

**Comparação
entre
testes de heteroscedasticidade**

Bianca de Carla Hernandez Matos

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

Área de concentração: Estatística
Orientadora: Prof^a. Dr^a. Silvia L. de Paula Ferrari

– Durante o desenvolvimento deste trabalho, a autora recebeu apoio financeiro da CAPES –

São Paulo
2003

**Comparação
entre
testes de heteroscedasticidade**

Este exemplar corresponde à redação
final da dissertação devidamente corrigida
e defendida por Bianca de Carla Hernandez Matos
e aprovada pela comissão julgadora:

São Paulo, 28 de março de 2003.

Banca examinadora:

- Prof^ª Dr^ª Silvia L. de Paula Ferrari (orientadora) (IME-USP)
- Prof^ª Dr^ª Mônica Carneiro Sandoval (IME-USP)
- Prof. Dr. Filidor Vilca Labra (IMECC-UNICAMP)

Aos meus pais.

Agradecimentos

Primeiramente, ao meu pai, Hildemar, e à minha falecida mãe, Maria de Lourdes, por todo o incentivo que sempre me deram, sem o qual eu não teria conseguido concluir esta etapa tão importante.

Ao meu irmão Heitor, pelo apoio e compreensão.

À professora Silvia Lopes de Paula Ferrari, pela orientação competente, e por seu importante papel em minha formação como Estatística.

Aos professores do Departamento de Estatística do IME-USP, pelos valiosos ensinamentos recebidos ao longo dos últimos anos.

Ao professor Abraham Laredo Sicsú, pelo incentivo e confiança.

À amiga Audrey Cysneiros, por suas inúmeras contribuições a esse trabalho.

Aos amigos Fábio, Jacqueline, Kelly, Marcos e Priscila, pela relação gratificante que mantemos desde os tempos da graduação.

Ao Leonardo, por abrir mão diversas vezes de minha companhia, para que eu pudesse me dedicar a esse trabalho.

À Universidade de São Paulo e à CAPES, por me concederem a oportunidade de cursar o Mestrado, e pelos recursos disponibilizados.

Resumo

O modelo de regressão linear usual supõe que os erros, que constituem seu componente aleatório, são homoscedásticos. Em muitas aplicações práticas, no entanto, essa suposição não é satisfeita. Na presença de heteroscedasticidade, as inferências baseadas no modelo homoscedástico podem ser muito imprecisas. Diversos testes de heteroscedasticidade são propostos na literatura. Esses testes baseiam-se em aproximações assintóticas que conduzem a resultados confiáveis em grandes amostras, mas não necessariamente em amostras pequenas a moderadas. Nesse trabalho, 28 diferentes testes de heteroscedasticidade são comparados com relação a tamanho e poder, em amostras finitas. Tais comparações foram feitas com o uso de simulações de Monte Carlo.

Abstract

The usual linear regression model assumes that the errors, which are the random component, are homoskedastic. In many practical applications, however, this assumption is not satisfied. In the presence of heteroskedasticity, inferences based on the homoskedastic model can be very inaccurate. Several heteroskedasticity tests are proposed in the literature. Such tests rely on asymptotic approximations which can lead to reliable results in large samples but not necessarily in small to moderate samples. In this work, 28 different heteroskedasticity tests are compared with respect to size and power in finite samples. Such comparisons are performed using Monte Carlo simulations.

Sumário

1	Introdução	1
1.1	O modelo de regressão linear	2
1.2	Estimação pelo método de máxima verossimilhança	4
1.3	Organização da dissertação	6
2	Testes de heteroscedasticidade	7
2.1	Testes construtivos	8
2.1.1	Teste da razão de verossimilhanças	8
2.1.2	Teste da razão de verossimilhanças com correção de Bartlett	9
2.1.3	Teste da razão de verossimilhanças condicional	9
2.1.4	Teste da razão de verossimilhanças modificado	10
2.1.5	Teste da razão de verossimilhanças modificado com correção de Bartlett	11
2.1.6	Teste da razão de verossimilhanças residual	11
2.1.7	Teste score	12
2.1.8	Teste score com correções tipo-Bartlett	14
2.1.9	Teste score com correção no valor crítico	15
2.1.10	Teste score com bootstrap	17
2.1.11	Teste score modificado	17
2.1.12	Teste score studentizado	18
2.1.13	Teste score studentizado modificado	18
2.1.14	Teste score residual	18
2.2	Testes não construtivos	19
2.2.1	Teste de White	19
2.2.2	Teste de Goldfeld-Quandt	20
2.2.3	Teste de Levene	21
2.2.4	Teste de Szroeter	22
2.2.5	Teste de Ramsey	22
2.3	Escolha das covariáveis no modelo da variância	23
3	Avaliação numérica	25
3.1	Detalhes metodológicos	25
3.2	Tamanhos dos testes	27
3.3	Poder dos testes	37
3.4	Impacto da heteroscedasticidade sobre os testes t usuais envolvendo β	46

3.5	Robustez dos testes à má especificação da função cedástica	49
3.6	Conclusões	54
4	Aplicação	56
A	Gráficos de resíduos heteroscedásticos	61
B	Programa	64
	Referências bibliográficas	74

Capítulo 1

Introdução

O modelo de regressão linear é amplamente utilizado em aplicações práticas com o objetivo de relacionar uma variável resposta a uma ou mais variáveis explicativas. Esse modelo supõe que os erros, que representam seu componente aleatório, têm variância constante, propriedade conhecida como *homoscedasticidade*. Esse termo deriva de duas raízes gregas: *homós*, que significa “semelhante” ou “igual”, e *skedánummi*, que significa “espalhar” ou “dispersar”.

Em muitas aplicações práticas a suposição de homoscedasticidade pode não ser adequada. Lyon e Tsai (1996) citam um exemplo econométrico em que o objetivo é relacionar, por meio de um modelo de regressão, o retorno de um ativo financeiro ao retorno de mercado. Segundo eles, é empiricamente bem conhecido que os erros de tal modelo de regressão freqüentemente (mas não sempre) exibem heteroscedasticidade, termo utilizado para designar a situação em que a homoscedasticidade não é satisfeita. Ali e Giacotto (1984) comentam que tal problema é comum em dados de corte transversal ¹ em que haja diferença nos “tamanhos” das unidades amostrais. Esse problema foi encontrado por Prais e Houthakker (1955) e Jorgenson (1965) em estudos sobre orçamento familiar utilizando indivíduos com diversas rendas e tamanhos familiares, e por Meyer e Kuh (1957), em estudos do comportamento dos investimentos de empresas de diversos tamanhos.

Sob as hipóteses usuais do modelo de regressão, a estimação dos parâmetros é geralmente feita por meio do método de mínimos quadrados ordinários (MQO). Esse método, além de apresentar facilidade computacional, fornece estimadores com propriedades bastante desejáveis, tais como não-viciosidade, consistência e eficiência. Sob a presença de heteroscedasticidade, esses estimadores mantêm-se não-viciados, mas são ineficientes, isto é, não possuem a menor variância possível para estimadores não-viciados. O maior problema, no entanto, é o fato de que nessa situação o estimador usual da matriz de covariâncias dos estimadores MQO torna-se não consistente. Nesse caso, inferências baseadas em tal estimador podem ser muito imprecisas. Portanto, em situações práticas faz-se necessário avaliar se a suposição de homoscedasticidade é satisfeita pelos dados.

Uma maneira usual de efetuar a verificação da homoscedasticidade é a análise de gráficos de

¹Dados de corte transversal são aqueles em que um certo número de unidades amostrais são observadas em um dado ponto no tempo, em contraste com os dados de séries temporais, em que uma única unidade amostral é observada em diversos pontos no tempo.

resíduos. Os resíduos do modelo, ou alguma versão padronizada, são plotados contra variáveis preditoras ou valores ajustados. Caso esse gráfico apresente o aspecto de uma banda horizontal, não há evidências de que os erros do modelo sejam heteroscedásticos. A existência de algum padrão de alteração da variabilidade dos resíduos, como por exemplo um formato de megafone, pode indicar a existência de heteroscedasticidade. O problema desses gráficos, no entanto, é que a interpretação dos mesmos é subjetiva. Como avaliar se um aumento ou diminuição na variabilidade dos resíduos exibido pelo gráfico é forte o bastante para levar a concluir pela heteroscedasticidade? Para tanto, faz-se necessário o uso de testes de hipóteses para testar a hipótese nula de homoscedasticidade.

Na literatura são propostos diversos testes de heteroscedasticidade. Segundo Harrison e McCabe (1979), no entanto, “na literatura econométrica empírica, em contraste com a preocupação com a correlação serial, os resultados de testes de heteroscedasticidade raramente são reportados”. Comentário semelhante é encontrado em Evans e King (1985), segundo os quais “testes de rotina para a heteroscedasticidade não parecem tão prevalentes quanto para a correlação serial, apesar das conseqüências potencialmente mais sérias de ignorá-la”. Uma possível explicação apontada por Harrison e McCabe para tal omissão era o alto esforço computacional envolvido na aplicação dos testes em sua época. Hoje em dia, porém, a omissão parece se manter, apesar dos grandes avanços dos recursos computacionais. Em grande parte dos textos sobre regressão não são abordados testes de heteroscedasticidade, assim como em vários dos pacotes estatísticos mais conhecidos nenhum desses testes está implementado.

Os testes de heteroscedasticidade são divididos por alguns autores em não-constructivos (ou gerais), para alternativas indefinidas ou muito amplas, e constructivos (ou específicos), os quais requerem a especificação da forma da heteroscedasticidade. Alguns exemplos de testes do primeiro tipo são os propostos por Goldfeld e Quandt (1965) e White (1980). Do segundo tipo, alguns exemplos são o teste da razão de verossimilhanças e o teste escore de Breusch e Pagan (1979).

Para grande parte dos testes, a distribuição exata da estatística não é conhecida, sendo que somente se conhece sua distribuição assintótica. Espera-se que a distribuição assintótica produza resultados confiáveis para tamanhos amostrais grandes, o que muitas vezes não está disponível em aplicações práticas.

O objetivo desse trabalho é avaliar e comparar os desempenhos dos diversos testes propostos na literatura em amostras finitas, além de avaliar a robustez dos mesmos à violação de outras suposições do modelo de regressão linear, tais como a normalidade dos erros. Alguns autores já se dedicaram à questão de comparação de testes, como por exemplo Ali e Giaccotto (1984), Griffiths e Surekha (1986) e Lyon e Tsai (1996). No entanto, cada um desses considerou somente um pequeno subconjunto dos testes existentes, além de não incluírem os testes propostos mais recentemente, tais como os testes com correções de Bartlett e tipo-Bartlett.

1.1 O modelo de regressão linear

O modelo de regressão linear pode ser definido como

$$y = X\beta + u, \tag{1.1}$$

onde y é um vetor aleatório de dimensão n , β é um vetor de dimensão p de parâmetros desconhecidos, X é uma matriz $n \times p$ de variáveis preditoras não aleatórias de posto p e $u = (u_1, \dots, u_n)'$ é um vetor de erros aleatórios de dimensão n tal que $E(u) = 0$ e $\text{Cov}(u) = \Omega = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$, com $0 < \sigma_t^2 < \infty$ representando a variância (desconhecida) do erro u_t . Se $\sigma_t^2 = \sigma^2$, uma constante finita positiva, para $t = 1, \dots, n$, o modelo é homoscedástico. Caso contrário, o modelo é heteroscedástico.

Sob o modelo (1.1), o estimador de mínimos quadrados ordinários (MQO) de β , ou seja,

$$\hat{\beta} = (X'X)^{-1}X'y$$

é não viciado, consistente e tem distribuição assintoticamente normal, mesmo se o modelo for heteroscedástico. No entanto, sob heteroscedasticidade esse estimador não é eficiente.

A matriz de covariância de $\hat{\beta}$ deve também ser estimada para serem feitas inferências sobre β . Esta é dada por

$$\Psi = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

que se reduz a $\sigma^2(X'X)^{-1}$ caso o modelo seja homoscedástico.

Se o modelo for homoscedástico, o estimador usual de Ψ é

$$\hat{\Psi} = s^2(X'X)^{-1}, \quad (1.2)$$

onde s^2 é o quadrado médio do resíduo da regressão, dado por $s^2 = \hat{u}'\hat{u}/(n-p)$, com $\hat{u} = (I-H)y$ representando o vetor de resíduos ordinários, sendo $H = X(X'X)^{-1}X'$. Este é um estimador consistente sob a suposição de homoscedasticidade mas não o é sob heteroscedasticidade. Logo, o uso do estimador usual na presença de heteroscedasticidade pode conduzir a inferências inválidas a respeito do vetor de parâmetros β .

Alternativas ao estimador usual de Ψ quando uma inspeção nos dados indica que um modelo heteroscedástico é mais adequado que o modelo mais simples, homoscedástico, são o estimador de White (1980) e várias de suas versões modificadas (MacKinnon e White, 1985), que são consistentes embora possam ter um viés pronunciado em algumas situações (Chesher e Jewitt, 1987). Recentemente, Cribari-Neto, Ferrari e Cordeiro (2000) obtiveram uma correção para o viés do estimador de White que possibilita a definição de estimadores modificados com viés substancialmente reduzido. O trabalho de Cribari-Neto, Ferrari e Cordeiro (2000) foi estendido para estimadores alternativos ao de White por Cribari-Neto e Galvão (2002). Embora a correção de viés seja efetiva, testes de hipóteses tipo Wald sobre o vetor de parâmetros β baseados nos estimadores corrigidos são muito vulneráveis à presença de pontos de alavanca na matriz X e os testes tendem a rejeitar a hipótese nula com uma frequência bem maior do que o nível de significância nominal dos testes (Cribari-Neto, Ferrari e Cordeiro, 2000; Cribari-Neto e Galvão, 2001; ver também Long e Ervin, 2000 e Cribari-Neto, Ferrari e Oliveira, 2002).

