



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE ESTATÍSTICA E MATEMÁTICA APLICADA
CURSO DE GRADUAÇÃO EM ESTATÍSTICA

DANIEL JACINTO DOS SANTOS FILHO

UTILIZAÇÃO DA ESTATÍSTICA GRADIENTE E SEU REFINAMENTO
VIA BOOTSTRAP EM MODELOS LINEARES SIMÉTRICOS

FORTALEZA

2021

DANIEL JACINTO DOS SANTOS FILHO

UTILIZAÇÃO DA ESTATÍSTICA GRADIENTE E SEU REFINAMENTO VIA
BOOTSTRAP EM MODELOS LINEARES SIMÉTRICOS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Estatística do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Estatística.

Orientador: Prof. Dr. Juvêncio Santos Nobre

FORTALEZA

2021

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Biblioteca Universitária
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

- F498u Filho, Daniel Jacinto dos Santos.
Utilização da estatística Gradiente e seu refinamento via Bootstrap em modelos lineares simétricos /
Daniel Jacinto dos Santos Filho. – 2021.
45 f. : il.
- Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, Centro de Ciências,
Curso de Estatística, Fortaleza, 2021.
Orientação: Prof. Dr. Juvêncio Santos Nobre.
1. Bootstrap. 2. Modelo linear simétrico. 3. Teste Gradiente. 4. Testes de hipóteses assintóticos. I. Título.
CDD 519.5
-

DANIEL JACINTO DOS SANTOS FILHO

UTILIZAÇÃO DA ESTATÍSTICA GRADIENTE E SEU REFINAMENTO VIA
BOOTSTRAP EM MODELOS LINEARES SIMÉTRICOS

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Estatística do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Estatística.

Aprovada em: / / 2021

BANCA EXAMINADORA

Prof. Dr. Juvêncio Santos Nobre (Orientador)
Universidade Federal do Ceará (UFC)

Profa. Dra. Silvia Maria de Freitas
Universidade Federal do Ceará (UFC)

Prof. Dr. João Maurício Araújo Mota
Universidade Federal do Ceará (UFC)

Prof. Dr. Gualberto Segundo Agamez Montalvo
Universidade Federal do Ceará (UFC)

Aos meus pais, pela capacidade de acreditar
e de investir em mim.

AGRADECIMENTOS

À Deus pelo dom da vida, por guiar os meus passos e ter me sustentado durante todo esse período. À Ele seja dado toda honra e toda glória !

Aos meus pais por todo apoio, carinho, amor e torcida pelas minhas conquistas. Amo vocês, Daniel (Pai), Fabiana (Mãe) e Débora (Irmã).

A todos os professores do Departamento de Estatística e Matemática Aplicada (DEMA). Aos ensinamentos do Prof. Welliandre, meu primeiro orientador (Iniciação a docência), ao Prof. Júlio, pela oportunidade e ensinamentos no PET-Estatística e ao Prof. Juvêncio pelos incentivos e pela oportunidade na iniciação científica que me trouxeram diversos conhecimentos inclusive para esta monografia. E entre outros professores/pesquisadores que contribuíram de forma significativa para meu crescimento e amadurecimento no curso como João Maurício, Gualberto, Sílvia, Ronald, José Ailton, José Roberto, Leandro, Jacqueline e Rafael Farias.

A todos funcionários do DEMA, tanto da parte administrativa como da limpeza/manutenção do departamento, por contribuírem de forma direta e indireta.

A todos os colegas da graduação, em especial, Matheus Castro (*Bora monitor!*), Keyliane Travassos (*I love Maranguape*), Caroline Duarte (*Aqui joga basquete!*), Luan Prudêncio (*O dramático de Tapuio*), Rickson Lívio (*The boy*), Daniel Amaral (*Machine Learning*), Alberto Rodrigues (*Stonks!*), Jamilly Santos (*I love Amontada*), Yohanna Gonçalves (*Ei juvenil!*), Juan Lucas (*Quer dindin ?!*), Pietro Esteves (*I love Rio de Janeiro*), Robert Oliveira (*Ein Dani, me ajuda nessa questão*).

“O temor do Senhor é o princípio do conhecimento; os loucos desprezam a sabedoria e a instrução. Filho meu, ouve a instrução de teu pai, e não deixes o ensinamento de tua mãe, porque serão como diadema gracioso em tua cabeça, e colares ao teu pescoço.”

(Provérbios 1:7-9)

RESUMO

A classe dos modelos lineares simétricos é uma classe ampla de modelos de regressão que engloba modelos com distribuições de caudas mais pesadas, bem como as de caudas mais leves que a distribuição Normal, como por exemplo t-Student, Cauchy, Slash e Exponencial Potência. Essa classe é uma alternativa ao modelo de regressão linear usual diante da presença de observações extremas, que podem influenciar de forma desproporcional na inferência dos parâmetros do modelo adotado. Neste trabalho, será proposto um refinamento das estatísticas dos testes assintóticos clássicos e do teste Gradiente em modelos lineares simétricos, utilizando o procedimento *Bootstrap*. Para mostrar a importância da proposta, realizou-se um estudo de simulação a fim de comparar o desempenho das estatísticas dos testes assintóticos com relação à aproximação dos seus respectivos tamanhos aos níveis de significância adotados em amostras de diferentes tamanhos. Tal estudo mostrou, de modo geral, que as estatísticas obtiveram bons resultados utilizando o refinamento *Bootstrap* para tamanhos amostrais pequenos e moderados. Com intuito de ilustrar a utilização dos procedimentos aqui adotados, realizou-se uma aplicação a um conjunto de dados reais, assumindo a distribuição t-Student com 4 graus de liberdade.

Palavras-chave: Bootstrap. Modelo linear simétrico. Teste de hipóteses assintótico. Teste Gradiente.

ABSTRACT

The class of symmetric linear models is a broad class of regression models that encompasses models with heavier tails distributions, as well as lighter tails than the Normal distribution, such as t-Student, Cauchy, Slash and Power Exponential. This class is an alternative to the usual linear regression model in the face of the presence of extremes, which can cause the disproportionate shape in the inference of the adopted parameters. In this work, we will propose a refinement of the statistics of the classic asymptotic tests and the Gradient test in symmetric linear models, using the procedure *Bootstrap*. To show the importance of the proposal, a simulation study to compare the performance of the statistics of the asymptotic tests in relation to the approximation of their respective levels of significance adopted in different sizes. This study revealed, in general, that the statistics obtained good results using the *Bootstrap* refinement for small and moderate sample sizes. In order to illustrate the use of the procedures adopted here, if an application to a set of real data, assuming the t-Student distribution with 4 degrees of freedom.

Keywords: Bootstrap. Symmetric linear model. Gradient test. Asymptotic hypothesis tests.

LISTA DE FIGURAS

Figura 1 – Gráfico de quantis com envelope simulado com 95% de confiança para os resíduos padronizados do modelo (1.1).	13
Figura 2 – Densidades de algumas distribuições simétricas.	17
Figura 3 – Representação geométrica da estatística do teste da razão de verossimilhanças generalizada.	27
Figura 4 – Representação geométrica da estatística do teste de Wald.	27
Figura 5 – Representação geométrica da estatística do teste Escore.	28
Figura 6 – Representação geométrica da estatística do teste Gradiente.	28
Figura 7 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 1.	34
Figura 8 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade para situação 1.	34
Figura 9 – Taxas do poder empírico para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o <i>Bootstrap</i> paramétrico para situação 1.	35
Figura 10 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 2.	35
Figura 11 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o <i>Bootstrap</i> paramétrico para situação 2.	36
Figura 12 – Taxas do poder empírico para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o <i>Bootstrap</i> paramétrico para situação 2.	37
Figura 13 – Taxas de rejeição empíricas para o modelo Cauchy.	38
Figura 14 – Taxas de rejeição empíricas para o modelo Cauchy, utilizando o <i>Bootstrap</i> paramétrico.	38
Figura 15 – Taxas do poder empírico para o modelo Cauchy, utilizando o <i>Bootstrap</i> paramétrico.	39

LISTA DE TABELAS

Tabela 1 – Algumas distribuições pertencentes à classe simétrica.	17
Tabela 2 – Expressões para $W_g(u)$, d_g , e_g e ξ para algumas distribuições simétricas.	20
Tabela 3 – Conjunto de dados com relação à resistência de vigas de madeira. . . .	40
Tabela 4 – Estimativas dos parâmetros de localização e de escala do modelo (5.1).	40
Tabela 5 – Valores das estatísticas do testes assintóticos e seus respectivos valores p e valores p <i>bootstrap</i> (p_{boot}) para modelo 5.1	41
Tabela 6 – Taxas de rejeição empíricas (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 1.	47
Tabela 7 – Taxas de rejeição empíricas (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o <i>Bootstrap</i> paramétrico para situação 1.	47
Tabela 8 – Taxas do poder empírico (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o <i>Bootstrap</i> paramétrico para situação 1. . . .	48
Tabela 9 – Taxas de rejeição (%) empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 2.	49
Tabela 10 – Taxas de rejeição empíricas (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o <i>Bootstrap</i> paramétrico para situação 2.	49
Tabela 11 – Taxas do poder empírico (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o <i>Bootstrap</i> paramétrico para situação 2. . . .	50
Tabela 12 – Taxas de rejeição empíricas (%) para o modelo Cauchy.	50
Tabela 13 – Taxas de rejeição empíricas (%) para o modelo Cauchy, utilizando o <i>Bootstrap</i> paramétrico.	51
Tabela 14 – Taxas do poder empírico (%) para o modelo Cauchy, utilizando o <i>Bootstrap</i> paramétrico.	51

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Organização do trabalho	15
2	MODELOS LINEARES SIMÉTRICOS	16
2.1	Classe simétrica de distribuições	16
2.2	Especificação do modelo	18
3	TESTES DE HIPÓTESES ASSINTÓTICOS	21
3.1	Considerações Gerais	21
3.2	Estatística Gradiente	23
3.3	Propriedades da estatística Gradiente	24
3.4	Comparação com os testes Clássicos	25
3.5	Interpretação geométrica dos testes assintóticos	26
3.6	Testes de hipóteses em modelos lineares simétricos	28
4	REFINAMENTO VIA <i>BOOTSTRAP</i>	31
4.1	Resultados Numéricos	32
5	APLICAÇÃO	40
6	CONCLUSÕES	42
	Referências	43
	APÊNDICES	47
	APÊNDICE A– Resultados das simulações	47
A.1	Resultados das simulações considerando a distribuição t-Student com $\nu = 4$ graus de liberdade.	47
A.2	Resultados das simulações considerando a distribuição Cauchy.	50
	ANEXOS	51

1 INTRODUÇÃO

Em análise de regressão, no modelo clássico, é comum assumir que a fonte de variação satisfaça a suposição de normalidade, contudo isso pode não ser razoável na modelagem de determinados fenômenos de interesse no contexto prático. Um procedimento comum, era tentar obter transformações com o objetivo de melhorar a aproximação para a distribuição normal, mas dado a dificuldade de interpretação dos modelos transformados e a gênese de novas classes de distribuições, esse procedimento ficou em total desuso.

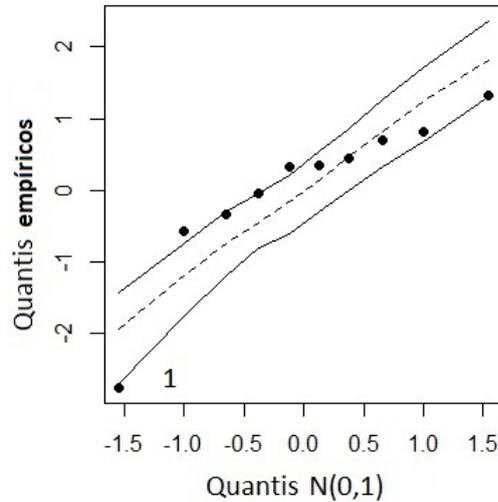
Para ilustrar tal situação, considere o conjunto de dados apresentado em Draper e Stoneman (1966) que trata do interesse em modelar a resistência (y_i) de vigas de madeira em função da gravidade específica (x_{1i}) e do teor de umidade do material (x_{2i}). Uma primeira proposta de modelagem poderia ser um modelo de regressão linear múltiplo (MRLM) com a seguinte forma funcional

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \quad i = 1, \dots, n, \quad (1.1)$$

considerando válidas todas as pressuposições associadas ao MRLM adicionado da suposição de normalidade, ou seja, que $\epsilon_1, \dots, \epsilon_n \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$. Após o ajuste por intermédio dos procedimentos usuais, pode-se ter interesse em avaliar a qualidade do ajuste através de técnicas de diagnóstico - veja Paula (2004). Utilizando através da análise de resíduos do modelo proposto (1.1) temos fortes indícios de que a suposição de normalidade não é adequada (Figura 1), além disso percebe-se que a observação 1 (Figura 1) apresenta um comportamento desproporcional quando comparado com as demais, podendo influenciar as estimativas e conseqüentemente a inferência dos parâmetros do modelo, bem como a qualidade de predição e o comprimento dos intervalos de confiança.

Nesse sentido, a fim de obter modelos que acomodem melhor essas observações, é comum considerar alternativas que utilizam como base distribuições com caudas mais pesadas/leves que a Normal. Aqui, será utilizado uma ampla classe de distribuições simétricas que, além da Normal, possui como casos particulares as distribuições t-Student, Cauchy, Logística, Exponencial potência, Slash, Normal contaminada, dentre outras.

Figura 1 – Gráfico de quantis com envelope simulado com 95% de confiança para os resíduos padronizados do modelo (1.1).



Fonte: O autor

A classe das distribuições simétricas já vem sendo utilizada em trabalhos há pelo menos 40 anos. Por exemplo, Chmielewski (1981) faz uma revisão bibliográfica mostrando diferentes aplicações dessa classe de distribuições. Lange et al. (1989) propõem obter as estimativas de máxima verossimilhança dos coeficientes da regressão linear e não-linear em modelos com erros t-Student, Little (1988) e Yamaguchi (1990) estendem para o caso da distribuição Normal Contaminada. Taylor (1992) propõe o ajuste de modelos de regressão sob suposição da distribuição exponencial potência. Arellano-Valle (1994) trata dos modelos com erros nas variáveis utilizando a distribuição t-Student. Cysneiros (2004) trata sobre métodos de validação em modelos lineares simétricos, em Cysneiros et al. (2005) podem ser vistos mais conceitos e resultados em modelos lineares simétricos.

Com exceção de casos extremamente simples, não é possível fazer uso de testes exatos para essa classe de modelos, seja pela complexidade das hipóteses, seja pelo desconhecimento da distribuição amostral das estatísticas e, por essa razão, são considerados para os testes de hipóteses para a classe de modelos lineares simétricos as estatísticas de testes assintóticos como por exemplo a razão de verossimilhanças generalizadas, Wald e Escore.

