq 1$ . Temos que

$$1 \geq \sum_{i \geq 1} \mathbb{P}(\{\theta_i\}|x) \geq \sum_{i \geq 1} \mathbb{P}(\{\theta_1\}|x) = \infty,$$

o que é absurdo. Assim,  $\exists \theta^* \in \cup_{i \in \mathbb{N}} A_i$  tal que  $\sup_{\theta \in \cup_{i \in \mathbb{N}} A_i} \mathbb{P}(\{\theta\}|x) = \mathbb{P}(\{\theta^*\}|x)$ . Deste modo,

$$\begin{aligned} ev_x(\cup_{i \in \mathbb{N}} A_i) &= \mathbb{P}\left(\{\theta : f(\theta|x) \leq \sup_{\theta \in \cup_{i \in \mathbb{N}} A_i} \mathbb{P}(\{\theta\}|x)\}|x\right) = \mathbb{P}(\{\theta : f(\theta|x) \leq \mathbb{P}(\{\theta^*\}|x)\}|x) \\ &= ev_x(\{\theta^*\}), \end{aligned}$$

Seja  $i_0 \in \{1, \dots, n\}$  tal que  $\theta^* \in A_{i_0}$ . Um argumento análogo mostra que  $ev_x(\{\theta^*\}) = ev_x(A_{i_0})$  e, portanto,  $ev_x(\cup_{i \in \mathbb{N}} A_i) = ev_x(A_{i_0}) \leq \sup_{i \in \mathbb{N}} ev_x(A_i)$ . Além disso,  $ev_x(\cup_{i \in \mathbb{N}} A_i) \geq \sup_{i \in \mathbb{N}} ev_x(A_i)$ . Daí segue (2).

Mostraremos agora que a classe  $\mathcal{L}$  definida no Teorema 2.5 é como a descrita no Exemplo 2.8. De fato, sejam  $A \in \sigma(\Theta)$  e  $x \in \chi$ . Temos que

$$\begin{aligned} \mathcal{L}(A)(x) &\equiv \min_{\theta_0 \in A} \mathcal{L}(\{\theta_0\})(x) = \min_{\theta_0 \in A} \mathbb{I}(ev_x(\{\theta_0\}) \leq c) = \mathbb{I}(ev_x(\{\theta_0\}) \leq c, \forall \theta_0 \in A) = \\ &= \mathbb{I}(\sup_{\theta_0 \in A} ev_x(\{\theta_0\}) \leq c) = \mathbb{I}(ev_x(A) \leq c), \end{aligned}$$

em que a última igualdade segue de (2).

- Suponhamos que  $\Theta$  é não enumerável.

Seja  $\{A_i\}_{i \in I} \subseteq \sigma(\Theta)$  tal que  $\cup_{i \in I} A_i \in \sigma(\Theta)$  e  $x \in \mathcal{X}$ . Suponhamos que  $\sup_{\theta \in \Theta} f(\theta|x) < \infty$ . Queremos mostrar inicialmente que  $\sup_{i \in I} ev_x(A_i) = ev_x(\cup_{i \in I} A_i)$ . Seja

$$a_i = \sup_{\theta \in A_i} f(\theta|x) < \infty, \quad \forall i \in I.$$

Como  $(a_i)_{i \in I}$  é limitada superiormente,  $\exists a_0 = \sup_{i \in I} a_i$ .

Seja  $a^* = \sup_{\theta \in \cup_{i \in I} A_i} f(\theta|x)$ . Mostraremos agora que  $a_0 = a^*$ .

1.  $a_0 \leq a^*$ .

De fato, temos que  $\forall i \in I$ ,  $a_i = \sup_{\theta \in A_i} f(\theta|x) \leq \sup_{\theta \in \cup_{i \in I} A_i} f(\theta|x) = a^*$ . Assim,  $a_0 \leq a^*$ .

2.  $a_0 \geq a^*$ .