Uma outra maneira de lidar com o problema de heteroscedasticidade é através da determinação de uma forma funcional que relacione as variâncias $\sigma_1^2, \dots, \sigma_n^2$ com alguns parâmetros desconhecidos, que não dependem do vetor de parâmetros de regressão β , e algumas variáveis auxiliares. Em outras palavras, admite-se que

$$\sigma_t^2 = \sigma^2 w_t = \sigma^2 w(z_t, \delta) > 0, \quad (1.3)$$

onde z_t é um vetor $k \times 1$ de variáveis auxiliares, σ^2 é uma constante positiva desconhecida e finita e δ é um vetor $k \times 1$ de parâmetros desconhecidos. A matriz de dimensão $n \times k$ de variáveis auxiliares,

com t -ésima linha dada por z_t' , será denotada por Z . A função $w_t = w(z_t, \delta)$ é denominada por alguns autores função cedástica ou função de pesos. Assim, tem-se que

$$\text{Cov}(u) = \sigma^2 W, \quad (1.4)$$

onde $W = \text{diag}(w_1, \dots, w_n)$. Duas formas que comumente se assumem para σ_t^2 são os chamados modelos aditivo e multiplicativo, dados respectivamente por

$$\sigma_t^2 = \sigma^2(1 + z_t'\delta)^2 \quad (1.5)$$

e

$$\sigma_t^2 = \sigma^2 \exp(z_t'\delta). \quad (1.6)$$

1.2 Estimação pelo método de máxima verossimilhança

Sob a suposição de normalidade dos erros e com variâncias dadas por (1.3), o logaritmo da função de verossimilhança é dado por

$$\begin{aligned} L = L(y; \delta, \sigma^2, \beta) = & -\frac{n}{2} \log(\sigma^2) - \frac{1}{2} \sum_{t=1}^n \log w(z_t, \delta) \\ & - \frac{1}{2\sigma^2} (y - X\beta)' W^{-1} (y - X\beta) + \text{constante}. \end{aligned} \quad (1.7)$$

Se δ for conhecido, os estimadores de máxima verossimilhança (EMVs) de β e σ^2 são dados por

$$\hat{\beta}_\delta = (X'W^{-1}X)^{-1}X'W^{-1}y, \quad (1.8)$$

e

$$\hat{\sigma}_\delta^2 = \frac{1}{n} (y - X\hat{\beta}_\delta)' W^{-1} (y - X\hat{\beta}_\delta). \quad (1.9)$$

No entanto, usualmente δ é desconhecido e os EMVs de β , σ^2 e δ não têm forma fechada e vêm como solução das equações em (1.8), (1.9) e $\partial L_p / \partial \delta = 0$, onde L_p é o logaritmo da função de verossimilhança perfilada, dado por

$$L_p = L_p(y; \delta) = L(y; \delta, \hat{\sigma}_\delta^2, \hat{\beta}_\delta) = -\frac{n}{2} \log \hat{\sigma}_\delta^2 - \frac{1}{2} \sum_{t=1}^n \log w(z_t, \delta) + \text{constante}. \quad (1.10)$$

Harvey (1976) apresentou o algoritmo scoring (Rao, 1965) para obter os estimadores de máxima verossimilhança dos parâmetros no caso do modelo multiplicativo dado em (1.6). Reescrevendo tal modelo como $\sigma_t^2 = \exp(z_t^{*'} \delta^*)$, com $z_t^{*'} = (1, z_t')$ e $\delta^{*'} = (\log \sigma^2, \delta')$, temos que

$$\begin{aligned} \hat{\beta}^{(m+1)} &= \hat{\beta}^{(m)} + \left\{ E \left[\begin{array}{c} \frac{\partial^2 L}{\partial \beta \partial \beta'} \\ \beta = \hat{\beta}^{(m)} \\ \delta^* = \hat{\delta}^{*(m)} \end{array} \right] \right\}^{-1} \left[\begin{array}{c} \frac{\partial L}{\partial \beta} \\ \beta = \hat{\beta}^{(m)} \\ \delta^* = \hat{\delta}^{*(m)} \end{array} \right] \\ &= \hat{\beta}^{(m)} + \left[\sum_{t=1}^n \exp(-z_t^{*'} \hat{\delta}^{*(m)}) x_t x_t' \right]^{-1} \sum_{t=1}^n x_t \exp(-z_t^{*'} \hat{\delta}^{*(m)}) (y_t - x_t' \hat{\beta}^{(m)}) \end{aligned}$$

e

$$\begin{aligned}\widehat{\delta}^{*(m+1)} &= \widehat{\delta}^{*(m)} + \left\{ E \left[\frac{\partial^2 L}{\partial \delta^* \partial \delta^{*'}} \right]_{\substack{\beta = \widehat{\beta}^{(m)} \\ \delta^* = \widehat{\delta}^{*(m)}}} \right\}^{-1} \left[\frac{\partial L}{\partial \delta^*} \right]_{\substack{\beta = \widehat{\beta}^{(m)} \\ \delta^* = \widehat{\delta}^{*(m)}}} \\ &= \widehat{\delta}^{*(m)} + \left(\sum_{t=1}^n z_t^* z_t^{*'} \right)^{-1} \sum_{t=1}^n z_t^* \left[\exp \left(-z_t^{*'} \widehat{\delta}^{*(m)} \right) \left(y_t - x_t' \widehat{\beta}^{(m)} \right)^2 - 1 \right],\end{aligned}$$

onde $\widehat{\beta}^{(m)}$ e $\widehat{\delta}^{*(m)}$ são as estimativas de β e δ^* , respectivamente, na m -ésima iteração e x_t ($p \times 1$) é o vetor de variáveis explicativas da t -ésima observação, correspondendo à t -ésima linha da matriz X . Como valores iniciais, pode-se tomar $\widehat{\beta}^{(0)}$ igual ao estimador de MQO de β . Além disso, pode-se tomar o primeiro elemento de $\widehat{\delta}^{*(0)}$ igual ao logaritmo do EMV de σ^2 no modelo homoscedástico, e os demais elementos desse vetor iguais a zero.

Denotemos por $\widehat{\theta} = (\widehat{\beta}_\delta, \widehat{\sigma}_\delta^2, \widehat{\delta})'$ o EMV de $\theta = (\beta, \sigma^2, \delta)'$. A distribuição assintótica de $\widehat{\theta}$ é normal multivariada com média θ e matriz de covariâncias igual ao inverso da matriz de informação de Fisher, dada por $I = -E(\partial^2 L / \partial \theta \partial \theta')$. Assim, tem-se que assintoticamente

$$\widehat{\theta} = \begin{pmatrix} \widehat{\beta}_\delta \\ \widehat{\sigma}_\delta^2 \\ \widehat{\delta} \end{pmatrix} \sim \mathcal{N}(\theta, I^{-1}),$$

em que

$$I = -E \left(\frac{\partial^2 L}{\partial \theta \partial \theta'} \right) = \begin{pmatrix} \frac{X'W^{-1}X}{\sigma^2} & 0 & 0 \\ 0 & \frac{n}{2\sigma^4} & \frac{1}{2\sigma^2} \mathbf{1}'W^{-1}D \\ 0 & \frac{1}{2\sigma^2} D'W^{-1}\mathbf{1} & \frac{1}{2} D'W^{-2}D \end{pmatrix}, \quad (1.11)$$

sendo D uma matriz $n \times k$ com a t -ésima linha dada por $\partial w(z_t, \delta) / \partial \delta'$ e 0 denotando matrizes ou vetores nulos de dimensões apropriadas. Como a matriz I é bloco diagonal, é fácil ver que assintoticamente

$$\widehat{\beta}_\delta \sim \mathcal{N}(\beta, \sigma^2 (X'W^{-1}X)^{-1}).$$

Esse estimador coincide com o estimador de mínimos quadrados generalizados de β , calculado utilizando-se \widehat{W} como estimativa da matriz de pesos, isto é, a matriz W calculada em $\widehat{\delta}$. Ao contrário do estimador de MQO, $\widehat{\beta}_\delta$ é assintoticamente eficiente para β . O EMV da matriz de covariâncias assintótica de $\widehat{\beta}_\delta$ é dado por

$$\widehat{\text{Cov}}(\widehat{\beta}_\delta) = \widehat{\sigma}_\delta^2 (X' \widehat{W}^{-1} X)^{-1}. \quad (1.12)$$

Hipóteses sobre o parâmetro β podem ser testadas com o uso de testes da razão de verossimilhanças ou de Wald, por exemplo.

Nos modelos heteroscedásticos não há uma medida de qualidade de ajuste análoga ao R^2 do modelo homoscedástico. Uma possibilidade simples seria considerar

$$R_*^2 = 1 - \frac{(y - X\hat{\beta}_\delta)'(y - X\hat{\beta}_\delta)}{\sum_{t=1}^n (y_t - \bar{y})^2}.$$

com $\bar{y} = n^{-1} \sum_{t=1}^n y_t$, isto é, a mesma expressão de R^2 no modelo homoscedástico, substituindo o estimador de MQO de β por $\hat{\beta}_\delta$. Essa medida, no entanto, não está limitada no intervalo entre 0 e 1 e, além disso, não diminui, necessariamente, ao serem excluídas variáveis do modelo, como ocorre com R^2 . Logo, deve ser vista com cautela, não podendo ser usada confiavelmente para comparar modelos.

1.3 Organização da dissertação

Essa dissertação de mestrado é constituída por quatro capítulos. No segundo capítulo, são apresentados 28 testes de heteroscedasticidade propostos na literatura. Tais testes podem ser divididos em construtivos e não-construtivos, sendo que o primeiro grupo inclui o teste da razão de verossimilhanças e o teste score, além de versões modificadas dos mesmos. No terceiro capítulo, são avaliados os comportamentos dos testes propostos com relação a tamanho e poder, sob diversas situações, por meio de simulações de Monte Carlo. Procurar-se-á identificar os testes que tenham tamanhos empíricos próximos aos nominais e, ao mesmo tempo, tenham um poder alto. No quarto capítulo, é apresentada a aplicação dos testes propostos a um conjunto de dados.

Foi desenvolvido na linguagem de programação R um programa para cálculo das estatísticas dos testes de heteroscedasticidade tratados nesse trabalho. Esse programa é apresentado no Apêndice B e está disponível no site <http://sites.uol.com.br/biancamatos>.

Capítulo 2

Testes de heteroscedasticidade

Conforme mencionado no Capítulo 1, os testes de heteroscedasticidade são divididos por alguns autores em construtivos e não-construtivos. Nas Seções 2.1 e 2.2 serão apresentados os principais testes de cada tipo tratados na literatura. A Seção 2.3 traz um procedimento para escolha das covariáveis que devem compor o vetor z_t no caso do modelo multiplicativo.

Inicialmente, introduziremos algumas notações que serão utilizadas nas próximas seções. Em geral, será empregado o sub-índice 0 para explicitar que a quantidade em questão é obtida a partir do modelo homoscedástico. Assim, $\hat{\beta}_0$ e $\hat{\sigma}_0^2$ denotarão, respectivamente, os estimadores de máxima verossimilhança de β e σ^2 no modelo homoscedástico, dadas por

$$\hat{\beta}_0 = (X'X)^{-1}X'y \quad (2.1)$$

e

$$\hat{\sigma}_0^2 = \hat{u}'_0\hat{u}_0/n, \quad (2.2)$$

em que \hat{u}_0 é um vetor $n \times 1$ com t -ésimo elemento dado por $\hat{u}_{t0} = y_t - x'_t\hat{\beta}_0$, o t -ésimo resíduo ordinário do modelo homoscedástico, sendo y_t a observação da variável resposta para a t -ésima unidade amostral e x_t ($p \times 1$) seu respectivo vetor de variáveis independentes, correspondendo à t -ésima linha da matriz X . As estimativas de máxima verossimilhança de β e σ^2 em função do verdadeiro valor do parâmetro δ , dadas em (1.8) e (1.9), serão denotadas por $\hat{\beta}_\delta$ e $\hat{\sigma}_\delta^2$, respectivamente. As estimativas desses parâmetros, quando tratadas como funções de $\hat{\delta}$, sendo $\hat{\delta}$ a estimativa de máxima verossimilhança de δ , serão denotadas por $\hat{\beta}_{\hat{\delta}}$ e $\hat{\sigma}_{\hat{\delta}}^2$. As funções de peso $w(z_t, \delta) = w_t$, quando calculadas em $\hat{\delta}$, serão denotadas por $w(z_t, \hat{\delta}) = \hat{w}_t$. A matriz chapéu, que exerce um importante papel em modelos de regressão, será denotada por

$$H = X(X'X)^{-1}X', \quad (2.3)$$

e o t -ésimo elemento de sua diagonal por h_{tt} , $t = 1, \dots, n$.

As estatísticas dos testes serão denotadas por siglas que explicitem qual o teste em questão. Por exemplo, RV denota a estatística do teste da razão de verossimilhanças e S denota a estatística do teste score. Para denotar as estatísticas das versões modificadas desses testes, serão acrescentadas letras que sugiram a modificação em questão. Assim, RVB denota a estatística do teste da razão de verossimilhanças com correção de Bartlett, por exemplo.

Por último, o vetor unitário $n \times 1$ será denotado por $\mathbf{1} = (1, \dots, 1)'$ e a matriz identidade $n \times n$, por I .

2.1 Testes construtivos

Nos testes construtivos é necessário estabelecer uma forma para a heteroscedasticidade sob a hipótese alternativa. Assumiremos que a matriz de covariâncias dos erros do modelo (1.1) é dada por (1.4). Desta forma, as hipóteses de homoscedasticidade e de heteroscedasticidade são dadas, respectivamente, por

$$H_0 : \delta = \delta_0 \quad \text{e} \quad H_1 : \delta \neq \delta_0, \quad (2.4)$$

onde δ_0 é um vetor $k \times 1$ de constantes tal que $w(z_t, \delta_0) = 1$ para $t = 1, \dots, n$.

No desenvolvimento de alguns testes apresentados a seguir, assume-se uma forma mais restritiva que (1.4) para a variância dos erros dada por

$$\sigma_t^2 = h(z_t^{*'} \delta^*) = h(\gamma + z_t' \delta), \quad (2.5)$$

onde $h(\cdot)$ é qualquer função independente de t que assuma valores positivos e possua primeira e segunda derivadas, $z_t^{*'} = (1, z_t')$, com $z_t' = (z_{t1}, \dots, z_{tk})$ e $\delta^{*'} = (\gamma, \delta')$, com γ escalar e $\delta' = (\delta_1, \dots, \delta_k)$. Nesse caso, as variâncias dos erros são função de uma função linear das covariáveis em z_t^* e as hipóteses de homoscedasticidade e de heteroscedasticidade são dadas, respectivamente, por

$$H_0 : \delta = 0 \quad \text{e} \quad H_1 : \delta \neq 0,$$

onde 0 denota um vetor nulo de dimensão k , de forma que, sob H_0 , $\sigma_t^2 = h(\gamma)$, $t = 1, \dots, n$. É importante notar que esse modelo não apresenta a constante multiplicativa σ^2 e, portanto, pode haver modelos dessa forma que não sejam caso particular do modelo (1.4). Alguns dos modelos mais utilizados na prática, no entanto, podem ser escritos de ambas as formas. É o caso do modelo aditivo (1.5), em que a variância dos erros é dada por (2.5) com $h(x) = x^2$, $z_t^{*'} = (1, z_t')$ e $\delta^{*'} = (\sigma, \sigma\delta)$ e do modelo multiplicativo (1.6), com $h(x) = \exp(x)$, $z_t^{*'} = (1, z_t')$ e $\delta^{*'} = (\log \sigma^2, \delta')$.

Os modelos considerados nesse trabalho não compreendem a situação em que a variância dos erros é uma função da média da variável resposta, isto é, $\sigma_t^2 = f(x_t' \beta) = f(\mu_t)$. Tal situação é tratada, por exemplo, em Bickel (1978) e Carroll e Ruppert (1981).

A categoria dos testes construtivos inclui o teste da razão de verossimilhanças e o teste escore. Os demais testes consistem, na maioria dos casos, em versões modificadas desses dois testes originais. Em tais versões, busca-se melhorar as propriedades dos testes originais com a utilização de funções de verossimilhança alternativas, ou correções das estatísticas dos testes ou dos valores críticos assintóticos.

2.1.1 Teste da razão de verossimilhanças

Seja $\hat{\delta}$ a estimativa de máxima verossimilhança de δ , obtida maximizando a log-verossimilhança perfilada dada em (1.10). Sob a suposição de normalidade dos erros, a estatística do teste da razão de verossimilhanças é dada por

$$RV = -2(L_p(y; \delta_0) - L_p(y; \hat{\delta})). \quad (2.6)$$

Uma dificuldade imposta por esse teste é a necessidade de efetuar a estimação tanto sob a hipótese nula quanto sob a alternativa. Rutemiller e Bowers (1968) derivaram essa estatística assumindo

que a variância dos erros é da forma (1.3) com $w_t = z_t'\delta$, e Harvey (1976) derivou-a no caso do modelo multiplicativo dado em (1.6). Nesse último caso, tem-se que

$$RV = n \log \left(\frac{\widehat{\sigma}_0^2}{\widehat{\sigma}_\delta^2} \right) - \sum_{t=1}^n z_t' \widehat{\delta}.$$

Essa estatística, assim como as suas versões modificadas apresentadas a seguir, tem distribuição assintótica χ^2 com k graus de liberdade sob H_0 , a hipótese nula de homoscedasticidade, onde k é a dimensão do vetor δ . Sabe-se que essa aproximação tem erro de ordem $O(n^{-1})$ e pode ser muito ruim caso a amostra não seja suficientemente grande.

