

Daniel da Silva Oliveira

# Teste da Razão de Verossimilhança via Bootstrap em Modelos Não Encaixados

Niterói - RJ, Brasil

19 de janeiro de 2017

**Daniel da Silva Oliveira**

**Teste da Razão de Verossimilhança  
via Bootstrap em Modelos Não  
Encaixados**

**Trabalho de Conclusão de Curso**

Monografia apresentada para obtenção do grau de Bacharel em  
Estatística pela Universidade Federal Fluminense.

Orientador: Prof. Wilson Calmon Almeida dos Santos

Niterói - RJ, Brasil

19 de janeiro de 2017

**Daniel da Silva Oliveira**

**Teste da Razão de Verossimilhança via  
Bootstrap em Modelos Não Encaixados**

Monografia de Projeto Final de Graduação sob o título “*Teste da Razão de Verossimilhança via Bootstrap em Modelos Não Encaixados*”, defendida por Daniel da Silva Oliveira e aprovada em 19 de janeiro de 2017, na cidade de Niterói, no Estado do Rio de Janeiro, pela banca examinadora constituída pelos professores:

---

**Prof. Dr. Wilson Calmon Almeida dos Santos**  
Departamento de Estatística – UFF

---

**Prof. Dr. Jony Arrais Pinto Junior**  
Departamento de Estatística – UFF

---

**Prof. Dr. Diogo Bravo Marinho Braga**  
Universidade Federal Fluminense

Niterói, 19 de janeiro de 2017

O48 Oliveira, Daniel da Silva

Teste da razão de verossimilhança via bootstrap em modelos não encaixados. / Daniel da Silva Oliveira. – Niterói, RJ: [s.n.], 2017.

47f.

Orientador: Prof. Dr. Wilson Calmon Almeida

TCC (Bacharelado em Estatística) – Universidade Federal Fluminense, 2017.

1.Bootstrap (Estatística). 2.Modelos estatísticos. I. Título.

CDD 519.2

# Resumo

Um dos principais usos do Teste da Razão de Verossimilhança consiste em comparar modelos encaixados, ou seja, o modelo sob a hipótese nula é um caso restrito do modelo sob a hipótese alternativa. Neste contexto, a execução do teste transcorre de forma simples pois sob certas condições [Casella [1] e Bickel [2]], a razão de verossimilhança, quando devidamente transformada, possui distribuição assintótica qui-quadrado sob a hipótese nula. Suponha agora uma comparação de modelos em que tais modelos são ditos separados ou não encaixados no sentido de que um membro arbitrário do primeiro não pode ser obtido como um limite de membros do outro. Em outras palavras, um dos modelos não pode ser obtido por meio de restrições ao espaço paramétrico do outro. Então, seria incorreto utilizar o teste da razão de verossimilhança da maneira usual para comparar os modelos, como discutido em Cox [3] e [4]. Uma primeira alternativa seria, como apontado nos trabalhos citados, considerar uma transformação da estatística de teste da razão de verossimilhança que possui distribuição assintótica normal-padrão. Aqui, investigaremos uma alternativa à proposta de Cox em que técnicas de *bootstrap* serão empregadas para se derivar a distribuição [em amostras finitas] da estatística do teste. O *bootstrap* permite implementar um teste de hipóteses em que poucas suposições são feitas sobre o mecanismo gerador dos dados e sem apelo à qualquer aproximação assintótica. Neste trabalho discutimos como utilizar o *bootstrap* para este fim. Através do método, comparamos os modelos contínuos Lognormal e Gama e os modelos discretos Poisson e Geométrica, e analisamos sua performance. Tais comparações, feitas por meio de simulações, geraram resultados satisfatórios. O teste utilizado consegue atingir [para as comparações tratadas] poder empírico alto e comparável aos obtidos pelo método de Cox.

Palavras-chaves: Bootstrap (Estatística); Modelos estatísticos.

# Dedicatória

Dedico este trabalho à minha família e aos amigos que fizeram parte da minha jornada, me fortalecendo a cada dia. Em especial, dedico ao meu pai, à minha mãe e ao meu irmão por me apoiarem em cada decisão que tomei e me darem todo suporte necessário. Dedico, ainda, a todos as pessoas que fizeram parte da minha formação educacional até aqui.

# Agradecimentos

Agradeço primeiramente a Deus por ter me dado saúde, sabedoria e força para chegar a esse momento da minha vida. Agradeço à minha mãe, Denilza da Silva Oliveira, e ao meu pai, Jorge Nicolau de Oliveira, por tornarem este sonho possível. Obrigado pelos seus ensinamentos, pelo apoio e por lutarem para que eu tivesse condições de chegar aqui. Vocês e meu irmão, Jorge Alexandre da Silva Oliveira, tornaram o caminho mais fácil com atitudes que espero retribuir. Obrigado aos meus avós, tias e tios, primas e primos, por terem complementado a educação de meus pais.

Grato pelas amizades que o curso me proporcionou, como Amanda Araujo, Andrine Mendonça, Bruno Nobrega, Deborah Pereira, Hugo Fiares, Pedro Romeiro, Rayssa Paz e Thaylla Carolina. Grandes amigos que foram incentivadores do meu sucesso. Dentre estes, agradeço imensamente à Andrine e ao Bruno por estarem do meu lado durante estes quatro anos, me ajudando a me superar. Não sei se conseguiria sem vocês.

Obrigado às escolas por que passei e a seus respectivos professores que me tornaram apto ao curso de Graduação em Estatística da Universidade Federal Fluminense. Aos professores do Departamento de Estatística que me capacitaram ao título de Bacharel em Estatística. Conheci excelentes professores ao longo da graduação. Professores que me acompanharam desde o início da graduação e que criei uma admiração. A cada um, muito obrigado. Em especial, agradeço ao meu orientador deste trabalho de conclusão de curso, Wilson Calmon Almeida dos Santos, não só pelo professor que foi para mim, mas também por ter me confiado a responsabilidade de ser seu primeiro orientando.

Muito agradecido, também, às pessoas que conheci no período estágio, ainda que por pouco tempo. Roberta, Marcia, Lennon e Cintia, o que me ensinaram será muito útil daqui em diante na minha vida profissional.

Agradeço ao Escotismo (e aos meus irmãos escoteiros) por sua doutrina tão presente em mim. Tais ensinamentos fundamentais para minha formação de caráter. Busquei neste trabalho, assim como me ensinaram, “fazer o melhor possível”.

Enfim, obrigado a todas as pessoas que estão presentes em minha vida, ou passaram por ela, deixando um pouco de si comigo.

# Sumário

Lista de Figuras

Lista de Tabelas

<b>1</b>	<b>Introdução</b>	p. 11
<b>2</b>	<b>Objetivos</b>	p. 13
2.1	Objetivos Gerais . . . . .	p. 13
2.2	Objetivos Específicos . . . . .	p. 13
<b>3</b>	<b>Materiais e Métodos</b>	p. 14
3.1	Modelos, Parâmetros e Testes de Hipótese . . . . .	p. 14
3.1.1	Função de Distribuição Empírica e Estimador Plug-In . . . . .	p. 17
3.1.2	Teste de Hipótese . . . . .	p. 19
3.2	Testes da Razão de Verossimilhança . . . . .	p. 22
3.3	Modelos Não Encaixados . . . . .	p. 24
3.4	Método de Bootstrap . . . . .	p. 26
3.5	Testes baseados na Razão de Verossimilhança no contexto de Modelos Não Encaixados . . . . .	p. 31
3.5.1	A solução de Cox . . . . .	p. 31
3.5.2	O Algoritmo de Bootstrap aplicado a Famílias Não Encaixadas . . . . .	p. 32
<b>4</b>	<b>Análise dos Resultados</b>	p. 34
4.1	Lognormal x Gama . . . . .	p. 35



4.2	Gama x Lognormal . . . . .	p. 38
4.3	Geométrica x Poisson . . . . .	p. 39
4.4	Poisson x Geométrica . . . . .	p. 41
4.5	Análise Geral . . . . .	p. 42
4.6	Bootstrap x Cox . . . . .	p. 43
<b>5</b>	<b>Conclusão</b>	p. 45
	<b>Referências</b>	p. 47

# Lista de Figuras

1	Histogramas da Distribuição da Estatística o Teste quando $X_i \sim Lognormal(\theta_1; \theta_2^2)$	p. 36
2	Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando $X_i \sim Lognormal(\theta_1; \theta_2^2)$	p. 36
3	Histogramas da Distribuição da Estatística o Teste quando $X_i \sim Gama(\beta_1; \beta_2)$	p. 39
4	Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando $X_i \sim Gama(\beta_1; \beta_2)$	p. 39
5	Histogramas da Distribuição da Estatística o Teste quando $X_i \sim Geo(p)$	p. 40
6	Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando $X_i \sim Geo(p)$	p. 40
7	Histogramas da Distribuição da Estatística o Teste quando $X_i \sim Poi(\lambda)$	p. 41
8	Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando $X_i \sim Poi(\lambda)$	p. 41

# Lista de Tabelas

1	Tipos de Erros de Testes de Hipóteses . . . . .	p. 21
2	Poder do teste de Cox: Exponencial x Lognormal . . . . .	p. 32
3	Erros do Tipo II para o teste Lognormal x Gama com variância = 1 . . . . .	p. 37
4	Erros do Tipo II para o teste Lognormal x Gama com média = 1 . . . . .	p. 38
5	Erros do Tipo II para o teste Gama x Lognormal . . . . .	p. 38
6	Erros do Tipo II para o teste Geométrica x Poisson . . . . .	p. 41
7	Erros do Tipo II para o teste Poisson x Geométrica . . . . .	p. 42
8	Erros do Tipo II para Exponencial x Lognormal: Bootstrap x Cox . . . . .	p. 44

# 1 Introdução

É necessidade do meio econômico, social, científico, industrial explorar e compreender seus componentes. Por vezes, por exemplo, se torna interessante analisar estatisticamente a durabilidade de um equipamento ou a eficácia de um medicamento através de modelos. Isso proporciona um melhor entendimento dos fatores que interferem positiva e negativamente nestas variáveis cujo comportamento gostaríamos de compreender.

Estatística é uma ciência que analisa dados, com base em uma amostra, os sintetizando em informações e respostas. Usualmente, classes específicas de distribuições de probabilidade [Modelos] são utilizadas para a interpretação dos dados. Mas como escolher dentre tantos modelos e, mais restritamente, parametrizações possíveis? Se for de interesse, por exemplo, estudar a espera de pessoas por um atendimento em um hospital, podemos considerar o tempo de espera de uma amostra de pacientes e adotar um modelo que se adeque bem às características dos dados. É possível, ao menos teoricamente, que hajam dois ou mais modelos aparentemente adequados, por exemplo, Log-Normal e Gama. Deveríamos, então, escolher o melhor.

Tal escolha, geralmente, passa pela realização de Testes de Hipóteses na busca por um modelo mais específico ou melhor representativo. Dos testes possíveis, o Teste da Razão de Verossimilhança (TRV) é um dos mais utilizados na comparação de um modelo Irrestrito (classe de distribuição mais geral) contra um Restrito (classe de distribuição mais específica), no caso em que os modelos são encaixados.

Quando o TRV compara modelos encaixados, o menor é visto como um modelo restrito, mais específico que o alternativo. Por exemplo, podemos citar um teste feito para avaliar um modelo  $Normal(\mu, \sigma^2)$  com  $\theta = (\mu, \sigma^2) \in \Theta = (-\infty, +\infty) \times (0, +\infty)$  versus um modelo  $Normal(\mu, 1)$ , onde o novo espaço paramétrico poderia ser expresso da forma  $\Theta = (-\infty, +\infty) \times \{1\}$ . Agora, quando os modelos comparados são não encaixados, ou seja, um não pode ser escrito em função do outro [como antes feito], não se faz possível a utilização do teste (pelo menos não da maneira convencional).

Cox [3] e [4] sugere uma solução para o problema. Ele utiliza uma padronização do dobro do simétrico do logaritmo da razão de verossimilhança cuja distribuição assintótica é normal. O fato de a distribuição ser conhecida é o que viabiliza o teste, podendo se encontrar quantis da distribuição. O empecilho desta abordagem é a necessidade de amostras significativamente grandes, sendo este um método assintótico. Além disso, a padronização envolve cálculos complexos de valores esperados e variâncias do logaritmo da razão de verossimilhança. Um exemplo de aplicação aparece em Pereira [5], comparando, através do método tratado por Cox, modelos Exponencial e Log-Normal.

Uma outra possível solução, com a qual trabalharemos, é o método de *bootstrap*. Tratado em Efron & Tibshirani [6], os autores mostram a capacidade do método na estimação de distribuições de estatísticas. Sua praticidade está na abordagem computacional. O método consiste na obtenção de estimativas [replicações] da estatística de interesse. Sendo este mecanismo capaz de aproximar uma distribuição (ou uma característica desta) a qualquer estatística, poderia também aproximar a uma estatística do teste da razão de verossimilhança. Planejamos, neste trabalho, aplicar o método de *Bootstrap* ao Teste da Razão de Verossimilhança, tal como foi idealizado por Pesaran [7].

O Capítulo 2 deste trabalho trata do que pretendemos alcançar, onde enunciamos nossos objetivos [gerais e específicos]. Já no Capítulo 3 mostramos toda a metodologia empregada, bem como a teoria que a cerca. Introduzimos conceitos importantes e primordiais na Seção 3.1 como os de Modelo e Teste de Hipótese. Em 3.2 demonstramos como é executado um Teste da Razão de Verossimilhança e sua aplicação na comparação de modelos encaixados. Descrevemos e exemplificamos o que seriam modelos não encaixados em 3.3. A metodologia de *bootstrap* pode ser encontrada de forma geral na Seção 3.4. No Capítulo 4 exibimos e avaliamos os resultados do trabalho. Realizamos o Teste da Razão de Verossimilhança por meio do método de *bootstrap* para diferentes pares de modelos não encaixados, e avaliamos a eficiência de sua aplicação. Por fim, concluimos o trabalho no Capítulo 5 resumindo todas informações obtidas pelos resultados, e pontuando facilidades e dificuldades na implementação e utilização do Teste da Razão de Verossimilhança via *Bootstrap* para Modelos Não Encaixados.

## 2 Objetivos

### 2.1 Objetivos Gerais

Apresentar um importante uso do método computacional de *bootstrap* na comparação de modelos em Testes da Razão de Verossimilhança para Modelos Não Encaixados.