Suponhamos, pelo absurdo, que  $a_0 < a^*$ . Então, existe  $t \in f(\cup_{i \in I} A_i|x) \equiv \{f(\theta|x) : \theta \in \cup_{i \in I} A_i\}$  tal que  $a_0 < t \leq a^*$ . Logo, existem  $i_0 \in I$  e  $\theta_0 \in A_{i_0}$  tais que  $a_0 < f(\theta_0|x) \leq a^*$  (pois  $f(\cup_{i \in I} A_i|x) = \cup_{i \in I} f(A_i|x)$ ). Assim,  $a_0 < f(\theta_0|x) \leq a_{i_0} \leq a^*$ , de modo que  $a_0 < a_{i_0}$ , contradição.

De 1. e 2., segue que  $a_0 = a^*$ .

Para verificar que  $\sup_{i \in I} ev_x(A_i) = ev_x(\cup_{i \in I} A_i)$ , notamos que

1.  $\sup_{i \in I} ev_x(A_i) \leq ev_x(\cup_{i \in I} A_i)$ .

De fato, basta notar que  $ev_x(A_i) \leq ev_x(\cup_{i \in I} A_i)$ ,  $\forall i \in I$ , como mostrado no Exemplo 2.3.

2.  $\sup_{i \in I} ev_x(A_i) \geq ev_x(\cup_{i \in I} A_i)$ .

Temos que  $ev_x(\cup_{i \in I} A_i) = \mathbb{P}(\{\theta : f(\theta|x) \leq a^*\}|x) = \mathbb{P}(\{\theta : f(\theta|x) < a^*\}|x)$ , por hipótese.

Vale também que

$$\{\theta : f(\theta|x) < a^*\} = \bigcup_{n \geq 1} \{\theta : f(\theta|x) \leq a^* - \frac{1}{n}\} = \bigcup_{n \geq 1} \{\theta : f(\theta|x) \leq a_0 - \frac{1}{n}\},$$

onde a última igualdade decorre do fato que  $a_0 = a^*$ .

Assim,

$$ev_x(\cup_{i \in I} A_i) = \mathbb{P} \left( \bigcup_{n \geq 1} \{\theta : f(\theta|x) \leq a_0 - \frac{1}{n}\} | x \right) = \lim_{n \rightarrow \infty} \mathbb{P}(\{\theta : f(\theta|x) \leq a_0 - \frac{1}{n}\} | x).$$

Mas, para todo  $n \geq 1$ ,  $\exists i_n \in I$  tal que  $a_0 - \frac{1}{n} < a_{i_n} \leq a_0$ . Logo, temos que  $\forall n \geq 1$ ,  $\{\theta : f(\theta|x) \leq a_0 - \frac{1}{n}\} \subseteq \{\theta : f(\theta|x) \leq a_{i_n}\}$ , de modo que  $\mathbb{P}(\{\theta : f(\theta|x) \leq a_0 - \frac{1}{n}\}|x) \leq \mathbb{P}(\{\theta : f(\theta|x) \leq a_{i_n}\}|x) = ev_x(A_{i_n})$ . Assim,

$$ev_x(\cup_{i \in I} A_i) = \lim_{n \rightarrow \infty} \mathbb{P}(\{\theta : f(\theta|x) \leq a_0 - \frac{1}{n}\}|x) \leq \lim_{n \rightarrow \infty} \sup_{i \in I} ev_x(A_{i_n}) \leq \sup_{i \in I} ev_x(A_i).$$

De 1. e 2., segue que  $ev_x(\cup_{i \in I} A_i) = \sup_{i \in I} ev_x(A_i)$  (3).