2.1.2 Teste da razão de verossimilhanças com correção de Bartlett

A estatística RV pode ser modificada por um fator de ajuste de Bartlett, de forma a melhorar a aproximação de sua distribuição à distribuição qui-quadrado. A estatística corrigida, dada por

$$RVB = \frac{RV}{1+c},$$

onde c vem de $E(RV) = k\{(1+c) + O(n^{-3/2})\}$ e é de ordem n^{-1} , tem distribuição χ_k^2 com erro de ordem n^{-2} (Barndorff-Nielsen e Hall, 1988). A quantidade c pode depender de parâmetros desconhecidos. Nesse caso, os parâmetros devem ser substituídos por seus respectivos estimadores de máxima verossimilhança e a ordem de aproximação permanece em $O(n^{-2})$.

Sob heteroscedasticidade multiplicativa, dada por (1.6), Cordeiro (1993) obteve o fator de ajuste de Bartlett c :

$$c = \frac{1}{k} \left[h(X, R) - \frac{3(p+1)^2 - 1}{6n} \right],$$

onde

$$\begin{aligned} h(X, R) &= -\frac{1}{2} \text{tr}(A_d^2) + \frac{1}{3} \mathbf{1}' A^{(3)} \mathbf{1} + \frac{1}{2} \mathbf{1}' A_d A A_d \mathbf{1} + \text{tr}(A_d H_d) - \mathbf{1}' (H^{(2)} \odot A) \mathbf{1} \\ &+ \frac{1}{2} \mathbf{1}' H_d A H_d \mathbf{1} + \mathbf{1}' A_d A H_d \mathbf{1}, \end{aligned}$$

onde $A = R(R'R)^{-1}R'$, com $R = [\mathbf{1}, Z]$, H é a matriz chapéu dada em (2.3) e $\mathbf{1}$ é o vetor unitário de dimensão n . O índice d indica que uma matriz diagonal foi formada a partir da matriz não-indexada, substituindo-se todos os elementos fora da diagonal por zero. Além disso, o símbolo \odot denota o produto de Hadamard (direto) de matrizes de mesma dimensão, $M^{(2)} = M \odot M$ e $M^{(3)} = M^{(2)} \odot M$, onde M é uma matriz dada.

2.1.3 Teste da razão de verossimilhanças condicional

Sabe-se que o teste da razão de verossimilhanças pode ter um desempenho ruim, mesmo com um número pequeno de parâmetros de perturbação. Como as estatísticas $\widehat{\sigma}_\delta^2$ e $\widehat{\beta}_\delta$, dadas em (1.9) e (1.8), são conjuntamente suficientes e completas para os parâmetros de perturbação σ^2 e β para

cada valor fixado de δ , Honda (1989) adaptou as sugestões de Cox e Hinkley (1974) para obter o logaritmo da função de verossimilhança perfilada condicional, dado por

$$\begin{aligned} CL_p(y; \delta) &= CL_p(y|\hat{\beta}_\delta, \hat{\sigma}_\delta^2; \delta) = L_p(y; \delta) + \frac{1}{2}\{(p+2)\log \hat{\sigma}_\delta^2 - \log |X'W^{-1}X|\} \\ &= -\left(\frac{n-p-2}{2}\right)\log \hat{\sigma}_\delta^2 - \frac{1}{2}\sum_{t=1}^n \log w_t - \frac{1}{2}\log |X'W^{-1}X| + \text{constante.} \end{aligned}$$

A estatística do teste da razão de verossimilhanças condicional, proposta por Lyon e Tsai (1996), é dada por

$$RVC = -2\{CL_p(y; \delta_0) - CL_p(y; \hat{\delta})\}. \quad (2.7)$$

Sob a hipótese nula de homoscedasticidade, RVC tem distribuição χ^2 com k graus de liberdade.

2.1.4 Teste da razão de verossimilhanças modificado

Cox e Reid (1987) propuseram um ajuste para a função de verossimilhança perfilada no caso em que o parâmetro de interesse e o parâmetro de perturbação são ortogonais¹. No caso considerado aqui, o parâmetro de interesse δ é ortogonal a β , mas não a σ^2 . Simonoff e Tsai (1994) obtiveram a transformação

$$\gamma = \sigma^2 \left(\prod_{t=1}^n w_t \right)^{1/n},$$

de tal forma que γ é ortogonal a δ e β . Então, eles adaptaram a equação (9) de Cox e Reid (1987) e obtiveram uma função de verossimilhança modificada. O logaritmo dessa função é dado por

$$\begin{aligned} ML_p(y; \delta) &= ML_p(y|\hat{\beta}_\delta, \hat{\gamma}_\delta; \delta) \\ &= -\left(\frac{n-p-2}{2}\right)\log \hat{\gamma}_\delta - \frac{1}{2}\log |X'W^{-1}X| - \frac{p}{2n}\sum_{t=1}^n \log w_t + \text{constante} \\ &= CL_p(y; \delta) + \frac{1}{n}\sum_{t=1}^n \log w_t + \text{constante,} \end{aligned} \quad (2.8)$$

onde $\hat{\gamma}_\delta$ é o estimador de máxima verossimilhança de γ para o valor dado de δ .

A estatística do teste da razão de verossimilhanças modificado é dada por

$$RVM = -2\{ML_p(y; \delta_0) - ML_p(y; \hat{\delta})\}. \quad (2.9)$$

¹Seja θ um vetor de parâmetros de dimensão g , particionado em $(\theta'_1, \theta'_2, \dots, \theta'_l)'$, em que θ_i tem dimensão g_i , com $\sum_{i=1}^l g_i = g$. Diz-se que os subvetores de parâmetros θ_i e θ_j , $i \neq j$, são ortogonais se $I_{ij} = -E(\partial L^2 / \partial \theta_i \partial \theta'_j) = 0$, sendo L o logaritmo da função de verossimilhança. Nesse caso, os estimadores de máxima verossimilhança de θ_i e θ_j são assintoticamente independentes.

É fácil ver que

$$RVM = RVC + \frac{2}{n} \sum_{t=1}^n \log \hat{w}_t,$$

onde \hat{w}_t é w_t calculado em $\hat{\delta}$.

Apesar de essa estatística ter distribuição assintótica χ_k^2 com erro de ordem $O(n^{-1})$, assim como a estatística do teste da razão de verossimilhanças original, espera-se que este teste produza inferências mais precisas que o original, já que o mesmo reduz os efeitos de parâmetros de perturbação em tais inferências.

Uma forma alternativa ao teste proposto por Simonoff e Tsai (1994) é obtida substituindo-se $\hat{\delta}$ em (2.9) por $\bar{\delta}$, com $\bar{\delta}$ denotando o valor que maximiza o logaritmo da função de verossimilhança modificada (2.8). Essa estatística será denotada por RVM_1 .

2.1.5 Teste da razão de verossimilhanças modificado com correção de Bartlett

DiCiccio e Stern (1994) mostraram que é possível aplicar uma correção de Bartlett para o teste da razão de verossimilhanças modificado com estatística RVM_1 , conforme dado acima, isto é, encontrar um fator de correção c usado para definir a estatística corrigida

$$RVMB_1 = \frac{RVM_1}{1+c}, \quad (2.10)$$

a qual tem distribuição χ^2 com k graus de liberdade com erro de ordem $O(n^{-2})$.

Ferrari, Cysneiros e Cribari-Neto (2002) obtiveram c no caso do modelo multiplicativo (1.6). Esse fator é dado por

$$c = \frac{1}{k} \left[-\frac{1}{2} \text{tr}(J_d J_d) + \frac{1}{2n} k^2 + \frac{1}{2} \mathbf{1}' J_d J J_d \mathbf{1} + \frac{1}{3} \mathbf{1}' J^{(3)} \mathbf{1} - \frac{2}{n} k + \frac{1}{n} \mathbf{1}' J^{(2)} \mathbf{1} \right],$$

onde $J = V(V'V)^{-1}V'$, com $V = (z_1 - \bar{z}, \dots, z_n - \bar{z})'$, em que $\bar{z} = n^{-1} \sum_{t=1}^n z_t$ e $\mathbf{1}$ é o vetor unitário. As notações referentes às operações matriciais são as mesmas empregadas no caso do teste da razão de verossimilhanças com correção de Bartlett.

2.1.6 Teste da razão de verossimilhanças residual

Verbyla (1993) usou a função de verossimilhança residual para obter o teste escore residual. Lyon e Tsai (1996) utilizaram essa função de verossimilhança e propuseram o teste da razão de verossimilhanças residual. Se tanto σ^2 quanto δ forem tratados como parâmetros de interesse, tem-se que a função de verossimilhança residual é a função de verossimilhança perfilada condicional dada a estimativa de máxima verossimilhança de β . A função de log-verossimilhança resultante é

$$\begin{aligned} RL_p(y; \sigma^2, \delta) &= RL_p(y | \hat{\beta}_\delta) \\ &= CL_p(y; \delta) - \log \hat{\sigma}_\delta^2 \\ &= ML_p(y; \delta) - \log \hat{\gamma}_\delta. \end{aligned} \quad (2.11)$$

A estatística do teste da razão de verossimilhanças residual é dada por

$$RVR = -2\{RL_p(y; \hat{\sigma}_0^2, \delta_0) - RL_p(y; \hat{\sigma}_\delta^2, \hat{\delta})\},$$

onde $\hat{\sigma}_0^2$, dado em (2.2), é a estimativa de máxima verossimilhança de σ^2 no modelo homoscedástico e $\hat{\sigma}_\delta^2$ é a estimativa de máxima verossimilhança desse parâmetro quando $\delta = \hat{\delta}$, dada por (1.9), com a matriz W avaliada em $\hat{\delta}$.

Usando a equação (2.11), temos que

$$RVR = RVC + 2(\log \hat{\sigma}_0^2 - \log \hat{\sigma}_\delta^2).$$

Além disso, é fácil mostrar que

$$RVR = RVM + 2(\log \hat{\gamma}_0 - \log \hat{\gamma}_\delta),$$

em que $\hat{\gamma}_0$ e $\hat{\gamma}_\delta$ são as estimativas de máxima verossimilhança de γ quando $\delta = \delta_0$ e $\delta = \hat{\delta}$, respectivamente. Sob a hipótese nula de homoscedasticidade, RVR tem distribuição assintótica χ^2 com k graus de liberdade.

2.1.7 Teste escore

Seja $L(\theta)$ o logaritmo de uma função de verossimilhança dependendo de um vetor de parâmetros $\theta = (\theta_1', \theta_2')'$ de dimensão g , sendo $\theta_1 = (\theta_1, \dots, \theta_{g_1})'$ o vetor de parâmetros de interesse e $\theta_2 = (\theta_{g_1+1}, \dots, \theta_g)'$ o vetor de dimensão $g_2 = g - g_1$ de parâmetros de perturbação. Sejam ainda $U = (U_1', U_2')'$ o vetor escore, com $U_i = \partial L(\theta) / \partial \theta_i$ ($i = 1, 2$) e

$$I = \begin{pmatrix} I_{11} & I_{12} \\ I_{12}' & I_{22} \end{pmatrix}$$

a matriz de informação de Fisher, com $I_{ij} = -E(\partial^2 L(\theta) / \partial \theta_i \partial \theta_j)$, ($i, j = 1, 2$) e inversa dada por

$$I^{-1} = \begin{pmatrix} I_{11}^* & I_{12}^* \\ I_{12}^{*'} & I_{22}^* \end{pmatrix}.$$

A estatística escore, também conhecida como estatística de multiplicadores de Lagrange, proposta por Silvey (1959) e Rao (1947) para testar a hipótese $H_0 : \theta_1 = \theta_1^0$ contra a alternativa $H_1 : \theta_1 \neq \theta_1^0$, onde θ_1^0 é um vetor fixado de dimensão g_1 , é dada por

$$S = \hat{U}_1' \hat{I}_{11}^* \hat{U}_1,$$

onde os circunflexos indicam que o vetor escore e a matriz de informação de Fisher são avaliados em $\theta = (\theta_1^{0'}, \hat{\theta}_2^{0'})'$, onde $\hat{\theta}_2^0$ é a estimativa de máxima verossimilhança de θ_2 sob H_0 . Essa estatística tem distribuição nula assintótica χ^2 com g_1 graus de liberdade.

Esse teste é particularmente útil em situações em que a estimação sob a hipótese alternativa é computacionalmente trabalhosa, uma vez que só requer a estimação sob a hipótese nula. Assim, em nosso caso a aplicação desse teste requer somente a estimação dos parâmetros no modelo homoscedástico.

Godfrey (1978), Breusch e Pagan (1979) e Cook e Weisberg (1983) obtiveram, independentemente, o teste escore para testar a hipótese nula de homoscedasticidade. A diferença entre os três trabalhos consistiu na escolha do modelo para as variâncias dos erros. Godfrey (1978) admitiu o modelo multiplicativo dado em (1.6), enquanto Breusch e Pagan (1979) assumiram a forma dada em (2.5). Cook e Weisberg (1983), por sua vez, admitiram o modelo mais geral, dado em (1.3).

Considerando o modelo mais geral (1.3), o logaritmo da função de verossimilhança é apresentado em (1.7) e as hipóteses de homoscedasticidade e de heteroscedasticidade são dadas em (2.4). As estimativas de máxima verossimilhança de β e σ^2 restritas sob H_0 correspondem às estimativas de máxima verossimilhança desses parâmetros no modelo homoscedástico, sendo dadas por (2.1) e (2.2), respectivamente. Nesse caso, o vetor escore referente a δ avaliado nas estimativas de máxima verossimilhança sob H_0 é dado por

$$\widehat{U}_\delta = \frac{1}{2}D'(f - 1),$$

onde D é uma matriz $n \times k$ com t -ésima linha dada por $\partial w(z_t, \delta)/\partial \delta'$, onde os elementos são avaliados em δ_0 , e f é um vetor $n \times 1$ com t -ésimo elemento dado por $\widehat{\sigma}_0^{-2}\widehat{u}_{t0}^2$.

A matriz de informação de Fisher é dada em (1.11) e pode ser particionada em

$$I = \begin{pmatrix} I_{\beta\beta} & 0 & 0 \\ 0 & I_{\sigma^2\sigma^2} & I_{\sigma^2\delta} \\ 0 & I'_{\sigma^2\delta} & I_{\delta\delta} \end{pmatrix},$$

onde $I_{\beta\beta}$, $I_{\sigma^2\sigma^2}$ e $I_{\delta\delta}$ referem-se, respectivamente, às segundas derivadas em relação aos elementos de β , σ^2 e δ e $I_{\sigma^2\delta}$ refere-se às derivadas mistas em relação a σ^2 e aos elementos de δ . Essas submatrizes são dadas por

$$I_{\beta\beta} = \frac{1}{\sigma^2}X'W^{-1}X, \quad I_{\sigma^2\sigma^2} = \frac{n}{2\sigma^4}, \quad I_{\delta\delta} = \frac{1}{2}D'W^{-2}D, \quad I_{\sigma^2\delta} = \frac{1}{2\sigma^2}1'W^{-1}D.$$

A submatriz da inversa da matriz de informação de Fisher correspondente ao vetor de parâmetros δ avaliada nas estimativas dos parâmetros restritas a H_0 é dada por

$$\widehat{I}_{\delta\delta}^* = 2(\bar{D}'\bar{D})^{-1},$$

onde $\bar{D} = D - 11'D/n$ é uma matriz $n \times k$ obtida da matriz D , subtraindo de cada elemento a média de sua coluna. Assim, é fácil mostrar que a estatística escore é dada por

$$S = \frac{1}{2}f'\bar{D}(\bar{D}'\bar{D})^{-1}\bar{D}'f = \frac{v'\bar{D}(\bar{D}'\bar{D})^{-1}\bar{D}'v}{2\widehat{\sigma}_0^4},$$

onde v é o vetor $n \times 1$ com t -ésimo elemento dado por $v_t = \widehat{u}_{t0}^2$. Essa estatística corresponde à metade da soma de quadrados da regressão de f sobre D , considerando um modelo com intercepto, da forma $f = \lambda_0\mathbf{1} + \lambda_1 D + \varepsilon$. Sob H_0 , S tem distribuição χ^2 com k graus de liberdade, onde k é a dimensão do vetor δ , com erro de ordem $O(n^{-1})$, de forma que $P(S \leq q) = P(\chi_k^2 \leq q) + O(n^{-1})$.