### 2.2 Objetivos Específicos

- Estabelecer um passo-a-passo para a implementação computacional do teste da razão de verossimilhança em modelos não encaixados por meio do procedimento de *bootstrap*.
- Fazer comparações de modelos separados [*Log-normal* com *Gama* e *Geométrica* com *Poisson*], obtendo dados simulados de um dos dois modelos do par e aplicando o Teste da Razão de Verossimilhança via *bootstrap*, e avaliar, através de dados simulados, a performance [probabilidade empírica de ocorrer o erro do tipo II] do teste.
- Analisar se o teste da razão de verossimilhança via *bootstrap* é uma boa alternativa ao método proposto por Cox para a comparação de modelos não encaixados. Avaliar, através de simulações, os poderes empíricos [ou a probabilidade de ocorrer o erro do tipo II] originados pelas duas metodologias, quando se compara modelos não encaixados [*Exponencial* e *Lognormal*].

## 3 Materiais e Métodos

A metodologia empregada no trabalho, assim como a teoria por trás dela, será explicada neste capítulo. Os dados utilizados são provindos de simulações. Utilizamos o programa *Rstudio* não só para simular os dados de distribuições, mas também para implementar o método [algoritmo de *bootstrap*] e produzir gráficos. Utilizamos funções do pacote *bmle* para estimar [pelo método de máxima verossimilhança] valores para os parâmetros e para capturar estimativas para a *log-verossimilhança*.

### 3.1 Modelos, Parâmetros e Testes de Hipótese

Muitos estudos e experimentos do cotidiano produzem **dados** cuja análise é, boa parte das vezes, a motivação principal da própria realização do estudo ou experimento. Dados podem corresponder a vetores, matrizes ou *arrays* contendo números e/ou códigos e até mesmo funções. Com frequência, estatísticos e usuários da estatística objetivam extrair informações úteis a partir dos dados, utilizando, para isso, todo o conhecimento e experiência que possuem. Naturalmente, são feitas e respondidas perguntas como: (i) “qual é a duração média de um certo equipamento eletrônico?”; (ii) “quanto se espera vender de um produto específico em um mercado delimitado?”; (iii) “qual é o risco de investir em uma dada ação na bolsa de valores?”; (iv) “qual é a chance de chover o suficiente nas proximidades do reservatório da cantareira para que não haja racionamento de água em São Paulo no ano que vem?” (v) “qual é a probabilidade de o atual presidente vencer em uma possível tentativa de reeleição?” e “qual é a margem de confiança?”. Não raro, é formulado algum modelo estatístico [*doravante*, modelo] que auxilia na produção de respostas a perguntas como estas.

Neste trabalho empregaremos o termo modelo no mesmo sentido utilizado por Bickel [2] e seguiremos suas convenções de notação e definições teóricas. Sem perda de generalidade, suponha que os dados correspondem a um vetor numérico  $\mathbf{x} = (x_1, \dots, x_n)$  e

que  $\mathbf{x} = \mathbf{X}(\omega)$  é o resultado de um experimento aleatório cujo espaço amostral é  $\Omega$ . Ao escrevermos  $\mathbf{x} = \mathbf{X}(\omega)$  estamos admitindo que  $\mathbf{x}$  é a realização de um vetor aleatório  $\mathbf{X} = (X_1, \dots, X_n)$  - ou que, para  $i$  entre 1 e  $n$ , cada  $x_i = X_i(\omega)$  é a realização da variável aleatória [v.a.]  $X_i$ .

**Definição 3.1.1.** *Um modelo  $\mathcal{P}$  é uma classe de distribuições de probabilidades sobre  $\mathbb{R}^n$  que, por hipótese, poderiam corresponder à distribuição de  $\mathbf{X}$ .*

Comumente, se assume que a distribuição verdadeira de  $\mathbf{X}$ , denotada por  $P^*$  é um elemento de  $\mathcal{P}$ . Neste caso, dizemos que o modelo é corretamente especificado. Mesmo que o modelo seja corretamente especificado, não é possível conhecer  $P^*$ . Por isso, utilizamos os dados para produzir uma estimativa de  $P^*$  ou de alguma característica sua, informalmente chamada de parâmetro. Abaixo apresentamos uma **definição formal de parâmetro**.

**Definição 3.1.2.** *Um parâmetro é uma função  $\theta$  do modelo  $\mathcal{P}$  para um espaço  $\Theta$ , chamado espaço paramétrico:*

$$\theta : \mathcal{P} \rightarrow \Theta \quad (3.1)$$

$$P \in \mathcal{P} \mapsto \theta(P). \quad (3.2)$$

Repare que o parâmetro, como definido acima, é uma função que depende da distribuição. Por isso, dizemos que o parâmetro é uma característica da distribuição. Nos acostumamos, porém, a chamar de parâmetro o valor assumido pela função  $\theta$  quando a distribuição é a verdadeira  $P^*$ , que denotaremos por  $\theta^* = \theta(P^*)$ . Um dos objetivos mais importantes em estatística é utilizar os dados para inferir o valor de  $\theta^*$ .

Formalmente, descreve-se um modelo através de uma **parametrização**. Se a função  $\theta$  for bijetiva, podemos usar a função inversa para parametrizar o modelo. Dessa forma, para cada  $\theta \in \Theta$  representará um elemento  $P = P_\theta$  do modelo  $\mathcal{P}$ . Fixado o conjunto  $\Theta$ , podemos escrever o modelo  $\mathcal{P}$  de forma mais interessante via:

**Definição 3.1.3.** *Um modelo  $\mathcal{P}$  é uma classe de distribuições da forma  $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ .*



Para efeito de ilustração, considere que  $\Theta$  representa o conjunto dos números reais estritamente positivos, ou seja,  $\Theta = (0, +\infty)$ . Se  $n = 1$  e a cada  $\theta > 0$  associamos a distribuição representada pela função de densidade probabilidade<sup>1</sup>  $f_{Exp(\theta)}(x) = \theta e^{-\theta x} \mathbb{I}(x > 0)$ , então, o modelo em questão assumiria que  $X_1$  é uma v.a. com distribuição exponencial cujo valor esperado é um número desconhecido  $E[X_1] = \theta^{-1}$ . Neste caso, escrevemos:  $\mathcal{P} = \{Exp(\theta) : \theta > 0\}$ , onde  $Exp(\theta)$  representa a distribuição exponencial associada à densidade  $f_{Exp(\theta)}$ .

A prática de modelar, no sentido aqui abordado, exige a criatividade do modelista em escolher uma classe  $\mathcal{P}$  apropriada, o que pode envolver a adoção de hipóteses como “ $\mathbf{X}$  é vetor aleatório contínuo” ou “as componentes de  $\mathbf{X}$  - *i.e.*, as v.a.’s  $X_1, \dots, X_n$  - são independentes e identicamente distribuídas [i.i.d.]”. Estas hipóteses, por sua vez, podem produzir representações alternativas para o modelo. Ao saber, no caso da hipótese de continuidade, poderíamos representar cada distribuição  $P$  do modelo por uma densidade [conjunta] e descrever o modelo, sem perda alguma, por meio das funções de densidade.

Repare que no caso da segunda hipótese [ $X_1, \dots, X_n$  i.i.d.] cada distribuição conjunta corresponde ao produto das distribuições marginais e as marginais são todas iguais. Não há perda nenhuma, portanto, em descrever os elementos do modelo por meio da distribuição marginal de  $X_1$ . De agora adiante, exceto quando comentado o contrário, admitiremos que as amostras são i.i.d. e utilizaremos  $P^* = P_{\theta^*}$  para denotar a distribuição [marginal] da primeira componente da amostra  $X_1$ . Diremos que  $X_1, \dots, X_n$  formam uma **amostra** de  $P^*$ .  $P^*$ , a distribuição verdadeira e, tipicamente, desconhecida será chamada de População. Ainda assim, podemos manter a representação do modelo via  $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ . Só que agora a distribuição  $P_\theta$  é definida sobre  $\mathbb{R}$ , independentemente do tamanho  $n$  da amostra. Voltando ao exemplo da exponencial, se  $n > 1$  e  $X_1, \dots, X_n$  são i.i.d. e, para algum  $\theta > 0$ , a densidade de  $X_1$  é  $f_{Exp(\theta)}(x) = \theta e^{-\theta x}$ ,  $x > 0$ ; continuamos escrevendo o modelo através de  $\mathcal{P} = \{Exp(\theta) : \theta > 0\}$ .

Assim como não conhecemos a distribuição verdadeira de  $X_1, \dots, X_n$  [ou de apenas  $X_1$ , no caso i.i.d.], denotada genericamente por  $P^*$ , não conhecemos o valor verdadeiro do parâmetro  $\theta^* = \theta(P^*)$  associado a tal distribuição. Um dos objetivos da inferência consiste em usar funções dos dados para produzir um valor aproximado para o parâmetro ou de uma parte sua. Tal valor aproximado é chamado de estimativa e pode ser pensado como a realização de uma v.a. chamada de **Estatística**. Definiremos Estatística na

<sup>1</sup>Usaremos a notação  $\mathbb{I}$  para representar a função indicadora:  $\mathbb{I}(S) = 1$ , se  $S$  é sentença verdadeira e 0, caso contrário.

sequência.

**Definição 3.1.4.** *Uma Estatística é uma função da amostra da forma  $T = T(X_1, \dots, X_n)$ . E sendo  $X_1 = x_1, \dots, X_n = x_n$  uma realização da amostra [amostra observada], calcula-se  $t = T(x_1, \dots, x_n)$  uma estimativa.*

Quando a estatística é empregada com a finalidade de produzir um valor aproximado para um parâmetro desconhecido [novamente, uma estimativa], então, ela é dita um **Estimador**. Como exemplo de estatística podemos citar a média amostral, definida por  $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ .

**Definição 3.1.5.** *Estimadores são Estatísticas utilizadas para aproximar [formalmente, estimar] o valor de um parâmetro.*

Ainda sobre a média amostral, quando  $X_1, \dots, X_n$  são variáveis aleatórias i.i.d.,  $\bar{X}$  é um estimador para a média populacional  $\mu = E[X_1]$ . Sob a condição de que a variância de  $X_1$  [ $\sigma^2$  ou variância populacional] seja finita,  $\bar{X}$  é um estimador *não-viesado* [valor esperado do estimador equivale ao parâmetro a ser estimado] de  $\mu^*$  e consistente no sentido de que converge em probabilidade para  $\mu^*$ , ou seja, para um  $n$  arbitrariamente grande, a probabilidade da diferença  $|\bar{X} - \mu^*|$  ser maior que qualquer número positivo tende a zero.

### 3.1.1 Função de Distribuição Empírica e Estimador Plug-In

Uma outra estatística “importante” é a Função de Distribuição Empírica [F.D.E.] definida por

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(X_i \leq x). \quad (3.3)$$

A F.D.E. é utilizada para estimar a Função de Distribuição Acumulada [F.D.A.] de  $X_1$ . Sabe-se que, se  $X_1$  possui F.D.A.  $F$ ,  $F(x) = P(X_1 \leq x)$  denota a probabilidade de tal v.a. assumir valor menor ou igual a  $x$ . Mais ainda, se  $X_1, \dots, X_n$  são i.i.d.,  $\hat{F}(x)$  é um estimador não-viesado e consistente, convergindo em probabilidade para  $F(x)$ . Se  $X_1 = x_1, \dots, X_n = x_n$  são realizações destas variáveis, a distribuição empírica de  $X_1$  associa a probabilidade de  $\frac{1}{n}$  de ocorrer cada valor  $x_i$ ,  $\forall i = 1, \dots, n$ , com a função de probabilidade podendo ser descrita da forma:

$$\hat{P}(X_i = x) = \#\{x_i; x_i = x\}/n, \quad (3.4)$$

onde  $\#A$  representa a cardinalidade do conjunto  $A$ , ou seja, o número de elementos deste conjunto.

Parâmetro, como já mencionado, é uma função  $\theta(\cdot)$  que depende da distribuição, sendo  $\theta^* = \theta(P^*)$  o valor verdadeiro do parâmetro para a distribuição  $P^*$ . A F.D.A.  $F^*$  de uma variável  $X_1$  é capaz de decrever a distribuição que a representa, tornando possível encontrar o valor verdadeiro do parâmetro via  $\theta^* = \theta(F^*)$ . Para fins de estimação de  $\theta^*$ , o método *plug-in* sugere trocarmos a F.D.A.  $F^*$  desconhecida por uma estimativa sua - por exemplo, pela Função de Distribuição Empírica [F.D.E.]. O chamado estimador *plug-in*, como pode ser visto em Efron & Tibshirani [6], é obtido por  $\hat{\theta}^* = \theta(\hat{F}^*)$ , onde  $\hat{F}^*$  é uma F.D.E. baseada numa amostra.

Afim de estimar a média populacional ( $\mu^* = E[X_1]$ ) de uma variável  $X_1$ , uma amostra  $\mathbf{x} = (x_1, \dots, x_n)$  pode ser obtida para definir uma distribuição empírica. Assim, uma estimativa, obtida pelo método *plug-in*, para  $\mu^*$  é dada pela esperança de  $X_1$  baseada em sua distribuição empírica, ou seja,

$$\hat{\mu}^* = \hat{E}[X_1] \quad (3.5)$$

$$= \frac{1}{n}x_1 + \frac{1}{n}x_2 + \dots + \frac{1}{n}x_n \quad (3.6)$$

$$= \frac{1}{n} \sum_{i=1}^n x_i \quad (3.7)$$

$$= \bar{\mathbf{x}}. \quad (3.8)$$

De maneira geral, usando a F.D.E., o estimador *plug-in* para o  $i$ -ésimo momento populacional equivale ao mesmo  $i$ -ésimo momento amostral, ou seja,  $\hat{E}[X_1^i] = \frac{1}{n} \sum_{j=1}^n X_j^i$ . Dessa forma, características que dependem de momentos populacionais são facilmente estimadas por este método.