Uma outra estratégia que vem sendo utilizada por apresentar uma maior simplicidade em seu algebrismo, não envolvendo o uso da matriz de informação Fisher, nem a operação de inversas de matrizes, é a utilização da estatística do teste Gradiente (TERRELL, 2002), como pode ser visto nos trabalhos de Lemonte (2012) e Medeiros e

Ferrari (2017). Tal estatística de teste também vem sendo utilizada em outras aplicações, em Lemonte (2016) pode-se encontrar uma variedade de aplicações do teste Gradiente como em modelos Birnbaum-Sauders e em modelos lineares generalizados, correção tipo-Bartlett à estatística gradiente em modelos não lineares da família exponencial é apresentado em Lôbo (2016), detalhes sobre uma estatística robusta do teste Gradiente pode ser visto em Lemonte (2016), Saulo et al. (2021) trata da aplicação da estatística Gradiente em um modelo tobit baseado na classe de distribuições log-simétricas.

Geralmente, sob tamanhos amostrais pequenos ou até mesmo moderados, as estatísticas de testes podem não apresentar um grau de aproximação satisfatório para a distribuição assintótica de referência. Um procedimento para contornar tal situação é fazer uso de métodos de refinamento, que consistem em modificar essas estatísticas com objetivo de tornar seus resultados mais precisos. Nessa linha, podem ser consideradas as correções de Bartlett e tipo-Bartlett (CORDEIRO; CRIBARI-NETO, 2014), e também a Correção de Skovgaard (SKOVGAARD, 1996). Todavia, essas ferramentas envolvem correções algébricas baseadas nos cumulantes e que nem sempre são de fácil trato algébrico. Uma alternativa, que por vezes se mostra eficiente e muito mais simples, é a utilização do método *Bootstrap* (EFRON, 1979), que baseia-se na reamostragem para estimar a distribuição amostral de alguma estatística de interesse e, dessa forma, realizar algumas inferências importantes como a construção de intervalos de confiança e testes de hipóteses, por exemplo.

Dessa forma, o objetivo deste trabalho é utilizar a estatística de teste Gradiente na classe de modelos lineares simétricos e, posteriormente, fazer uso do refinamento *Bootstrap* para tamanhos amostrais pequenos e moderados. Especificamente, também, deseja-se:

- Apresentar do ponto de vista algébrico e interpretar graficamente como as estatísticas dos testes assintóticos funcionam, como fizeram Buse (1982) e Muggeo e Lovison (2014).
- Apresentar de forma introdutória a classe de modelos lineares simétricos e a utilização do teste gradiente nessa classe de modelos.
- Comparar por meio de simulações via Monte Carlo e *Bootstrap* resultados empíricos da taxa de rejeição e poder dos principais testes assintóticos, quando utilizados em modelos lineares simétricos.

1.1 Organização do trabalho

Este trabalho está organizado em seis capítulos. O Capítulo 2 é destinado a apresentar brevemente os modelos lineares simétricos. Neste capítulo, apresenta-se a forma funcional do modelo e o método de estimação mais comumente utilizado. No Capítulo 3, é feita uma breve introdução dos testes de hipóteses assintóticos clássicos e do Gradiente, tanto do ponto de vista algébrico, como sua interpretação gráfica. Além disso, será apresentada a obtenção das estatísticas de teste associadas aos testes assintóticos usuais para a classe de modelos simétricos. No Capítulo 4 discutimos sobre alguns métodos de refinamento e apresentamos o refinamento via o procedimento *Bootstrap*, e também são realizados alguns estudos de simulação dos testes assintóticos em modelos lineares simétricos. Uma aplicação dos procedimentos aqui utilizados a um conjunto de dados reais é encontrada no Capítulo 5. Por fim no Capítulo 6 serão feitas as principais conclusões.

2 MODELOS LINEARES SIMÉTRICOS

Aqui será apresentada brevemente a classe do modelo linear simétrico bem como seu método de estimação, métodos diagnósticos e os aspectos computacionais. Primeiramente será abordada a classe de distribuições simétricas e algumas de suas contribuições.

2.1 Classe simétrica de distribuições

Uma variável aleatória Y pertence à classe simétrica de distribuições (CYSNEIROS et al., 2005), isto é $Y \sim S(\mu, \phi)$, se sua função densidade é dada por

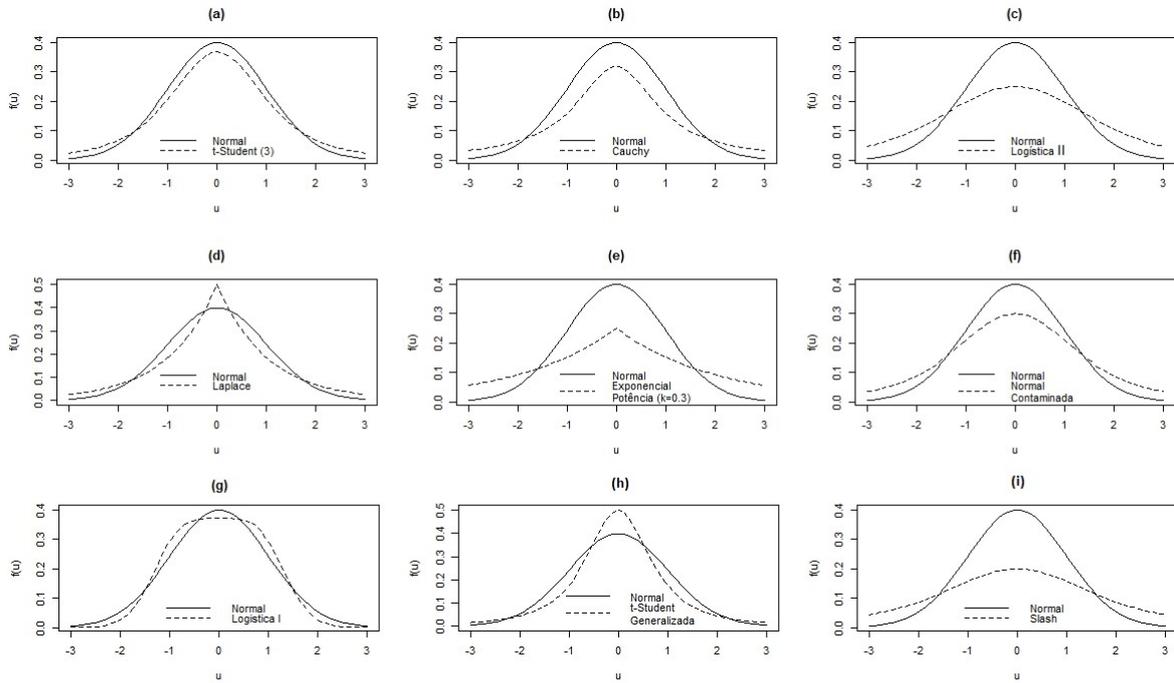
$$f(y; \mu, \phi) = \frac{1}{\sqrt{\phi}} g \left\{ \frac{(y - \mu)^2}{\phi} \right\}, \quad y \in \mathbb{R},$$

para alguma função $g(\cdot)$ denominada função geradora de densidade, sendo $g(u) > 0$ para $u > 0$ e $\int_0^\infty u^{-1/2} g(u) du = 1$, uma condição necessária para que $f(y; \mu, \phi)$ seja uma legítima função densidade de probabilidade. Os parâmetros μ e ϕ , tais que $\mu \in \mathbb{R}$ e $\phi > 0$, são denominados respectivamente de parâmetros de posição e de escala.

Pode-se citar alguns exemplos de distribuições pertencentes a essa classe como Normal, Normal contaminada, Cauchy, t-Student, t-Student generalizada, Slash, Logística tipos I e II, exponencial dupla, exponencial potência entre outras.

Na Tabela 1 podemos observar as funções geradoras de algumas distribuição dessa classe, para mais detalhes veja também Cysneiros et al. (2005). Na Figura 2 estão representadas algumas distribuições da classe simétrica, com $\mu = 0$ e $\phi = 1$, em comparação à distribuição Normal Padrão. Cabe destacar diversos autores que contribuíram para o desenvolvimento dessa classe de distribuições como Cambanis et al. (1981), Berkane e Bentler (1986), Fang et al. (1990), Anderson e Fang (1990), Rao (1990), Cysneiros et al. (2005) e entre outros.

Figura 2 – Densidades de algumas distribuições simétricas.



Fonte: O autor

Tabela 1 – Algumas distribuições pertencentes à classe simétrica.

Distribuição	Função geradora de densidade $g(u)$
Normal	$\frac{1}{\sqrt{2\pi}} \exp(-u/2)$
Normal Contaminada	$(1 - \epsilon) \frac{\exp(-u/2)}{\sqrt{2\pi}} + \epsilon \frac{\exp(-u/(2\sigma^2))}{\sigma\sqrt{2\pi}}, \sigma > 0, 0 \leq \epsilon \leq 1$
Cauchy	$\frac{1}{\pi(1+u)}$
t-Student	$\frac{1}{\text{Beta}(1/2, \nu/2)} \nu^{\nu/2} (\nu + u)^{-(\nu+1)/2}, \nu > 0$
t-Student generalizada	$\frac{1}{\text{Beta}(1/2, r/2)} s^{r/2} (s + u)^{-(r+1)/2}, s, r > 0$
Logística tipo I	$\frac{c}{\exp(u)(1+\exp(-u))^2}, c \approx 1,4843$
Logística tipo II	$\frac{1}{\exp(u^{1/2})(1+\exp(-u^{1/2}))^2}$
Logística generalizada	$\frac{\alpha}{\text{Beta}(m, m)} \frac{\exp(-m\alpha\sqrt{u})}{(1+\exp(-\alpha\sqrt{u}))^{2m}}, \alpha, m > 0$
Exponencial potência	$\frac{\exp(-\frac{1}{2}u^{1/(k+1)})}{\Gamma(1+\frac{1+k}{2})2^{1+(1+k)/2}}, -1 < k \leq 1$
Slash	$\frac{\varphi(0)-\varphi(u)}{u^2}, u \neq 0$

Nota: $\varphi(\cdot)$ densidade normal padrão

Fonte: Adaptado de Cysneiros et al. (2005)

2.2 Especificação do modelo

O modelo de regressão linear simétrico (CYSNEIROS et al., 2005) é definido por

$$y_i = \mu_i(\boldsymbol{\beta}, \mathbf{x}_i) + \epsilon_i, \quad i = 1, \dots, n, \quad (2.1)$$

em que $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^\top$ é o vetor que contém os valores das variáveis explicativas para o i -ésimo indivíduo, $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^\top$ é o vetor de parâmetros (desconhecidos) de regressão do modelo e a fonte de variação $\epsilon_i \stackrel{\text{iid}}{\sim} S(0, \phi)$. Usualmente, assume-se que $\mu_i(\boldsymbol{\beta}, \mathbf{x}_i)$ seja uma função contínua e diferenciável em relação ao vetor $\boldsymbol{\beta}$. No caso linear, tem-se que $\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta}$, sendo \mathbf{X} uma matriz $n \times p$ de posto completo cujas linhas são dadas pelos vetores \mathbf{x}_i^\top , $i = 1, \dots, n$. A densidade de Y_i , denotada por $S \sim (\mu_i, \phi)$, é dada por

$$f_{Y_i}(y_i) = \frac{1}{\sqrt{\phi}} g(u_i),$$

em que $u_i = \frac{(y_i - \mu_i)^2}{\phi}$, tendo, caso a variável Y_i possua os dois primeiros momentos finitos, $\mathbb{E}(Y_i) = \mu_i = \mathbf{x}_i^\top \boldsymbol{\beta}$ e $\text{Var}(Y_i) = \xi\phi$, com $\xi > 0$ uma constante específica de cada distribuição. Expressões para ξ são apresentados na Tabela 2.

Portanto, sob as suposições acima, o logaritmo da função de verossimilhança de $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \phi)^\top$, denotado por $\ell(\boldsymbol{\theta})$ é dado por

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2} \log \phi + \sum_{i=1}^n \log g(u_i).$$

Para mais detalhes dessa função veja Cysneiros et al. (2005).

A função score de $\boldsymbol{\beta}$ e de ϕ são dadas respectivamente por (CYSNEIROS et al., 2005)

$$\mathbf{U}_\beta(\boldsymbol{\theta}) = \frac{1}{\phi} \mathbf{X}^\top \mathbf{D}(\mathbf{v})(\mathbf{y} - \boldsymbol{\mu}) \quad (2.2)$$

e

$$U_\phi(\boldsymbol{\theta}) = \frac{1}{2\phi} \left[\frac{1}{\phi} Q - n \right], \quad (2.3)$$

em que $\mathbf{y} = (y_1, \dots, y_n)^\top$, $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)^\top$, $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)^\top$, $\mathbf{D}(\mathbf{v}) = \text{diag}(v_1, \dots, v_n)$, com $v_i = -2W_g(u_i)$, sendo $W_g(u_i) = \frac{d \log g(u)}{du}$ (Tabela 2) e $Q = (\mathbf{y} - \boldsymbol{\mu})^\top \mathbf{D}(\mathbf{v})(\mathbf{y} - \boldsymbol{\mu})$.

O estimador de máxima verossimilhança de $\boldsymbol{\theta}$, $\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\beta}}^\top, \hat{\phi})^\top$ é obtido resolvendo simultaneamente o sistema de equações $\mathbf{U}_\beta(\boldsymbol{\theta}) = \mathbf{0}$ e $U_\phi(\boldsymbol{\theta}) = 0$, sendo

$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{D}(\mathbf{v}) \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{v}) \mathbf{y}$ e $\hat{\phi} = \frac{1}{n} Q$. Percebe-se na Tabela 2 que, no caso da distribuição Normal, $v_i = 1$ para todo i , o que faz com que os estimadores possuam forma analítica fechada. No caso geral, um processo iterativo, como por exemplo, o scoring de Fisher pode ser utilizado para obter as estimativas de máxima verossimilhança de $\boldsymbol{\beta}$ e ϕ . Tal procedimento é dado por

$$\boldsymbol{\beta}^{(m+1)} = (\mathbf{X}^\top \mathbf{D}(\mathbf{v}^{(m)}) \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{D}(\mathbf{v}^{(m)}) \mathbf{y}$$

e

$$\phi^{(m+1)} = \frac{1}{n} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}^{(m+1)})^\top \mathbf{D}(\mathbf{v}^{(m+1)}) (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}^{(m+1)}) \quad (m = 0, 1, \dots).$$

Conforme sugerido por Cysneiros et al. (2005), os valores iniciais para a estimativa do vetor de parâmetros $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \phi)^\top$ são dados usando mínimos quadrados sendo, respectivamente, $\boldsymbol{\beta}^{(0)} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$ e $\phi^{(0)} = [n^{-1} (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}^{(0)})^\top (\mathbf{y} - \mathbf{X} \boldsymbol{\beta}^{(0)})]^{1/2}$.

A matriz de informação de Fisher para $\boldsymbol{\theta}$ pode expressa por (para mais detalhes veja (CYSNEIROS et al., 2005))

$$\mathbf{I}_F(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{I}_F(\boldsymbol{\beta}) & \mathbf{0} \\ \mathbf{0} & I_F(\phi) \end{bmatrix}, \quad (2.4)$$

com $\mathbf{I}_F(\boldsymbol{\beta}) = \frac{4d_g}{\phi} \mathbf{X}^\top \mathbf{X}$ e $I_F(\phi) = \frac{n}{4\phi^2} (4f_g - 1)$, sendo $f_g = \mathbb{E}[W_g^2(U^2)U^4]$ e $U \sim S(0, 1)$. Uma vez que a matriz de informação é bloco diagonal, temos que os parâmetros $\boldsymbol{\beta}$ e ϕ são ortogonais, o que acaba facilitando o processo de inferência. Na Tabela 2 encontram-se as expressões para d_g e f_g considerando algumas distribuições simétricas.