No caso particular em que as variâncias são da forma considerada por Breusch e Pagan (1979), isto é, $\sigma_t^2 = h(z_t'\delta)$, com $z_{t1} = 1$, $t = 1, \dots, n$, é possível mostrar que a estatística do teste escore simplifica-se em

$$S = \frac{1}{2}(f - \mathbf{1})'Z(Z'Z)^{-1}Z'(f - \mathbf{1}) = \frac{e'Z(Z'Z)^{-1}Z'e}{2\widehat{\sigma}_0^4}, \quad (2.12)$$

onde $Z = (z_1, \dots, z_n)'$ e e é um vetor $n \times 1$ com t -ésimo elemento dado por $\widehat{u}_{t0}^2 - \widehat{\sigma}_0^2$. Nota-se que nesse caso a estatística não depende da particular escolha da função $h(\cdot)$, e corresponde à metade da soma de quadrados da regressão de f sobre Z . Essa é uma característica bastante desejável, uma vez que, em geral, é difícil determinar em situações práticas a função cedástica $h(\cdot)$.

Embora este teste tenha um bom desempenho em amostras grandes, os resultados de simulação em Breusch e Pagan (1979), Godfrey (1978), Griffiths e Surekha (1986) e Honda (1988) mostram que o mesmo rejeita a hipótese nula de homoscedasticidade com uma frequência menor do que o esperado com base em seu tamanho nominal em amostras pequenas.

2.1.8 Teste escore com correções tipo-Bartlett

Cribari-Neto e Ferrari (1995) utilizaram resultados de Cordeiro e Ferrari (1991) para obter uma correção tipo-Bartlett para a estatística S . Na construção desse teste, assume-se que as variâncias dos erros são da forma (2.5). A estatística corrigida é dada por

$$SB = S \{1 - (c_1 + c_2S + c_3S^2)\},$$

onde

$$c_1 = \frac{A_1 - A_2 + A_3}{12k}, \quad c_2 = \frac{A_2 - 2A_3}{12k(k+2)}, \quad c_3 = \frac{A_3}{12k(k+2)(k+4)}.$$

Os coeficientes A 's são dados em Honda (1988), e podem ser escritos matricialmente como

$$A_1 = \frac{1}{n} \left\{ 24k(p-1) - 24n \text{tr}(J_d H_d) + 6n \mathbf{1}' H_d J H_d \mathbf{1} + 12n \mathbf{1}' (J \odot H^{(2)}) \mathbf{1} \right\}, \quad (2.13)$$

$$A_2 = \frac{1}{n} \left\{ -24k(k+2) + 36n \text{tr}(J_d^{(2)}) - 24n \mathbf{1}' J_d J H_d \mathbf{1} \right\} \quad (2.14)$$

e

$$A_3 = \frac{1}{n} \left\{ 24n \mathbf{1}' J_d J J_d \mathbf{1} + 16n \mathbf{1}' J^{(3)} \mathbf{1} \right\}, \quad (2.15)$$

onde H é a matriz chapéu dada em (2.3), $\mathbf{1}$ é o vetor unitário de dimensão n e J é a mesma do caso do teste da razão de verossimilhanças modificado com correção de Bartlett. As notações referentes às operações matriciais coincidem com as utilizadas no teste da razão de verossimilhanças com correção de Bartlett.

A estatística proposta tem uma distribuição nula χ^2 com k graus de liberdade, com erro de ordem $O(n^{-3/2})$. Assim, a estatística com correção de Bartlett converge para uma distribuição qui-quadrado sob a hipótese nula mais rapidamente que a estatística escore original, e portanto, deve resultar em tamanhos empíricos mais próximos aos nominais em amostras finitas.

Cordeiro e Ferrari (1994) mostraram que o coeficiente A_3 em (2.15) é não-negativo. Então, se $A_3 > 0$ ou se $A_3 = 0$ e $A_2 > 0$, valores altos de S podem, em amostras finitas, conduzir a valores

negativos de SB . Uma estatística alternativa que procura contornar este problema foi sugerida por Cribari-Neto e Ferrari (1995) e é dada por

$$SB_1 = \begin{cases} SB, & \text{se } S < \bar{t}_s \\ S, & \text{caso contrário,} \end{cases}$$

onde \bar{t}_s é uma constante positiva convenientemente escolhida. A justificativa para essa proposta é o fato de que, quando S for grande, maior que \bar{t}_s digamos, há fortes evidências contra a hipótese nula, e portanto nenhuma correção é necessária. O valor \bar{t}_s pode ser escolhido como sendo, por exemplo, o quantil de 99,5% da distribuição qui-quadrado com k graus de liberdade.

Cribari-Neto e Ferrari (1995) mostraram que no modelo de regressão com uma única covariável (x_t) e $\sigma_t^2 = h(\alpha_1 + \alpha_2 x_t)$, o teste corrigido é afetado pelas medidas de assimetria e curtose amostrais da covariável. Nesse caso, os A 's dados em (2.13)–(2.15) simplificam-se em

$$A_1 = -6(4\gamma_{2x} - 3\gamma_{1x}^2 - 4), \quad A_2 = 12(3\gamma_{2x} - 2\gamma_{1x}^2 - 6) \quad \text{e} \quad A_3 = 40\gamma_{1x}^2,$$

onde $\gamma_{1x} = \bar{s}_{3x}\bar{s}_{2x}^{-3/2}$ e $\gamma_{2x} = \bar{s}_{4x}\bar{s}_{2x}^{-2}$, com $\bar{s}_{jx} = n^{-1} \sum_{t=1}^n (x_t - \bar{x})^j$, são as medidas usuais de assimetria e curtose amostrais da covariável x .

Recentemente, foram desenvolvidas formas alternativas para a correção tipo-Bartlett, equivalentes à de Cordeiro e Ferrari (1991), e que são transformações monótonas da estatística escore original. Seguindo Cordeiro e Ferrari (1991), Kakizawa (1996) sugeriu uma estatística escore modificada alternativa dada por

$$SB_2 = SB + \frac{1}{4} \left\{ c_1^2 S + 2c_1 c_2 S^2 + \left(2c_1 c_3 + \frac{4}{3} c_2^2 \right) S^3 + 3c_2 c_3 S^4 + \frac{9}{5} c_3^2 S^5 \right\}.$$

Cordeiro, Ferrari e Cysneiros (1998), por outro lado, apresentaram uma forma alternativa dada por

$$SB_3 = \begin{cases} SB, & \text{se } c_2 = c_3 = 0 \\ \sqrt{\frac{\pi}{3c_3}} \exp\left(\frac{c_2^2}{3c_3} - c_1\right) \left\{ \Phi\left(\sqrt{6c_3}S + \sqrt{\frac{2}{3c_3}}c_2\right) - \Phi\left(\sqrt{\frac{2}{3c_3}}c_2\right) \right\}, & \text{se } c_3 > 0 \\ \frac{1}{2c_2} \exp(-c_1) \{1 - \exp(-2c_2 S)\}, & \text{se } c_3 = 0 \text{ e } c_2 \neq 0, \end{cases} \quad (2.16)$$

em que $\Phi(\cdot)$ é a função de distribuição de uma variável aleatória com distribuição normal padrão. Ambas as estatísticas têm, sob H_0 , distribuição χ^2 com k graus de liberdade com erro de ordem $O(n^{-3/2})$.

2.1.9 Teste escore com correção no valor crítico

Harris (1985) propôs um procedimento geral de correção de tamanho do teste escore, em que, ao invés de se corrigir a estatística, corrige-se o valor crítico do teste. Honda (1988) utilizou esses resultados para obter uma correção para o valor crítico do teste escore de heteroscedasticidade. No desenvolvimento desse teste assume-se que a variância dos erros é da forma (2.5).

O valor crítico corrigido é dado por

$$q^* = q \{1 + (c_1 + c_2q + c_3q^2)\},$$

onde q é tal que $P(\chi_k^2 \geq q) = \alpha$, em que α denota o nível de significância nominal do teste e χ_k^2 é uma variável aleatória com distribuição qui-quadrado com k graus de liberdade. Os coeficientes c_1 , c_2 e c_3 são os mesmos do teste escore com correção de Bartlett.

Nas simulações de Honda (1988), o teste corrigido apresentou tamanhos mais próximos dos valores nominais que o teste original e, além disso, em geral apresentou um maior poder.

Cribari-Neto e Ferrari (2001) desenvolveram formas alternativas, equivalentes ao valor crítico ajustado q^* , que são transformações monótonas dos valores críticos não modificados. O primeiro valor crítico modificado alternativo segue argumentos utilizados em Kakizawa (1996) e é dado por

$$q_1^* = q^* + \frac{1}{4} \left\{ c_1^2 q + 2c_1 c_2 q^2 + \left(2c_1 c_3 + \frac{4}{3} c_2^2 \right) q^3 + 3c_2 c_3 q^4 + \frac{9}{5} c_3^2 q^5 \right\}.$$

O segundo valor crítico modificado alternativo é baseado em alguns argumentos utilizados por Cordeiro, Ferrari e Cysneiros (1998) e é dado por

$$q_2^* = \begin{cases} q^*, & \text{se } c_2 = c_3 = 0 \\ \frac{1}{2} \sqrt{\frac{\pi}{-3c_3}} \exp \left\{ \frac{c_2^2 - 3c_1 c_3}{-3c_3} \right\} \{ \operatorname{erf}(\sqrt{-3c_3}q - K) - \operatorname{erf}(-K) \}, & \text{se } c_3 \neq 0 \\ \frac{\exp(c_1) \{ \exp(2c_2 q) - 1 \}}{2c_2}, & \text{se } c_3 = 0 \text{ e } c_2 \neq 0, \end{cases}$$

onde $K = c_2/\sqrt{-3c_3}$, $\operatorname{erf}(\cdot)$ denota a chamada função erro, dada por $\operatorname{erf}(z) = 2\pi^{-\frac{1}{2}} \int_0^z e^{-t^2} dt$, $z \in \mathcal{C}$, onde \mathcal{C} é o conjunto dos números complexos.

O terceiro valor crítico modificado alternativo é baseado na idéia utilizada por Fujisawa (1997) e pode ser escrito como

$$q_3^* = \begin{cases} q^*, & \text{se } c_1 \geq 0 \text{ e } c_2 \geq 0 \\ \frac{1}{1-c_1} \{ q + c_2 q^2 + c_3 q^3 \}, & \text{se } c_1 < 0 \text{ e } c_2 \geq 0 \\ \frac{1}{2c_2} \left\{ 1 - \sqrt{1 - 4c_2(q + c_1 q + c_3 q^3)} \right\}, & \text{se } c_1 \geq 0 \text{ e } c_2 < 0 \\ \frac{1}{2c_2} \left\{ 1 - c_1 - \sqrt{(1-c_1)^2 - 4c_2(q + c_3 q^3)} \right\}, & \text{se } c_1 < 0 \text{ e } c_2 < 0. \end{cases}$$

Cysneiros, Cribari-Neto e Ferrari (2002) propuseram um outro valor crítico modificado obtido da transformação inversa de SB_3 dada em (2.16). Este é dado por

$$q_4^* = \begin{cases} q^*, & \text{se } c_2 = c_3 = 0 \\ \frac{1}{\sqrt{6c_3}} \left\{ \Phi^{-1} \left[q \left(\sqrt{\frac{\pi}{3c_3}} \exp \left(\frac{c_2^2}{3c_3} - c_1 \right) \right)^{-1} + \Phi(\sqrt{2}L) \right] - \sqrt{2}L \right\}, & \text{se } c_3 \neq 0 \\ \frac{-1}{2c_2} \log \left\{ q \left[\exp \left(-\frac{c_1}{2c_2} \right) + 1 \right]^{-1} \right\}, & \text{se } c_3 = 0 \text{ e } c_2 \neq 0, \end{cases}$$

onde $L = c_2/\sqrt{3c_3}$.

2.1.10 Teste escore com bootstrap

Além das correções analíticas para os valores críticos do teste escore mencionadas anteriormente, pode-se estimar esses valores por meio do método de bootstrap (Efron, 1979). Como em geral a distribuição exata dos erros do modelo não é conhecida, considerar-se-á o método de bootstrap não-paramétrico. Seja \hat{u} ($n \times 1$) o vetor observado de resíduos de MQO. O método é dado pelos seguintes passos:

1. Retirar uma amostra aleatória $u^* = (u_1^*, \dots, u_n^*)'$ com reposição de \hat{u} .
2. Formar uma amostra de bootstrap $y^* = X\hat{\beta}_0 + u^*$, em que X é a matriz de covariáveis original e $\hat{\beta}_0$ é a estimativa de MQO de β na amostra original.
3. Calcular o valor da estatística do teste escore para a nova amostra, tendo como vetor de valores da variável dependente y^* e matrizes de covariáveis X , para o modelo da média e Z , para o modelo da variância.
4. Repetir os passos 1 a 3 um número B de vezes, sendo B um número alto.
5. Tomar o quantil de ordem α da distribuição dos B valores gerados para a estatística como o valor crítico do teste escore, sendo α o nível de significância nominal do teste.

Esse teste foi considerado por Godfrey e Orme (1999) e Cribari-Neto e Zarkos (1999). Os primeiros tomaram $B = 400$, enquanto que os últimos tomaram $B = 500$. Vale observar que o método de bootstrap pode ser empregado para estimar os valores críticos de qualquer um dos testes aqui tratados. No entanto, devido ao grande esforço computacional requerido para efetuar as simulações de Monte Carlo envolvendo tal teste, restringir-se-á ao teste escore. Algumas simulações feitas envolvendo correções de bootstrap para os demais testes considerados nesse trabalho sugeriram que, com a correção, todos apresentam desempenhos razoavelmente próximos em amostras finitas.

2.1.11 Teste escore modificado

Simonoff e Tsai (1994) utilizaram o fato de que a estatística do teste escore é uma aproximação de primeira ordem para a estatística da razão de verossimilhanças avaliada sob a hipótese nula para propor uma modificação para o teste escore. A estatística do teste da razão de verossimilhanças modificado consiste na estatística RV original, mais um termo de ajuste. A modificação do teste escore consiste em somar a S uma aproximação de primeira ordem para tal termo de ajuste.