**Exemplo 3.1.6.** *Considere a necessidade de estimar a variância  $\sigma^2 = E[X_1^2] - (E[X_1])^2$  de uma população. Sendo  $X_1, \dots, X_n$  v.a.'s i.i.d. e  $x_1, \dots, x_n$  valores por estas assumidos, o método *plug-in* estima  $\sigma^2$  por:*

$$\hat{\sigma}^2 = \hat{E}[X_1^2] - (\hat{E}[X_1])^2 \quad (3.9)$$

$$= \frac{1}{n} \sum_{i=1}^n x_i^2 - \left( \frac{1}{n} \sum_{i=1}^n x_i \right)^2 \quad (3.10)$$

$$= \frac{1}{n} \left( \sum_{i=1}^n x_i^2 - n \bar{\mathbf{x}}^2 \right). \quad (3.11)$$

### 3.1.2 Teste de Hipótese

Como dito anteriormente, na prática da estatística muitas vezes são feitas hipóteses sobre a distribuição que supostamente teria gerado os dados. Na realidade, mesmo que exista uma distribuição verdadeira  $P_{\theta^*}$ , dificilmente a análise vai conduzir à sua descoberta. Pior ainda, raramente se pode ter certeza de que a distribuição escolhida para representá-la é igual à verdadeira. Todavia, estamos interessados em escolher um bom modelo  $\mathcal{P}$  e encontrar uma distribuição  $\hat{P}$ , elemento de tal modelo, que aproxime a distribuição verdadeira  $P^*$ . Suposições sobre o modelo podem surgir pela experiência de estudos anteriores e, eventualmente, pela consulta à especialistas sobre as variáveis envolvidas no estudo. Não raro, do confronto de propostas, teorias e ideias emerge a necessidade de confrontar diferentes modelos ou restringí-los, por exemplo. **Testes de Hipóteses** podem ser utilizados para lidar com essa averiguação.

**Definição 3.1.7.** *Testes de hipóteses são regras de decisão que permitem rejeitar ou não rejeitar uma determinada hipótese  $H_0$  de interesse [chamada hipótese nula] quando confrontada a uma segunda hipótese  $H_1$  [chamada hipótese alternativa]. Tipicamente, depois de observar os dados  $x_1, \dots, x_n$ , se calcula uma estimativa  $t = T(x_1, \dots, x_n)$  de uma estatística de teste  $T = T(X_1, \dots, X_n)$  e, dependendo do valor  $t$  encontrado, se decide a favor ou contra  $H_0$ .*

A veracidade de qualquer afirmação à respeito de um parâmetro, por exemplo, ou da distribuição de uma v.a. pode ser confrontada através de um teste de hipótese. Muitos dos Testes de Hipótese decidem, por exemplo, entre duas hipóteses exclusivas sobre o valor do parâmetro  $\theta$ . Considera-se, por exemplo, uma hipótese nula da forma  $\theta \in \Theta_0$ , sendo  $\Theta_0 \subset \Theta$ , contra uma hipótese alternativa da forma  $\theta \in \Theta - \Theta_0$ . A aplicação do teste depende da escolha de uma estatística  $T(\mathbf{X})$  e da formulação de uma regra de decisão pré-definida baseada nela. Tal estatística recebe o nome de **estatística do teste** e é uma função da amostra - portanto, uma estatística, de fato.

As respostas possíveis do teste são “rejeita” ou “não rejeita” a hipótese nula. A particularidade de cada teste está em como o mesmo opta pela resposta. Esta decisão é tomada com base na amostra coletada ou, mais precisamente, na estimativa  $t = T(\mathbf{x})$  que ela origina. Tipicamente, valores extremos da estatística do teste, baseados em  $H_0$ , resultam em sua rejeição. Na sequência, definiremos a **região de rejeição**  $[RR]$  de um teste.

**Definição 3.1.8.** *Dado um teste de hipótese  $\Psi$ , a região de rejeição  $RR_\Psi$  é o subconjunto*

de elementos  $\mathbf{x} \in \text{Im}\mathbf{X}$  que retornam, como resultado de  $\Psi(\mathbf{x})$ , a rejeição de  $H_0$ .

$$RR_\Psi = \{\mathbf{x} \in \text{Im}\mathbf{X} : \Psi(\mathbf{x}) = \text{“rejeita”}\} \quad (3.12)$$

Assim, dada uma realização  $\mathbf{x} = (x_1, \dots, x_n)$  da amostra  $X_1, \dots, X_n$ , o teste para uma hipótese  $H_0$  é uma função  $\Psi$  tal que:

$$\Psi(\mathbf{x}) = \begin{cases} \text{“rejeita”} & , \quad \text{se } \mathbf{x} \in RR \\ \text{“não rejeita”} & , \quad \text{caso contrário.} \end{cases} \quad (3.13)$$

Como dissemos, em um teste de hipóteses convencional existem dois tipos de hipóteses: nula e alternativa. A hipótese nula ( $H_0$ ) é a afirmação a ser posta a prova e, pelo menos à princípio, é assumida como verdade. A hipótese alternativa ( $H_1$ ) sugere algo contrário a  $H_0$  ou sem pontos de interseção. Normalmente, as duas hipóteses são complementares. Se  $H_0 : \theta \in \Theta_0$ ,  $H_1 : \theta \in \Theta_0^c$  onde  $\Theta_0^c$  é o conjunto complementar a  $\Theta_0$ . O teste é realizado sobre a hipótese nula e, por serem exclusivas, a não rejeição desta implica na rejeição de  $H_1$ .

Suponha que, por exemplo, há o interesse em testar se a média  $\mu$  de uma população é igual a 0, com base em uma amostra  $X_1, \dots, X_n$ . Considere ainda que a condição para o teste não rejeitar a hipótese nula de que  $\mu = 0$  seja a média amostral estar a uma distância menor que  $c$  de 0. Portanto, dada uma realização da amostra  $\mathbf{x} = x_1, \dots, x_n$ , para este teste especificamente a região de rejeição seria da forma  $RR = \{\mathbf{x} \in \text{Im}\mathbf{X} : \bar{\mathbf{x}} \leq -c \text{ ou } \bar{\mathbf{x}} \geq c\}$ , sendo  $\bar{X}$  a estatística do teste.

Nunca é garantida a confiabilidade na resposta de um teste. Pensando nas suas possíveis respostas, podem ocorrer dois tipos de erros. É possível que  $H_0$  seja verdadeira e o teste a rejeite. Ou ainda, que a hipótese seja falsa e o teste resulte na não rejeição. A Tabela 1 classifica essas ocorrências.

**Definição 3.1.9.** *Erro do Tipo I ocorre quando rejeitamos a hipótese nula, sendo ela a hipótese verdadeira.*

**Definição 3.1.10.** *Erro do Tipo II é a não rejeição de  $H_0$ , quando sua afirmação não corresponde com a verdade.*

A resposta do teste deverá depender só e somente da amostra observada.  $RR$  é limitada com base em quantis da distribuição de  $T$ , em função de um valor  $\alpha$  [ $\alpha \in (0, 1)$ ] que é chamado de nível de significância do teste e representa um limite superior para a

Tabela 1 – Tipos de Erros de Testes de Hipóteses

		Decisão do Teste	
		Rejeita $H_0$	Não Rejeita $H_0$
Possibilidades	$H_0$ é VERDADEIRA	Erro do Tipo I	Correto
	$H_0$ é FALSA	Correto	Erro do Tipo II

probabilidade de erro do tipo I. No exemplo acima, se  $RR = \{\mathbf{x} \in Im\mathbf{X} : \bar{\mathbf{X}} \leq Q_{1-\alpha}\}$  [ $Q_{1-\alpha}$  sendo o quantil  $(1 - \alpha)\%$  de  $\bar{X}$  sob a hipótese nula], o nível de significância seria de  $\alpha\%$ . Note que quanto maior o nível de significância menor o quantil que delimita a  $RR$  e, conseqüentemente, maior é a região de rejeição. Devido a essa propriedade,  $\alpha$  é também chamado de tamanho do teste. Seu cálculo é associado ao Erro do Tipo I e será definido a seguir.

**Definição 3.1.11.** *Tamanho do teste é definido pela maior probabilidade de rejeitar a hipótese nula dado que ela seria verdadeira, descrita por*

$$\alpha = \sup_{\theta \in \Theta_0} P(\mathbf{X} \in RR). \quad (3.14)$$

Uma outra maneira usual para responder a um teste de hipótese é através do uso do **p-valor**. O critério de decisão do teste depende de um valor escolhido para  $\alpha$  [usualmente, 0,05]. Dada a condição de a hipótese nula estar correta, buscamos encontrar a distribuição da estatística do teste  $T$ . Dessa forma, às vezes, é possível através de alguns cálculos probabilísticos, obter o p-valor. Se o p-valor for menor ou igual a  $\alpha$  rejeita-se  $H_0$ . Caso contrário, a hipótese não pode ser rejeitada. Definimos p-valor em seguida.

**Definição 3.1.12.** *P-valor [ou valor-p] é a probabilidade de se obter, com a estatística do teste, uma estimativa extrema em relação à observada. Ou ainda, p-valor é o menor nível de significância para o qual a hipótese nula deveria ser rejeitada com a amostra observada, no teste previamente escolhido.*

Uma das maneiras de comparar diferentes testes é pela comparação de suas potências [poderes].

**Definição 3.1.13.** *Poder do teste é a probabilidade de “rejeitar”  $H_0$  dado que é a hipótese falsa, ou seja, a capacidade de acertar tomando a decisão de rejeição, ou ainda,  $P(\mathbf{x} \in RR | \theta \in \Theta_0^c)$*

## 3.2 Testes da Razão de Verossimilhança

Testes da razão de verossimilhança estão naturalmente associados aos chamados estimadores de máxima verossimilhança. Como afirmado em Casella [1], tal teste é usualmente empregado para testar a hipótese nula de que um determinado parâmetro da população satisfaz ou não uma certa restrição ou ainda, como sugerido em Bickel [2], para testar a hipótese nula de se um modelo restrito é ou não apropriado. Neste segundo caso, poderíamos dizer que o teste também cumpre a função de comparar modelos.

Considere as variáveis aleatórias  $X_1, \dots, X_n$  com função de densidade conjunta dada por  $f_n(X_1, \dots, X_n|\theta)$  em que  $\theta$  representa um possível valor para o parâmetro (escalar ou vetorial), ou seja, pertencente ao espaço paramétrico  $\Theta$ . Uma vez coletada uma amostra  $\mathbf{x} = (x_1, \dots, x_n)$  pode-se, então, definir a função de verossimilhança como:

$$L : \Theta \longrightarrow R \quad (3.15)$$

$$L(\theta|\mathbf{x}) \mapsto f_n(x_1, \dots, x_n|\theta). \quad (3.16)$$

Portanto, sendo  $x_1, \dots, x_n$  os valores assumidos pelas v.a.'s  $X_1, \dots, X_n$  em um experimento, a função de verossimilhança é uma função que depende somente do parâmetro  $\theta$ . Portanto, a notação será simplificada tendo  $L(\theta) = L(\theta|x_1, \dots, x_n)$ . E, sendo  $X_1, \dots, X_n$  i.i.d. e  $f$  a função de densidade marginal de cada uma destas variáveis, a obtemos por:

$$L(\theta) = f_n(x_1, \dots, x_n|\theta) = \prod_{i=1}^n f(x_i|\theta). \quad (3.17)$$

O estimador de máxima verossimilhança (EMV) de  $\theta$  é o valor que maximiza  $L(\theta)$ .

**Definição 3.2.1.** *Assumindo  $\mathbf{x} = x_1, \dots, x_n$  uma realização de  $X_1, \dots, X_n$ , um teste de razão de verossimilhança (TRV) para  $H_0 : \theta \in \Theta_0$  contra  $H_1 : \theta \in \Theta_0^c$  é qualquer teste que tem uma região de rejeição da forma  $\{\mathbf{x} \in Im \mathbf{X} : \lambda(\mathbf{x}) \leq c\}$ . Sendo que a estatística do teste é*

$$\lambda(\mathbf{X}) = \frac{\sup_{\Theta_0} L(\theta|\mathbf{X})}{\sup_{\Theta} L(\theta|\mathbf{X})}. \quad (3.18)$$

A lógica de TRV's pode ser melhor compreendida na situação em que  $f(\mathbf{x}|\theta)$  é a função de probabilidade de uma variável aleatória discreta. Neste caso, o numerador de  $\lambda(\mathbf{x})$  representa a probabilidade máxima de  $X_1, \dots, X_n$  assumirem os valores da amostra observada  $x_1, \dots, x_n$ , sendo válida a hipótese nula. O denominador de  $\lambda(\mathbf{x})$  é a probabili-

dade máxima de  $X_1, \dots, X_n$  assumirem os valores da amostra observada  $x_1, \dots, x_n$ , sendo a hipótese nula verdadeira ou não.

A razão deveria ser pequena se a hipótese nula não é verdadeira. Nesta situação, o TRV indica que  $H_0$  deve ser rejeitada e  $H_1$  aceita como verdadeira.

Existe uma correspondência entre TRV's e EMV's. Suponha que existam  $\hat{\theta}$  e  $\hat{\theta}_0$ , respectivamente estimadores de  $L(\theta)$  em  $\Theta$  [Estimador Irrestrito] e em  $\Theta_0$  [Estimador Restrito]. A estatística do TRV é obtida por:

$$\lambda(\mathbf{X}) = \frac{L(\hat{\theta}_0|\mathbf{X})}{L(\hat{\theta}|\mathbf{X})}. \quad (3.19)$$

Ilustraremos o teste da razão de verossimilhança na sequência através de um exemplo específico.

**Exemplo 3.2.2.** *Suponha  $X_1, \dots, X_n$  i.i.d. e assuma um modelo exponencial. Através de um TRV, deseja-se testar a hipótese  $H_0 : \theta = \theta_0$  contra  $H_0 : \theta \neq \theta_0$  com base nos valores observados  $\mathbf{x} = (x_1, \dots, x_n)$ . Sob o modelo considerado, a densidade marginal pode ser escrita como:  $f_{Exp(\theta)}(x|\theta) = \theta \exp\{-\theta x\} \mathbb{I}(x > 0)$ , com  $\Theta = (0, +\infty)$ . Sendo assim, a função de verossimilhança  $L(\theta)$  é dada por:*

$$L(\theta) = \prod_{i=1}^n f(x_i|\theta) \quad (3.20)$$

$$= \prod_{i=1}^n \theta \exp\{-\theta x_i\} \quad (3.21)$$

$$= \theta^n \exp\left\{-\theta \sum_{i=1}^n x_i\right\} \quad (3.22)$$

$$= \theta^n \exp\left\{-\theta n \bar{\mathbf{x}}\right\}. \quad (3.23)$$

Para encontrar a estatística  $\lambda(\mathbf{x}) = \frac{\sup_{\Theta_0} L(\theta|\mathbf{x})}{\sup_{\Theta} L(\theta|\mathbf{x})}$  precisamos maximizar a função acima para o caso Restrito no numerador, e para o caso Irrestrito no denominador. No primeiro caso é simples pois  $\Theta_0 = \{\theta_0\}$ , o que faz com que o a função no espaço restrito tenha somente um valor, ou ainda,  $\sup_{\Theta_0} L(\theta|\mathbf{x}) = L(\theta_0|\mathbf{x})$ .