Com relação aos métodos de diagnóstico, Galea et al. (2005) (apud PENAGOS, 2005) apresentam um resíduo para os modelos simétricos baseado na definição geral de Cox e Snell (1968). Penagos (2005) propõe uma definição geral dos resíduos para modelos simétricos de forma que a proposta de Galea et al. (2005) seja um caso particular. Galea et al. (2003) e Galea et al. (2005) estudam a influência local em modelos simétricos lineares e não lineares, respectivamente, enquanto que Galea et al. (2005) estudam a influência global em modelos lineares simétricos.

Tabela 2 – Expressões para $W_g(u)$, d_g , e_g e ξ para algumas distribuições simétricas.

Distribuição	$W_g(u)$	d_g	f_g	ξ
Normal	$-\frac{1}{2}$	$\frac{1}{4}$	$\frac{3}{4}$	1
t-Student	$-\frac{\nu+1}{2(\nu+u)}$	$\frac{\nu+1}{4(\nu+3)}$	$\frac{3(\nu+1)}{4(\nu+3)}$	$\frac{\nu}{\nu-2}$, $\nu > 2$
t-Student generalizada	$-\frac{r+1}{2(s+u)}$	$\frac{r(r+1)}{4s(r+3)}$	$\frac{3(r+1)}{4(r+3)}$	$\frac{s}{r-2}$, $s > 0, r > 2$
Logística-I	$-\tanh(u/2)$	$\approx 0,3693$	$\approx 1,0034$	$\approx 0,7956$
Logística-II	$-\frac{\exp(-\sqrt{u})-1}{-2\sqrt{u}[1+\exp(-\sqrt{u})]}$	$\frac{1}{12}$	$\approx 0,6074$	$\frac{\pi^2}{3}$
Logística generalizada	$-\frac{\alpha m[\exp(-\alpha\sqrt{u})-1]}{-2\sqrt{u}[1+\exp(-\alpha\sqrt{u})]}$	$\frac{\alpha^2 m^2}{4(2m+1)}$	$\frac{2m(2+m^2\psi'(m))}{4(2m+1)}$	$2\psi'(m)$
Exponencial potência	$-\frac{1}{2(1+k)u^{k/(k+1)}}$	$\frac{\Gamma\{(3-k)/2\}}{4(2^{k-1})(1+k)^2\Gamma\{(k+1)/2\}}$	$\frac{k+3}{4(k+1)}$	$2^{k+1}\frac{\Gamma\{3(k+1)/2\}}{\Gamma\{(k+1)/2\}}$

Fonte: Adaptado de Cysneiros et al. (2005)

Na literatura estatística podemos encontrar algumas extensões dos modelos lineares simétricos. Por exemplo, Lange et al. (1989) estudam modelos com erros sob distribuição t-Student considerando uma estrutura longitudinal; Arellano-Valle (1994) trata dos modelos elípticos com erros nas variáveis; uma comparação entre procedimentos clássicos e bayesianos pode ser encontrada em Kowalski et al. (1999); Lindsey (1999) propõe um classe de modelos sob erros exponencial potência no contexto de dados com medidas repetidas; Liu (2004) propõe alguns procedimentos de influência local em modelos de séries temporais com estrutura heteroscedástica; modelos não lineares simétricos homocedásticos e heterocedásticos são encontrados respectivamente em Braga (2007) e Brito (2009); recentemente Saulo et al. (2021) propuseram um modelo tobit baseado na classe de distribuições log-simétricas.

Recursos computacionais referentes ao ajuste e à análise de modelos lineares simétricos podem ser obtidos no *software* R, por exemplo, através das bibliotecas `elliptical` (CYSNEIROS, 2005) e `heavy` (OSORIO, 2019). A biblioteca `fitHeavyTail` (PALOMAR; ZHOU, 2020) trata de métodos robustos de estimativa para o vetor de médias e para a matriz de covariância de dados, sob distribuições multivariadas de caudas pesadas como t-Student e Cauchy.

3 TESTES DE HIPÓTESES ASSINTÓTICOS

No presente capítulo, serão abordados as principais estatísticas de testes de hipóteses em modelos lineares simétricos. Porém, antes, serão apresentados brevemente os principais testes assintóticos e suas propriedades tanto sob aspecto algébrico como geométrico.

3.1 Considerações Gerais

Em inferência estatística, os testes de hipóteses têm por objetivo tomar uma decisão acerca da família de distribuição da população com base em um conjunto de observações, sendo, nos casos paramétricos, caracterizados por avaliar o quão plausível é a hipótese de que um parâmetro θ da família pertença a um determinado subconjunto do espaço paramétrico Θ (LEHMANN; ROMANO, 2006). Podemos encontrar na literatura alguns procedimentos que fornecem estatísticas para testar tais hipóteses, como trata o lema de Neyman-Pearson, que possui ótimas propriedades em determinadas situações bem limitadas, como é o caso de obter testes mais poderosos (testes relacionados a hipóteses simples) - para maiores detalhes ver Lehmann e Romano (2006). Em situações mais gerais, quando se utilizam hipóteses compostas bilaterais, outros procedimentos são necessários, como o teste da razão de verossimilhanças. Contudo, o uso desses procedimentos exatos nem sempre é possível, seja por dificuldade na obtenção da distribuição da estatística de teste, ou pela complexidade das hipóteses.

Assim, uma alternativa é a utilização dos testes assintóticos, como os testes assintóticos clássicos, denominados de “*Holy Trinity*”, por Rao (2005), são eles: o teste da razão de verossimilhanças generalizada (WILKS, 1938), o teste de Wald (WALD, 1943) e o teste Escore (RAO, 1948); e por último o mais recente o teste Gradiente (TERRELL, 2002).

Definição: Considere X_1, \dots, X_n uma amostra aleatória da variável aleatória X com função densidade/função de probabilidade $f(x_i, \theta)$, sendo $\theta \in \Theta$, e supondo o interesse em testar $\mathcal{H}_0 : \theta = \theta_0$ contra $\mathcal{H}_1 : \theta \neq \theta_0$. Então as estatísticas dos testes da razão de verossimilhanças generalizada, Wald, Escore e Gradiente são definidas, respectivamente, por

$$S_{RV} = 2[\ell(\hat{\theta}) - \ell(\theta_0)], \quad (3.1)$$

$$S_W = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^\top \mathbf{I}_F(\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \quad (3.2)$$

$$S_R = \mathbf{U}^\top(\boldsymbol{\theta}_0) \mathbf{I}_F^{-1}(\boldsymbol{\theta}_0) \mathbf{U}(\boldsymbol{\theta}_0) \text{ e} \quad (3.3)$$

$$S_T = \mathbf{U}(\boldsymbol{\theta}_0)^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0), \quad (3.4)$$

em que $\boldsymbol{\theta}_0$ é um vetor p -dimensional, $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p)^\top$ é o estimador de máxima verossimilhança de $\boldsymbol{\theta}$ e $\ell(\hat{\boldsymbol{\theta}})$ é o logaritmo da função de verossimilhança aplicada em $\hat{\boldsymbol{\theta}}$.

Nos casos acima, a função escore, a informação observada e a informação de Fisher (sob as condições de regularidade de Frechet-Cramer-Rao) são, respectivamente, definidas por:

$$\mathbf{U}(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta}, \mathbf{x})}{\partial \boldsymbol{\theta}}, \quad \mathbf{J}(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta}, \mathbf{x})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top} \text{ e} \quad \mathbf{I}_F(\boldsymbol{\theta}) = -\mathbb{E}[\mathbf{J}(\boldsymbol{\theta})],$$

Sob certas condições de regularidade, as estatísticas (3.1-3.4) possuem distribuição assintótica qui-quadrado com p graus de liberdade, χ_p^2 , sendo os graus de liberdade determinados pelo número de restrições impostas à hipótese nula \mathcal{H}_0 . Assim, a regra de decisão a ser adotada é rejeitar a hipótese nula \mathcal{H}_0 ao nível de significância de $100\alpha\%$ quando o valor observado da estatística em questão for maior que o quantil $(1 - \alpha)$ da distribuição χ_p^2 , com $\alpha \in (0, 1)$.

Caso o vetor de parâmetros p -dimensional $\boldsymbol{\theta}$ seja particionado de forma que $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, em que $\boldsymbol{\theta}_1$ é um vetor de dimensão q e $\boldsymbol{\theta}_2$ um vetor de dimensão $p - q$, e supondo que exista o interesse em testar $\mathcal{H}_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_{10}$ contra $\mathcal{H}_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_{10}$ com $\boldsymbol{\theta}_{10}$ um vetor fixo q -dimensional, então as estatísticas dos testes assintóticos serão

$$S_{RV} = 2[\ell(\hat{\boldsymbol{\theta}}) - \ell(\tilde{\boldsymbol{\theta}})]$$

$$S_W = (\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}})^\top \mathbf{I}_F(\hat{\boldsymbol{\theta}})(\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}),$$

$$S_R = \mathbf{U}^\top(\tilde{\boldsymbol{\theta}}) \mathbf{I}_F^{-1}(\tilde{\boldsymbol{\theta}}) \mathbf{U}(\tilde{\boldsymbol{\theta}}) \text{ e}$$

$$S_T = \mathbf{U}(\tilde{\boldsymbol{\theta}})^\top (\hat{\boldsymbol{\theta}} - \tilde{\boldsymbol{\theta}}),$$

possuindo distribuição assintótica qui-quadrado com q graus de liberdade, χ_q^2 . Nesse caso, o estimador de máxima verossimilhança (EMV) $\hat{\boldsymbol{\theta}}$ também será particionado como

$\hat{\boldsymbol{\theta}} = (\hat{\boldsymbol{\theta}}_1^\top, \hat{\boldsymbol{\theta}}_2^\top)^\top$ e sob a restrição da hipótese \mathcal{H}_0 o EMV será $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\theta}_{10}^\top, \tilde{\boldsymbol{\theta}}_2^\top)^\top$, além disso o vetor de parâmetro $\boldsymbol{\theta}_2$ é considerado como um vetor de parâmetros de perturbação (*nuisance parameter*).

3.2 Estatística Gradiente

O livro de Lemonte (2016) é uma importante referência que trata dos principais resultados para a estatística do teste Gradiente, nele podemos encontrar uma gama de aplicações em alguns modelos de regressão. Em Mota (2017), também podemos encontrar de forma introdutória sobre a utilização dessa estatística de teste.

A estatística Gradiente proposta por Terrell (2002) foi construída a partir da combinação de duas outras estatísticas de teste a saber: Escore e Wald modificada, proposta por Hayakawa e Puri (1985). O arranjo delas contribuiu para formação de uma estatística de teste simplificada quando comparada com as demais, como veremos a seguir.

A estatística do teste Wald modificada (3.5), diferentemente da estatística do teste Wald, faz uso da informação de Fisher avaliada em $\boldsymbol{\theta}_0$,

$$S_W^* = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^\top \mathbf{I}_F(\boldsymbol{\theta}_0) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0). \quad (3.5)$$

A ideia apresentada por Terrell (2002) foi fazer uso de uma matriz simétrica \mathbf{L} de forma que \mathbf{L} seja a matriz raiz quadrada de $\mathbf{I}_F(\boldsymbol{\theta})$, ou seja, $\mathbf{L}^\top \mathbf{L} = \mathbf{I}_F(\boldsymbol{\theta})$. Logo, podemos expressar (3.3) e (3.5) da seguinte forma:

$$\begin{aligned} S_R &= \mathbf{U}^\top(\boldsymbol{\theta}_0) (\mathbf{L}^\top \mathbf{L})^{-1} \mathbf{U}(\boldsymbol{\theta}_0) \\ &= [(\mathbf{L}^{-1})^\top \mathbf{U}(\boldsymbol{\theta}_0)]^\top (\mathbf{L}^{-1})^\top \mathbf{U}(\boldsymbol{\theta}_0), \end{aligned}$$

e

$$\begin{aligned} S_W^* &= (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^\top (\mathbf{L}^\top \mathbf{L}) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \\ &= [\mathbf{L}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)]^\top \mathbf{L}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0). \end{aligned}$$

Das expressões anteriores podemos denotar $\mathbf{Z}_1 = (\mathbf{L}^{-1})^\top \mathbf{U}(\boldsymbol{\theta}_0)$ e $\mathbf{Z}_2 = \mathbf{L}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)$ vetores padronizados cuja esperança e variância são dadas, respectivamente, por

$$\begin{aligned}
\mathbb{E}[\mathbf{Z}_1] &= \mathbb{E}[(\mathbf{L}^{-1})^\top \mathbf{U}(\boldsymbol{\theta}_0)] & \mathbb{E}[\mathbf{Z}_2] &= \mathbb{E}[\mathbf{L}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)] \\
&= (\mathbf{L}^{-1})^\top \mathbb{E}[\mathbf{U}(\boldsymbol{\theta}_0)] = \mathbf{0}, & &= \mathbf{L}\mathbb{E}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)] = \mathbf{0}, \\
\text{Var}[\mathbf{Z}_1] &= \text{Var}[(\mathbf{L}^{-1})^\top \mathbf{U}(\boldsymbol{\theta}_0)] & \text{Var}[\mathbf{Z}_2] &= \text{Var}[\mathbf{L}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)] \\
&= (\mathbf{L}^{-1})^\top \text{Var}[\mathbf{U}(\boldsymbol{\theta}_0)] \mathbf{L}^{-1} & &= \mathbf{L}\text{Var}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)] \mathbf{L}^\top \\
&= (\mathbf{L}^{-1})^\top \mathbf{I}_F(\boldsymbol{\theta}_0) \mathbf{L}^{-1} & &= \mathbf{L}\mathbf{I}_F^{-1}(\boldsymbol{\theta}_0) \mathbf{L}^\top \\
&= (\mathbf{L}^{-1})^\top \mathbf{L}^\top \mathbf{L} \mathbf{L}^{-1} = \mathbf{I}_p, & &= \mathbf{L}(\mathbf{L}^\top \mathbf{L})^{-1} \mathbf{L}^\top = \mathbf{I}_p,
\end{aligned}$$

sendo $\mathbf{0}$ um vetor de zeros e \mathbf{I}_p uma matriz identidade de ordem p . Tem-se que os vetores \mathbf{Z}_1 e \mathbf{Z}_2 possuem cada um, assintoticamente, distribuição normal padrão multivariada, isto é, $\mathbf{Z}_1 \stackrel{a}{\sim} \mathcal{N}_p(\mathbf{0}, \mathbf{I}_p)$ e $\mathbf{Z}_2 \stackrel{a}{\sim} \mathcal{N}_p(\mathbf{0}, \mathbf{I}_p)$, além de serem independentes (LEMONTE, 2016). Logo, o produto interno de \mathbf{Z}_1 e \mathbf{Z}_2 possui distribuição assintótica χ_p^2 :

$$\begin{aligned}
\mathbf{Z}_1^\top \mathbf{Z}_2 &= ((\mathbf{L}^{-1})^\top \mathbf{U}(\boldsymbol{\theta}_0))^\top \mathbf{L}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \\
&= \mathbf{U}(\boldsymbol{\theta}_0)^\top \mathbf{L}^{-1} \mathbf{L}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \\
&= \mathbf{U}(\boldsymbol{\theta}_0)^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0).
\end{aligned}$$

Definição 2.1: A estatística gradiente, S_T , para testar $\mathcal{H}_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ contra $\mathcal{H}_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ é dada por:

$$S_T = \mathbf{U}(\boldsymbol{\theta}_0)^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0). \quad (3.6)$$

Observa-se que a nova estatística (3.6) possui uma forma mais simples que as estatísticas de testes anteriores, não envolvendo a obtenção da matriz de informação, nem esperada, nem observada, e muito menos operações como inversão de matrizes. Tais cálculos podem ser bastante tediosos, e por vezes não são possíveis de serem obtidos de forma analítica em modelos mais complexos.