A estatística resultante é dada por

$$SM = S + \sum_{j=1}^k \tau_j \sum_{t=1}^n h_{tt} a_{tj}, \quad (2.17)$$

onde h_{tt} é o t -ésimo elemento da diagonal da matriz H dada em (2.3),

$$a_{tj} = \frac{\partial w_t}{\partial \delta_j} - \frac{1}{n} \sum_{l=1}^n \frac{\partial w_l}{\partial \delta_j}$$

avaliado em $\delta = \delta_0$ e τ_j é o j -ésimo elemento do vetor $\tau = (\bar{D}'\bar{D})^{-1}\bar{D}'f$ de dimensão $k \times 1$, em que \bar{D} e f são os mesmos definidos no caso do teste escore. No caso do modelo multiplicativo dado em (1.6) com $k = 1$, o termo de ajuste somado a S em (2.17) é dado pela forma simples

$$\frac{[\sum_{t=1}^n h_{tt}(z_t - \bar{z})] [\sum_{t=1}^n \hat{u}_{t0}^2(z_t - \bar{z})]}{\sum_{t=1}^n \hat{\sigma}_0^2(z_t - \bar{z})^2}.$$

Sob a hipótese nula de homoscedasticidade, SM tem distribuição assintótica χ^2 com k graus de liberdade.

2.1.12 Teste escore studentizado

A estatística S tem o problema de depender fortemente da normalidade dos erros do modelo de regressão. Koenker (1981) propôs “studentizar” a estatística S no caso em que as variâncias dos erros são da forma (2.5), admitida por Breusch e Pagan (1979), substituindo $2\hat{\sigma}_0^4$ no denominador da expressão (2.12) por

$$\hat{\psi} = \frac{1}{n} e'e = \frac{1}{n} \sum_{t=1}^n (\hat{u}_{t0}^2 - \hat{\sigma}_0^2)^2.$$

A estatística do teste é dada por

$$SS = \frac{2\hat{\sigma}_0^4 S}{\hat{\psi}}.$$

Koenker (1981) mostrou que se os erros forem independentes e identicamente distribuídos, com uma distribuição que tenha o quarto momento finito, e satisfizerem algumas condições de regularidade, SS tem distribuição nula assintótica χ^2 com k graus de liberdade.

2.1.13 Teste escore studentizado modificado

Simonoff e Tsai (1994) propuseram modificar o teste escore studentizado (Koenker, 1981) adicionando a SS o mesmo termo de ajuste considerado em (2.17). A estatística resultante é dada por

$$SSM = SS + \sum_{j=1}^k \tau_j \sum_{t=1}^n h_{tt} a_{tj},$$

onde os valores envolvidos no termo de ajuste são os mesmos do caso do teste escore modificado. Sob a hipótese nula de homoscedasticidade, SSM tem distribuição assintótica χ^2 com k graus de liberdade.

2.1.14 Teste escore residual

Sob o logaritmo da função de verossimilhança residual (2.11), Verbyla (1993) obteve o teste escore residual para o modelo multiplicativo dado em (1.6). Lyon e Tsai (1996) adotaram essa

abordagem e obtiveram o teste escore residual para a função de pesos geral, cuja estatística é dada por

$$SR = \frac{1}{2} A' D \left[D' \left\{ M - \frac{(1-h)(1-h)'}{(n-p)} \right\} D \right]^{-1} D' A,$$

onde $A = v/s^2 - (1-h)$, sendo que v é o vetor de dimensão n de resíduos de MQO ao quadrado, M é uma matriz $n \times n$ com termos da diagonal iguais a $(1-h_{ii})^2$ e termos de fora da diagonal iguais a h_{ij}^2 , h_{ij} é o elemento (i, j) da matriz chapéu H , $h = (h_{11}, \dots, h_{nn})'$, $s^2 = n\hat{\sigma}_0^2/(n-p)$ e D é a mesma definida no caso do teste escore. Essa estatística leva em consideração pontos de alavanca, ao contrário de S . Verbyla (1993) afirma que possivelmente o teste baseado em SR , assim como aquele baseado em S , é conservativo. Sob a hipótese nula de homoscedasticidade, SR tem distribuição assintótica χ^2 com k graus de liberdade.

2.2 Testes não construtivos

Nos testes não construtivos, ao contrário do que ocorre nos testes construtivos vistos anteriormente, não é necessário conhecer a forma exata das variâncias σ_t^2 sob a hipótese alternativa. Nesses testes são feitas somente algumas suposições amplas a respeito da estrutura das variâncias, como por exemplo, a suposição de que, sob a hipótese alternativa, as observações podem ser ordenadas conforme suas variâncias, isto é, de forma que $\sigma_1^2 \leq \sigma_2^2 \leq \dots \leq \sigma_n^2$, ocorrendo ao menos uma desigualdade.

Desta forma, estes testes levam a concluir simplesmente pela presença ou ausência de heteroscedasticidade nos erros. No caso de presença, não trazem informações a respeito da forma que as variâncias devem assumir, o que seria necessário para efetuar a estimação eficiente dos parâmetros de regressão por meio do método de máxima verossimilhança.

2.2.1 Teste de White

White (1980) apresenta um estimador para a matriz de covariâncias dos parâmetros de regressão, consistente tanto sob homoscedasticidade quanto sob heteroscedasticidade dos erros. Esse estimador não depende da especificação de um modelo para a variância, nem faz suposições sobre a distribuição dos erros.

Nesse artigo são apresentados dois testes de heteroscedasticidade, os quais consistem em uma comparação dos elementos do novo estimador e do estimador usual da matriz de covariâncias. Na presença de homoscedasticidade, os dois estimadores serão aproximadamente iguais, mas caso contrário, diferirão.

O primeiro teste, apresentado no Teorema 2 de White (1980), não faz suposições sobre a distribuição dos erros nem sobre a forma da heteroscedasticidade, supondo somente que algumas condições de regularidade estejam satisfeitas. A estatística desse teste é dada por

$$W_1 = v' \tilde{D} (\tilde{D}' S \tilde{D})^{-1} \tilde{D}' v, \quad (2.18)$$

onde $S = \text{diag}(s_1, \dots, s_n)$, com $s_t = e_t^2 = (\hat{u}_{t0}^2 - \hat{\sigma}_0^2)^2$, $t = 1, \dots, n$, v é o vetor de resíduos de MQO ao quadrado e $\tilde{D} = (I - 11'/n)D^*$, em que D^* é uma matriz $n \times p(p+1)/2$ contendo as funções polinomiais de segunda ordem das p variáveis independentes, isto é

$$D^* = \begin{bmatrix} x_{11}^2 & x_{11}x_{12} & \dots & x_{1,p-1}x_{1p} & x_{1p}^2 \\ x_{21}^2 & x_{21}x_{22} & \dots & x_{2,p-1}x_{2p} & x_{2p}^2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ x_{n1}^2 & x_{n1}x_{n2} & \dots & x_{n,p-1}x_{np} & x_{np}^2 \end{bmatrix}. \quad (2.19)$$

O teste do Corolário 1 de White (1980) também não faz suposições a respeito da forma da heteroscedasticidade, mas introduz uma suposição a respeito da distribuição dos erros. Supõe-se que $E(u_t^4) = \mu_4$, para todo $t = 1, \dots, n$, ou seja, que os erros têm todos a mesma curtose. Com isso, obtém-se uma estatística mais simples que W_1 , dada por

$$W_2 = nR^2,$$

onde R^2 é o coeficiente de determinação da regressão dos quadrados dos resíduos sobre os $p(p+1)/2$ quadrados e produtos cruzados das variáveis explicativas originais, isto é,

$$\hat{u}_{t0}^2 = \alpha_0 + \sum_{j=1}^p \sum_{k=j}^p \alpha_l x_{tj} x_{tk}, \quad (2.20)$$

onde l é um índice que varia entre 1 e $p(p+1)/2$. Esse teste coincide com o SS , no caso em que a função de pesos envolve os termos polinomiais de segunda ordem de p variáveis independentes.

Sob a hipótese nula de homoscedasticidade, W_1 e W_2 têm distribuição assintótica χ^2 com $p(p+1)/2$ graus de liberdade. Para aplicação desses testes é necessário que $n \geq p(p+1)/2$.

É importante notar que caso a matriz de covariáveis X tenha um termo constante, digamos, $x_{t1} = 1$, $t = 1, \dots, n$, o cálculo de ambas as estatísticas apresentará problemas de singularidade. No primeiro teste, a matriz \tilde{D} em (2.18) apresentará a coluna correspondente à constante nula, de tal forma que a linha e a coluna de $\tilde{D}'S\tilde{D}$ correspondentes à constante também serão nulas, e no segundo teste, a regressão (2.20) apresentará uma constante redundante. Nesse caso, deve-se eliminar a constante redundante em (2.20), ou a coluna de \tilde{D} e a linha e a coluna de $\tilde{D}'S\tilde{D}$ correspondentes à constante, e diminuir os graus de liberdade em uma unidade. Esse mesmo problema pode ocorrer caso X contenha um termo constante e termos polinomiais. Por exemplo, se $x_{t1} = 1$ e $x_{t3} = x_{t2}^2$, então $x_{t1}x_{t3} = x_{t2}^2$. Nesse caso, deve-se proceder da mesma forma, isto é, eliminar o termo redundante em (2.20) ou a coluna de \tilde{D} e linha e coluna de $\tilde{D}'S\tilde{D}$ correspondentes ao termo redundante, e reduzir um grau de liberdade.

Segundo White (1980), espera-se que os testes que supõem um modelo para a heteroscedasticidade tenham um poder maior que W_1 e W_2 em amostras finitas, mas que W_1 e W_2 tenham desempenho igual ou superior aos testes que especificam incorretamente a estrutura de variâncias.

2.2.2 Teste de Goldfeld-Quandt

O teste de Goldfeld-Quandt (Goldfeld e Quandt, 1965) é amplamente utilizado em aplicações econométricas, possivelmente devido à sua grande facilidade computacional. Na construção desse teste supõe-se que as observações podem ser ordenadas de forma crescente de acordo com suas variâncias. Por exemplo, no caso em que a variância dos erros cresce com uma das variáveis preditoras, deve-se ordenar as observações de acordo com tal variável. Feito isso, deve-se escolher

um certo número r de observações centrais para serem omitidas, e ajustam-se regressões separadas, por MQO às $(n-r)/2$ primeiras e $(n-r)/2$ últimas observações, satisfazendo a condição de que $(n-r)/2 > p+1$, o número de parâmetros a serem estimados. Denotando por S_1 e S_2 as somas de quadrados do resíduo das regressões baseadas nas observações com menores e maiores variâncias, respectivamente, têm-se que sob a suposição de normalidade dos erros,

$$GQ = \frac{S_2}{S_1}$$

tem distribuição exata F com $(n-r-2p)/2$ graus de liberdade no numerador e $(n-r-2p)/2$ graus de liberdade no denominador sob a hipótese nula de homoscedasticidade.

Os autores comentam que o poder desse teste depende do valor escolhido para r , o número de observações omitidas. Para valores muito altos de r o poder do teste deve ser baixo, mas não é óbvio que o poder do teste aumente monotonicamente conforme r tende a zero. A explicação para isso é o fato de que, quando r é muito pequeno, apesar de as duas regressões serem baseadas em mais observações, conduzindo a um maior poder, ocorre que as variâncias das observações dos dois grupos diferirão menos do que quando r é relativamente grande. Esse último efeito tende a diminuir o poder. Eles comentam também que o poder do teste dependerá da natureza da amostra de valores da variável de ordenação. Assim, se a variância de tal variável for pequena em relação à sua média, espera-se que o poder do teste seja baixo.

Por meio de simulação, os autores mostraram que a escolha de r não causou grande impacto ao poder do teste, a não ser nos casos em que r era muito grande quando comparado ao número total de observações.

2.2.3 Teste de Levene

Levene (1960) propôs um teste para a comparação de variâncias de g grupos de observações, situação típica no contexto de modelos de análise de variância. Neter et al (1996) propõem a utilização desse teste para testar a existência de heteroscedasticidade em modelos de regressão. Eles apresentam o teste com modificações sugeridas por Conover et al (1981) para aumentar a robustez à não-normalidade.

Esse teste supõe que as variâncias dos erros aumentam (ou diminuem) com uma variável preditora X . Para aplicá-lo, deve-se dividir as observações em dois grupos, sendo um deles formado pelas observações com valores relativamente baixos de X e o outro formado pelas observações com valores relativamente altos dessa variável. Denotemos por \hat{u}_{t1} o t -ésimo resíduo do grupo 1 e por \hat{u}_{t2} o t -ésimo resíduo do grupo 2, e por n_1 e n_2 , respectivamente, os números de observações nesses grupos. Além disso, sejam \tilde{u}_1 e \tilde{u}_2 as medianas dos resíduos dos dois grupos e d_{t1} e d_{t2} os desvios absolutos dos resíduos em relação à mediana de seu grupo, isto é, $d_{t1} = |\hat{u}_{t1} - \tilde{u}_1|$ e $d_{t2} = |\hat{u}_{t2} - \tilde{u}_2|$. A estatística do teste é dada por

$$LV = \frac{\bar{d}_1 - \bar{d}_2}{s_d \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

onde \bar{d}_1 e \bar{d}_2 são as médias amostrais de d_{t1} e d_{t2} , respectivamente, e s_d^2 é a variância ponderada dos desvios dada por

$$s_d^2 = \frac{\sum_{t=1}^{n_1} (d_{t1} - \bar{d}_1)^2 + \sum_{t=1}^{n_2} (d_{t2} - \bar{d}_2)^2}{n - 2}.$$

Fica claro, assim, que esse teste corresponde a um teste t de comparação das médias dos desvios absolutos dos dois grupos. Mostra-se que mesmo que os erros da regressão não sejam normais, LV segue, sob a hipótese nula de homoscedasticidade e com tamanhos amostrais n_1 e n_2 não muito pequenos, distribuição aproximadamente t -Student com $n - 2$ graus de liberdade.

2.2.4 Teste de Szroeter

Na construção do teste de Szroeter (Szroeter, 1978), assim como na do teste de Goldfeld-Quandt, supõe-se somente que as observações podem ser ordenadas de acordo com suas variâncias, não sendo necessário conhecer a forma exata das variâncias σ_t^2 sob a hipótese alternativa. As observações podem ser ordenadas, por exemplo, com base nos valores de alguma variável preditora, ou nos valores previstos da variável dependente. Também é feita a suposição de normalidade dos erros.

Seja $\hat{u}_t = y_t - x_t' \hat{\beta}$, onde $\hat{\beta}$ é algum estimador consistente de β (como por exemplo o estimador de MQO). A estatística do teste é dada por

$$SZ = n(\tilde{b} - \bar{b}) \left[2 \sum_{t=1}^n (b_t - \bar{b})^2 \right]^{-\frac{1}{2}},$$

onde $\tilde{b} = \sum_{t=1}^n p_t b_t$, $p_t = \hat{u}_t^2 / \sum_{t=1}^n \hat{u}_t^2$, $\bar{b} = n^{-1} \sum_{t=1}^n b_t$ e b_1, \dots, b_n é um conjunto de n escalares não-estocásticos com a propriedade $b_t \geq b_s$ se $t > s$. Sob a hipótese nula de homoscedasticidade, isto é, $H_0 : \sigma_t^2 = \sigma_{t-1}^2$ para todo $t = 2, \dots, n$, contra a hipótese alternativa $H_1 : \sigma_t^2 > \sigma_{t-1}^2$ para pelo menos um $t \in \{2, \dots, n\}$, SZ segue uma distribuição assintótica normal com média zero e variância unitária.

Se for tomado $b_t = t$, $t = 1, \dots, n$, a estatística do teste assume uma forma simples, dada por

$$SZ = \left(\frac{6n}{n^2 - 1} \right)^{\frac{1}{2}} \left(\tilde{b} - \frac{n+1}{2} \right),$$

onde $\tilde{b} = \sum_{t=1}^n t \hat{u}_t^2 / \sum_{t=1}^n \hat{u}_t^2$. Esta será a forma considerada nesse trabalho.

2.2.5 Teste de Ramsey

Ramsey (1969) propôs uma adaptação ao conhecido teste de Bartlett (Kendall e Stuart, 1961) de comparação das variâncias de g grupos no modelo de análise de variância, para testar a hipótese de homoscedasticidade no modelo de regressão. O teste foi denominado *BAMSET* (*Bartlett's M Specification Error Test*).