Para avaliar o denominador é necessário maximizar a função  $L(\theta)$  com respeito a  $\theta$  e isto pode ser feito através das condições de primeira ordem. Com um pouco de manipulação algébrica verificamos que o argumento que maximiza a função de verossimilhança



é o ponto

$$\theta^* = \frac{1}{\bar{\mathbf{x}}}. \quad (3.24)$$

Temos então que

$$\lambda(\mathbf{x}) = \frac{\theta_0^n \exp\{-\theta_0 n \bar{\mathbf{x}}\}}{\left(\frac{1}{\bar{\mathbf{x}}}\right)^n \exp\{-n\}} \quad (3.25)$$

$$= \left(\frac{\theta_0}{\bar{\mathbf{x}}}\right)^n \exp\{n(1 - \theta_0 \bar{\mathbf{x}})\} \quad (3.26)$$

$$= \left(\frac{\theta_0}{\bar{\mathbf{x}}} \exp\{1 - \theta_0 \bar{\mathbf{x}}\}\right)^n. \quad (3.27)$$

Como pode ser visto em Casella [1], para grandes amostras  $-2\log(\lambda(\mathbf{X}))$  segue uma distribuição Qui-quadrado ( $\chi^2$ ).

### 3.3 Modelos Não Encaixados

A escolha de um modelo na prática é subjetiva. Por vezes, diferentes famílias de distribuições são capazes de representar relativamente bem os dados. Digamos que haja o interesse em modelar o tempo entre as falhas mecânicas apresentadas por uma determinada aeronave. Digamos que  $X_1, \dots, X_n$  denotam, respectivamente, as v.a.'s associadas ao tempo até a primeira falha, entre a primeira e a segunda falha e, assim, sucessivamente. Se é razoável admitir que  $X_1, \dots, X_n$  são i.i.d., então, a descrição do modelo deveria especificar apenas uma distribuição para  $X_1$ . Não seria absurdo assumir que as v.a.'s são contínuas e que poderiam assumir qualquer valor não negativo.

Um modelo -  $\mathcal{P}_0$  - razoável é adotar a família de distribuições exponenciais. Tal modelo é bastante empregado na prática em contextos similares, porém, será razoável apenas se a variância populacional for igual ao quadrado da média populacional - uma característica das distribuições exponenciais. Se não há indícios que tal característica se verifique, poderia se considerar um modelo alternativo  $\mathcal{P}_1$  - em que se supõe que  $X_1$  tenha distribuição *Gama*( $\nu, \varphi$ ) para algum par  $\nu > 0$  e  $\varphi > 0$ .

Sob o primeiro modelo, a densidade de  $X_1$ , por exemplo, poderia ser escrita como:

$$f_{Exp(\beta)}(x_1) = \beta \exp\{-\beta x_1\} \mathbb{I}(x_1 > 0), \quad (3.28)$$

com  $0 < \beta < +\infty$ .

Sob o segundo modelo, todavia, a densidade de  $X_1$  seria dada por:

$$f_{Gama(\nu, \varphi)}(x_1) = \frac{\varphi^\nu}{\Gamma(\nu)} x_1^{\nu-1} \exp\{-\varphi x_1\} \mathbb{I}(x_1 > 0) \quad (3.29)$$

com o vetor paramétrico  $\theta = (\nu, \varphi)$ ,  $0 < \nu, \varphi < \infty$ .

Note que a densidade de  $X_1$  no modelo  $\mathcal{P}_0$  é um caso particular da densidade de  $X_1$  no modelo  $\mathcal{P}_1$  quando  $\nu = 1$  e  $\varphi = \beta$ . De fato, as distribuições [elementos] que compõem o modelo  $\mathcal{P}_0$  pertencem ao modelo  $\mathcal{P}_1$  - ou seja  $\mathcal{P}_0 \subset \mathcal{P}_1$ .

Quando isso ocorre, dizemos que os modelos são encaixados [o modelo  $\mathcal{P}_0$  é um sub-modelo do modelo  $\mathcal{P}_1$ ]. Em casos como este, diz-se, usualmente, que  $\mathcal{P}_0$  é um modelo restrito do modelo [irrestrito]  $\mathcal{P}_1$ . Repare que o modelo  $\mathcal{P}_0$  pode ser obtido do modelo  $\mathcal{P}_1$  por restringirmos o espaço paramétrico. Basta definirmos  $\Theta = (0, +\infty) \times (0, +\infty)$  e  $\Theta_0 = \{1\} \times (0, +\infty)$ . Neste caso,  $\mathcal{P}_0 = \{P_\theta \in \mathcal{P}_1; \theta \in \Theta_0\}$ . Aqui  $\theta = (\nu, \varphi)$ .

Modelos encaixados são vistos com frequência na prática. Em princípio, para compará-los podemos utilizar, dentre outras alternativas, o Teste da Razão de Verossimilhança, tal como visto na seção anterior. A comparação se torna mais complexa, porém, quando os modelos não são encaixados.

Suponha agora, ilustrativamente, que  $\mathcal{P}_1$  continue denotando o modelo gama acima apresentado, porém, que este é confrontado com uma alternativa [diga-se de passagem, também razoável]  $\mathcal{P}_0$  redefinido como a classe das distribuições lognormais. Agora, as densidades associadas com o modelo  $\mathcal{P}_0$  assumiriam a forma:

$$f_{Lognormal(\mu, \sigma^2)}(x_1) = \frac{1}{x_1 \sqrt{2\pi\sigma^2}} \exp\left\{ \frac{1}{2\sigma^2} - (\log(x_1) - \mu)^2 \right\} \mathbb{I}(x_1 > 0). \quad (3.30)$$

Neste caso, não existe uma maneira de escrever uma classe de distribuições em função da outra restringindo o espaço paramétrico. Estaríamos lidando, neste caso, com dois modelos não encaixados.

**Definição 3.3.1.** *Modelos Não Encaixados são duas classes de distribuições tais quais uma não pode originar a outra por meio de restrições no espaço paramétrico.*

Considere agora a necessidade de um grande empresário em estudar suas vendas

diárias. Novamente, é aceitável supor que as variáveis  $X_i \forall i = 1, \dots, n$ , que representam a receita da empresa no dia  $i$ , são variáveis aleatórias i.i.d. e podem ser modeladas por meio de uma família Log-normal  $\mathcal{P}_0$ . Mesmo que o modelo admita para a v.a. somente valores no conjunto  $(0; +\infty)$  que não inclui o 0, seriam possíveis valores em sua vizinhança. Pelas mesmas características, seria aceitável o modelo  $\mathcal{P}_1$  Exponencial.

Ainda assim, outra família de distribuições que talvez representasse bem a aleatoriedade esperada para a receita seria Weibull, definindo  $\mathcal{P}_2$ . Se  $X_1, \dots, X_n$  admitem esse modelo, tem-se para  $X_1$  a seguinte função de densidade:

$$f_{Weibull(\gamma, \delta)}(x_1) = \frac{\gamma}{\delta^\gamma} x_1^{\gamma-1} \exp\left\{-\left(\frac{x_1}{\delta}\right)^\gamma\right\} \mathbb{I}(x_1 \geq 0), \quad (3.31)$$

onde o vetor  $\boldsymbol{\theta} = (\gamma, \delta) \in \Theta_2 = (0, +\infty) \times (0, +\infty)$ . Tal comparação intra-modelos é também sugerida e estudada em outros trabalhos como em Pereira [8]. Embora  $\mathcal{P}_1$  e  $\mathcal{P}_2$  sejam modelos não encaixados, uma restrição paramétrica da forma  $\Theta_2 := \{1\} \times (0, +\infty)$  faz das famílias Exponencial e Weibull encaixadas.

As classes de distribuições Exponencial, Log-Normal e Weibull são constantemente utilizadas nos ramos da saúde e industrial em análises de sobrevivência ou de confiabilidade, respectivamente, como visto no trabalho de Barbosa [9]. As três classes são utilizadas para modelar curvas que representam a morte (ou falha) de indivíduos (ou equipamentos). Sendo possível a utilização de diferentes modelos para este fim, faz-se necessário averiguar qual melhor se adequa em cada caso. Existem estudos específicos nessas áreas e não iremos nos aprofundar. Porém, isto mostra o quão comum são comparações de modelos não encaixados.

Como já dito, casos como este (de modelos não encaixados) inviabilizam o método da Razão de Verossimilhanças como método comparativo, ao menos quando empregado da maneira usual. Isso ocorre pois, nos casos de famílias não encaixadas, a distribuição exata da estatística do teste  $\lambda(\mathbf{X})$  (ou de uma função dela) é desconhecida, o que impossibilita a obtenção de quantis e definição de região de rejeição, por exemplo. Veremos por aqui uma possível solução através de um método computacional chamado *bootstrap*.

## 3.4 Método de Bootstrap

*“The Bootstrap is a computer-based method of statistical inference that can answer many real statistical questions without formulas”*

A frase acima, extraída do prefácio de Efron & Tibshirani [6] ilustra a essência do método de *bootstrap*, conforme esperamos explicar ao longo desta seção, ainda que de forma sucinta.

Como antes, assumimos que uma amostra de v.a.'s  $Y_1, \dots, Y_n$  i.i.d. de uma população desconhecida  $P^*$  é considerada e que há o interesse em realizar inferências sobre um parâmetro [também desconhecido]  $\theta^* = \theta(P^*)$ . Repare que, como convencionado,  $\theta^*$  é o valor assumido pela função  $\theta(\cdot)$  quando o argumento é a distribuição verdadeira  $P^*$ . A função  $\theta$  pode representar, por exemplo, o inverso do valor esperado da distribuição:

$$\theta = \frac{1}{E[Y]}, \text{ se } Y \sim P. \quad (3.32)$$

No exemplo dado anteriormente em que, por hipótese,  $Y_1, \dots, Y_n$  possuem distribuição exponencial, a função  $\theta$  é definida via:

$$\theta(P_\theta) = \left( \int_{-\infty}^{\infty} y f_{Exp(\theta)}(y) dy \right)^{-1} = \left( \int_0^{\infty} y (\theta y^{-\theta y}) dy \right)^{-1}. \quad (3.33)$$

Se a distribuição  $P^*$  fosse conhecida [ou de forma equivalente, a densidade associada  $f_{Exp(\theta^*)}$ ], então, conheceríamos o valor verdadeiro de  $\theta$ , que é  $\theta^*$ .

Na ausência de conhecimento sobre a verdadeira distribuição, coletamos dados  $y_1, \dots, y_n$  que, por hipótese, são realizações de  $Y_1, \dots, Y_n$  e utilizamos alguma estatística  $T = T(Y_1, \dots, Y_n)$  para obter uma estimativa de  $\nu^* \equiv \frac{1}{\theta^*}$ , denotada por  $t = T(y_1, \dots, y_n)$ . Digamos que para uma certa amostra observada  $t$  assume o valor 1,0325. Algumas perguntas naturais aparecem: i) “a estimativa gerada é boa?”; ii) “podemos dizer que o valor verdadeiro de  $\nu^*$  é 1?”; iii) “o valor verdadeiro de  $\nu^*$  deve estar entre 1 e 1,05?”.

Formalmente, as perguntas acima são reformuladas pelos estatísticos utilizando, respectivamente, conceitos de “precisão”, “testes de hipótese” e “intervalos de confiança”. Ocorre que para avaliar a precisão da estimativa, realizar um teste de hipótese ou construir um intervalo de confiança é necessário conhecer [ou aproximar, estimar] a distribuição da estatística  $T$  ou algumas características suas. Para ilustrar tal afirmação utilizaremos, a seguir, um exemplo ainda mais simples do que o acima.

Considere que  $Z_1, \dots, Z_n$  são i.i.d. e que  $Z_1 \sim N(\mu^*, 1)$ . Repare que quando a variância da distribuição normal é conhecida [igual a 1, neste caso], basta conhecer a

média para definir a distribuição. O parâmetro agora é denotado por  $\mu$ . Temos, então:

$$\mu = \mu(N(\mu, 1)) = \int_{-\infty}^{+\infty} z \left( \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-\mu)^2} \right) dz. \quad (3.34)$$

Sabemos que a média amostral,  $\bar{Z} = \frac{Z_1 + \dots + Z_n}{n}$  gera boas estimativas do valor esperado de  $\mu^*$ , se, de fato,  $Z_1, \dots, Z_n$  são i.i.d., com  $Z_1 \sim N(\mu^*, 1)$ . De fato,  $\bar{Z}$  é um estimador não viesado [no sentido de que  $E(\bar{Z}) = \mu^*$ ], sua precisão é dada por  $n [Var(\bar{Z}) = \frac{Var(Z_1)}{n} = \frac{1}{n}]$  e consistente no sentido de que  $\bar{Z}$  converge em probabilidade para  $\mu^*$ . Se uma estimativa de  $\bar{Z}$ , baseada em uma amostra observada, assume o valor 0,0325, tais propriedades não são suficientes para garantir que o verdadeiro valor de  $\mu^*$  é diferente ou igual a 0 [sequer igual a 0,0325]. Como sabemos, porém, soma de normais independentes também possui distribuição normal e, dessa forma:

$$\bar{Z} \sim N\left(\mu^*, \frac{1}{n}\right). \quad (3.35)$$

Repare que, por conhecer a distribuição de  $\bar{Z}$  [ao menos de um número desconhecido  $\mu^*$ ] podemos avaliar se  $\mu^* = 0$  é uma hipótese sustentada pelos dados observados ou não. Sob a hipótese  $H_0 : \mu^* = 0$ , escrevemos:

$$\bar{Z} \sim N\left(0, \frac{1}{n}\right), \quad (3.36)$$

o que implica que, quando o tamanho da amostra for  $n = 258^2 = 38\,416$  [ilustrativamente] temos:

$$P\left(\left|\frac{\bar{Z}}{\sqrt{\frac{1}{258^2}}}\right| \leq 2,58\right) = 99\% \text{ ou } P(|\bar{Z}| > 0,01) = 1\%. \quad (3.37)$$

Ora, então, ao observarmos a estimativa 0,0325 da média amostral, ou nos deparamos com uma amostra rara ou a hipótese nula  $H_0 : \mu^* = 0$  não é razoável.

