

---

INFERÊNCIA EM MODELOS HETEROSCEDÁSTICOS  
NA PRESENÇA DE PONTOS DE ALAVANCA

TATIENE CORREIA DE SOUZA

Orientador: Prof. Dr. Francisco Cribari Neto  
Co-orientador: Prof. Dr. Klaus Leite P. Vasconcellos  
Área de concentração: Estatística Matemática

Dissertação submetida como requerimento parcial para obtenção do  
grau de Mestre em Estatística pela Universidade Federal de Pernambuco

Recife, dezembro de 2003

---

Universidade Federal de Pernambuco  
Mestrado em Estatística

04 de dezembro de 2003

(data)

Nós recomendamos que a dissertação de mestrado de autoria de

Tatiene Correia de Souza

intitulada

Inferência em modelos heteroscedásticos na presença de pontos de alavanca

seja aceita como cumprimento parcial dos requerimentos para o grau de Mestre em Estatística.



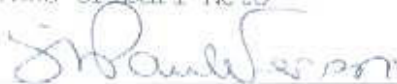
Coordenador da Pós-Graduação em Estatística

Banca Examinadora:



Francisco Cribari Neto

orientador



Silvia Lopes de Paula Ferrari (USP)



Viviana Giampaoli

Este documento será anexado à versão final da dissertação.

Aos meus pais.

## Agradecimentos

A Deus, por me cercar de pessoas maravilhosas e por me abençoar a cada dia.

Aos meus pais, que são as pessoas que mais amo, mais importantes da minha vida e aos quais eu dedico todas as minhas conquistas.

A minha querida e amada, Vitória, por existir, pelos inúmeros momentos de felicidade e por me ensinar mais uma forma de amar.

Ao meu orientador, professor e amigo, Francisco Cribari Neto, pela seriedade, dedicação, estímulo, confiança, competência e por estar sempre presente nos momentos difíceis, a minha eterna gratidão.

Aos meus irmãos que tanto amo, Hélio e Humberto.

A Antônio Carlos, pelo carinho, compreensão, estímulo e por seu amor incondicional.

Aos meus familiares, em especial, ao meu avô Severino, às minhas avós Rilsa e Josefa, aos meus tios Herton, Adriana e Elizabete pelo incentivo e confiança.

Ao excelente professor Klaus Vasconcellos, por sua amizade, carinho e pela forma competente com que forneceu ótimas sugestões para realização deste trabalho.

A Tarciana Liberal, por ser uma grande amiga e por todos os momentos inesquecíveis que passamos juntas.

A Andréa, por ser tão meiga e amiga.

A minha inesquecível amiga, Valéria, a excelente secretária do Mestrado em Estatística.

A minha querida amiga, Sílvia, pelos momentos de alegria e por sua enorme amizade.

Aos meus grandes amigos que conquistei durante o mestrado: Felipe, pelo imenso carinho; Sílvia, pelos momentos de confidências e companheirismo; Keila, por me mostrar o quanto podemos ser fortes e por sua enorme confiança; Moisés, um grande irmão por quem tenho um enorme carinho; Gilson, por sua seriedade e amizade; Raydonal, por estar sempre disposto a ajudar; João Marcelo, por seu espírito fraterno e carinho; Patrícia, por sua tranquilidade; Bartolomeu, por sua alegria e a Cristina Moraes, por seu espírito materno.

Aos meus grandes amigos, Andréa, Angela, Ernando, José Ramos, Juliana, Patrícia, Tarciana, pelo amor, carinho e apoio que sempre estão dispostos a me oferecer.

A minha querida e eterna professora Cristina, por ser uma verdadeira mãe para mim.

A Claudia, por estar sempre disposta a me ajudar, por seu enorme carinho, amizade e apoio.

Aos professores Manoel e Sylvio, pelo incentivo.

Aos professores do Programa de Mestrado em Estatística da UFPE, pela socialização de seus conhecimentos.

A Adriana, Cícero e Antônio, funcionários do Departamento de Estatística, pelo carinho e amizade.

Aos meus colegas de mestrado Lenaldo, Sandra Maria, Sandra Rêgo, Gecynalda, Cherubino e Renata, em especial a André, Junior e Tatiane.