Em sua versão original, o teste baseia-se em um tipo alternativo de resíduos proposto por Theil (1965). Tais resíduos possuem a propriedade de terem distribuição normal e serem homoscedásticos e não-correlacionados caso os erros do modelo (1.1) tenham essas propriedades, ao contrário dos resíduos de MQO. Estudos de simulação de Ramsey e Gilbert (1972) e Ali e Giaccotto (1984), no entanto, mostraram que com o uso de resíduos de MQO em tal teste obtém-se

um melhor desempenho, além da maior facilidade computacional. Assim, será apresentada aqui a versão baseada nos resíduos de MQO.

Esse teste supõe que os erros do modelo (1.1) sejam normais, mas não exige que as observações sejam ordenadas segundo suas variâncias. As observações devem ser divididas em três grupos com tamanhos aproximadamente iguais, denotados por n_1 , n_2 e n_3 . A estatística do teste é dada por

$$RS = n \log \hat{\sigma}_0^2 - \sum_{i=1}^3 n_i \log \hat{\sigma}_{0i}^2$$

onde $\hat{\sigma}_0^2$ é dado por (2.2) e $\hat{\sigma}_{0i}^2 = n_i^{-1} \sum_{t \in C_i} \hat{u}_{t0}^2$, onde C_i é o conjunto dos índices das observações no grupo i , $i = 1, 2, 3$.

Sob a hipótese nula de homoscedasticidade, RS tem distribuição assintótica χ^2 com 2 graus de liberdade. Caso as observações possam ser ordenadas segundo suas variâncias antes da composição dos grupos, é de se esperar que esse teste tenha um maior poder do que no caso em que as observações não são ordenadas.

2.3 Escolha das covariáveis no modelo da variância

Uma questão importante na aplicação dos testes construtivos tratados nesse capítulo é a escolha das k covariáveis que devem compor o vetor z_t , $t = 1, \dots, n$, em (1.3) ou (2.5). Talvez a forma mais intuitiva de se fazer tal escolha seja a seguinte: avaliar os gráficos de dispersão dos resíduos, ou alguma versão padronizada dos mesmos, versus as covariáveis que se imagina que possam afetar a variância dos erros do modelo de regressão, incluindo em z_t as covariáveis para as quais os gráficos exibirem alguma tendência de alteração da variabilidade dos pontos. Esses gráficos são eficazes no caso em que as variâncias são afetadas por apenas uma covariável ($k = 1$). No entanto, caso $k \geq 2$, devem ser vistos com cuidado, conforme ilustrado pelo exemplo a seguir. Suponhamos que as variâncias dos erros de um determinado modelo de regressão sejam determinados pelo modelo multiplicativo (1.6) com $k = 2$. Nesse caso,

$$\sigma_t^2 = \sigma^2 \exp(\delta_1 z_{1t} + \delta_2 z_{2t}).$$

Suponhamos ainda que $\delta_1 > 0$ e $\delta_2 > 0$. É fácil observar que as variâncias σ_t^2 , nesse caso, aumentarão com o aumento de z_{t1} , mantendo-se z_{t2} constante. Da mesma forma, aumentarão com o aumento de z_{t2} , mantendo-se z_{t1} constante. Nesse caso, os gráficos mencionados podem não ser adequados à seleção de variáveis, uma vez que as covariáveis não consideradas no gráfico não são mantidas constantes.

Verbyla (1993) utilizou resultados de Harvey (1976) e Cook e Weisberg (1983) e sugeriu um procedimento gráfico útil para auxiliar na escolha das covariáveis no caso do modelo multiplicativo (1.6). Tem-se que $\hat{u}_{t0}^2 \sim \nu_t^2 \chi_1^2$. Cook e Weisberg (1983) mostraram que

$$\nu_t^2 = (1 - h_{tt})^2 \sigma_t^2 + \sum_{j \neq t} h_{tj}^2 \sigma_j^2,$$

onde h_{tj} são os elementos da matriz chapéu (2.3). Em geral, os termos h_{tj} ($t \neq j$), são pequenos, de forma que $\nu_t^2 \approx (1 - h_{tt})^2 \sigma_t^2$. Além disso, Harvey (1976) mostrou que, assintoticamente,

$$E \left[\log \hat{u}_{t0}^2 + \log \frac{1}{2} - \psi \left(\frac{1}{2} \right) \right] = \log \nu_t^2,$$

$$\text{Var} \left[\log \hat{u}_{t0}^2 + \log \frac{1}{2} - \psi \left(\frac{1}{2} \right) \right] = \frac{\pi^2}{2},$$

onde $\psi(\cdot)$ é a função digama, dada por $\psi(s) = \Gamma'(s)/\Gamma(s)$, em que $\Gamma(\cdot)$ é a função gama. Com base nesses resultados, e considerando que $\log(1/2) - \psi(1/2) = 1,27036$, Verbyla (1993) mostrou que no caso do modelo multiplicativo (1.6),

$$E \left\{ \log \left[\frac{\hat{u}_{t0}^2}{(1 - h_{tt})^2} \right] + 1,27036 \right\} \approx z_t' \delta, \quad (2.21)$$

e que $\log(\hat{u}_{t0}^2/(1 - h_{tt})^2)$ tem variância constante. Uma vez que essa quantidade é linear nas covariáveis em z_t , pode-se utilizar os métodos convencionais de análise de regressão para escolha das covariáveis. Uma dificuldade prática ocorre se $\hat{u}_{t0} = 0$ para um ou mais valores de t , caso em que não é possível calcular o logaritmo. Verbyla (1993) sugeriu que tais pontos fossem omitidos no procedimento.

Uma sugestão para a escolha das covariáveis em z_t é a utilização de gráficos de variável adicionada (Atkinson, 1985). Na construção desses gráficos é necessário ajustar modelos de regressão tanto da variável resposta em (2.21), isto é,

$$r_t = \log[\hat{u}_{t0}^2/(1 - h_{tt})^2] + 1,27036, \quad (2.22)$$

quanto de uma dada variável z_{tj} , em função das outras covariáveis sob consideração. Em cada ajuste, devem ser obtidos os resíduos, os quais refletirão as porções de cada variável que não são explicadas pelas demais variáveis. O gráfico desses resíduos, um contra o outro, compõe o gráfico de variável adicionada. Assim, esses gráficos são mais adequados que os gráficos de resíduos versus covariáveis mencionados anteriormente, uma vez que refletem a importância de z_{tj} na explicação das variâncias σ_i^2 , dadas as demais variáveis. Se z_{tj} não for importante na presença das demais variáveis, o gráfico exibirá o formato de uma banda horizontal. Caso contrário, o formato obtido sugerirá a natureza da relação entre σ_i^2 e z_{tj} .

Capítulo 3

Avaliação numérica

Nesse capítulo serão avaliadas numericamente as propriedades, em amostras finitas, dos testes de heteroscedasticidade apresentados no capítulo anterior. Buscar-se-á identificar, entre os 28 testes considerados, quais apresentam melhores desempenhos em pequenas amostras, sob diversas condições.

Mas, o que vem a ser um teste com bom desempenho? Pode-se considerar que um dado teste de heteroscedasticidade tem um bom desempenho em amostras finitas se, nessa situação, apresentar baixas probabilidades de erros dos tipos I e II, correspondentes à rejeição da hipótese de homoscedasticidade quando os dados são na verdade homoscedásticos, e não rejeição da hipótese de homoscedasticidade quando os dados são heteroscedásticos, respectivamente. Os testes que apresentam taxa de rejeição superior ao nível de significância nominal são ditos anticonservativos, sendo, portanto, inadequados para uso em pequenas amostras. Esses testes levam à rejeição da hipótese nula de homoscedasticidade erroneamente com probabilidade mais alta que o desejado, induzindo a que sejam feitas, desnecessariamente, análises mais sofisticadas que as usuais baseadas no método de MQO. O oposto disso são os testes conservativos. À primeira vista, pode parecer uma característica desejável o fato de um teste ser conservativo, uma vez que nessa situação o erro tipo I ocorre com baixa probabilidade. No entanto, em geral testes conservativos apresentam um baixo poder, não sendo capazes de detectar níveis baixos ou moderados de heteroscedasticidade. Assim, busca-se identificar testes que tenham tamanhos empíricos razoavelmente próximos (ou menores) ao nível de significância nominal e, ao mesmo tempo, tenham um alto poder.

3.1 Detalhes metodológicos

Os resultados numéricos apresentados a seguir baseiam-se no modelo de regressão linear da forma

$$y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_{t,p-1} x_{t,p-1} + u_t,$$

$t = 1, \dots, n$, em que u_1, \dots, u_n são não-correlacionados, cada u_t tendo média zero e variância dada por $\sigma_t^2 = \sigma^2 w(z_t, \delta)$, com $z_t = (z_{t1}, \dots, z_{tk})'$ e $\delta = (\delta_1, \dots, \delta_k)'$.

Para a obtenção dos resultados, utilizou-se a linguagem de programação matricial Ox (Doornik, 2001). Em todas as simulações tomou-se $\beta_0 = \beta_1 = \dots = \beta_p = 1$ e $\sigma^2 = 1$. O número de replicações de Monte Carlo foi fixado em 10.000 e o número de replicações de bootstrap (B),

empregado no caso do teste *escore*, foi igual a 500. As maximizações não-lineares necessárias para o cálculo de estimativas de máxima verossimilhança foram feitas utilizando-se o algoritmo quasi-Newton BFGS, disponível em uma função pré-definida da linguagem `0x`. A matriz Z foi formada a partir das colunas $2, \dots, k+1$ de X , sendo Z a matriz com t -ésima linha dada por z'_t e X a matriz com t -ésima linha dada por $x'_t = (1, x_{t1}, \dots, x_{t,p-1})$. Essa escolha deve-se ao fato de que, em situações práticas, em geral as covariáveis que explicam a variância dos erros são um subconjunto, ou a totalidade, das covariáveis que afetam a média da variável resposta. Nos casos em que $k > p-1$, as colunas extra de Z foram obtidas independentemente a partir da distribuição $\mathcal{U}(0, 1)$.

Nos casos em que $k > 1$, ocorrem dificuldades para se ordenar as observações segundo suas variâncias, conforme requerido em vários testes não-constructivos. Se, por exemplo, as variâncias forem da forma $\sigma_t^2 = h(z'_t \delta)$, sendo $h(\cdot)$ uma função não decrescente, as observações devem ser ordenadas de acordo com os valores de $z'_t \delta$. No entanto, os parâmetros em δ são desconhecidos, inviabilizando na prática a correta ordenação das observações. Uma possível forma de viabilizar a aplicação prática desses testes quando $k > 1$ seria ordenar as observações apenas de acordo com uma das variáveis explicativas das variâncias (z_{t1}). Esse foi o procedimento adotado nesse trabalho. Será necessário avaliar se esse procedimento produz resultados satisfatórios, ou seja, se os testes são robustos a uma ordenação incorreta das observações. É de se esperar que a ordenação das observações não afete o tamanho dos testes. No entanto, é possível que o poder desses testes seja afetado pela ordenação incorreta.

No caso do teste de Goldfeld-Quandt, foram eliminadas cerca de um terço das observações. Assim, para $n = 20, 30, 50$ e 100 os números de observações eliminadas (r) foram, respectivamente, 6, 10, 16 e 34 observações. Além disso, no teste de Levene as observações foram divididas em dois grupos de igual tamanho.

Em uma primeira etapa, foram feitas simulações para avaliação dos tamanhos empíricos dos testes. Em tais simulações, tomou-se $\delta_1 = \dots = \delta_k = 0$, de forma a gerar dados homoscedásticos. Os resultados são apresentados na seção 3.2.

Nas simulações desta etapa buscou-se variar os elementos que podem influenciar o tamanho dos testes em amostras finitas. Tais elementos consistem no tamanho amostral (n), número de parâmetros de perturbação ($p+1$), número de parâmetros de interesse (k), valores das variáveis independentes e distribuição da variável dependente. Foram considerados níveis de significância nominais (α) de 10%, 5%, 1% e 0,5%. É importante notar que os tamanhos observados dos testes estão sujeitos a erros amostrais, uma vez que foram obtidos a partir de um número finito de replicações (10.000). Ao nível de 0,5%, um teste é significativamente anticonservativo¹ se sua taxa de rejeição observada exceder 10,8%, 5,6%, 1,3% e 0,7% respectivamente, para níveis de significância nominais de 10%, 5%, 1% e 0,5%. Por outro lado o teste será significativamente conservativo se suas taxas de rejeição forem inferiores a 9,2%, 4,4%, 0,7% e 0,3%, para os níveis de significância considerados. Os resultados apresentados basearam-se no modelo multiplicativo dado por (1.6).

A segunda etapa consistiu no estudo do poder dos testes sob diversas situações. Nesta etapa,

¹No sentido de que a probabilidade observada excede o valor crítico do teste de comparação de proporções com hipóteses $H_0 : \pi = \alpha$ versus $H_1 : \pi > \alpha$, sendo π a probabilidade real de rejeição da hipótese nula de homoscedasticidade por um determinado teste, em uma certa situação, e α o nível de significância nominal adotado.

novamente assumiu-se o modelo multiplicativo.

Todas as entradas das tabelas apresentadas correspondem a porcentagens. Além disso, os resultados relativos ao teste escore com correção no valor crítico são denotados por siglas que envolvem um asterisco sobrescrito. Assim, S^* , S_1^* , S_2^* , S_3^* e S_4^* denotam os resultados relativos ao teste escore com valores críticos modificados q^* , q_1^* , q_2^* , q_3^* e q_4^* , respectivamente, e S_{BO}^* corresponde aos resultados considerando os valores críticos obtidos pelo método de bootstrap.

3.2 Tamanhos dos testes

A Tabela 3.1 traz os tamanhos observados dos testes para tamanhos amostrais $n = 20, 30, 50$ e 100, fixando-se $p = 5$, $k = 2$, covariáveis geradas independentemente a partir da distribuição $U(0, 1)$ e erros com distribuição normal.

Observa-se que os tamanhos empíricos dos testes aproximam-se dos níveis de significância nominais com o aumento de n , conforme esperado.

Destaca-se o fato de que o teste da razão de verossimilhanças (RV) é altamente anticonservativo para tamanhos amostrais pequenos. No caso em que $n = 20$ e ao nível nominal de 5%, por exemplo, a taxa de rejeição observada (24,6%) é quase 5 vezes maior que a nominal. O teste mantém-se anticonservativo até mesmo para $n = 100$. A correção de Bartlett aplicada a esse teste (RVB) reduz as taxas de rejeição, porém não o bastante, de forma que para amostras com tamanhos menores ou iguais a 50, o teste manteve-se anticonservativo. O teste da razão de verossimilhanças modificado proposto por Simonoff e Tsai (RVM) foi eficiente em reduzir as taxas de rejeição do teste original. No entanto, em amostras com $n = 20$ e $n = 30$ a correção ocorreu de forma exagerada, tendo resultado em testes altamente conservativos. A formulação alternativa desse teste (RVM_1) apresentou melhores resultados, tendo sido menos conservativa. A aplicação da correção de Bartlett a esse teste ($RVMB_1$) produziu resultados bastante favoráveis, de forma que as taxas de rejeição empíricas são próximas às nominais mesmo em amostras pequenas, tais como $n = 20$ e $n = 30$. O teste da razão de verossimilhanças condicional (RVC) mostrou-se um pouco anticonservativo para tamanhos amostrais menores ou iguais a 50, enquanto que o teste da razão de verossimilhanças residual (RVR) apresentou resultados bastante satisfatórios até mesmo para tamanhos amostrais pequenos.