O exemplo simplório apresentado acima ilustra a aplicação de um teste de hipótese onde a hipótese nula  $H_0 : \mu^* = 0$  deveria ser rejeitada. Além disso, poderíamos dizer que 0,0325 é uma boa estimativa do verdadeiro valor de  $\mu^*$ , pois, a variância de  $\bar{X}$  é bem pequena  $\frac{1}{258^2}$  e seu valor esperado é  $\mu^*$ . É fácil ver ainda que  $[0,0225 ; 0,0425]$  é um intervalo de confiança com coeficiente de confiança de 99% para o verdadeiro valor  $\mu^*$ .

Infelizmente, para outras estatísticas e modelos a realização das análises feitas acima pode ser bem mais complexa. Voltando, ilustrativamente, ao exemplo inicial em que

$Y_1, \dots, Y_n$  possuem distribuição exponencial com parâmetro  $\theta^*$ , alguém poderia sugerir a adoção da estatística  $T = \bar{Y}$  para a estimação de  $\lambda^* \equiv \frac{1}{\theta^*}$ , pois, como vimos na seção 3.2,  $(\bar{Y})^{-1}$  é EMV de  $\theta^*$  e, pela invariância do método de máxima verossimilhança,  $\bar{Y}$  é EMV de  $\lambda^* \equiv \frac{1}{\theta^*}$ . Como verificado por Bibinger [10], não é fácil derivar a densidade de  $\sum_{i=1}^n Y_i$  e, portanto, de  $T = \bar{Y}$ , o que dificulta a realização de testes de hipótese envolvendo  $\lambda^*$  ou a construção de intervalos de confiança para tal parâmetro, por exemplo. Poderíamos, se o tamanho da amostra fosse suficientemente grande, apelar para o Teorema Central do Limite [ver James [11]]. Contudo, e se  $n$  não for tão grande assim? Uma alternativa está na adoção do procedimento de *bootstrap*.

Como apontado em Wasserman [12] o *bootstrap* permite estimar, através de simulações, a distribuição de uma estatística  $T = T(X_1, \dots, X_n)$  ou qualquer característica sua - em particular, quantis que seriam úteis para a construção de intervalos de confiança para o parâmetro estimado por  $T$  ou a realização de testes de hipótese envolvendo tal parâmetro. Para entender como o procedimento é implementado, suponha que  $X_1, \dots, X_n$  são v.a.'s i.i.d. com distribuição comum  $P_{\theta^*}$ . Seja  $\psi^*$  uma característica da distribuição verdadeira da estatística  $T$  que pode ser escrita como  $\psi^* = \psi(F_T^*)$ , onde  $F_T^*$  é a F.D.A. verdadeira de  $T$ . O *bootstrap* é realizado, então, da seguinte maneira:

**Passo 1:** Obtenha  $\hat{P}$ , uma estimativa de  $P$ , a distribuição de  $X_1$ , com base nos valores observados  $x_1, \dots, x_n$ ;

**Passo 2:** Sorteie [simule] uma amostra aleatória simples  $X_1^{boot}, \dots, X_n^{boot}$  da distribuição estimada  $\hat{P}$  - esta é chamada de “amostra de *bootstrap*”;

**Passo 3:** Compute  $t^*$  via  $t^* = T(x_1^{boot}, \dots, x_n^{boot})$  sendo  $x_1^{boot}, \dots, x_n^{boot}$  os valores obtidos na simulação para  $X_1^{boot}, \dots, X_n^{boot}$ ;

**Passo 4:** Repita apenas os Passos 2 e 3 acima por mais  $B - 1$  vezes e armazene os diferentes valores computados  $T(x_1^{boot}, \dots, x_n^{boot})$  no vetor:

$$t_1^{boot}, \dots, t_B^{boot}; \quad (3.38)$$

**Passo 5:** Finalmente, estime  $\psi^*$  via

$$\hat{\psi} = \psi \left( \hat{F}_T^{boot} \right), \quad (3.39)$$

$$\text{onde } \hat{F}_T^{boot}(t) = \frac{1}{B} \sum_{r=1}^B \mathbb{I}(t_r^{boot} \leq t) \quad (3.40)$$

$$\text{é a F.D.E. baseada em } t_1^{boot}, \dots, t_B^{boot}. \quad (3.41)$$

Como apontado em Efron & Tibshirani [6], se o modelo para  $X_1, \dots, X_n$  é paramétrico<sup>2</sup> da forma  $\mathcal{P} = \{P_\theta; \theta \in \Theta\}$ , então, no Passo 1 pode-se obter  $\hat{P}$  via  $\hat{P} = P_{\hat{\theta}}$ , sendo  $\hat{\theta}$  uma estimativa do verdadeiro parâmetro  $\theta^*$ . Neste caso, dizemos que o **Bootstrap** é **paramétrico** e as simulações do Passo 2 podem ser feitas por utilizar algum algoritmo computacional de simulação quando a distribuição é conhecida. Se não há um modelo paramétrico para  $X_1, \dots, X_n$  [ou não há um modelo confiável ou é difícil gerar as simulações computacionalmente], então, pode-se lançar mão do **Bootstrap não paramétrico** em que  $\hat{P}$  é a distribuição gerada pela F.D.E. construída a partir dos dados originais  $x_1, \dots, x_n$ . Neste segundo caso, o passo dois consiste apenas em sortear, com reposição [reamostrar] valores da amostra observada  $x_1, \dots, x_n$ . Em Efron & Tibshirani [6] e Davison & Hinkley [13] é possível encontrar uma série de exemplos e aplicações do método de bootstrap em cada um dos contextos [paramétrico e não paramétrico], mas, a essência é sempre a mesma: construir uma distribuição para a estatística  $T = T(X_1, \dots, X_n)$  com base em  $B$  diferentes valores encontrados para ela<sup>3</sup>, da forma  $t^{boot} = T(x_1^{boot}, \dots, x_n^{boot})$  em que  $x_1^{boot}, \dots, x_n^{boot}$  correspondem a valores sorteados ou simulados da distribuição estimada de  $X_1$ .

Devemos reforçar que a função  $T(\dots)$  é conhecida. Então, implementar o Passo 3 é relativamente simples, mesmo que a função seja complexa e que seja difícil avaliar a distribuição exata de  $T$ . Se  $T$  fosse, por exemplo, a média amostral [utilizada nas ilustrações feitas acima], então, no passo 3 definiríamos:  $t^* = \frac{1}{n} (x_1^{boot} + \dots + x_n^{boot})$ . Da mesma forma, a reconstrução da distribuição de  $T$  via *bootstrap* que aparece no Passo 5 poderia ser feita sempre da mesma forma, independentemente de se os dados  $x_1, \dots, x_n$  são gerados por uma distribuição exponencial, normal, ou outra. Bastaria computar a F.D.E. baseada em  $t_1^{boot}, \dots, t_B^{boot}$ , denotada por  $\hat{F}_T^{boot}$ . Repare ainda que no Passo 5  $\hat{\psi}$  é estimado por aplicar a função  $\psi$  em  $\hat{F}_T^{boot}$ . Por exemplo, se  $\psi(F)$  denota a variância associada a distribuição cuja F.D.A. é  $F$ ,  $\hat{\psi} = \psi \left( \hat{F}_T^{boot} \right)$  é obtido pelo cálculo da variância de uma

<sup>2</sup>Paramétrico, no sentido de que  $\Theta$  tem dimensão finita como aparece em Bickel e Doksum, 2001.

<sup>3</sup>Tais valores também são chamados de replicações de *Bootstrap*.

v.a. que possui F.D.A. exatamente igual a  $\widehat{F}_T^{boot}$ , ou seja:

$$\frac{1}{B} \sum_{r=1}^B \left( t_r^{boot} - \overline{t^{boot}} \right)^2, \text{ onde } \overline{t^{boot}} = \frac{1}{B} \sum_{r=1}^B t_r^{boot}. \quad (3.42)$$

De maneira similar, podemos computar uma estimativa do  $p$ -quantil da distribuição de  $T$  por considerar o quantil empírico baseado em  $t_1^{boot}, \dots, t_B^{boot}$ .

Não é difícil reproduzir a análise para estatísticas mais complexas [por exemplo, a estatística de teste da razão de verossimilhança]. Além disso, a complexidade do procedimento pouco depende da complexidade dos modelos.

## 3.5 Testes baseados na Razão de Verossimilhança no contexto de Modelos Não Encaixados

O que dificulta, por vezes, a análise do TRV que compara modelos não encaixados é a incapacidade da aproximação de uma distribuição de probabilidade à estatística do teste, vista na equação 3.19, ou a uma função desta. Entretanto, para representar a comparação modelos não encaixados, essa estatística será denotada de forma diferente. Suponha um TRV que compara dois modelos  $\mathcal{P}_0$  e  $\mathcal{P}_1$  não encaixados cujas respectivas f.d.p.'s marginais que os representam são  $f(\cdot|\theta)$  e  $g(\cdot|\beta)$ . Dessa forma, um TRV que confronta as hipóteses  $H_0 : P^* = \mathcal{P}_0$  contra  $H_1 : P^* = \mathcal{P}_1$  poderia ter como estatística do teste

$$\lambda(\mathbf{X}) = \frac{L_f(\widehat{\theta})}{L_g(\widehat{\beta})} = \frac{f(X_1|\widehat{\theta})_{\mathbf{X}} \dots_{\mathbf{X}} f(X_n|\widehat{\theta})}{g(X_1|\widehat{\beta})_{\mathbf{X}} \dots_{\mathbf{X}} g(X_n|\widehat{\beta})}, \quad (3.43)$$

onde  $\widehat{\theta}$  e  $\widehat{\beta}$  são EMV's. Denotaremos  $\ell(\theta) = \log(L(\theta))$  e, por consequência,  $\log(\lambda(\mathbf{X})) = \ell_f(\widehat{\theta}) - \ell_g(\widehat{\beta})$ . Nesta seção discutiremos duas abordagens possíveis para comparar modelos não encaixados através de tal estatística.

### 3.5.1 A solução de Cox

Cox [3] propõe uma alternativa baseada em uma transformação da estatística de teste. A fim de aplicar tais resultados ao TRV e encontrar uma distribuição conhecida de  $\lambda(\mathbf{X})$ , Cox faz uso de transformações à estatística da forma

$$T(\mathbf{X}) = \ell_f(\widehat{\theta}) - \ell_g(\widehat{\beta}). \quad (3.44)$$



Para chegar a estatística do teste, Cox inicia definindo os estimadores de máxima verossimilhança  $\hat{\theta}$  e  $\hat{\beta}$ . Em seguida, computa as *log*-verossimilhanças  $[\ell_f(\hat{\theta})$  e  $\ell_g(\hat{\beta})]$  de cada modelo em função de seu respectivo EMV. Considere a estatística dada pela equação 3.44. A ideia do método é aproximar uma distribuição Normal [que depende dos parâmetros média e variância] à essa estatística. Logo, o terceiro passo [potencialmente problemático, como sugere o autor] seria obter uma aproximação da média de  $\ell_f(\hat{\theta}) - \ell_g(\hat{\beta})$  via  $E_{\hat{\theta}}[\ell_f(\hat{\theta}) - \ell_g(\hat{\beta})]$  [valor esperado calculado sob a função  $\ell_f(\hat{\theta})$ ]. Por último, deve ser calculada a variância assintótica da estatística  $T(\mathbf{X}) - E_{\hat{\theta}}[\ell_f(\hat{\theta}) - \ell_g(\hat{\beta})]$ .

De tal forma, o autor chega a um resultado que designa a  $T(\mathbf{X})$  uma distribuição assintótica Normal que não depende do parâmetro, com média 0 e variância (ou matriz de covariâncias, no caso multivariado) também estimável. É notável a complexidade existente em alguns cálculos. Calcular, por exemplo, o valor esperado que a transformação necessita se torna muito difícil para alguns modelos ou distribuições.

Em um de seus trabalhos, Pereira [5] gera resultados do método de Cox. O autor encontra, por meio de simulações, o poder empírico do teste para alguns pares de modelos não encaixados. Dentre estes, há o teste que confronta a hipótese nula, de pertencer a um modelo Exponencial, com a hipótese alternativa, que afirma que a família Lognormal representa melhor os dados. Computam-se poderes do teste para quando os dados são provenientes de distribuições *Lognormal*(0; 0, 2), *Lognormal*(0; 1) e *Lognormal*(0; 2). Mostraremos a seguir os poderes empíricos encontrados por Pereira [5], dado um nível de significância de 5%. Comentaremos à respeito destes resultados no Capítulo 4.

Tabela 2 – Poder do teste de Cox: Exponencial x Lognormal

$n$	<i>Lognormal</i> (0; 0, 2)	<i>Lognormal</i> (0; 1)	<i>Lognormal</i> (0; 2)
20	100%	37,2%	55,6%
50	100%	82,6%	90,6%
100	100%	99,4%	98,6%
150	100%	100%	100%
200	100%	100%	100%

### 3.5.2 O Algoritmo de Bootstrap aplicado a Famílias Não Encaixadas

O trabalho estuda uma alternativa à abordagem de Cox [já proposta por Pesaran [7]] utilizando o método de *bootstrap*, visto na seção 3.4. O procedimento faz jus a uma eventual impossibilidade na aproximação de uma distribuição conhecida à estatística do

teste. O utilizaremos para estimar quantis da distribuição de  $T(\mathbf{X})$  por simulações e, portanto, poder responder se a estimativa da estatística de teste obtida seria ou não discrepante, considerando sua distribuição empírica. Em outras palavras, se deveríamos ou não rejeitar a hipótese nula.