3.3 Propriedades da estatística Gradiente

No trabalho inicial de Terrell (2002) são citadas algumas propriedades da estatística Gradiente. Dentre estas propriedades, está sua distribuição assintótica qui-quadrado como visto anteriormente; para uma demonstração mais formal consultar Lemonte (2016). Além disso, diferentemente das demais estatísticas, a estatística gradiente não é claramente não negativa, embora seja quase-certamente, conforme a seguir.

Teorema 2.2: Se o logaritmo da função de verossimilhança $l(\boldsymbol{\theta})$ for unimodal e diferenciável em $\boldsymbol{\theta}_0 \in \Theta$, então, quase certamente

$$S_T = \mathbf{U}(\boldsymbol{\theta}_0)^\top (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \geq 0$$

Uma demonstração mais formal pode ser vista em Lemonte (2016).

Uma outra característica da estatística Gradiente está relacionada à possibilidade de melhorar sua aproximação para a distribuição χ^2 usando uma estimativa menos viesada para $\boldsymbol{\theta}$ no lugar de $\hat{\boldsymbol{\theta}}$.

Teorema 2.3: Seja $\bar{\boldsymbol{\theta}}$ um estimador não viesado de $\boldsymbol{\theta} \in \Theta$, então

$$\mathbb{E}[\mathbf{U}(\boldsymbol{\theta})^\top (\bar{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)] = p$$

Para mais detalhes dessa demonstração pode-se consultar Lemonte (2016).

Diferentemente dos testes assintóticos clássicos anteriores, uma outra propriedade da estatística gradiente é a sua não invariância sob reparametrizações não lineares do vetor de parâmetro $\boldsymbol{\theta}$ (LEMONTE, 2016). Ou seja, o valor observado da estatística pode variar de acordo como a hipótese nula é formulada. Mesmo assim, ainda é possível melhorar seu comportamento escolhendo uma parametrização na qual a estimativa de máxima verossimilhança seja menos viesada.

3.4 Comparação com os testes Clássicos

Devido aos resultados e às propriedades da estatística Gradiente, Rao (2005) menciona: “The suggestion by Terrell is attractive as it is simple to compute. It would be of interest to investigate the performance of the [gradient] statistic.” Nessa linha de pensamento, trabalhos como de Lemonte e Ferrari (2012) e de Lemonte (2016) comparam o desempenho das estatísticas dos testes assintóticos as hipóteses alternativas locais, afim de saber qual deles é mais poderoso, caso algum se sobressaia.

Uma maneira de avaliar tal desempenho é através das hipóteses alternativas locais de Pitman. Como pode ser visto em Lemonte (2016), as hipóteses alternativas locais de Pitman são uma sequência de hipóteses do tipo

$$\mathcal{H}_{1n} : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_{10} + n^{-1/2}\boldsymbol{\epsilon},$$

com n representando o tamanho amostral e $\boldsymbol{\epsilon} = n^{1/2}(\boldsymbol{\theta}_1 - \boldsymbol{\theta}_{10})$. Tal sequência converge para a hipótese nula \mathcal{H}_0 a uma taxa de convergência de $O(n^{-1/2})$.

Sob a hipótese alternativa \mathcal{H}_{1n} , Lemonte e Ferrari (2012) obtêm a expansão assintótica para a estatística do teste Gradiente, que possui distribuição assintótica qui-quadrado não central com q graus de liberdade. Resultados similares foram obtidos por Hayakawa (1975) para a estatística do teste da razão de verossimilhanças generalizadas e para a estatística do teste Wald, e por Harris e Peers (1980) para estatística do teste Escore (LEMONTE, 2016).

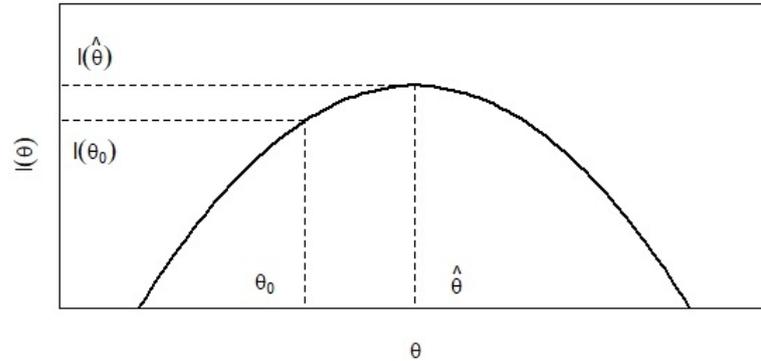
Harris e Peers (1980) publicaram um estudo sobre o poder local até uma ordem $O(n^{-1/2})$ para as estatísticas de teste da razão de verossimilhanças generalizadas, Wald e Escore, tal estudo mostrou que nenhum teste assintótico é uniformemente melhor que o outro até segunda ordem. Lemonte (2016) estende esse estudo incluindo o teste Gradiente e conclui que as aproximações assintóticas das estatísticas dos testes assintóticos são idênticas até primeira ordem sob a hipótese nula e sob alternativas locais de Pitman. Tais testes apresentam o mesmo tamanho até um erro de ordem $O(n^{-1})$, porém as funções poder diferem até um erro de ordem $O(n^{-1/2})$. Dessa forma, o desempenho dos testes assintóticos podem ser comparados através das expansões de suas funções poder, ignorando o termo de ordem $O(n^{-1/2})$. Portanto, o teste que usa a estatística Gradiente, que é muito simples de calcular, pode ser uma alternativa interessante aos testes clássicos, dado que nenhum dos quatro testes é uniformemente superior em termos de poder local (LEMONTE, 2016).

3.5 Interpretação geométrica dos testes assintóticos

Como pode ser observado em Buse (1982) e em Muggeo e Lovison (2014), as estatísticas dos testes assintóticos possuem uma importante interpretação geométrica, o que nos ajuda a entender as similaridades e as diferenças entre elas. Tais autores, para uma melhor compreensão visual, apresentam o caso uniparamétrico. Contudo, cabe destacar que as interpretações valem também para o caso multiparamétrico, com as devidas extensões adequadas.

Como pode ser observado na Figura 3, a estatística do teste da razão de verossimilhanças generalizada considera que a distância do logaritmo das funções de verossimilhanças sob o modelo restrito (sob \mathcal{H}_0) e irrestrito dependam de θ_0 e $\hat{\theta}$. Veja que quanto maior a diferença $(\hat{\theta} - \theta_0)$ maior será também a diferença $\ell(\hat{\theta}) - \ell(\theta_0)$, o que dá indícios contra a hipótese \mathcal{H}_0 .

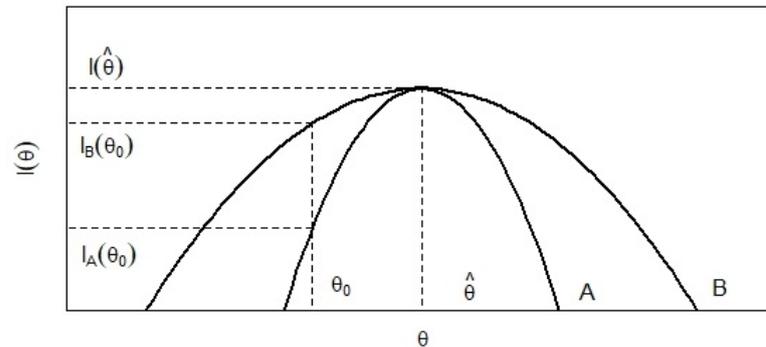
Figura 3 – Representação geométrica da estatística do teste da razão de verossimilhanças generalizada.



Fonte: Adaptado de Muggeo e Lovinson (2014)

A estatística do teste Wald avalia a adequabilidade das hipóteses por meio do uso da distância quadrática $(\hat{\theta} - \theta)^2$. Contudo, analisando a Figura 4, duas funções de verossimilhanças, sendo uma delas menos favorável à hipótese nula \mathcal{H}_0 (caso A), podem fornecer o mesmo resultado para a distância $(\hat{\theta} - \theta)^2$. Desse modo, se faz necessário a ponderação da distância quadrática por meio da utilização da curvatura do logaritmo da verossimilhança $J(\theta) = -\frac{d^2}{d\theta^2} \ell(\theta)$, avaliada em $\hat{\theta}$. Usualmente, utiliza-se também a curvatura média - informação de Fisher $I_F(\theta)$ - avaliada em $\hat{\theta}$.

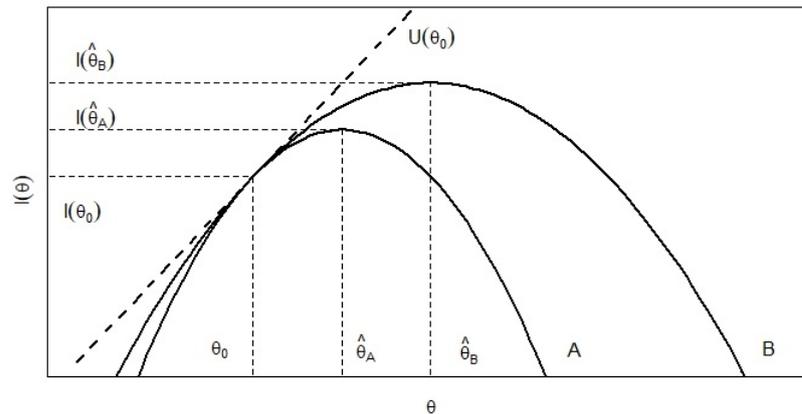
Figura 4 – Representação geométrica da estatística do teste de Wald.



Fonte: Adaptado de Muggeo e Lovinson (2014)

Com relação à estatística do teste Escore, note que se θ_0 estiver próximo de $\hat{\theta}$, então a função escore avaliada em θ_0 , $U(\theta_0)$ estará próximo de zero. Como não é de interesse o sinal da função escore, $U(\theta_0)$, então uma solução seria o uso de $U(\theta_0)^2$. Todavia, como pode ser visto pela Figura 5, duas funções de verossimilhanças sendo uma delas menos favorável a hipótese nula \mathcal{H}_0 (caso B) podem ter o mesmo valor de $U(\theta_0)^2$. Uma solução, então, seria ponderar $U(\theta_0)^2$ pelo o inverso da curvatura do logaritmo da verossimilhança $J(\theta)$ avaliada θ_0 , podendo ser utilizado também a curvatura média $I_F(\theta)$.

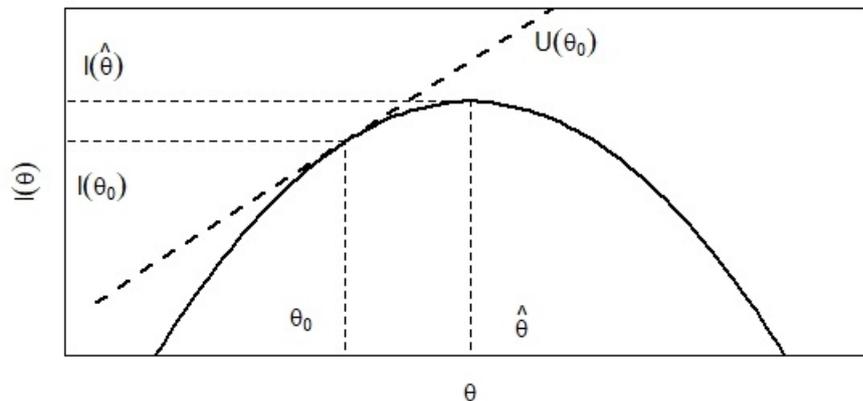
Figura 5 – Representação geométrica da estatística do teste Escore.



Fonte: Adaptado de Muggeo e Lovinson (2014)

Diferente das estatísticas dos testes Wald e Escore, a estatística do teste Gradiente não é considerada uma distância quadrática, ela mede a projeção ortogonal do vetor Escore, avaliada em θ_0 , com relação ao vetor $(\hat{\theta} - \theta_0)$ (LEMONTE, 2016). Observe que tal estatística utiliza simplesmente a distância $(\hat{\theta} - \theta_0)$, e para anular o sinal dessa diferença usa como ponderação a função escore $U(\theta_0)$ tendo em vista que $(\hat{\theta} - \theta_0)$ e $U(\theta_0)$ possuem o mesmo sinal.

Figura 6 – Representação geométrica da estatística do teste Gradiente.



Fonte: Adaptado de Muggeo e Lovinson (2014)

3.6 Testes de hipóteses em modelos lineares simétricos

Pode-se encontrar algumas referências que tratam sobre testes de hipóteses na classe dos modelos simétricos. Braga (2007) trata dos refinamentos para testes de hipóteses em modelos de regressão não-lineares simétricos; o trabalho de Brito (2009) é uma extensão da referência anterior, sendo aplicado ao caso heteroscedástico. Araújo (2012) trata da

utilização da verossimilhança perfilada (e sua forma modificada) em modelos não-lineares simétricos. Medeiros e Ferrari (2017) obtêm um refinamento para os testes assintóticos clássicos e Gradiente em modelos simétricos e log-simétricos.