O teste escore (S) mostrou-se um pouco conservativo para tamanhos amostrais pequenos. A correção tipo-Bartlett (SB), assim como as correções monótonas (SB_1, SB_2 e SB_3) e correções do valor crítico (S^* , S_1^* , S_2^* , S_3^* e S_4^*) tiveram efeitos muito semelhantes sobre a estatística do teste escore original. Em todos os casos, houve pequenas melhorias, tendo esses se mostrado menos conservativos que o original. A aplicação da correção de bootstrap ao valor crítico do teste escore (S_{BO}^*) também trouxe pequenas melhorias em relação ao teste escore original, sendo esse teste, no entanto, ligeiramente conservativo para tamanhos amostrais pequenos. O teste escore modificado (SM) apresentou, em todos os casos, um desempenho praticamente igual ao do teste original. O teste escore studentizado apresentou resultados razoáveis, mesmo para tamanhos amostrais pequenos como $n = 20$ ou $n = 30$, ao nível de significância nominal de 5%. Ao nível de 10%, no entanto, mostrou-se ligeiramente anticonservativo, enquanto que aos níveis de 1% ou 0,5% mostrou-se conservativo. A versão modificada desse teste (SSM) apresentou desempenho quase idêntico ao original. O teste escore residual, por sua vez, mostrou-se em geral conservativo.

Com relação aos testes não-constructivos, chama a atenção o fato de que o primeiro teste de

Tabela 3.1.: Tamanhos observados dos testes com $p = 5$, $k = 2$, covariáveis uniformes e erros normais

Estatística	$n = 20$			$n = 30$			$n = 50$			$n = 100$					
	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%			
<i>RV</i>	34,2	24,6	11,5	21,8	13,8	4,3	2,8	16,1	9,0	2,4	1,4	11,9	6,7	1,5	0,8
<i>RVB</i>	21,3	13,3	4,7	13,9	7,4	1,7	0,9	11,3	5,7	1,3	0,6	9,9	5,3	1,0	0,5
<i>RVM</i>	4,4	1,9	0,2	7,7	3,4	0,5	0,1	9,0	4,2	0,8	0,4	9,4	4,9	0,8	0,4
<i>RVM₁</i>	7,5	3,7	0,5	8,6	4,0	0,6	0,2	9,4	4,4	0,8	0,4	9,4	4,9	0,8	0,4
<i>RVMB₁</i>	8,6	4,3	0,7	9,5	4,7	0,8	0,3	9,9	4,8	1,0	0,4	9,7	5,1	0,9	0,4
<i>RVC</i>	13,0	7,8	2,4	11,6	6,5	1,6	0,8	11,3	6,5	1,5	0,7	10,4	5,3	1,1	0,5
<i>RVR</i>	7,8	4,0	0,8	9,8	4,9	0,9	0,4	10,2	5,0	1,0	0,4	9,8	5,2	1,0	0,5
<i>S</i>	8,1	3,5	0,4	9,1	4,3	0,6	0,2	9,3	4,3	0,7	0,3	9,6	4,8	0,9	0,4
<i>SB</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>SB₁</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>SB₂</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>SB₃</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>S*</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>S*₁</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>S*₂</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>S*₃</i>	8,2	3,7	0,6	9,4	4,5	0,6	0,2	9,7	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>S*₄</i>	8,3	3,8	0,6	9,4	4,5	0,6	0,2	9,8	4,5	0,7	0,3	9,7	5,0	0,9	0,4
<i>S*_{BO}</i>	8,7	3,9	0,5	9,8	4,4	0,6	0,2	10,1	4,6	0,7	0,3	10,1	5,0	1,0	0,5
<i>SM</i>	8,1	3,7	0,5	9,1	4,3	0,6	0,2	9,3	4,2	0,7	0,3	9,5	4,8	0,9	0,4
<i>SS</i>	11,3	4,7	0,5	11,2	5,0	0,5	0,2	10,6	4,6	0,6	0,2	9,8	4,8	0,7	0,4
<i>SSM</i>	11,3	4,6	0,5	10,9	5,1	0,6	0,2	10,5	4,6	0,6	0,2	9,8	4,8	0,7	0,3
<i>SR</i>	6,9	2,8	0,4	8,0	3,7	0,5	0,2	9,0	4,0	0,6	0,3	9,3	4,8	0,9	0,4
<i>W₁</i>	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,9	0,2	0,0	0,0	2,9	0,7	0,0	0,0
<i>W₂</i>	0,0	0,0	0,0	6,3	1,4	0,0	0,0	9,3	3,9	0,5	0,2	9,5	4,8	1,1	0,5
<i>GQ</i>	10,7	5,5	1,2	10,2	5,0	1,1	0,6	10,5	5,1	1,1	0,5	10,0	5,1	1,0	0,5
<i>LV</i>	10,9	5,0	1,0	10,5	5,0	1,0	0,4	10,2	5,0	1,0	0,5	10,2	5,0	1,1	0,6
<i>SZ</i>	10,1	4,7	0,9	10,8	5,3	0,9	0,4	10,9	5,3	0,9	0,4	10,3	5,3	1,0	0,5
<i>RS</i>	11,8	5,9	1,3	12,0	6,3	1,3	0,7	11,4	5,9	1,4	0,7	10,7	5,8	1,1	0,5

White (W_1) é extremamente conservativo, apresentando baixíssimas taxas de rejeição da hipótese nula, mesmo para $n = 100$. Percebe-se, assim, que embora esse teste apresente características altamente desejáveis em aplicações práticas, não requerendo a especificação da forma da variância e da distribuição dos erros, seu desempenho é bastante ruim em pequenas amostras, uma vez que as baixas taxas de rejeição devem refletir em baixo poder do teste. O teste W_2 , no entanto, é altamente conservativo para amostras de tamanhos menores ou iguais a 30, mas para $n = 50$ ou $n = 100$ apresenta desempenho mais aceitável. Os testes de Goldfeld–Quandt (GQ), Levene (LV) e Szroeter (SZ) apresentaram bons resultados mesmo em amostras de tamanho $n = 20$. Por último, nota-se que o teste de Ramsey (RS) é ligeiramente anticonservativo para todos os tamanhos amostrais considerados.

Na Tabela 3.2 são apresentados resultados obtidos variando p , a dimensão de β . Foram considerados os valores $p = 2, 4, 6$ e 8 . Fixou-se $n = 30$, $k = 2$, as covariáveis em X foram geradas independentemente a partir da distribuição $\mathcal{U}(0, 1)$ e os erros foram gerados a partir da distribuição normal.

É possível observar que com o aumento de p , os testes RV e RVB tendem a se tornar cada vez mais anticonservativos. Para $p = 8$ e $\alpha = 5\%$, por exemplo, as taxas de rejeição desses testes são iguais a $28,3\%$ e $15,8\%$, respectivamente, excedendo em muito o nível nominal. O teste RVC apresentou tendência semelhante para valores de p menores ou iguais a 6, porém suas taxas de rejeição regrediram para $p = 8$. O aumento de p teve o efeito de tornar RVM e RVR mais conservativos, sendo que nesse último isso se deu sobretudo para $p = 8$. Nos casos dos testes RVM_1 e RVM_{B_1} , no entanto, o aumento de p não causou alterações substanciais nas taxas de rejeição. O primeiro teste, em todos os casos, mostrou-se ligeiramente conservativo, enquanto o segundo mostrou resultados muito favoráveis, até mesmo para $p = 8$.

As taxas de rejeição do teste score (S) não foram fortemente afetadas pelo aumento de p , o mesmo podendo ser dito com relação às versões desse teste com correções tipo–Bartlett ou correções no valor crítico ($SB, SB_1, SB_2, SB_3, S^*, S_1^*, S_2^*, S_3^*$ e S_4^*). Mais uma vez, todos esses testes apresentaram taxas de rejeição praticamente idênticas. As correções foram efetivas em tornar o teste score menos conservativo, nos casos em que $p = 2$ e $p = 4$. Nos demais casos, a alteração obtida foi muito pequena. Para valores de p iguais a 6 e 8, o teste score com valor crítico corrigido por bootstrap (S_{BO}^*) tornou-se ligeiramente conservativo. Os testes score studentizado (SS) e score residual (SR) também não foram muito afetados pelo aumento de p , sendo que o segundo manteve-se, em todos os casos, conservativo. Mais uma vez podemos notar que as modificações do teste score e score studentizado propostas por Simonoff e Tsai (SM e SSM) não tiveram praticamente nenhum efeito sobre as taxas de rejeição dos testes originais.

Com relação aos testes de White (W_1 e W_2), observa-se a tendência de se tornarem cada vez mais conservativos com o aumento de p . Para $p = 6$, ambos apresentaram taxas de rejeição da hipótese nula extremamente baixas. Vale lembrar que no caso em que $p = 8$, não foi possível calcular as estatísticas dos testes, já que nesse caso $p(p + 1)/2 = 36$, excedendo n . Os demais testes não construtivos não foram muito afetados pelo aumento de p , exceto SZ , que se tornou ligeiramente conservativo para $p = 8$. Mais uma vez, GQ, LV e SZ apresentaram taxas de rejeição razoavelmente próximas aos níveis nominais. O teste de Ramsey, conforme observado anteriormente, foi em todos os casos ligeiramente anticonservativo.

Na Tabela 3.3 são apresentados resultados envolvendo o aumento de k , o número de parâmetros de interesse. Foram considerados os valores $k=1, 2, 3$ e 4 . Além disso, fixou-se $n = 30$, $p = 5$,

Tabela 3.2: Tamanhos observados dos testes, com $n = 30$, $k = 2$, covariáveis uniformes e erros normais

Estatística	$p = 2$			$p = 4$			$p = 6$			$p = 8$						
	10%	5%	1%	0,5%	10%	5%	1%	0,5%	10%	5%	1%	0,5%				
<i>RV</i>	14,3	7,9	2,0	1,0	19,2	11,4	3,6	2,2	24,4	16,7	6,6	4,4	38,3	28,3	14,4	10,6
<i>RVB</i>	10,7	5,4	1,1	0,4	12,6	6,7	1,8	0,9	16,0	9,2	2,5	1,5	24,1	15,8	5,7	3,7
<i>RVM</i>	8,6	4,2	0,6	0,2	8,4	4,0	0,7	0,3	7,6	3,5	0,6	0,2	4,1	1,7	0,2	0,1
<i>RVM₁</i>	8,7	4,3	0,6	0,2	8,8	4,4	0,8	0,3	9,1	4,6	0,8	0,3	8,4	3,9	0,7	0,3
<i>RVMB₁</i>	9,8	4,8	0,9	0,3	9,8	4,9	1,1	0,5	10,0	5,1	0,9	0,5	9,3	4,5	0,8	0,4
<i>RVC</i>	11,7	6,7	1,8	1,0	13,3	7,7	2,1	1,2	14,9	9,3	3,1	1,9	12,0	6,9	1,7	0,8
<i>RVR</i>	10,4	5,2	1,0	0,4	10,3	5,3	1,3	0,6	9,8	5,2	1,0	0,5	6,3	3,0	0,5	0,2
<i>S</i>	8,3	3,9	0,7	0,4	9,1	4,5	0,7	0,3	9,6	5,0	0,9	0,4	8,4	3,6	0,7	0,3
<i>SB</i>	9,8	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,4	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>SB₁</i>	9,8	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,4	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>SB₂</i>	9,9	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,4	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>SB₃</i>	9,9	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,4	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>S*</i>	9,9	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,4	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>S*₁</i>	9,9	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,3	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>S*₂</i>	9,9	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,3	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>S*₃</i>	9,8	4,7	0,8	0,5	9,9	5,0	0,9	0,4	9,4	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>S*₄</i>	9,9	4,8	0,9	0,5	9,9	5,0	0,9	0,4	9,4	4,8	0,9	0,4	8,1	3,5	0,7	0,3
<i>S*_{BO}</i>	10,3	4,6	0,8	0,5	10,2	4,8	0,9	0,4	9,8	4,4	0,7	0,3	9,1	4,1	0,7	0,4
<i>SM</i>	8,2	3,9	0,7	0,4	9,1	4,4	0,7	0,3	9,6	4,9	0,9	0,4	8,4	3,6	0,7	0,4
<i>SS</i>	10,4	4,6	0,7	0,3	11,0	4,6	0,7	0,2	11,9	5,3	0,6	0,3	10,4	4,3	0,5	0,1
<i>SSM</i>	10,3	4,6	0,7	0,3	10,9	4,8	0,7	0,2	11,9	5,2	0,6	0,3	10,1	4,3	0,5	0,2
<i>SR</i>	8,4	4,0	0,7	0,5	8,8	4,3	0,6	0,3	8,2	4,0	0,6	0,3	7,9	3,4	0,6	0,3
<i>W₁</i>	9,7	3,0	0,1	0,0	1,3	0,2	0,0	0,0	0,0	0,0	0,0	0,0	—	—	—	—
<i>W₂</i>	9,2	4,3	0,8	0,4	8,7	3,4	0,2	0,1	0,4	0,0	0,0	0,0	—	—	—	—
<i>GQ</i>	10,4	5,2	1,0	0,5	10,1	5,0	0,9	0,5	10,3	5,1	1,1	0,6	9,2	4,4	0,9	0,4
<i>LV</i>	9,6	4,4	0,6	0,3	9,5	4,2	0,6	0,3	9,5	4,5	0,7	0,4	9,2	4,2	0,8	0,3
<i>SZ</i>	10,1	4,7	0,8	0,3	9,6	4,4	0,7	0,3	10,2	4,9	0,9	0,4	8,4	3,8	0,6	0,3
<i>RS</i>	12,3	6,4	1,4	0,7	11,5	6,3	1,4	0,8	11,2	6,1	1,5	0,9	12,1	6,4	1,4	0,7

as covariáveis foram geradas a partir da distribuição $\mathcal{U}(0,1)$ e os erros a partir da distribuição normal.

Os comportamentos verificados foram bastante similares aos obtidos no caso do aumento de p . Os testes que tiveram suas taxas de rejeição mais afetadas pelo aumento de k foram RV , RVB e RVC , tornando-se mais anticonservativos, e RVM , tendendo a se tornar mais conservativo. Novamente, observou-se que para qualquer valor de k o teste SR foi um pouco conservativo, os testes de White (W_1 e W_2) foram altamente conservativos e que o teste de Ramsey (RS) foi ligeiramente anticonservativo. Os demais apresentaram desempenhos razoáveis, tendo tamanhos empíricos próximos aos níveis nominais, ou então sendo ligeiramente conservativos.

Nas simulações apresentadas até esse ponto, as características das covariáveis e da variável resposta foram mantidas em condições ideais, sem a ocorrência de pontos de alavanca, “outliers” ou pontos influentes nos dados, problemas típicos em aplicações práticas. Assim, os resultados a seguir contemplam situações mais realísticas, com a variação da distribuição das covariáveis e da variável resposta.