Suponha a mesma comparação  $H_0$  x  $H_1$ . Uma realização da amostra serve como base para estimar  $\hat{\theta}$ , estimador de máxima verossimilhança de  $\theta$ . Este procedimento faz possível, por meio da abordagem paramétrica do *bootstrap*, chegar a uma distribuição estimada para, através dessa distribuição, simular amostras e obter estimativas da estatística do teste. Com base nas estimativas, obtém-se a distribuição empírica da estatística. Avanços tecnológicos possibilitam simular amostras de distribuições desejadas, algo que é crucial para o desenvolvimento do método de *bootstrap* paramétrico e o inviabilizava décadas atrás. O algoritmo a seguir mostra o passo-a-passo para a implementação teste baseado no método.

1. Com base nos dados  $\mathbf{x}$ , uma realização de  $\mathbf{X}$ , calcular  $\hat{\theta}$  e  $\hat{\beta}$ , e definir uma estimativa  $T(\mathbf{x})$ ;
2. Simular  $B$  amostras de *bootstrap* de tamanho  $n$  a partir da distribuição associada a  $f(\cdot, \hat{\theta})$ . Cada amostra será denotada por  $\mathbf{x}_b^*$ ,  $b = 1, \dots, B$ ;
3. Para cada amostra  $\mathbf{x}_b^*$  estimar  $\hat{\theta}_b$  e  $\hat{\beta}_b$  por máxima verossimilhança e obter uma estimativa de *bootstrap* [replicação] da estatística  $T(\mathbf{X})$ , denotada por  $T_b^*$ ;
4. Considere  $Q_p(\mathbf{v})$  uma função que calcula o  $p$ -quantil do vetor  $\mathbf{v}$ . Com base no vetor de replicações  $\mathbf{t}^* = T_1^*, \dots, T_B^*$ , a região de rejeição  $RR$  do teste, utilizando um nível de significância de  $\alpha$ , é da forma:

$$RR = \begin{cases} (-\infty; Q_\alpha(\mathbf{t}^*)] & , \text{ se o teste é unilateral à esquerda;} \\ (-\infty; Q_{\alpha/2}(\mathbf{t}^*)] \cup [Q_{1-\alpha/2}(\mathbf{t}^*); +\infty) & , \text{ se o teste é bilateral} \end{cases}$$

## 4 Análise dos Resultados

Os testes em questão foram implementados e avaliados, como já dito, através de simulações. Para cada teste, a hipótese  $H_0$  considera a distribuição dos dados proveniente de uma determinada família de distribuição  $\mathcal{P}_0$ . Em contrapartida, a hipótese  $H_1$  sugere que a distribuição é original de outra classe de distribuições  $\mathcal{P}_1$ . E tais distribuições ( $P_\psi \in \mathcal{P}_0$  e  $P_\tau \in \mathcal{P}_1$ ), indexadas por seus respectivos vetores paramétricos  $\psi$  e  $\tau$ , possuem f.d.p [eventualmente, funções de probabilidade],  $f$  e  $g$  respectivamente. Na condição dos modelos em teste serem separados, independente da restrição a  $\psi$  ou a  $\tau$ ,  $f$  não pode ser escrita como  $g$  ou o contrário para escolhas arbitrárias de  $\psi$  e  $\tau$ .

Iniciamos pela simulação de amostras, escolhida uma distribuição para os dados. Tal distribuição foi escolhida dentre os dois modelos sugeridos pelas hipóteses nula e alternativa e associada a um vetor paramétrico ( $\psi_0$  ou  $\tau_0$ ) pré-definido. 200 amostras foram simuladas de cada distribuição escolhida e para cada uma delas foi implementado o teste via *bootstrap*. Admitindo que não conhecemos a procedência da amostra, damos início ao teste.

Para cada amostra simulada, estimou-se, via Máxima Verossimilhança, valores dos parâmetros associados aos modelos  $\mathcal{P}_0$  e  $\mathcal{P}_1$ . O pacote *bml* do *R* foi utilizado para esta finalidade. A estimativa da estatística de teste utilizada depende destas funções aplicadas ao seu respectivo EMV. Dessa forma, foi possível obter uma estimativa da estatística do teste calculada por:

$$t_{calc} = T(\mathbf{x}) = \ell_0(\hat{\psi}) - \ell_1(\hat{\tau}), \quad (4.1)$$

sendo  $L_f$  [ $L_g$ ] a verossimilhança associada a densidade  $f$  [respectivamente,  $g$ ],  $\ell_0(\cdot) = \log(L_f(\cdot))$  e  $\ell_1(\cdot) = \log(L_g(\cdot))$ .

O passo seguinte foi a utilização do *bootstrap* para encontrar uma distribuição empírica da estatística acima, ou seja, a implementação do algoritmo visto em 3.5.2 no contexto do teste. Foram sorteadas  $B$  amostras de tamanho  $n$  da distribuição  $P_{\hat{\psi}} \in \mathcal{P}_0$ . Para

cada amostra  $\mathbf{x}_b$  de *bootstrap* [ $b = 1, \dots, B$ ] foram calculados os EMV's  $\hat{\psi}_b$  e  $\hat{\tau}_b$  com base nos modelos  $\mathcal{P}_0$  e  $\mathcal{P}_1$ , respectivamente, e uma estimativa  $T(\mathbf{x}_b) = \ell(\hat{\psi}_b) - \ell(\hat{\tau}_b)$ . Considere o vetor  $\mathbf{t} = (t_1 = T(\mathbf{x}_1), t_2 = T(\mathbf{x}_2), \dots, t_B = T(\mathbf{x}_B))$ . Calculamos os devidos quantis destas estimativas de *bootstrap* [vetor  $\mathbf{t}$ ], para definir a região de rejeição do teste [conforme visto no último item do algoritmo em 3.5.2] e decidimos, baseados na estimativa  $t_{calc}$ , se rejeitaríamos ou não  $H_0$ . Por exemplo, se o teste fosse unilateral à esquerda de significância  $\alpha$ , a região de rejeição seria da forma  $RR = (-\infty; Q_\alpha\{\mathbf{t}\}]$ , com  $Q_\alpha\{\mathbf{t}\}$  sendo  $\alpha$ -quantil empírico baseado nos valores  $t_1, \dots, t_B$ . Assim, se  $t_{calc}$  fosse menor que  $Q_\alpha\{\mathbf{t}\}$  rejeitaríamos  $H_0$ .<sup>1</sup>

Para cada tipo de teste (bilateral e unilateral à esquerda) a região de rejeição é feita de uma forma. Dois tipos de teste, bilateral e unilateral à esquerda, foram avaliados neste trabalho para entender qual melhor se adequaria. Tendo escolhido o modelo  $\mathcal{P}_1$  para gerar cada uma das 200 amostras, conseguimos avaliar “empiricamente” a probabilidade de cometer erro do tipo II - que decresce com a potência do teste - ao menos para algumas especificações de parâmetros em  $\mathcal{P}_1$ .

## 4.1 Lognormal x Gama

Nesta primeira situação as famílias separadas em teste foram lognormal e gama, onde  $H_0$  diz que os dados são provenientes do primeiro modelo [lognormal] e  $H_1$  afirma o segundo modelo [gama] como o verdadeiro. Consideramos as funções de densidade  $f$  e  $g$  que descrevem, respectivamente, uma  $Gama(\beta_1; \beta_2)$  e uma  $Lognormal(\theta_1; \theta_2^2)$  da forma:

$$f(x_i) = \frac{1}{x_i \sqrt{2\pi\theta_2^2}} \exp\left\{ \frac{1}{2\theta_2^2} - (\log(x_i) - \theta_1)^2 \right\} \mathbb{I}(x_i > 0) \quad (4.2)$$

$$\text{e} \quad (4.3)$$

$$g(x_i) = \frac{1}{\beta_2^{\beta_1} \Gamma(\beta_1)} x_i^{\beta_1-1} \exp\left\{ -\frac{x_i}{\beta_2} \right\} \mathbb{I}(x_i > 0). \quad (4.4)$$

No teste proposto em Cox [3], é utilizada uma distribuição assintótica normal para a estatística de teste padronizada. Para checar se os quantis da distribuição normal poderiam ser utilizados [no lugar dos quantis obtidos via bootstrap], analisamos a distribuição das replicações de Bootstrap em alguns casos de forma ilustrativa. Para isso, selecionamos três exemplos de amostras simuladas com distribuição Lognormal e realizamos o teste uma vez para cada uma afim de obter a distribuição de 200 estimativas da estatística do

<sup>1</sup>O teste bilateral possui região de rejeição da forma  $RR = (-\infty; Q_{\alpha/2}\{\mathbf{t}\}] \cup [Q_{1-\alpha/2}\{\mathbf{t}\}; +\infty)$ .

teste calculadas sobre a hipótese nula. As Figuras 1 e 2 ilustram os resultados obtidos.

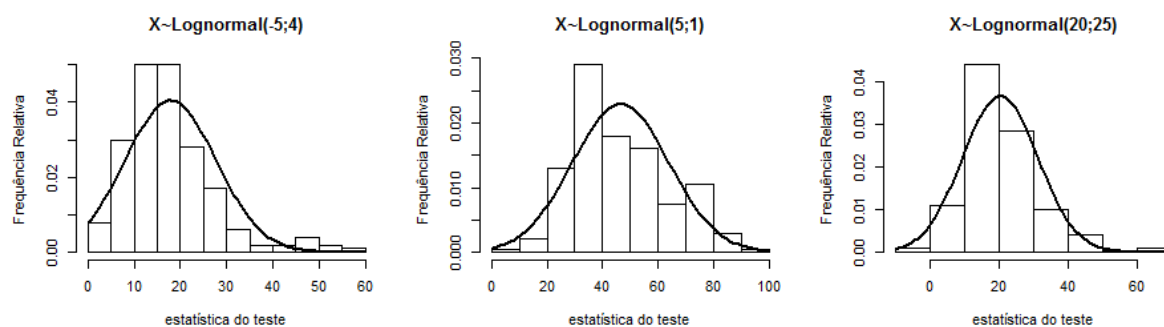


Figura 1 – Histogramas da Distribuição da Estatística o Teste quando  $X_i \sim \text{Lognormal}(\theta_1; \theta_2^2)$

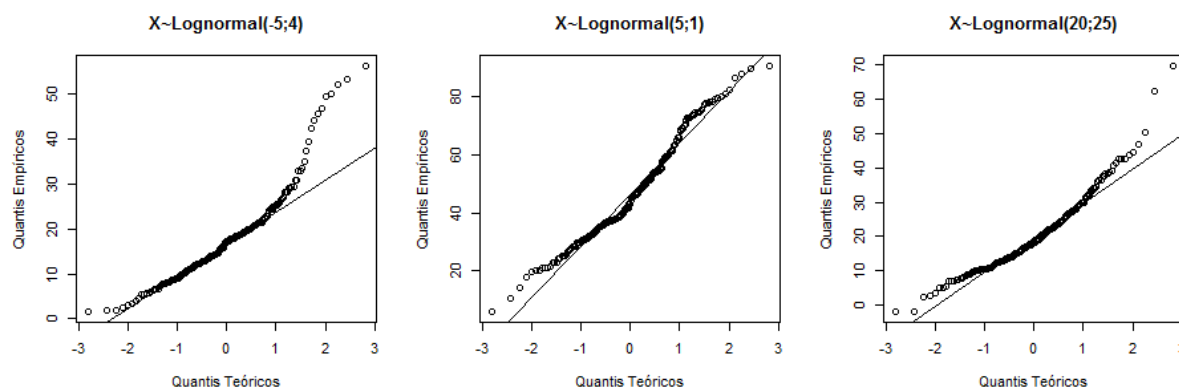


Figura 2 – Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando  $X_i \sim \text{Lognormal}(\theta_1; \theta_2^2)$

Os histogramas da Figura 1 ilustram a distribuição estimada via *Bootstrap* para a estatística de teste em cada caso. As curvas sobrepostas simulam a distribuição Normal mais próxima. Ainda que alguns dos histogramas se adequem à curva na parte central, apresentam uma assimetria considerável, o que já descaracteriza o modelo. Além disso, na Figura 2 vemos gráficos que comparam os quantis das estimativas com os da respectiva distribuição Normal. Observamos uma grande discrepância das distribuições obtidas via *bootstrap* com a normal nas caudas, indicada pela distância dos pontos à reta em suas extremidades. Ressalta-se que para a aplicação do teste as caudas são de suma importância, pois, as utilizamos para definir a regra de rejeição do teste.

A fim de avaliar empiricamente a probabilidade do erro tipo II nesta configuração do teste, realizamos o teste 200 vezes e, para cada uma das 200 amostras simuladas de uma distribuição  $Gama(\beta_1; \beta_2)$ , avaliamos em quantas destas não se rejeitou [erroneamente] a hipótese nula. Para isso, comparamos, em cada caso, a estimativa da estatística do teste

da amostra original com outras  $B = 200$  estimativas de amostras de *bootstrap*, originais de uma  $Lognormal(\hat{\theta}_1; \hat{\theta}_2^2)$  (distribuição estimada da hipótese nula, onde  $\hat{\theta}_1$  e  $\hat{\theta}_2$  são os EMV's de  $\theta_1$  e  $\theta_2$ , respectivamente).

Para simular os dados usamos duas configurações. Uma em que a média populacional foi mantida constante [igual a 1] e outra em que a variância foi mantida constante [igual a 1]. Sabe-se que em situações do cotidiano não é regra ocorrer tais restrições. Porém, propriedades de transformação da distribuição Gama tornam padronizações viáveis. Ao multiplicarmos uma Gama pelo inverso da média, produzimos uma nova Gama com valor esperado 1. Se, ao invés disso, multiplicamos a Gama pelo inverso de seu desvio-padrão, geramos uma nova Gama com variância unitária. A seguir, nas Tabelas 3 e 4 exibimos os erros tipo II adquiridos para cada parametrização e tamanho da amostra da distribuição Gama dos dados.