De modo geral, considerando o modelo (2.1) e supondo interesse no teste $\mathcal{H}_0 : \boldsymbol{\beta} = \boldsymbol{\beta}_0$ contra $\mathcal{H}_1 : \boldsymbol{\beta} \neq \boldsymbol{\beta}_0$ as estatísticas de teste de Wald, Escore e Gradiente são, respectivamente

$$S_{RV} = 2[\ell(\hat{\boldsymbol{\beta}}, \hat{\phi}) - \ell(\boldsymbol{\beta}_0, \phi)] \quad (3.7)$$

$$S_W = \frac{4d_g}{\hat{\phi}} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^\top (\mathbf{X}^\top \mathbf{X}) (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \quad (3.8)$$

$$S_R = \frac{1}{4\phi d_g} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top \mathbf{D}(\mathbf{v}) \mathbf{X} (\mathbf{X}^\top \mathbf{X}) \mathbf{X}^\top \mathbf{D}(\mathbf{v}) (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \quad (3.9)$$

$$S_T = \frac{1}{\hat{\phi}} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^\top \mathbf{D}(\mathbf{v}) \mathbf{X} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0). \quad (3.10)$$

Em Lemonte (2012) obtêm-se as expressões das estatísticas para testar uma partição do vetor de parâmetros de interesse. Ou seja, considerando $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)^\top$, em que $\boldsymbol{\beta}_1 = (\beta_1, \dots, \beta_q)^\top$ e $\boldsymbol{\beta}_2 = (\beta_{q+1}, \dots, \beta_p)^\top$, conseqüentemente $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$, sendo \mathbf{X}_1 a matriz $n \times q$ e \mathbf{X}_2 a matriz $n \times (p - q)$; e o vetor escore sendo $\mathbf{U}_\beta = (\mathbf{U}_{\beta_1}^\top, \mathbf{U}_{\beta_2}^\top)^\top$, com $\mathbf{U}_{\beta_1} = \frac{1}{\phi} \mathbf{X}_1^\top \mathbf{D}(\mathbf{v}) (\mathbf{y} - \boldsymbol{\mu})$ e $\mathbf{U}_{\beta_2} = \frac{1}{\phi} \mathbf{X}_2^\top \mathbf{D}(\mathbf{v}) (\mathbf{y} - \boldsymbol{\mu})$. A informação de Fisher para $\boldsymbol{\beta}$ nesse caso será

$$\mathbf{K}_{\beta\beta} = \begin{bmatrix} \mathbf{K}_{\beta_1\beta_1} & \mathbf{K}_{\beta_1\beta_2} \\ \mathbf{K}_{\beta_2\beta_1} & \mathbf{K}_{\beta_2\beta_2} \end{bmatrix} = \frac{\alpha_{2,0}}{\phi} \begin{bmatrix} \mathbf{X}_1^\top \mathbf{X}_1 & \mathbf{X}_1^\top \mathbf{X}_2 \\ \mathbf{X}_2^\top \mathbf{X}_1 & \mathbf{X}_2^\top \mathbf{X}_2 \end{bmatrix},$$

em que $\alpha_{r,s} = \mathbb{E}(t^{(r)}(z)z^s)$ para $r, s = 0, 1, 2, 3$, com $t(z) = \log g(z^2)$ e $t^{(r)}(z) = \frac{d^r t(z)}{dz^r}$, tais expressões podem ser encontradas em Lemonte (2012). Portanto, para testar uma partição de $\boldsymbol{\beta}$, isto é, $\mathcal{H}_0 : \boldsymbol{\beta}_2 = \boldsymbol{\beta}_{20}$ versus $\mathcal{H}_1 : \boldsymbol{\beta}_2 \neq \boldsymbol{\beta}_{20}$ com $\boldsymbol{\beta}_{20}$ um vetor fixo com dimensão $p - q$, as estatísticas dos testes Wald, Escore e Gradiente são respectivamente

$$S_{RV} = 2[\ell(\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\phi}) - \ell(\tilde{\boldsymbol{\beta}}_1, \boldsymbol{\beta}_{20}, \tilde{\phi})] \quad (3.11)$$

$$S_W = -\frac{\alpha_{2,0}}{\hat{\phi}} (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_{20})^\top (\mathbf{R}^\top \mathbf{R}) (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_{20}) \quad (3.12)$$

$$S_R = -\frac{1}{\tilde{\phi} \alpha_{2,0}} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})^\top \mathbf{D}(\mathbf{v}) \mathbf{X}_2 (\mathbf{R}^\top \mathbf{R})^{-1} \mathbf{X}_2^\top \mathbf{D}(\mathbf{v}) (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}}) \quad (3.13)$$

$$S_T = \frac{1}{\tilde{\phi}} (\mathbf{y} - \mathbf{X}\tilde{\boldsymbol{\beta}})^\top \mathbf{D}(\mathbf{v}) \mathbf{X}_2 (\hat{\boldsymbol{\beta}}_2 - \boldsymbol{\beta}_{20}). \quad (3.14)$$

Nos casos acima, temos que $\mathbf{R} = \mathbf{X}_2 - \mathbf{X}_1\mathbf{C}$, em que $\mathbf{C} = (\mathbf{X}_1^\top \mathbf{X}_1)^{-1} \mathbf{X}_1^\top \mathbf{X}_2$. Sob hipótese nula \mathcal{H}_0 as estatísticas dos testes acima convergem para uma distribuição qui-quadrado com $p - q$ graus de liberdade χ_{p-q}^2 .

Para o parâmetro ϕ obtém-se também as expressões dos testes assintóticos veja Lemonte (2012). Supondo interesse em testar $\mathcal{H}_0 : \phi = \phi_0$ contra $\mathcal{H}_1 : \phi \neq \phi_0$, e que $\phi_0 > 0$, as estatísticas de Wald, Escore e Gradiente são dadas respectivamente

$$S_{RV} = 2[\ell(\hat{\boldsymbol{\beta}}, \hat{\phi}) - \ell(\tilde{\boldsymbol{\beta}}, \phi_0)] \quad (3.15)$$

$$S_W = n(1 - \alpha_{2,2}) \left(\frac{\hat{\phi} - \phi_0}{\hat{\phi}} \right)^2 \quad (3.16)$$

$$S_R = \frac{n \left(\frac{\bar{m}(\tilde{\boldsymbol{\beta}}, \phi_0)}{\phi_0} - 1 \right)^2}{1 - \alpha_{2,2}} \quad (3.17)$$

$$S_T = n \left(\frac{\bar{m}(\tilde{\boldsymbol{\beta}}, \phi_0)}{\phi_0} - 1 \right) \left(\frac{\hat{\phi} - \phi_0}{\phi_0} \right). \quad (3.18)$$

Nesta situação o parâmetro $\boldsymbol{\beta}$ é considerado como um parâmetro de incômodo/perturbação (nuisance parameter). O estimador $\tilde{\boldsymbol{\beta}}$ representa o estimador de máxima verossimilhança de $\boldsymbol{\beta}$ quando a hipótese \mathcal{H}_0 for verdade, além disso $\bar{m}(\boldsymbol{\beta}, \phi) = \frac{\sum_{i=1}^n v_i u_i}{n}$, em que $v_i = -2W_g(u_i)$ é dado na Tabela 2 e $u_i = \frac{(y_i - \mu_i)^2}{\phi}$. As estatísticas dos testes acima sob a suposição da hipótese nula \mathcal{H}_0 convergem para a uma distribuição qui-quadrado com 1 grau de liberdade χ_1^2 .

4 REFINAMENTO VIA *BOOTSTRAP*

Como visto nos capítulos anteriores, as estatísticas do teste da razão de verossimilhanças generalizada, Wald, Escore e Gradiente possuem o mesmo comportamento assintótico, ou seja, utilizando grandes amostras, esses testes convergem assintoticamente para uma distribuição qui-quadrado, tanto sob a hipótese nula como sob as hipóteses alternativas de Pitman. Por conta disso, para pequenas amostras, essas estatísticas podem não possuir resultados satisfatórios para aproximação qui-quadrado. Sendo assim, com o objetivo de melhorar essa aproximação diversos métodos de aperfeiçoamento foram propostos, no caso, utilizando formas modificadas dessas estatísticas.

Como exemplo desses métodos, pode-se citar a correção de Bartlett e tipo-Bartlett (CORDEIRO; CRIBARI-NETO, 2014) e a correção de Skovgaard (SKOVGAARD, 1996). Aplicações da correção de Bartlett para as estatísticas do teste da razão de verossimilhanças generalizada, Wald e Escore são discutidos em Cordeiro e Cribari-Neto (2014), e para a estatística gradiente podem ser encontrados em Lemonte (2016). Para os testes de hipóteses em modelos lineares simétricos cabe citar novamente os trabalhos de Braga (2007), Brito (2009), Araújo (2012) e Medeiros e Ferrari (2017) que fazem uso desses métodos no refinamento de estatísticas de testes em outras classes de modelos simétricos.

Uma outra opção utilizada em refinamentos das estatísticas dos testes de hipóteses sem a necessidade de realizar cálculos analíticos complexos é o uso do método *Bootstrap* proposto por Efron (1979). Tal procedimento pode ser aplicado a diversos problemas estatísticos como na estimação de erro-padrão e viés, na construção de intervalos de confiança e na realização de testes de hipóteses, para mais detalhes veja Tibshirani e Efron (1993) e Davison e Hinkley (1997). Podemos encontrar algumas publicações sobre a utilização do *Bootstrap* em refinamento de estatísticas de testes de hipóteses. Stein (2008) faz uso do refinamento *Bootstrap* em testes da razão de verossimilhanças usando a proposta de Barndorff-Nielsen (1983) em modelos lineares mistos, Lozano (2015) utiliza em modelos não lineares generalizados em séries de potência, Lôbo (2016) utiliza em modelos não lineares da família exponencial e Nobre et al. (2020) também utiliza o refinamento *Bootstrap* em testes para componentes de variância em modelos de efeitos aleatórios usando estatísticas U .

Como pode ser visto em Tibshirani e Efron (1993), o método *Bootstrap* consiste em estimar a função de distribuição F a partir da função de distribuição empírica \hat{F} de

uma amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$, e a partir de reamostras $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^\top$ geradas com base em \hat{F} estimar a distribuição amostral de alguma estatística de interesse. Quando a função de distribuição for parametrizada por θ , temos que $\hat{F}(y) = F_{\hat{\theta}}(y)$ e assim o *Bootstrap* é dito ser paramétrico. No caso não-paramétrico, a distribuição empírica $\hat{F}(y)$ designa probabilidade $1/n$ para cada valor da amostra y_i , $i = 1, \dots, n$. Outras versões do *Bootstrap* podem ser encontradas em Chernick (2011).

No presente trabalho obteremos o refinamento das estatísticas dos testes de hipóteses vistas na seção 3.6, utilizando o *Bootstrap* paramétrico. Assim, seja a amostra $\mathbf{y} = (y_1, \dots, y_n)^\top$ de variáveis aleatórias independentes tal que $y_i \sim S(\mu_i, \phi)$, e considere os seguintes passos:

1. Calcula-se o valor da estatística de teste de interesse, \mathcal{S} (seção 3.6).
2. Geram-se B amostras *Bootstrap* (\mathbf{y}^*) de tamanho n , de modo que $\mathbf{y}^* \sim S(\hat{\mu}_i, \hat{\phi})$ com $\hat{\mu}_i$ e $\hat{\phi}$ sendo os estimadores de máxima verossimilhança, sendo $\hat{\mu}_i = \mu(\hat{\beta}, \mathbf{x}_i)$.
3. De cada amostra *Bootstrap* (\mathbf{y}^*) obtida no passo 2, calcula-se a estatística de teste de interesse obtida no passo 1, obtendo-se assim B valores do teste *Bootstrap*, \mathcal{S}^* .
4. Ordenam-se os B valores \mathcal{S}^* obtidos, de forma que a posição $(1 - \alpha) \cdot B$, com $\alpha \in (0, 1)$ represente uma estimativa *Bootstrap* do quantil $(1 - \alpha)$ da distribuição da estatística de interesse.

A região de rejeição do teste pode ser definida a partir do quantil *Bootstrap* obtido no passo 4, de forma que a hipótese nula \mathcal{H}_0 seja rejeitada quando a estatística de teste S for maior que o quantil *Bootstrap*, ou seja, $\mathcal{S} > \mathcal{S}_{(1-\alpha) \cdot B}^*$. A regra de decisão poderá ser expressa também em termos do valor p *Bootstrap* (p_{boot}) que nesse caso é dado por

$$p_{\text{boot}} = \frac{\#\{\mathcal{S}_b^* > \mathcal{S}\}}{B}, \quad b = 1, \dots, B,$$

com $\#$ representando a cardinalidade do conjunto. Assim a hipótese nula \mathcal{H}_0 é rejeitada se p_{boot} for menor que o nível de significância fixado α , no caso $p_{\text{boot}} < \alpha$.

4.1 Resultados Numéricos

Foram realizadas algumas simulações para verificar o desempenho das estatísticas dos testes de hipóteses assintóticos em alguns modelos simétricos. Para tal objetivo comparou-se as taxas de rejeição empíricas, ou seja, a proporção de vezes que a hipótese

nula foi rejeitada sendo ela verdadeira, e o poder empírico, que é a proporção de vezes que a hipótese nula foi rejeitada sob a validade/veracidade da hipótese alternativa. Espera-se que as taxas de rejeição empíricas dos testes estejam próximas ao nível de significância fixado.

Os modelos foram simulados utilizando tamanhos amostrais $n = 10, 50, 100$ e 200 ; aqui utilizamos 10.000 simulações de Monte Carlo. Além disso foram fixados os níveis de significância de $\alpha = 1\%, 5\%$ e 10% . Para obter um refinamento via *Bootstrap* utilizando a proposta paramétrica - dado que é conhecido a distribuição a ser utilizada -, foi escolhido o tamanho de amostras *Bootstrap* como sendo $B = 10.000$ afim de comparar com as simulações de Monte Carlo. Obtido os valores *Bootstrap* dos testes, procedeu-se para o cálculo das taxas de rejeição empírica, como forma de comparar com os níveis de significância fixados.

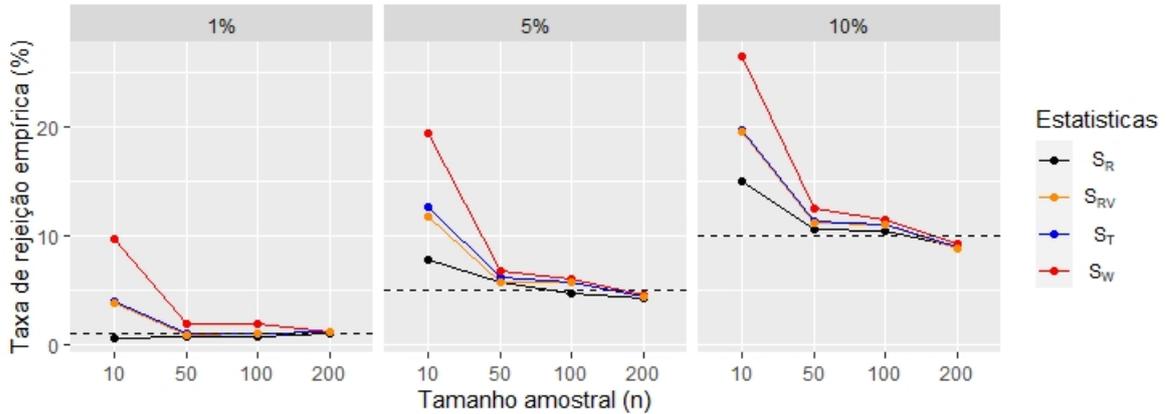
Para auxílio computacional utilizou-se o software R através da biblioteca `elliptical` (CYSNEIROS, 2005).

- Modelo t-Student

- Situação 1

Considerando o seguinte modelo $\mu_i(\beta, x_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$ com a fonte de variação $\epsilon \sim t\text{-Student}(4)$. Os parâmetros do modelo foram simulados tendo os seguintes valores $\beta_0 = 4$, $\beta_1 = 5$ e $\beta_2 = 2$. As variáveis explicativas foram obtidas a partir da variável aleatória $\mathcal{U}(0, 1)$. Aqui será proposto testar $\mathcal{H}_0 : \beta_0 = 4$ contra $\mathcal{H}_1 : \beta_0 \neq 4$. Na Figura 7 e na Tabela 6 (Apêndice A.1) podem ser observadas as taxas de rejeição empíricas da hipótese nula \mathcal{H}_0 . Os valores obtidos, na maioria dos casos, são superiores aos níveis de significância adotados sobretudo para tamanhos amostrais pequenos o que demonstra a necessidade de refinamento para essas situações.