Finalmente, mostraremos que a classe  $\mathcal{L}$  definida no Teorema 2.5 é como a descrita no Exemplo 2.8. De fato, sejam  $A \in \sigma(\Theta)$  e  $x \in \chi$ . Temos que

$$\begin{aligned} \mathcal{L}(A)(x) &\equiv \min_{\theta_0 \in A} \mathcal{L}(\{\theta_0\})(x) = \min_{\theta_0 \in A} \mathbb{I}(ev_x(\{\theta_0\}) \leq c) = \mathbb{I}(ev_x(\{\theta_0\}) \leq c, \forall \theta_0 \in A) = \\ &= \mathbb{I}(\sup_{\theta_0 \in A} ev_x(\{\theta_0\}) \leq c) = \mathbb{I}(ev_x(A) \leq c), \end{aligned}$$

em que a última igualdade segue de (3).

## Referências Bibliográficas

- [1] M. A. Betz e J. R. Levin. *Coherent analysis-of-variance hypothesis-testing strategies: A general model*. Journal of Educational Statistics (1982) **7:3**, 193-206. [6](#)
- [2] G. Casella e R. L. Berger. *Statistical inference*. Duxbury Advanced Series, Pacific Grove, second edition, 2002. [2](#), [3](#), [18](#), [21](#)
- [3] M. H. DeGroot. *Optimal statistical decisions*. McGraw-Hill, New York, 1970. [3](#)
- [4] R. Engelking. *General topology*. Heldermann, Berlin, 1989. [34](#), [40](#), [41](#)
- [5] V. Fossaluzza. *Testes de hipóteses em eleições majoritárias*. Dissertação de mestrado, Universidade de São Paulo, 2008. [6](#)
- [6] R. Fraïssé. *Theory of Relations, Volume 145 (Studies in Logic and the Foundations of Mathematics)*. Elsevier Science B. V., Amsterdam, 2000. [39](#), [41](#)
- [7] K. R. Gabriel. *Simultaneous test procedures - some theory of multiple comparisons*. The Annals of Mathematical Statistics (1969) **40:1**, 224-250. [6](#), [16](#)
- [8] H. U. Gerber, B. P. K. Leung e E. S. W. Shiu. *Indicator Function and Hattendorff Theorem*. North American Actuarial (2003) **7:1**, 38-47. [1](#)
- [9] M. Lavine e M. J. Schervish. *Bayes factors: what they are and what they are not*. The American Statistician (1999) **53**, 119-122. [6](#), [14](#)
- [10] E. L. Lehmann. *A theory of some multiple decision problems, I*. The Annals of Mathematical Statistics (1957) **28:1**, 1-25. [6](#)
- [11] E. L. Lehmann. *A theory of some multiple decision problems, II*. The Annals of Mathematical Statistics (1957) **28:3**, 547-572. [6](#)
- [12] S. K. Lin, C. K. Chen, D. Ball, H. C. Liu e E. W. Loh. *Gender-specific contribution of the GABAA subunit genes on 5q33 in methamphetamine use disorder*. Pharmacogenomics Journal (2003) **3**, 349-355. [5](#)
- [13] J. M. Nicolau. *Sistema eleitoral e reforma política*. Foglio, Rio de Janeiro, 1993. [15](#)
- [14] C. D. Paulino e J. M. Singer. *Análise de dados categorizados*. Edgard Blucher, 2006. [5](#), [6](#)

- [15] C. A. B. Pereira e J. M. Stern. *Evidence and credibility: a full Bayesian test of precise hypothesis*. Entropy (1999) **1**, 99-110. [7](#)
- [16] D. M. Rom e B. Holland. *A new closed multiple testing procedure for hierarchical families of hypotheses*. Journal of Statistical Planning and Inference (1995) **46:3**, 265-275. [6](#)
- [17] H. Royden. *Real Analysis*. Prentice Hall, New Jersey, third edition, 1988. [3](#)
- [18] G. M. Silva. *Monotonicidade em testes de hipóteses*. Dissertação de mestrado, Universidade de São Paulo, 2010. [2](#), [6](#), [7](#)
- [19] M. J. Schervish. *P values: what they are and what they are not*. The American Statistician (1996) **50**, 203-206. [6](#)