Tabela 3 – Erros do Tipo II para o teste Lognormal x Gama com variância = 1

$n$	$\beta_1$	$\beta_2$	Bilateral	Unilateral à Esquerda
150	0,04	5	0%	0%
150	0,16	2,5	0%	0%
150	0,75	16/9	4%	3%
100	1	1	3%	1,5%
200	4	0,5	20,5%	12%
300	4	0,5	7,5%	5,5%
100	25	0,2	84%	75%
200	25	0,2	78,5%	74%
300	25	0,2	67,5%	54%
400	25	0,2	65,5%	53,5%
500	25	0,2	55,5%	42,5%
1000	25	0,2	27,5%	20%
1500	25	0,2	12,5%	5,5%

Tais tabelas mostram a probabilidade empírica de ocorrer um erro do tipo II com base em uma amostra de tamanho  $n$  de uma distribuição  $Gama(\beta_1; \beta_2)$ . Além da dificuldade já mencionada no cálculo da estatística do teste, percebemos que o teste erra mais quando o parâmetro  $\beta_1$  é mais discrepante. Efeito esse que pode ser diminuído com um aumento significativo no tamanho da amostra. Porém, nos primeiros valores das duas tabelas, quando  $\beta_1$  assume valores “pequenos”, o teste apresenta resultados mais satisfatórios. Vemos, em casos com  $\beta_1 < 5$ , erros menores que 20%, mesmo com amostras de tamanho 100. Em todos os casos, o teste cometeu mais erros ou errou de forma equivalente no teste bilateral do que no unilateral à esquerda.

Tabela 4 – Erros do Tipo II para o teste Lognormal x Gama com média = 1

$n$	$\beta_1$	$\beta_2$	Bilateral	Unilateral à Esquerda
100	0,2	5	0%	0%
100	0,5	2	0%	0%
100	0,75	4/3	0%	0%
100	1	1	3%	1,5%
100	1,25	0,8	3,5%	0%
100	1,5	2/3	13%	7%
150	1,5	2/3	3,5%	0,5%
100	2	0,5	22,5%	13,5%
200	2	0,5	4,5%	0%
100	5	0,2	53%	41,5%
200	5	0,2	29%	19%
300	5	0,2	13,5%	10%
400	5	0,2	6,5%	3%

## 4.2 Gama x Lognormal

Nesta próxima seção abordaremos um teste parecido com o da seção anterior, porém, com a inversão de suas hipóteses. Agora, a hipótese nula afirma que a população segue um modelo gama, enquanto a alternativa admite o modelo lognormal. Assim como na situação anterior, calculamos erros do tipo II empiricamente baseado em 200 execuções do teste na mesma configuração, e, também, utilizando  $B = 200$  replicações de *bootstrap* para verificar a distribuição da estatística do teste. Agora, tivemos de escolher pares  $\theta = (\theta_1; \theta_2)$  [média e desvio-padrão] e computar os erros do tipo II dada uma amostra de cada distribuição  $Lognormal(\theta_1; \theta_2^2)$ . Admita a mesma parametrização utilizada anteriormente e as mesmas funções de densidade para cada classe de distribuições.

Tabela 5 – Erros do Tipo II para o teste Gama x Lognormal

$n$	$\theta_1$	$\theta_2$	Bilateral	Unilateral à Esquerda
100	-5	5	0%	0%
100	-1	1,5	0,5%	0,5%
100	-0,5	1,5	0%	0%
100	2	0,5	51,5%	37%
200	2	0,5	20,5%	14%
300	2	0,5	7,5%	4,5%
100	2	1	8%	5%
100	5	2	0,5%	0,5%
100	10	3	0%	0%
100	30	5	0%	0%

Os resultados mostram uma variação bem pequena no número de erros quando va-

riamos a média do logaritmo dos dados  $\theta_1$ . Em contrapartida, o desvio-padrão aparenta influenciar fortemente, sendo inversamente proporcional aos erros. Além disso, valores para  $\theta_2$ , quando menores que 1, chegam a induzir cometermos erros do tipo II em mais de 50% dos casos para amostras de tamanho 100. O tamanho da amostra também tem efeito positivo em relação ao erro, como já era esperado.

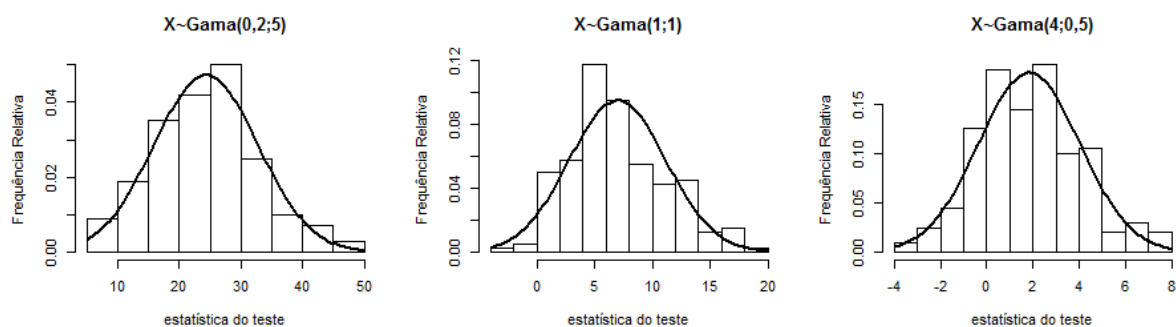


Figura 3 – Histogramas da Distribuição da Estatística o Teste quando  $X_i \sim Gama(\beta_1; \beta_2)$

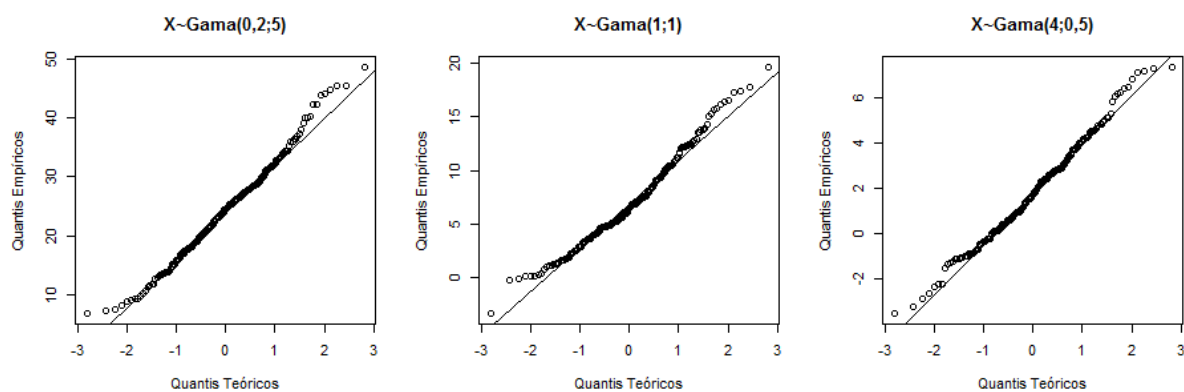


Figura 4 – Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando  $X_i \sim Gama(\beta_1; \beta_2)$

Analisamos, novamente, a distribuição da estatística de teste, agora, com dados provindos de distribuições Gama. Na Figura 3 a distribuição estimada via *Bootstrap* aparentemente não é bem aproximada pela distribuição normal. Mais uma vez, conforme pode ser visto na Figura 4, a discrepância é ainda maior nas caudas.

### 4.3 Geométrica x Poisson

Para comparar modelos discretos, realizamos o mesmo procedimento testando uma distribuição Geométrica contra uma Poisson. Temos que uma variável aleatória que possui a distribuição da hipótese nula [Geométrica] conta o número de “fracassos” obtidos antes



de ocorrer o primeiro “sucesso”. Enquanto o modelo Poisson conta quantas ocorrências de um evento de interesse ocorrem em um intervalo de tempo. As funções de probabilidade  $f$  e  $g$  que descrevem as famílias  $Geo(p)$  e  $Poi(\lambda)$ , respectivamente, aqui serão escritas por:

$$f(x_i) = p(1-p)^{x_i} \mathbb{I}\{x_i \in \mathbb{Z}^*\} \quad (4.5)$$

$$\text{e} \quad (4.6)$$

$$g(x_i) = \frac{\lambda^{x_i} \exp\{-\lambda\}}{x_i!} \mathbb{I}\{x_i \in \mathbb{Z}^*\} \quad (4.7)$$

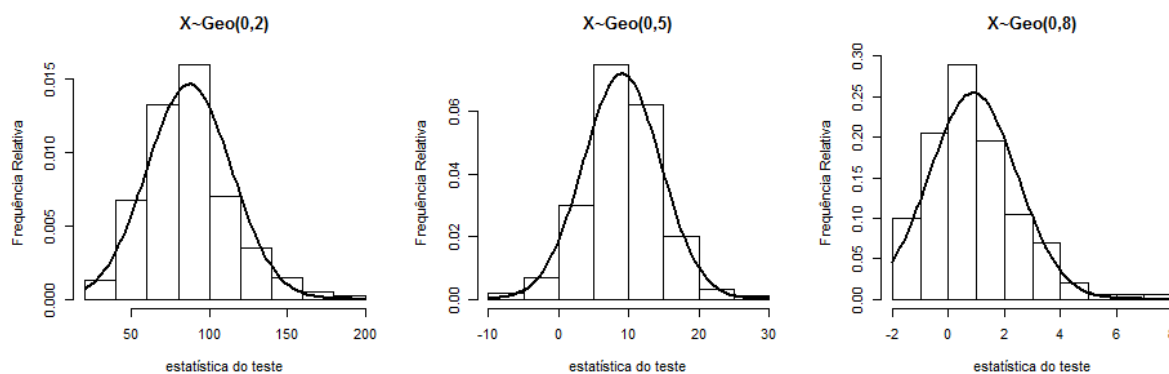


Figura 5 – Histogramas da Distribuição da Estatística o Teste quando  $X_i \sim Geo(p)$

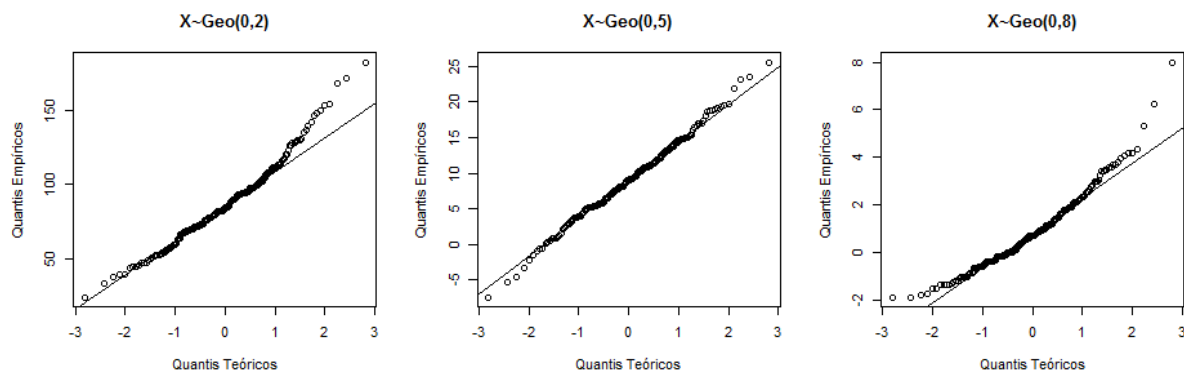


Figura 6 – Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando  $X_i \sim Geo(p)$

Vemos pela Figura 5, que os histogramas, ainda que acompanhem melhor a curva da normal, apresentam uma leve assimetria. E, assim como nos demais casos, notamos pela Figura 6 a incompatibilidade nos extremos da distribuição. Percebemos, pelos resultados vistos na Tabela 6, uma influência do valor assumido por  $\lambda$  sobre a eficácia do teste quando a hipótese nula é falsa. O parâmetro, que representa média e variância do modelo Poisson, aumenta a porcentagem de erro do tipo II quando tende a 0. Sabemos que a média e variância da distribuição Poisson assumem valor  $\lambda$ . Assim, podemos concluir,

Tabela 6 – Erros do Tipo II para o teste Geométrica x Poisson

$n$	$\lambda$	Bilateral	Unilateral à Esquerda
100	0,2	83,56%	72,60%
200	0,2	68%	57%
300	0,2	55,5%	41,5%
500	0,2	30%	23%
700	0,2	22%	14%
900	0,2	12%	6,5%
100	0,5	34%	25,5%
300	0,5	1,5%	0,5%
100	0,7	14%	6,5%
100	1	1,5%	1%
100	2	0%	0%
100	5	0%	0%

mais uma vez, que quanto maior a variância melhor a performance do teste.

### 4.4 Poisson x Geométrica

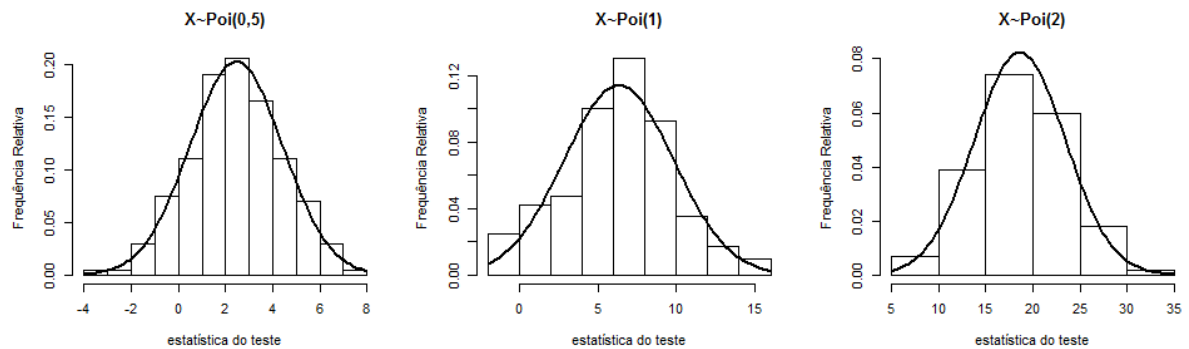


Figura 7 – Histogramas da Distribuição da Estatística o Teste quando  $X_i \sim Poi(\lambda)$

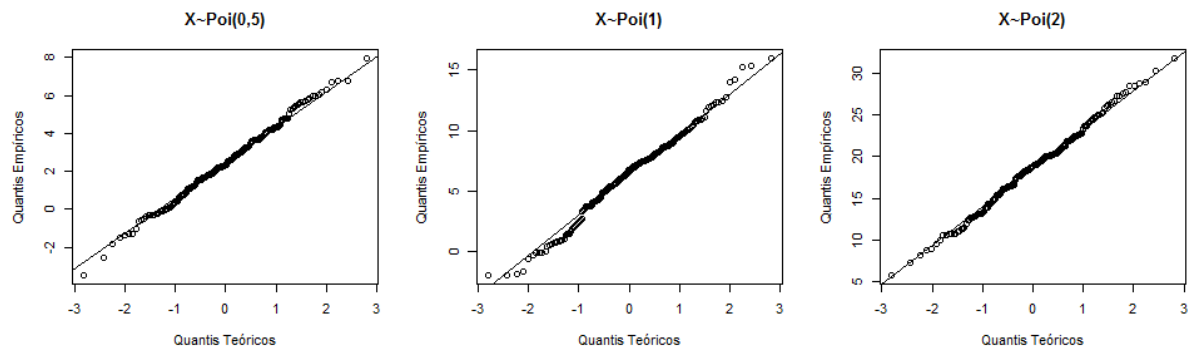


Figura 8 – Quantis Empíricos x Quantis Teóricos (Normal) da Estatística o Teste quando  $X_i \sim Poi(\lambda)$

Tabela 7 – Erros do Tipo II para o teste Poisson x Geométrica

$n$	$p$	Bilateral	Unilateral à Esquerda
100	0,2	0%	0%
100	0,3	0%	0%
100	0,4	0,5%	0%
100	0,5	2%	1%
100	0,6	10%	7,5%
100	0,7	34,5%	27%
200	0,7	12,5%	6%
300	0,7	3%	1,5%
100	0,8	59,5%	51,5%
200	0,8	45,5%	34%
300	0,8	33%	24%
400	0,8	21%	13%
500	0,8	12%	6,5%
600	0,8	8%	3%

Essa foi a situação em que a distribuição da estatística parece mais ter se aproximado da distribuição normal. As Figuras 7 e 8 mostram histogramas mais comportados e quantis mais próximos à reta, respectivamente. Nesta outra comparação, agora com o modelo Poisson representando a hipótese nula, os dados foram simulados através de distribuições Geométricas. Observamos que quando a probabilidade de “sucesso” tende a 1, o teste apresenta mais dificuldade de acertar. A média e a variância desta distribuição são calculadas por  $\frac{1-p}{p}$  e  $\frac{1-p}{p^2}$ , respectivamente, diminuindo para maiores valores de  $p$ . E, novamente, para menores variâncias temos uma maior porcentagem nos erros.