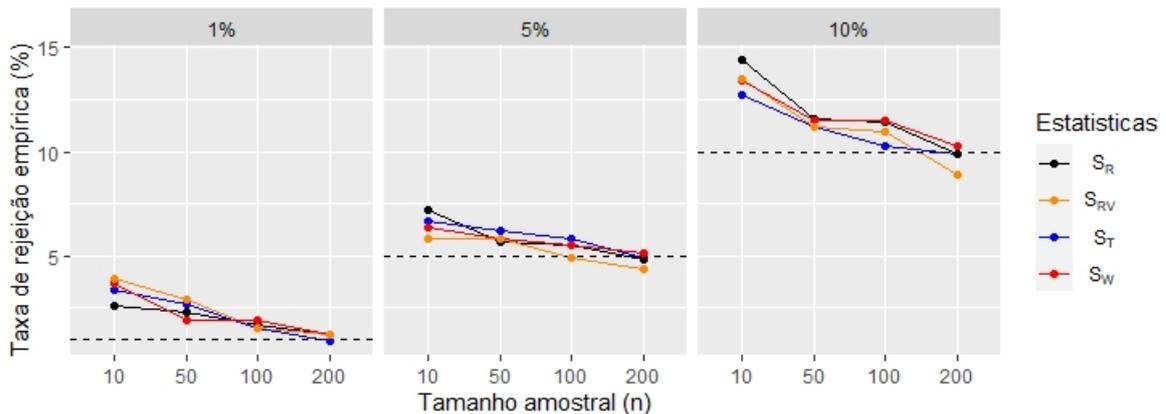
Figura 7 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 1.



Fonte: O autor

Desse modo, utilizou-se o *Bootstrap* paramétrico como método de refinamento para os tamanhos $n = 10, 25, 50, 100$ e utilizando $B = 10.000$ amostras *Bootstrap*. Na Figura 8 e na Tabela 7 (Apêndice A.1) tem-se os valores das taxas de rejeição empíricas obtidas através da simulação, observe que tais valores estão mais próximos dos níveis de significâncias.

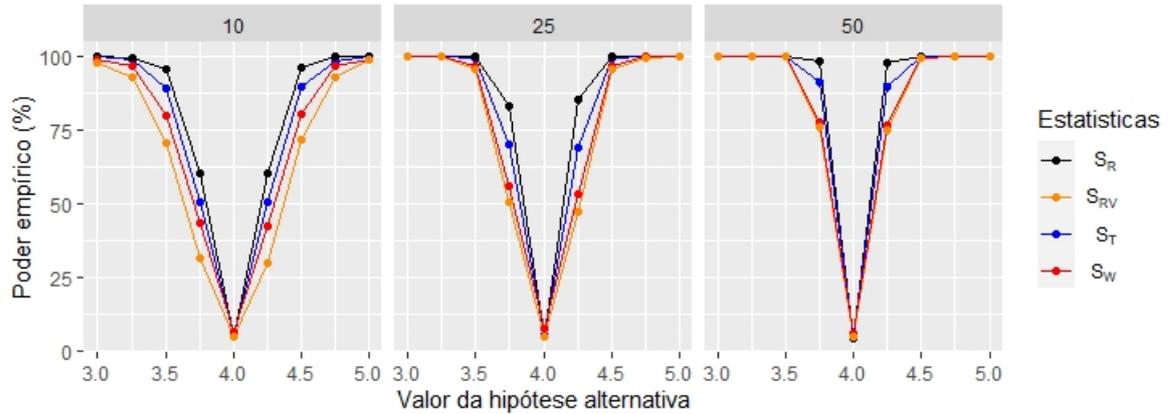
Figura 8 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade para situação 1.



Fonte: O autor

Em seguida, ainda utilizando o refinamento *Bootstrap* paramétrico obtemos os resultados para o poder empírico dos testes (veja Figura 9 e Tabela 8 (Apêndice A.1)). As hipóteses alternativas foram obtidas a partir de $\mathcal{H}_1 : \beta_0 = \gamma_0$ com γ_0 assumindo valores de 3 a 5 com incrementos de 0,25. Fixando $\alpha = 5\%$ e escolhendo os tamanhos de amostras $n = 10, 25$ e 50 observa-se que o teste Escore apresentou as maiores taxas de poder empírico para os respectivos tamanhos amostrais.

Figura 9 – Taxas do poder empírico para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o *Bootstrap* paramétrico para situação 1.

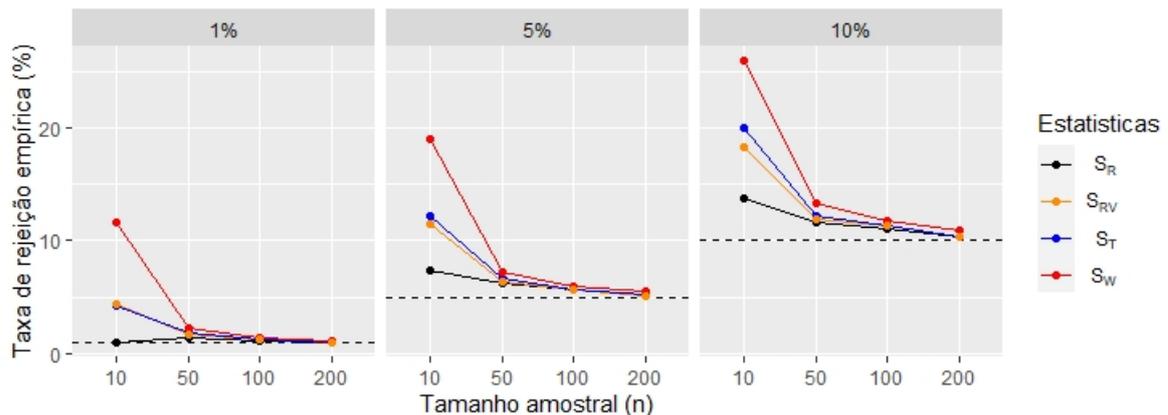


Fonte: O autor

- Situação 2

Aqui o modelo simulado foi $\mu_i(\beta, x_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i$ com a fonte de variação $\epsilon \sim t\text{-Student}(4)$. As variáveis explicativas foram obtidas respectivamente de $\mathcal{U}(0, 4; 0, 61)$ e $\mathcal{U}(8, 7; 11, 2)$, pois o modelo aqui simulado será base para tomada de decisão a ser utilizado no capítulo de aplicação. Os parâmetros do modelo foram simulados tendo os seguintes valores $\beta_0 = 9$, $\beta_1 = 9$ e $\beta_2 = 0$. Desse modo estamos interessados em testar a inclusão ou não do parâmetro de regressão β_2 , isto é, testar $\mathcal{H}_0 : \beta_2 = 0$ contra $\mathcal{H}_1 : \beta_2 \neq 0$.

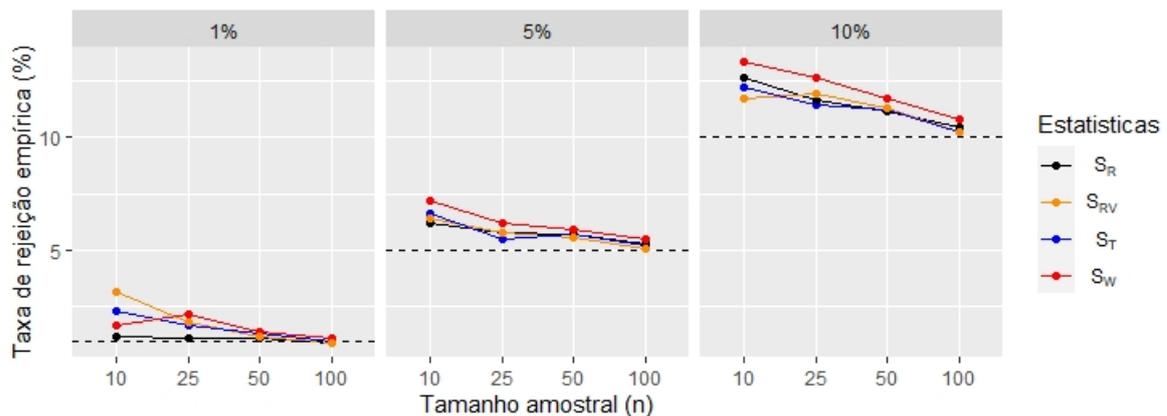
Figura 10 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 2.



Fonte: O autor

Na Figura 10 e na Tabela 9 (Apêndice A.1) observamos as taxas de rejeição empíricas da hipótese nula \mathcal{H}_0 . Percebe-se que de modo geral os testes apresentaram um comportamento liberal, ou seja, as taxas empíricas obtidas são superiores aos níveis de significância fixados, principalmente para os menores tamanhos amostrais o que demonstra uma necessidade de refinamento para tal situação. Assim, utilizamos um refinamento *Bootstrap* paramétrico seguindo os passos descritos anteriormente para $n = 10, 25, 50, 100$ e utilizando $B = 10.000$ amostras *Bootstrap*. Perceba na Figura 11 que utilizando o refinamento os valores das taxas de rejeição empíricas estão mais próximas dos níveis de significâncias adotados.

Figura 11 – Taxas de rejeição empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o *Bootstrap* paramétrico para situação 2.

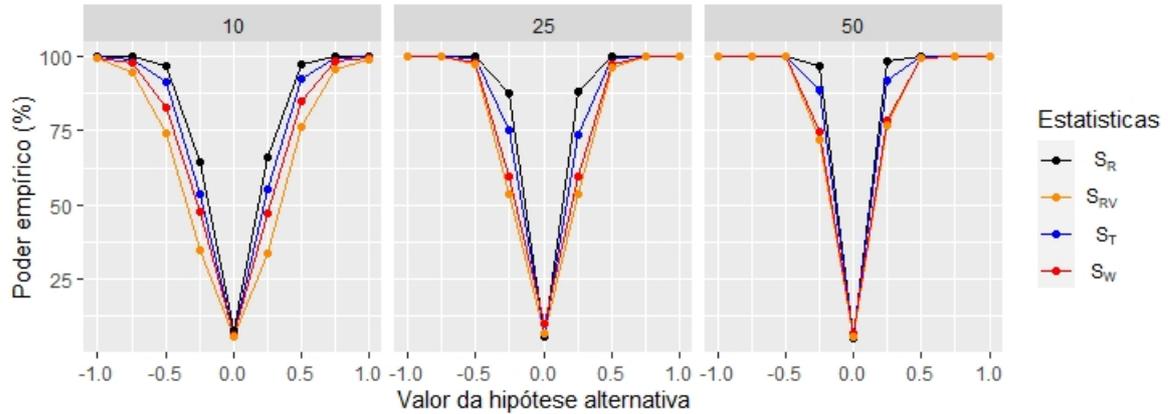


Fonte: O autor

Na Figura 12 e na Tabela 11 (Apêndice A.1), obtemos os resultados para o poder empírico dos testes utilizando o refinamento *Bootstrap*. As hipóteses alternativas foram obtidas a partir de $\mathcal{H}_1 : \beta_2 = \beta_{20}$ com β_{20} assumindo valores de -1 a 1 com incrementos de 0,25. No caso fixamos o nível de significância 5% e escolheu-se os tamanhos de amostras $n = 10, 25$ e 50.

Observa-se que o teste *Score* teve as maiores taxas de poder empírico para os respectivos tamanhos amostrais. No caso em que $n = 50$ os testes apresentaram um aumento na taxa de poder empírico especialmente para valores da hipótese alternativa próximos ao da hipótese nula \mathcal{H}_0 , tendo as estatísticas dos teste resultados próximos entre si.

Figura 12 – Taxas do poder empírico para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o *Bootstrap* paramétrico para situação 2.



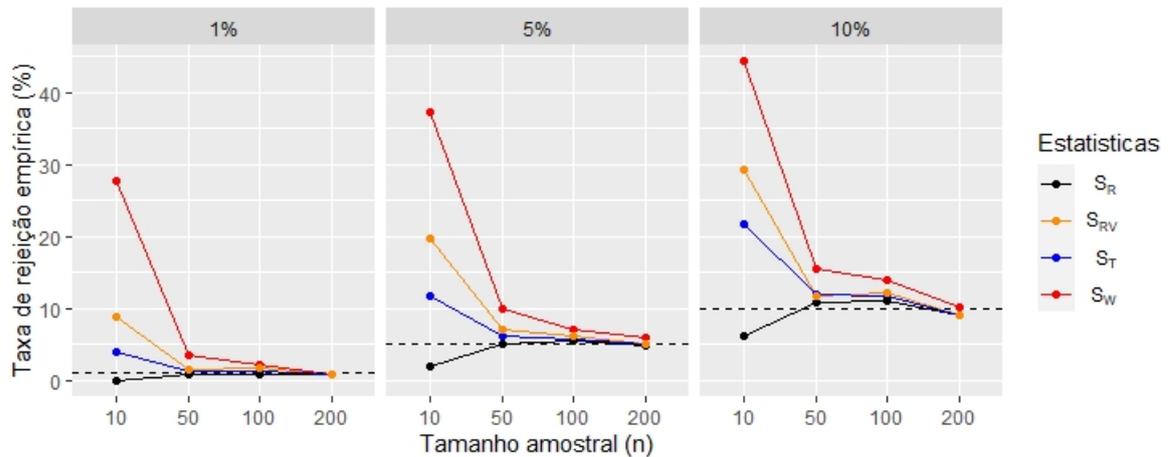
Fonte: O autor

– Modelo Cauchy

O modelo simulado neste cenário foi $\mu_i(\beta, x_i) = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}$, com $\beta_0 = 7$, $\beta_1 = 4$ e $\beta_2 = 0$ e a fonte de variação $\epsilon_i \sim \text{Cauchy Padrão}$. As variáveis explicativas foram obtidas a partir da variável aleatória $\mathcal{U}(0, 1)$. O teste aqui interessado é sobre a inclusão ou não do parâmetro de regressão β_2 , isto é estamos interessados em testar $\mathcal{H}_0 : \beta_2 = 0$ contra $\mathcal{H}_1 : \beta_2 \neq 0$.

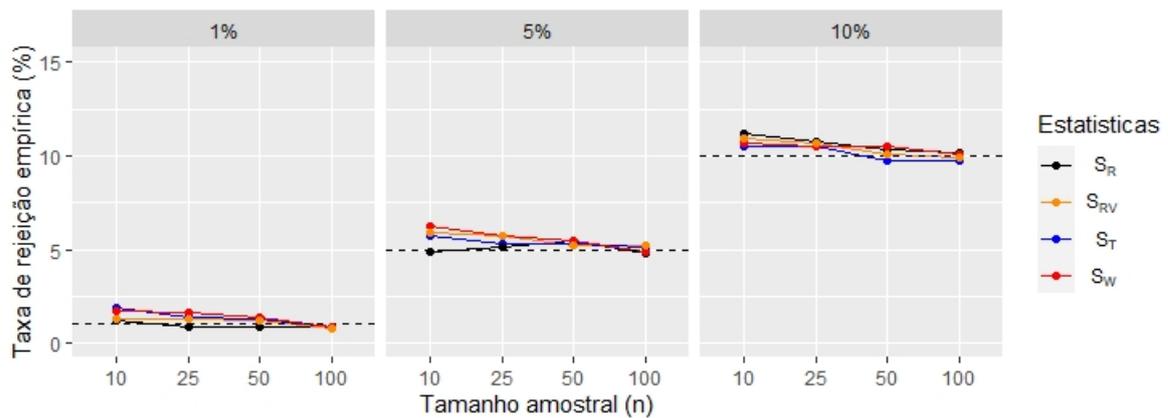
Os resultados da rejeição da hipótese nula quanto esta for verdadeira podem ser vistos na Figura 13 e na Tabela 13 (Apêndice A.2). De modo geral, as taxas de rejeição empíricas se aproximaram dos níveis de significância quando os tamanhos das amostras aumentaram. Porém para os tamanhos amostrais pequenos as taxas de rejeição dos testes estão bem distantes dos níveis de significância fixados, necessitando assim de um refinamento. Desse modo, utilizando o *Bootstrap* paramétrico obtemos um refinamento para as estatísticas considerando os tamanhos amostrais $n = 10, 25$ e 50 e $B = 10.000$ amostras *Bootstrap*. Na Figura 14, percebemos que as taxas de rejeição empíricas obtidas estão próximas dos níveis fixados, indicando um bom desempenho do procedimento.