À CAPES, pelo apoio financeiro.

## Resumo

Técnicas clássicas de regressão linear assumem que os erros, que representam a componente aleatória do modelo, têm variância constante, ou seja, assumem homoscedasticidade. Contudo, esta suposição é bastante forte e, em uma relevante parte dos problemas práticos, muito pouco razoável. A presente dissertação considera a estimação consistente da matriz de covariâncias do estimador de mínimos quadrados ordinários em um modelo de regressão linear sob heteroscedasticidade de forma desconhecida. O estimador mais usado é aquele proposto por Halbert White, conhecido como HC0. Consideramos também outros estimadores consistentes, a saber: o estimador HC3, que é uma aproximação do estimador jackknife, e o estimador HC4 proposto por Cribari-Neto (2004), que leva em consideração o efeito de pontos de alta alavancagem em amostras finitas. Dois estimadores consistentes obtidos a partir de esquemas de reamostragem de bootstrap são também considerados. Nós propomos, com base no estimador HC4, um novo estimador: HC5. Este estimador é o primeiro estimador na classe dos estimadores consistentes da matriz de covariâncias do estimador de mínimos quadrados a incorporar termos de descontos que se ajustam a variações no grau máximo de alavancagem dos dados. Nós apresentamos resultados de simulação de Monte Carlo sobre o desempenho de testes quasi- $t$  cujas estatísticas são baseadas nos diferentes estimadores consistentes. A avaliação é realizada tanto sob homoscedasticidade quanto sob heteroscedasticidade e os resultados revelam que o teste construído a partir do estimador HC5 tipicamente apresenta desempenho superior aos demais testes considerados. No que se refere a inferência via bootstrap, há muito pouco ganho em amostras finitas em se usar o esquema de reamostragem de bootstrap ponderado para realizar testes bootstrap, estimando-se valores  $p$  ou valores críticos, ao invés de se utilizar o bootstrap ponderado para estimação de erros-padrão a serem utilizados em estatísticas de teste convencionais. Nossos resultados também revelam que a presença de pontos de alta alavancagem exerce um papel importante no desempenho dos diferentes estimadores consistentes em amostras de tamanho típico. Algumas aplicações empíricas são, por fim, apresentadas.

## Abstract

The chief goal of this thesis is to study the finite-sample behavior of different heteroskedasticity-consistent covariance matrix estimators, under both constant and unequal error variances. We consider the estimator proposed by Halbert White (HC0), its variant known as HC3, and Wu's (1986) weighted bootstrap estimator. Recently proposed estimators, such as Cribari–Neto's (2004) HC4 and Cribari–Neto and Zarkos's (2004) inversely adjusted weighted bootstrap, are also considered. We propose a new covariance matrix estimator: HC5. It is the first consistent estimator to explicitly take into account the effect that the maximum level of leverage of the data has on the associated inference. Our numerical (Monte Carlo) results show that quasi- $t$  inference based on HC5 is typically more reliable than inference based on other covariance matrix estimators. We also present four applications to real data.

# Índice

1. Introdução .....	1
1.1. Considerações iniciais .....	1
1.2. Modelo e estimadores .....	2
1.3. Organização da dissertação .....	10
1.4. Plataforma computacional .....	11
2. Métodos bootstrap .....	12
2.1. Introdução .....	12
2.2. O método bootstrap .....	12
2.3. Bootstrap em modelos de regressão .....	14
3. Proposta de um novo estimador e sua avaliação numérica .....	17
3.1. Introdução .....	17
3.2. Proposta de um novo estimador .....	18
3.3. Avaliação numérica .....	20
3.4. Bootstrap na estatística de teste quasi- $t$ .....	44
4. Aplicações .....	49
4.1. Introdução .....	49
4.2. Atividade em instalações da marinha americana .....	50
4.3. Gasto com educação nos EUA - I .....	52
4.4. Gasto com educação nos EUA - II .....	58
4.5. Graus de prestígio de atividades no Canadá .....	60
5. Implementação dos estimadores consistentes .....	63
6. Conclusões e sugestões para trabalhos futuros .....	70
o Apêndice .....	72
o Referências .....	81