## 4.5 Análise Geral

Analisando o teste de maneira geral, baseado nos quatro casos ilustrativos que escolhemos para investigar, observamos que o teste de razão de verossimilhança via *bootstrap* apresenta boa performance. Assim como era esperado, obtivemos uma influência positiva com o aumento do tamanho da amostra da população, diminuindo os erros do teste.

Aparentemente, a variância (ou o desvio-padrão) populacional também influenciam nos resultados. Observamos que para menores valores dessa característica o teste comete mais equívocos. Quando observamos os valores de  $\beta_1 \times \beta_2^2$  em 4, ou os de  $\lambda$  em 6, vemos que quando eles se aproximam de 0, as proporções de erro se elevam. Em tese, este resultado não é necessariamente inesperado, pois a variabilidade da população influencia de forma direta na variabilidade da estatística do teste. Podemos notar isso ao observar

a maior variabilidade do histograma do primeiro gráfico da Figura 3 [em que a variância populacional é igual 0,  $2 \times 5^2 = 5$ ] em comparação aos demais [de variância 1]. Ou ainda, na Figura 5 em que vemos os histogramas mais estreitos conforme aumentam os valores para  $p$  [diminui a variância populacional]. O problema é que quanto menor a variância da estatística, mais seus dados ficam concentrados em torno de um valor, aproximando também seus quantis. Dessa forma, se torna mais difícil diferenciar o formato das distribuições geradas pelos modelos concorrentes.

Todos os testes foram implementados para o caso Bilateral e o Unilateral à Esquerda e, este último, apresentou em todos os casos chance menor [ou igual] de errar ao não rejeitar  $H_0$ . Isso nos faz supor que este seria o teste adequado à aplicação. De fato, isto parece razoável. A distribuição da estatística do teste sob a hipótese nula, bem como vimos nos histogramas, contém valores em sua maioria positivos. Em contrapartida, se a amostra é provinda do modelo descrito pela hipótese alternativa, produzirá uma estimativa para a estatística, provavelmente, negativa. Afinal, estimativas para a estatística do teste [como a vista na equação 4.1] 3.19 assumem valores “grandes” quando os dados se adequam melhor a  $\mathcal{P}_0$  [modelo de  $H_0$ ]. Do contrário, rejeitamos a hipótese nula. Portanto, se o teste rejeitar  $H_0$ , é mais provável que ele acerte sua afirmação com baixas estimativas.<sup>2</sup>

Simulamos amostras também do modelo descrito na hipótese nula apenas para checar se a Probabilidade de Erro do Tipo I é compatível com o nível de significância escolhido. Os resultados destas probabilidades foram satisfatórios, com valores iguais ou bem próximos a 5% [valor do nível de significância utilizado em todos os testes].

## 4.6 Bootstrap x Cox

Em 3.5.1 do trabalho tratamos da solução abordada por Cox para TRV's para modelos separados. A seção contava ainda com alguns resultados na Tabela 2 de Pereira [5], utilizando-se do método de Cox. Mostramos os resultados de Pereira para o poder empírico [ $1 - P(\text{erro tipo II})$ ] de testes da forma Exponencial x Lognormal. Para manter as notações prévias, lembramos que uma exponencial é uma  $Gama(1; \theta_2)$ . Aqui, iremos comparar suas probabilidades de ocorrer um erro do tipo II com os obtidos via *bootstrap*, sob as mesmas configurações.

As probabilidades de erro do tipo II para quando  $\theta_2 = \sqrt{0,2}$  não foram acrescentados

---

<sup>2</sup>Fizemos testes Unilaterais à Direita e computamos suas porcentagens de erro do tipo II, que chegavam quase sempre a 100% (ou próximo a isso).

Tabela 8 – Erros do Tipo II para Exponencial x Lognormal: Bootstrap x Cox

$n$	$\theta_2 = \sqrt{1}$			$\theta_2 = \sqrt{2}$		
	Bilateral	Unil. Esq.	Cox	Bilateral	Unil. Esq.	Cox
20	81%	67%	62,8%	56%	50%	44,4%
50	32,5%	18,5%	17,4%	14,5%	10,5%	9,4%
100	1,5%	0%	0,6%	1,5%	1,5%	1,4%
150	0%	0%	0%	0%	0%	0%
200	0%	0%	0%	0%	0%	0%

a Tabela 8 pois foram os mesmos, 0%, para todos os tamanhos de amostra e utilizando os dois métodos. Vemos que o método de Cox aparenta ter a mesma relação positiva de tamanho amostral com o poder empírico. Comparando os dois métodos, se analisarmos a metodologia de Cox e a via *bootstrap* [somente no caso unilateral à esquerda] vemos proximidade na probabilidade de cometer o erro do tipo II. Na situação cujo os dados são provenientes de uma *Lognormal*(0, 1) observamos, para uma amostra de tamanho 20, uma diferença de 4,2% na probabilidade de erro [ou no poder empírico]. Quando aumentamos para 50 o tamanho da amostra, essa diferença diminui para 1,1%. Mais ainda, o método que conta com o auxílio do *bootstrap* chega primeiro a 0% de erro, com  $n = 100$ .

Ao observarmos as probabilidades de erro do tipo II quando  $\theta_2^2 = 2$ , vemos o mesmo comportamento. Para amostras de tamanho 20, a diferença entre as probabilidades chega a 5,6%. As diferenças de poderes diminuem, assim como no caso anterior, conforme aumentamos o tamanho da amostra. As potências empíricas do caso bilateral são mais próximas às do unilateral à esquerda nesta configuração da distribuição populacional.

Além desses, outros resultados vistos no trabalho de Pereira [5] [tamanho empírico do teste, por exemplo] foram bem parecidos com os nossos, principalmente, para os testes unilaterais à esquerda. E, para todos os casos, sempre bem próximos ao nível de significância escolhido de 5%.

Podemos notar uma distância pequena entre os dois métodos, potencialmente falando. A maior diferença [para os dados utilizados] de chance de erro [Cox x Unilateral à Esquerda] não ultrapassa 6%. E sendo o método via *bootstrap*, em muitos casos, de mais fácil aplicação, talvez ele seja preferível nestas situações. Além disso, o TRV via *bootstrap* aparenta uma convergência de sua potência empírica para 100% [ou da probabilidade de erro do tipo II para 0%] mais rápida, ao passo que aumentamos o tamanho amostral. Acreditamos, então, que estes fatores somam a favor da metodologia por nós empregada.

## 5 Conclusão

Neste trabalho mostramos uma forma de utilizar o Teste da Razão de Verossimilhança para comparar modelos ditos não encaixados. A metodologia empregada contou com a atuação do método computacional bootstrap para estimar a distribuição da estatística do teste. Avaliamos o teste realizando o procedimento diversas vezes e computando a probabilidade empírica de erros cometidos. Aplicamos o teste a um par de modelos contínuos [Gama e Lognormal] e a outro de modelos discretos [Poisson e Geométrica]. Em ambas comparações os testes mostraram eficiência, conseguindo atingir um alto poder empírico [ou uma baixa probabilidade de erro do tipo II].

Nos casos considerados, a variância populacional afetou positivamente a performance do teste [quanto maior a variância populacional menor a probabilidade de cometer erros do tipo II ou maior a potência do teste]. E, apesar do teste apresentar potência menor nos casos em que a variância é pequena [próxima de 0], a potência aparenta crescer com o tamanho da amostra para uma configuração fixa de parâmetros.

Vimos que, em princípio, podemos utilizar a técnica *Bootstrap* e empregar testes bilaterais ou unilaterais à esquerda. Para o mesmo nível de significância, 5%, nossas simulações apontaram para uma melhor performance [maior potência ou menor erro do tipo II] para o teste unilateral à esquerda.

A praticidade do teste via Bootstrap apresenta um enorme ganho em relação à proposta de Cox pois não precisa de fundamentos teóricos à respeito da distribuição de uma estatística, ou cálculos complexos para estimações. Como vimos, na seção 4.6, ao menos para comparar os modelos Exponencial e Lognormal as potências foram bem parecidas. No método de Cox, se a comparação de modelos fosse, ao invés da dita anteriormente, entre o modelo Gama e o Lognormal certamente aumentaria o grau de dificuldade na definição da distribuição da estatística de teste. Entretanto, no teste via bootstrap a dificuldade de trocar o modelo Exponencial pelo Gama é praticamente nula. Além da simplicidade presente no método via Bootstrap, ele é facilmente adaptável a outras situações.

O TRV via *bootstrap* é de fácil implementação e depende de um algoritmo relativamente simples. Aparenta aplicabilidade a todo par de modelos que possam ser simulados computacionalmente. Pares de modelos contínuos ou discretos podem ser comparados, não importando quão complexa seja a expressão da densidade [desde que seja possível estimar os parâmetros por máxima verossimilhança e simular dados de uma distribuição qualquer do modelo sob a hipótese nula].

Em uma análise prática, para um dado conjunto de dados, recomendamos que se investigue, tal como feito aqui, uma análise da potência no teste, pelo menos para valores próximos dos estimados no modelo associado à hipótese nula. Na prática, não se escolhe, usualmente, o tamanho da amostra e é útil, portanto, ter uma ideia da potência do teste.

Antes de estimar os parâmetros de interesse em uma análise, por vezes, é conveniente escolher uma classe de distribuições de interesse [modelo]. A estimação pelo método de máxima verossimilhança, exige a definição de uma densidade ou probabilidade para representar a distribuição populacional. Por vezes, podemos escolher distintos modelos e, por isso, é fundamental utilizar procedimentos técnicos de escolha entre modelos, sejam eles encaixados ou não. Existem alguns métodos desenvolvidos para este fim, porém, poucos com a facilidade e abrangência do método aqui apresentado. Esse trabalho apresentou uma forma ainda pouco utilizada de uso do TRV para comparar modelos não encaixados. Mostramos que o Teste da Razão de Verossimilhança via *Bootstrap* é um método útil, simples e prático.

# Referências

- [1] CASELLA, G.; BERGER, R. L. *Statistical Inference*. 2. ed. [S.l.]: Duxbury, 2004.
- [2] BICKEL, P.; DOKSUM, K. *Mathematical Statistics: Basic Ideas and Selected Topics*. [S.l.]: Prentice Hall, 2001. (Holden-Day series in probability and statistics, v. 1). ISBN 9780138503635.
- [3] COX, D. R. Tests of separate families of hypotheses. In: *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*. [S.l.: s.n.], 1961. v. 1, p. 105–123.
- [4] COX, D. R. Further results on tests of separate families of hypotheses. *Journal of the Royal Statistical Society. Series B (Methodological)*, [Royal Statistical Society, Wiley], v. 24, n. 2, p. 406–424, 1962. ISSN 00359246.
- [5] PEREIRA, B. d. B. Empirical comparisons of some tests of separate families of hypotheses. *Metrika*, Springer, v. 25, n. 1, p. 219–234, 1978.
- [6] EFRON, B.; TIBSHIRANI, R. Miscellaneous, *An Introduction to the Bootstrap*. [S.l.]: Macmillan Publishers Limited. All rights reserved, 1993.
- [7] PESARAN, M. H.; WEEKS, M. Non-nested hypothesis testing: an overview. *A Companion to Theoretical Econometrics*, Basil Blackwell Oxford, p. 279–309, 2001.
- [8] PEREIRA, B. Choice of a survival model for patients with a brain tumour. *Metrika-International Journal for Theoretical and Applied Statistics*, v. 28, n. 1, p. 53–61, 1981.
- [9] BARBOSA, H. A. L.; OLIVEIRA, F. L. P. d. Modelagem e escolha entre embalagens usando técnicas de confiabilidade e análise de sobrevivência. 2011.
- [10] BIBINGER, M. Notes on the sum and maximum of independent exponentially distributed random variables with different scale parameters. *arXiv preprint arXiv:1307.3945*, 2013.
- [11] JAMES, B. R. *Probabilidade: um curso intermediário*. 3. ed. [S.l.]: Impa, 2004.
- [12] WASSERMAN, L. *All of Nonparametric Statistics (Springer Texts in Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. ISBN 0387251456.
- [13] DAVISON, A. C.; HINKLEY, D. V. *Bootstrap methods and their application*. [S.l.]: Cambridge university press, 1997.
- [14] COX, D. R. A return to an old paper: ‘tests of separate families of hypotheses’. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 75, n. 2, p. 207–215, 2013.