Figura 13 – Taxas de rejeição empíricas para o modelo Cauchy.



Fonte: O autor

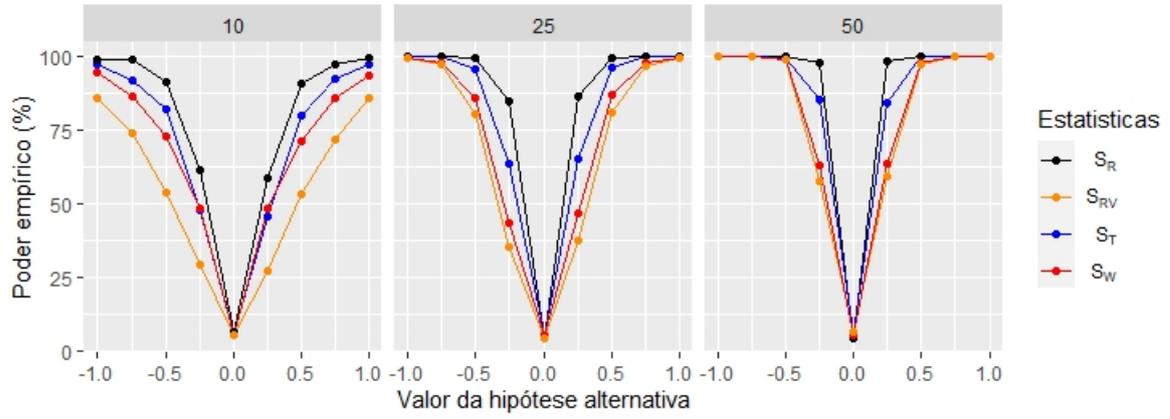
Figura 14 – Taxas de rejeição empíricas para o modelo Cauchy, utilizando o *Bootstrap* paramétrico.



Fonte: O autor

Na Figura 15 e na Tabela 14 (Apêndice A.2) estão representados os poderes empíricos dos testes utilizando o refinamento *Bootstrap*. No caso as hipóteses alternativas foram construída da forma $\mathcal{H}_1 : \beta_2 = \beta_{20}$ com β_{20} assumindo valores de -1 a 1 com incrementos de 0,25. Em concordância ao que foi verificado na Figura 13, no tamanho amostral $n = 10$ o teste Wald obteve maior taxa de poder empírico quando sob hipótese nula. Percebe-se que os valores das taxas de poder empírico das estatísticas aumentaram quando o tamanho amostral aumentou, sobretudo nos caso que hipótese alternativa assumiu valores próximos à hipótese nula. Nessa última situação, para $n = 50$ as estatísticas do teste Escore e do teste Gradiente obtiveram as maiores taxas de poder empírico, enquanto que os testes da razão de verossimilhanças generalizada e Wald apresentaram os menores valores sendo próximos entre si.

Figura 15 – Taxas do poder empírico para o modelo Cauchy, utilizando o *Bootstrap* paramétrico.



Fonte: O autor

5 APLICAÇÃO

Nesta seção será ilustrada a utilização das estatísticas dos testes assintóticos clássicos e Gradiente, utilizando o *Bootstrap* como método de refinamento. O conjunto de dados (Tabela 3) apresentado em Draper e Stoneman (1966) trata da resistência (y_i) de vigas de madeira em função da gravidade específica (x_{1i}) e do teor de umidade do material (x_{2i}).

Tabela 3 – Conjunto de dados com relação à resistência de vigas de madeira.

Resistência (y_i)	Gravidade específica (x_{1i})	Umidade (x_{2i})
11,14	0,499	11,1
12,74	0,558	8,9
13,13	0,604	8,8
11,51	0,441	8,9
12,38	0,550	8,8
12,60	0,528	9,9
11,13	0,418	10,7
11,70	0,480	10,5
11,02	0,406	10,5
11,41	0,467	10,7

Fonte: Adaptado de Draper e Stoneman (1966)

Em Lemonte (2012) pode ser encontrado uma análise deste conjunto de dados considerando que a fonte de variação seja distribuída conforme t-Student com $\nu = 4$ graus de liberdade. Tal autor considera o seguinte modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i, \quad i = 1, \dots, 10. \quad (5.1)$$

Na Tabela 4 são apresentadas as estimativas de máxima verossimilhança dos parâmetros do modelo e seus respectivos erros padrão.

Tabela 4 – Estimativas dos parâmetros de localização e de escala do modelo (5.1).

Parâmetros	Estimativas	Erro Padrão
β_0	9,05	1,29
β_1	9,23	1,21
β_2	-0,17	0,08
ϕ	0,15	0,04

Fonte: O autor

Com o objetivo de verificar se a variável umidade é significativa ou não no modelo adotado, prosseguiu-se com o seguinte teste de hipóteses $\mathcal{H}_0 : \beta_2 = 0$ contra

$\mathcal{H}_1 : \beta_2 \neq 0$. Na Tabela 5 encontram-se os valores calculados das estatísticas dos testes assintóticos e os respectivos valores p e valores p *bootstrap* (p_{boot}). No caso, escolheu-se o teste Gradiente, objeto deste trabalho e o teste Wald, o qual é fornecido na biblioteca `elliptical`. Para obtermos um refinamento *bootstrap* paramétrico iniciamos com as estimativas do modelo (Tabela 4) e geramos $B = 10.000$ amostras *bootstrap* e seguimos os passos descritos no Capítulo 4. Percebe-se que o teste Gradiente rejeita a hipótese nula \mathcal{H}_0 levando em conta níveis de significância maiores que $\alpha = 10\%$, tanto para a aproximação assintótica como para o refinamento *bootstrap*, enquanto que o teste Wald rejeita tal hipótese para as duas situações utilizando o nível de significância $\alpha = 5\%$.

Tabela 5 – Valores das estatísticas do testes assintóticos e seus respectivos valores p e valores p *bootstrap* (p_{boot}) para modelo 5.1

Estatística	Valor	valor p	valor p_{boot}
S_T	2,674	0,102	0,114
S_W	4,278	0,039	0,046

Fonte: O autor

Simulamos um modelo com estrutura semelhante na Seção (4.1), e observamos que utilizando o refinamento *bootstrap* paramétrico, o teste Gradiente apresentou taxas de rejeição empíricas mais próximas aos níveis de significância adotados quando comparado com o teste Wald; além disso, os valores das taxas do poder empírico foram maiores para o teste Gradiente. Assim, levando em conta o teste Gradiente e utilizando o refinamento *bootstrap* paramétrico temos que a hipótese nula \mathcal{H}_0 não é rejeitada a um nível de $\alpha = 10\%$, ou seja, a variável umidade não é significativa para o modelo considerado.

6 CONCLUSÕES

Os testes de hipóteses assintóticos são uma alternativa aos procedimentos exatos, quando em algumas situações não for possível ou for extremamente complicado de obtê-los. Nesse sentido, os testes da razão de verossimilhanças generalizada, Wald, Escore e Gradiente são amplamente utilizados. Sob situações regulares, tais procedimentos possuem comportamento assintótico similar, tanto sob a hipótese nula como sob sequência de hipóteses alternativas locais de Pitman. Além disso, autores como Lemonte (2016) e Lemonte e Ferrari (2012) concluem ainda que os quatro testes sob determinadas condições possuem idênticos poder local, o que faz do teste Gradiente, dado a facilidade em sua obtenção, ser uma alternativa frente aos demais.

Seguindo esse raciocínio a aplicação de tais testes assintóticos em modelos lineares simétricos mostram-se coerente com a proposta aqui apresentada. Viu-se através das simulações que de um modo geral as taxas de rejeição empíricas da hipótese nula \mathcal{H}_0 se aproximaram do nível de significância adotado para tamanhos amostrais moderados, o que indica uma boa aproximação para a distribuição qui-quadrado de referência. Com relação ao poder dos testes, observou-se que para tamanhos amostrais moderados os testes apresentavam taxas empíricas próximas, porém os testes Gradiente e Escore obtiveram taxas um pouco maiores, demonstrando que o teste Gradiente é uma alternativa mais fácil a ser considerada.

Com relação aos modelos lineares simétricos viu-se que essa classe de modelos é uma estratégia robusta diante da suposição de não normalidade e da presença de observações influentes. Diversas referências citadas tratam de extensões dos modelos lineares simétricos. Desse modo, como sugestão para trabalhos futuros poderíamos utilizar a proposta apresentada nesta monografia em modelos lineares simétricos heterocedásticos, em modelos lineares generalizados simétricos e em modelos lineares mistos elípticos, por exemplo.

REFERÊNCIAS

- ANDERSON, T. W.; FANG, K.-T. **Theory and applications of elliptically contoured and related distributions**. [S.l.], 1990. Citado na p. 16.
- ARAÚJO, M. C. d. **Verossimilhança perfilada nos modelos não lineares simétricos heteroscedásticos**. 2012. Dissertação de Mestrado – Universidade Federal de Pernambuco. Citado nas pp. 28, 31.
- ARELLANO-VALLE, R. **Distribuições elípticas: Propriedades, inferência e aplicações a modelos de regressão**. 1994. Tese (Doutorado) – Universidade de São Paulo. Citado nas pp. 13, 20.
- BARNDORFF-NIELSEN, O. On a formula for the distribution of the maximum likelihood estimator. **Biometrika**, v. 70, n. 2, p. 343–365, 1983. Citado na p. 31.
- BERKANE, M.; BENTLER, P. Moments of elliptically distributed random variates. **Statistics & probability letters**, v. 4, n. 6, p. 333–335, 1986. Citado na p. 16.
- BRAGA, K. S. P. R. **Aperfeiçoamento de testes de hipóteses para modelos não-lineares Simétricos**. 2007. Dissertação de Mestrado – Universidade Federal de Pernambuco. Citado nas pp. 20, 28, 31.
- BRITO, C. C. R. d. **Correção de barlett nos modelos não-lineares simétricos heteroscedásticos**. 2009. Dissertação de Mestrado – Universidade Federal de Pernambuco. Citado nas pp. 20, 28, 31.
- BUSE, A. The likelihood ratio, Wald, and Lagrange multiplier tests: An expository note. **The American Statistician**, v. 36, 3a, p. 153–157, 1982. Citado nas pp. 14, 26.
- CAMBANIS, S.; HUANG, S.; SIMONS, G. On the theory of elliptically contoured distributions. **Journal of Multivariate Analysis**, Elsevier, v. 11, n. 3, p. 368–385, 1981. Citado na p. 16.
- CHERNICK, M. R. **Bootstrap methods: A guide for practitioners and researchers**. [S.l.: s.n.], 2011. v. 619. Citado na p. 32.
- CHMIELEWSKI, M. Elliptically symmetric distributions: A review and bibliography. **International Statistical Review/Revue Internationale de Statistique**, p. 67–74, 1981. Citado na p. 13.
- CORDEIRO, G. M.; CRIBARI-NETO, F. **An introduction to Bartlett correction and bias reduction**. [S.l.]: Springer, 2014. Citado nas pp. 14, 31.
- COX, D. R.; SNELL, E. J. A general definition of residuals. **Journal of the Royal Statistical Society: Series B (Methodological)**, v. 30, n. 2, p. 248–265, 1968. Citado na p. 19.
- CYSNEIROS, F. J. **elliptical: Analysis in symmetrical linear and nonlinear regression models with unrestricted and restricted in the parameters**. [S.l.], 2005. R package version 2.0. Disponível em: <http://www.de.ufpe.br/~cysneiros/elliptical/elliptical.html>. Citado nas pp. 20, 33.
- CYSNEIROS, F. J. **Métodos restritos e validação de modelos simétricos de regressão**. 2004. Tese (Doutorado) – Universidade de São Paulo. Citado na p. 13.
- CYSNEIROS, F. J. A.; PAULA, G. A.; GALEA, M. Modelos simétricos aplicados. **Escola de Modelos de Regressão**, v. 9, 2005. Citado nas pp. 13, 16–20.

- DAVISON, A. C.; HINKLEY, D. V. **Bootstrap methods and their application**. [S.l.: s.n.], 1997. Citado na p. 31.
- DRAPER, N. R.; STONEMAN, D. M. Testing for the Inclusion of Variables in Einear Regression by a Randomisation Technique. **Technometrics**, v. 8, n. 4, p. 695–699, 1966. Citado nas pp. 12, 40.
- EFRON, B. Bootstrap Methods: Another Look at the Jackknife. **The Annals of Statistics**, Institute of Mathematical Statistics, v. 7, n. 1, p. 1–26, 1979. Citado nas pp. 14, 31.
- FANG, K.; KOTZ, S.; NG, K. Symmetric Multivariate and Related Distributions. **Monographs on Statistics and Applied Probability**, v. 36, 1990. Citado na p. 16.
- GALEA, M.; PAULA, G. A.; CYSNEIROS, F. J. On diagnostics in symmetrical nonlinear models. **Statistics & Probability Letters**, v. 73, n. 4, p. 459–467, 2005. Citado na p. 19.
- GALEA, M.; PAULA, G. A.; URIBE-OPAZO, M. On influence diagnostic in univariate elliptical linear regression models. **Statistical Papers**, v. 44, n. 1, p. 23–45, 2003. Citado na p. 19.
- HARRIS, P.; PEERS, H. The local power of the efficient scores test statistic. **Biometrika**, v. 67, n. 3, p. 525–529, 1980. Citado na p. 26.
- HAYAKAWA, T. The likelihood ratio criterion for a composite hypothesis under a local alternative. **Biometrika**, v. 62, n. 2, p. 451–460, 1975. Citado na p. 26.
- HAYAKAWA, T.; PURI, M. L. Asymptotic expansions of the distributions of some test statistics. **Ann. Inst. Statist. Math**, v. 37, Part A, p. 95–108, 1985. Citado na p. 23.
- KOWALSKI, J. et al. On the difference in inference and prediction between the joint and independent f-error models for seemingly unrelated regressions. **Communications in Statistics-Theory and Methods**, v. 28, n. 9, p. 2119–2140, 1999. Citado na p. 20.
- LANGE, K. L.; LITTLE, R. J.; TAYLOR, J. M. Robust statistical modeling using the t distribution. **Journal of the American Statistical Association**, v. 84, n. 408, p. 881–896, 1989. Citado nas pp. 13, 20.
- LEHMANN, E. L.; ROMANO, J. P. **Testing statistical hypotheses**. [S.l.]: Springer Science & Business Media, 2006. Citado na p. 21.
- LEMONTE, A. Local power properties of some asymptotic tests in symmetric linear regression models. **Journal of Statistical Planning and Inference**, v. 142, n. 5, p. 1178–1188, 2012. Citado nas pp. 13, 29, 30, 40.
- LEMONTE, A. The gradient test: another likelihood-based test. Academic Press, 2016. Citado nas pp. 14, 23–26, 28, 31.
- LEMONTE, A. J.; FERRARI, S. L. The local power of the gradient test. **Annals of the Institute of Statistical Mathematics**, v. 64, n. 2, p. 373–381, 2012. Citado nas pp. 25, 26.
- LINDSEY, J. Multivariate elliptically contoured distributions for repeated measurements. **Biometrics**, v. 55, n. 4, p. 1277–1280, 1999. Citado na p. 20.
- LITTLE, R. J. Robust estimation of the mean and covariance matrix from data with missing values. **Journal of the Royal Statistical Society: Series C (Applied Statistics)**, v. 37, n. 1, p. 23–38, 1988. Citado na p. 13.

- LIU, S. On diagnostics in conditionally heteroskedastic time series models under elliptical distributions. **Journal of Applied Probability**, p. 393–405, 2004. Citado na p. 20.
- LÔBO, T. d. S. **Correção tipo-Bartlett à estatística gradiente nos modelos não lineares da família exponencial**. 2016. Dissertação de Mestrado – Universidade Federal de Pernambuco. Citado nas pp. 14, 31.
- LOZANO, S. F. **Teste de razão de verossimilhanças Bootstrap e técnicas de diagnóstico em modelos em séries de potências não-lineares generalizados**. 2015. Dissertação de Mestrado – UNIVERSIDADE FEDERAL DE PERNAMBUCO. Citado na p. 31.
- MEDEIROS, F. M.; FERRARI, S. L. Small-sample testing inference in symmetric and log-symmetric linear regression models. **Statistica Neerlandica**, v. 71, n. 3, p. 200–224, 2017. Citado nas pp. 13, 29, 31.
- MOTA, A. A. X. **Estatística Gradiente: Conceitos e Aplicações**. [S.l.], 2017. Citado na p. 23.
- MUGGEO, V. M.; LOVISON, G. The “three plus one” likelihood-based test statistics: unified geometrical and graphical interpretations. **The American Statistician**, v. 68, n. 4, p. 302–306, 2014. Citado nas pp. 14, 26.
- NOBRE, J. S.; SINGER, J. M.; BATISTA, M. J. et al. Improved U -tests for variance components in one-way random effects models. **Brazilian Journal of Probability and Statistics**, v. 34, n. 3, p. 464–477, 2020. Citado na p. 31.
- OSORIO, F. A. **heavy: Robust estimation using heavy-tailed distributions**. [S.l.], 2019. R package version 0.38.196. Disponível em: <<https://CRAN.R-project.org/package=heavy>>. Citado na p. 20.
- PALOMAR, D. P.; ZHOU, R. **fitHeavyTail: Mean and Covariance Matrix Estimation under Heavy Tails**. [S.l.], 2020. R package version 0.1.2. Disponível em: <<https://CRAN.R-project.org/package=fitHeavyTail>>. Citado na p. 20.
- PAULA, G. A. **Modelos de regressão: com apoio computacional**. [S.l.]: IME-USP São Paulo, 2004. Citado na p. 12.
- PENAGOS, L. H. V. **Diagnóstico em modelos simétricos de regressão**. 2005. Dissertação de Mestrado – Universidade Federal de Pernambuco. Citado na p. 19.
- RAO, B. L. S. P. Remarks on univariate elliptical distributions. **Statistics & Probability Letters**, v. 10, n. 4, p. 307–315, 1990. Citado na p. 16.
- RAO, C. R. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. **Mathematical Proceedings of the Cambridge Philosophical Society**, Cambridge University Press, v. 44, n. 1, p. 50–57, 1948. Citado na p. 21.
- RAO, C. R. Score test: historical review and recent developments, p. 3–20, 2005. Citado nas pp. 21, 25.
- SAULO, H. et al. A class of asymmetric regression models for left-censored data. **Brazilian Journal of Probability and Statistics**, v. 35, n. 1, p. 62–84, 2021. Citado nas pp. 14, 20.
- SKOVGAARD, I. M. An explicit large-deviation approximation to one-parameter tests. **Bernoulli Society for Mathematical Statistics and Probability**, v. 2, n. 2, p. 145–165, 1996. Citado nas pp. 14, 31.

- STEIN, M. C. **Testes da Razão de Verossimilhanças em Modelos Lineares Mistos**. 2008. Dissertação de Mestrado – Universidade Federal de Minas Gerais. Citado na p. 31.
- TAYLOR, J. M. Properties of modelling the error distribution with an extra shape parameter. **Computational statistics & data analysis**, v. 13, n. 1, p. 33–46, 1992. Citado na p. 13.
- TERRELL, G. R. The gradient statistic. **Computing Science and Statistics**, v. 34, n. 34, p. 206–215, 2002. Citado nas pp. 13, 21, 23, 24.
- TIBSHIRANI, R. J.; EFRON, B. An introduction to the bootstrap. **Monographs on statistics and applied probability**, v. 57, p. 1–436, 1993. Citado na p. 31.
- WALD, A. Tests of statistical hypotheses concerning several parameters when the number of observations is large. **Transactions of the American Mathematical society**, v. 54, n. 3, p. 426–482, 1943. Citado na p. 21.
- WILKS, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. **The annals of mathematical statistics**, v. 9, n. 1, p. 60–62, 1938. Citado na p. 21.
- YAMAGUCHI, K. Generalized EM algorithm for model with contaminated error term. In: **PROCEEDINGS of the Seven Japan and Korea Joint Conference of Statistics**. [S.l.: s.n.], 1990. P. 107–114. Citado na p. 13.

APÊNDICE A – RESULTADOS DAS SIMULAÇÕES

A.1 Resultados das simulações considerando a distribuição t-Student com $\nu = 4$ graus de liberdade.

– Situação 1

Tabela 6 – Taxas de rejeição empíricas (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 1.

n	$\alpha\%$	S_T	S_R	S_W	S_{RV}
10	1	4,0	0,6	9,7	3,9
	5	12,7	7,8	19,4	11,8
	10	19,7	15,0	26,4	19,5
50	1	1,0	0,7	1,9	0,9
	5	6,2	5,7	06,8	5,8
	10	11,3	10,6	12,5	11,2
100	1	1,0	0,7	2,0	1,0
	5	5,8	4,8	6,1	5,8
	10	11,1	10,4	11,5	11,0
200	1	1,2	1,0	1,2	1,2
	5	4,4	4,3	4,6	4,4
	10	9,0	9,0	9,3	8,9

Fonte: O autor

Tabela 7 – Taxas de rejeição empíricas (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o *Bootstrap* paramétrico para situação 1.

n	$\alpha\%$	S_T	S_R	S_W	S_{RV}
10	1	3,4	2,6	3,7	3,9
	5	6,7	7,2	6,4	5,8
	10	12,7	14,4	13,4	13,5
25	1	2,7	2,3	1,9	2,9
	5	6,2	5,7	5,8	5,8
	10	11,2	11,6	11,5	11,2
50	1	1,5	1,7	1,9	1,5
	5	5,8	5,5	5,5	4,9
	10	10,3	11,4	11,5	11,0
100	1	0,9	1,2	1,2	1,2
	5	4,9	4,8	5,1	4,4
	10	9,9	9,9	10,3	8,9

Fonte: O autor

Tabela 8 – Taxas do poder empírico (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o *Bootstrap* paramétrico para situação 1.

n	Estatística	3,00	3,25	3,50	3,75	4,00	4,25	4,50	4,75	5,00
10	S_T	99,9	99,0	89,2	50,4	5,5	50,3	90,0	98,6	99,9
	S_R	100,0	99,8	95,7	60,5	4,9	60,4	96,1	99,9	99,9
	S_W	99,3	96,8	80,2	43,4	6,6	42,5	80,5	96,7	99,3
	S_{RV}	98,2	93,1	70,9	31,7	4,5	29,6	71,9	93,2	98,8
25	S_T	100,0	100,0	99,4	70,0	5,1	68,9	99,6	99,9	100,0
	S_R	100,0	100,0	100,0	83,5	4,5	85,6	99,9	100,0	100,0
	S_W	100,0	100,0	97,0	56,0	7,4	53,1	96,7	99,9	100,0
	S_{RV}	100,0	99,9	95,7	50,4	4,7	47,4	95,9	99,7	100,0
50	S_T	100,0	100,0	100,0	91,7	4,6	89,6	100,0	100,0	100,0
	S_R	100,0	100,0	100,0	98,3	4,1	97,9	100,0	100,0	100,0
	S_W	100,0	100,0	099,9	78,0	5,8	76,9	99,8	100,0	100,0
	S_{RV}	100,0	100,0	099,9	75,9	4,6	74,9	99,8	100,0	100,0

Fonte: O autor

– Situação 2

Tabela 9 – Taxas de rejeição (%) empíricas para o modelo t-Student com $\nu = 4$ graus de liberdade, considerando situação 2.

n	$\alpha\%$	S_T	S_R	S_W	S_{RV}
10	1	4,3	0,9	11,7	4,4
	5	12,1	7,3	19,0	11,5
	10	20,0	13,8	25,9	18,3
50	1	1,8	1,4	2,2	1,8
	5	6,6	6,2	7,2	6,4
	10	12,2	11,6	13,3	11,9
100	1	1,3	1,1	1,4	1,2
	5	5,7	5,7	5,9	5,6
	10	11,4	11,1	11,7	11,3
200	1	1,0	0,9	1,1	0,9
	5	5,2	5,1	5,5	5,1
	10	10,4	10,4	10,8	10,3

Fonte: O autor

Tabela 10 – Taxas de rejeição empíricas (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o *Bootstrap* paramétrico para situação 2.

n	$\alpha\%$	S_T	S_R	S_W	S_{RV}
10	1	2,3	1,2	1,7	3,2
	5	6,6	6,2	7,2	6,4
	10	12,2	12,6	13,3	11,7
25	1	1,7	1,1	2,2	1,8
	5	5,5	5,8	6,2	5,8
	10	11,4	11,6	12,6	11,9
50	1	1,3	1,1	1,4	1,2
	5	5,7	5,7	5,9	5,6
	10	11,2	11,1	11,7	11,3
100	1	1,0	0,9	1,1	0,9
	5	5,2	5,3	5,5	5,1
	10	10,2	10,4	10,8	10,2

Fonte: O autor

Tabela 11 – Taxas do poder empírico (%) para o modelo t-Student com $\nu = 4$ graus de liberdade, utilizando o *Bootstrap* paramétrico para situação 2.

n	Estatística	4,00	4,25	4,50	4,75	5,00	5,25	5,50	5,75	6,00
10	S_T	100,0	99,3	91,6	53,4	06,0	55,4	92,8	99,6	100,0
	S_R	100,0	99,9	97,0	64,7	7,5	66,1	97,6	100,0	100,0
	S_W	99,8	97,8	83,1	47,5	5,8	47,2	85,0	98,3	99,6
	S_{RV}	99,5	94,8	74,1	34,5	5,7	33,4	76,2	95,9	99,3
25	S_T	100,0	100,0	99,5	75,5	6,7	73,6	99,7	100,0	100,0
	S_R	100,0	100,0	100,0	87,8	5,7	88,4	100,0	100,0	100,0
	S_W	100,0	100,0	98,1	59,6	9,5	59,8	97,2	100,0	100,0
	S_{RV}	100,0	100,0	97,3	53,6	6,3	53,4	96,4	100,0	100,0
50	S_T	100,0	100,0	100,0	89,0	5,3	92,1	100,0	100,0	100,0
	S_R	100,0	100,0	100,0	97,0	4,8	98,4	100,0	100,0	100,0
	S_W	100,0	100,0	99,9	74,7	6,4	78,5	99,8	100,0	100,0
	S_{RV}	100,0	100,0	99,9	72,2	5,3	76,8	99,8	100,0	100,0

Fonte: O autor

A.2 Resultados das simulações considerando a distribuição Cauchy.

Tabela 12 – Taxas de rejeição empíricas (%) para o modelo Cauchy.

n	$\alpha\%$	S_T	S_R	S_W	S_{RV}
10	1	4,0	< 0,01	27,8	8,8
	5	11,7	2,0	37,2	19,8
	10	21,7	6,2	44,3	29,3
50	1	1,4	0,9	3,6	1,6
	5	6,3	5,1	10,0	7,0
	10	12,0	10,8	15,5	11,7
100	1	1,3	0,9	2,2	1,7
	5	5,8	5,5	7,1	6,2
	10	11,7	11,0	13,9	12,1
200	1	0,8	0,9	0,9	0,8
	5	5,2	4,8	6,0	5,2
	10	9,0	9,2	10,3	9,2

Fonte: O autor

Tabela 13 – Taxas de rejeição empíricas (%) para o modelo Cauchy, utilizando o *Bootstrap* paramétrico.

n	$\alpha\%$	S_T	S_R	S_W	S_{RV}
10	1	1,9	1,2	1,7	1,3
	5	5,7	4,9	6,2	5,9
	10	10,5	11,2	10,7	10,9
25	1	1,4	0,9	1,6	1,3
	5	5,3	5,1	5,7	5,7
	10	10,5	10,8	10,5	10,7
50	1	1,3	0,9	1,4	1,2
	5	5,3	5,5	5,5	5,2
	10	9,7	10,3	10,5	10,1
100	1	0,9	0,9	0,9	0,8
	5	5,1	4,8	4,9	5,2
	10	9,7	10,2	10,1	9,9

Fonte: O autor

Tabela 14 – Taxas do poder empírico (%) para o modelo Cauchy, utilizando o *Bootstrap* paramétrico.

n	Estatística	-1,00	-0,75	-0,50	-0,25	0,00	0,25	0,50	0,75	1,00
10	S_T	97,7	91,8	82,2	48,3	6,2	46,1	79,9	92,7	97,3
	S_R	99,1	98,9	91,4	61,8	6,5	58,8	90,7	97,6	99,4
	S_W	94,8	86,7	73,1	48,8	5,9	48,7	71,2	86,0	93,5
	S_{RV}	86,2	74,2	54,2	29,5	5,7	27,3	53,6	72,0	86,1
25	S_T	100,0	100,0	95,9	63,6	5,5	65,5	96,4	99,9	100,0
	S_R	100,0	100,0	99,8	85,1	4,7	86,6	99,7	100,0	100,0
	S_W	99,6	98,1	85,8	43,7	5,6	47,0	86,9	98,0	99,7
	S_{RV}	99,5	97,2	80,5	35,3	4,6	37,7	81,3	97,1	99,5
50	S_T	100,0	100,0	99,8	85,5	6,0	84,6	99,9	100,0	100,0
	S_R	100,0	100,0	100,0	98,1	4,5	98,5	99,9	100,0	100,0
	S_W	100,0	100,0	99,1	63,4	5,5	63,6	98,2	100,0	100,0
	S_{RV}	100,0	100,0	99,0	58,0	6,5	59,3	97,7	100,0	100,0

Fonte: O autor