Na Tabela 3.4 são apresentados resultados obtidos com diferentes configurações das covariáveis. Fixou-se $n = 30$ e $p = 5$. Foram considerados modelos com intercepto, de forma que a primeira coluna de X é unitária. As 4 demais colunas foram geradas de diferentes formas. Foram considerados os casos em que essas colunas foram geradas a partir das distribuições t_3 , $LN(0,1)$ (log-normal), normal multivariada (NM) com média nula e matriz de covariâncias Σ dada por

$$\Sigma = \begin{pmatrix} 1,0 & 0,7 & 0,0 & 0,0 \\ 0,7 & 1,0 & 0,0 & 0,0 \\ 0,0 & 0,0 & 1,0 & 0,7 \\ 0,0 & 0,0 & 0,7 & 1,0 \end{pmatrix},$$

e a matriz correspondente a um modelo de análise de variância (ANOVA) com um fator e cinco grupos, havendo em cada grupo 6 observações. O primeiro caso consiste em uma distribuição simétrica com caudas pesadas, enquanto o segundo consiste em uma distribuição assimétrica. No terceiro caso, ilustra-se o comportamento dos testes na presença de multicolinearidade moderada, uma vez que os pares x_{t2} e x_{t3} , e x_{t4} e x_{t5} , foram gerados com um coeficiente de correlação de 0,7. No último caso, têm-se uma situação em que há interesse em se testar a igualdade das variâncias dos grupos, isto é, $H_0 : \sigma_1^2 = \dots = \sigma_5^2 = \sigma^2$, em que σ_i^2 é a variância das observações do i -ésimo grupo. Nos três primeiros casos, fixou-se $k = 2$, e no último, tomou-se $k = 4$, para obter um modelo que admita variâncias diferentes para cada grupo.

No caso do modelo de ANOVA, cabem alguns comentários a respeito da aplicação de alguns testes não-constructivos. O primeiro deles refere-se ao teste de Goldfeld-Quandt. Esse teste supõe, conforme mencionado na seção 2.2.2, que as observações sejam ordenadas de acordo com suas variâncias, e que sejam excluídas r observações centrais. Feito isso, as observações são divididas em dois blocos, sendo o primeiro formado pelas primeiras $(n - r)/2$ primeiras observações, e o segundo formado pelas últimas $(n - r)/2$ observações. Assim, para aplicar o teste no contexto de modelos de análise de variância seria necessário classificar as observações em ordem crescente das variâncias dos grupos, eliminando-se, eventualmente, um ou mais grupos centrais. O problema ocorre no ajuste de modelos de regressão a cada um dos dois blocos de observações formados, o que é requerido para o cálculo da estatística desse teste. Para as observações do primeiro bloco, por exemplo as colunas de X que contiverem as variáveis indicadoras dos grupos incluídos no

Tabela 3.3: Tamanhos observados dos testes, com $n = 30$, $p = 5$, covariáveis uniformes e erros normais

Estatística	$k = 1$			$k = 2$			$k = 3$			$k = 4$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
<i>RV</i>	17,7	10,5	3,2	21,8	13,8	4,3	26,7	17,4	6,6	30,7	20,9	8,2
<i>RVB</i>	12,4	6,5	1,5	13,9	7,4	1,7	16,0	9,1	2,2	18,2	10,4	2,8
<i>RVM</i>	8,3	3,7	0,6	7,7	3,4	0,5	7,2	3,1	0,4	6,4	2,8	0,4
<i>RVM₁</i>	8,8	4,2	0,7	8,6	4,0	0,6	8,7	4,0	0,6	8,6	4,0	0,7
<i>RVMB₁</i>	9,6	4,7	0,8	9,5	4,7	0,8	9,6	4,8	0,7	9,8	4,5	0,8
<i>RVC</i>	11,8	6,7	1,6	11,6	6,5	1,6	13,5	8,0	2,0	13,2	8,1	2,3
<i>RVR</i>	9,9	5,0	0,9	9,8	4,9	0,9	9,6	4,8	0,7	9,3	4,5	0,8
<i>S</i>	9,7	4,3	0,6	9,1	4,3	0,6	9,0	4,4	0,7	8,5	4,3	0,9
<i>SB</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,6
<i>SB₁</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,6
<i>SB₂</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,6
<i>SB₃</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,6
<i>S*</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,7
<i>S*₁</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,7
<i>S*₂</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,7
<i>S*₃</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,7
<i>S*₄</i>	9,8	4,5	0,8	9,4	4,5	0,6	9,4	4,6	0,7	8,9	4,3	0,7
<i>S*_{BO}</i>	10,3	4,8	0,8	9,8	4,4	0,6	9,3	4,0	0,4	9,1	3,6	0,4
<i>SM</i>	9,4	4,4	0,7	9,1	4,3	0,6	8,9	4,4	0,7	8,7	4,3	1,0
<i>SS</i>	11,5	5,4	0,7	11,2	5,0	0,5	10,7	4,6	0,4	10,5	4,4	0,3
<i>SSM</i>	11,4	5,5	0,7	10,9	5,1	0,6	10,6	4,6	0,4	10,5	4,4	0,3
<i>SR</i>	8,8	3,9	0,6	8,0	3,7	0,5	8,2	3,8	0,6	7,9	4,0	0,8
<i>W₁</i>	0,3	0,0	0,0	0,3	0,0	0,0	0,3	0,0	0,0	0,3	0,0	0,0
<i>W₂</i>	6,3	1,4	0,0	6,3	1,4	0,0	6,3	1,4	0,0	6,3	1,4	0,0
<i>GQ</i>	10,2	5,0	1,1	10,2	5,0	1,1	10,2	5,1	1,0	10,2	5,0	1,1
<i>LV</i>	10,5	5,0	1,0	10,5	5,0	1,0	10,5	5,0	1,0	10,5	5,0	1,0
<i>SZ</i>	10,8	5,3	0,9	10,8	5,3	0,9	10,8	5,4	0,9	10,8	5,3	0,9
<i>RS</i>	12,0	6,3	1,3	12,0	6,3	1,3	12,0	6,3	1,3	12,0	6,3	1,3

segundo bloco serão constantes e iguais a zero, causando problemas de singularidade. Logo, esse teste, tal como apresentado originalmente por Goldfeld e Quandt (1965), não é adequado para aplicação no contexto de modelos de análise de variância. O segundo comentário refere-se aos testes de White. Nesse caso, a matriz de quadrados e produtos cruzados das variáveis em X , dada em (2.19), terá diversas colunas redundantes ou nulas. Essas colunas deverão ser removidas para cálculo dessas estatísticas, de forma que só restarão em (2.19) as colunas $(2, \dots, p)$ da matriz X . O número de graus de liberdade da distribuição assintótica será igual a $p - 1$.

É possível observar na Tabela 3.4 que as características das covariáveis que tiveram maior impacto sobre os tamanhos empíricos dos testes, em relação aos resultados obtidos com covariáveis uniformemente distribuídas reportados na Tabela 3.1, foram a curtose e assimetria. Os testes da razão de verossimilhanças (RV) e sua versão com correção de Bartlett (RVB) mostraram-se ainda mais anticonservativos que anteriormente, sobretudo sob assimetria das covariáveis. Para covariáveis com distribuição $LN(0, 1)$ e ao nível nominal $\alpha = 5\%$, por exemplo, foram obtidas taxas de rejeição de 35, 2% e 16, 2% para cada um deles, excedendo em muito o nível nominal. As versões modificadas desse teste (RVM , RVM_1 e $RVMB_1$), assim como RVR , tornaram-se conservativas sob curtose e assimetria das covariáveis. Dentre essas, a que apresentou resultados mais razoáveis em ambas as situações foi $RVMB_1$. O teste da razão de verossimilhanças condicional (RVC), por sua vez, mostrou-se fortemente afetado pela curtose das covariáveis, tendo sido muito conservativo no caso da distribuição t_3 . Sob assimetria das covariáveis, os resultados obtidos para este teste foram bastante razoáveis. Este teste também foi afetado pela multicolinearidade, tendo sido conservativo no caso de covariáveis $NM(0, \Sigma)$.

Com relação ao teste escore (S), houve também forte impacto da assimetria e curtose das covariáveis, que o tornaram altamente conservativo. Para a distribuição $LN(0, 1)$ e ao nível de significância nominal de 5%, por exemplo, sua taxa de rejeição foi de apenas 1, 1%. Nessas situações, ao contrário do que vinha sendo observado até então, as correções tipo-Bartlett e correções do valor crítico (SB , SB_1 , SB_2 , SB_3 , S^* , S_1^* , S_2^* , S_3^* e S_4^*) produziram resultados diferenciados entre si. Nota-se que essas correções foram eficazes em aproximar as taxas de rejeição dos níveis nominais, sobretudo aos níveis de significância nominais mais altos, de 10% e 5%. Dentre essas correções, as mais eficientes, no sentido de tornar o teste escore menos conservativo, foram SB_3 e S_4^* . A correção do valor crítico por meio de bootstrap conduziu a taxas de rejeição próximas às nominais no caso de curtose, sendo esse teste apenas um pouco conservativo no caso de assimetria das covariáveis. O teste escore studentizado (SS) também foi afetado pela assimetria e curtose das covariáveis, tendo se mostrado bastante conservativo. Assim como observado nas simulações anteriores, as correções propostas por Simonoff e Tsai (1994) aos testes escore (SM) e escore studentizado (SSM) não trouxeram praticamente nenhuma melhoria aos testes originais em termos de taxas de rejeição. O teste escore residual (SR) apresentou, tanto sob curtose quanto sob assimetria, tamanhos próximos aos nominais ao nível $\alpha = 5\%$. Para os demais níveis, foi por vezes ligeiramente conservativo ou anticonservativo.

Os testes de White, sobretudo W_1 , mantiveram, sob as diferentes configurações das covariáveis, seu comportamento altamente conservativo verificado anteriormente. A única exceção foi W_2 no caso do modelo de ANOVA, em que as taxas de rejeição não se distanciaram exageradamente dos níveis nominais. Os demais testes não construtivos não foram afetados pela curtose das covariáveis. No entanto, os testes de Levene (LV) e Szroeter (SZ) tornaram-se consideravelmente conservativos sob assimetria das covariáveis (sobretudo o primeiro). O teste de Ramsey (RS)

Tabela 3.4: Tamanhos observados dos testes para diversas distribuições das covariáveis, com $n = 30$, $p = 5$, $k = 2$ (exceto no modelo de ANOVA, em que $k = 4$) e erros normais

Estatística	t_3				$LN(0,1)$				$NM(0, \Sigma)$				ANOVA			
	10%	5%	1%	0,5%	10%	5%	1%	0,5%	10%	5%	1%	0,5%	10%	5%	1%	0,5%
<i>RV</i>	29,7	20,1	8,0	5,3	45,5	35,2	19,0	14,9	24,2	15,6	5,8	3,7	20,5	12,4	3,9	2,2
<i>RVB</i>	17,9	10,4	2,9	1,6	25,4	16,2	5,8	3,8	15,5	8,9	2,3	1,2	12,1	6,5	1,2	0,6
<i>RVM</i>	2,8	1,2	0,2	0,1	3,1	1,4	0,3	0,1	6,5	3,0	0,4	0,2	10,4	5,2	0,8	0,4
<i>RVM₁</i>	5,9	2,8	0,5	0,3	7,7	3,6	0,6	0,3	8,4	4,0	0,6	0,3	10,4	5,2	0,8	0,4
<i>RVMB₁</i>	7,8	3,8	0,8	0,4	7,8	3,6	0,6	0,3	9,7	4,7	0,9	0,4	10,4	5,2	0,8	0,4
<i>RVC</i>	2,9	1,3	0,3	0,1	9,1	4,7	1,3	0,7	6,5	2,9	0,5	0,2	11,4	6,4	1,3	0,7
<i>RVR</i>	4,0	2,0	0,4	0,2	4,5	2,3	0,5	0,3	8,7	4,2	0,8	0,3	13,1	7,1	1,5	0,7
<i>S</i>	4,4	2,0	0,4	0,2	2,7	1,1	0,2	0,1	7,9	4,0	0,7	0,4	13,4	7,2	1,9	1,1
<i>SB</i>	8,5	3,3	0,0	0,0	8,1	3,5	0,0	0,0	9,0	4,2	0,5	0,3	11,6	5,6	0,0	0,0
<i>SB₁</i>	8,8	3,5	0,2	0,2	8,2	3,6	0,1	0,1	9,0	4,2	0,5	0,4	11,6	5,7	1,1	1,1
<i>SB₂</i>	9,5	4,2	0,1	0,0	8,8	4,1	0,2	0,1	9,0	4,2	0,5	0,3	11,6	5,8	0,0	0,0
<i>SB₃</i>	10,3	4,8	0,0	0,0	9,8	4,9	0,5	0,0	9,1	4,3	0,6	0,3	11,7	5,8	0,7	0,0
<i>S*</i>	8,6	2,9	0,1	0,0	14,2	4,4	0,1	0,0	9,0	4,2	0,6	0,3	11,7	5,9	1,1	0,5
<i>S₁*</i>	7,9	2,6	0,0	0,0	12,8	3,7	0,0	0,0	8,9	4,1	0,6	0,3	11,6	5,9	1,0	0,3
<i>S₂*</i>	7,3	2,2	0,0	0,0	9,7	3,1	0,0	0,0	8,9	4,1	0,5	0,3	11,6	5,9	1,0	0,4
<i>S₃*</i>	6,6	2,5	0,2	0,0	4,3	1,5	0,2	0,1	8,9	4,1	0,6	0,3	12,4	6,4	1,4	0,8
<i>S₄*</i>	10,3	4,8	0,0	0,0	9,8	4,9	0,5	0,0	9,1	4,3	0,6	0,3	11,7	5,8	0,7	0,0
<i>S*_{BO}</i>	11,0	5,4	1,2	0,6	8,3	3,7	0,8	0,5	10,3	4,6	0,7	0,4	7,7	3,0	0,3	0,1
<i>SM</i>	4,1	1,9	0,5	0,2	2,3	1,2	0,3	0,2	7,7	3,8	0,7	0,4	13,4	7,2	1,9	1,1
<i>SS</i>	5,3	2,3	0,3	0,1	3,0	1,2	0,3	0,1	10,2	4,4	0,6	0,2	13,5	6,0	0,7	0,3
<i>SSM</i>	4,8	2,0	0,3	0,1	2,6	1,3	0,3	0,2	10,0	4,2	0,6	0,2	13,5	6,0	0,7	0,3
<i>SR</i>	9,6	5,4	1,7	1,1	7,6	4,5	1,6	1,0	8,6	4,3	0,9	0,5	7,7	3,6	0,8	0,4
<i>W₁</i>	0,4	0,0	0,0	0,0	0,3	0,0	0,0	0,0	0,2	0,0	0,0	0,0	6,6	1,9	0,2	0,0
<i>W₂</i>	5,3	1,4	0,0	0,0	4,7	1,3	0,0	0,0	5,8	1,3	0,0	0,0	13,5	6,0	0,7	0,3
<i>GQ</i>	10,4	5,3	1,0	0,4	10,1	4,7	1,0	0,5	9,7	4,9	0,9	0,4	—	—	—	—
<i>LV</i>	11,4	5,5	1,0	0,4	3,7	1,7	0,2	0,1	9,5	4,6	0,7	0,3	9,5	4,6	0,7	0,3
<i>SZ</i>	10,4	4,9	0,8	0,3	6,9	3,2	0,6	0,2	11,6	5,8	1,1	0,4	9,5	4,3	0,6	0,2
<i>RS</i>	12,1	6,5	1,6	0,8	14,2	8,0	2,1	1,2	11,5	6,3	1,4	0,7	10,6	5,4	1,0	0,5

manteve seu comportamento anticonservativo sob as diferentes configurações das covariáveis, com exceção do modelo de ANOVA.

Na Tabela 3.5 são apresentados os tamanhos dos testes variando-se a distribuição dos erros. Foram consideradas as distribuições t_3 e t_{10} , com caudas pesadas, e a distribuição χ_5^2 padronizada², assimétrica. Fixou-se $n = 30$, $p = 5$, $k = 2$, e as covariáveis foram geradas independentemente a partir da distribuição $U(0, 1)$. Conforme visto no Capítulo 2, a grande maioria dos testes supõem normalidade dos erros. As únicas exceções são o teste escore studentizado (SS), sua versão modificada (SSM), os testes de White (W_1 e W_2) e o teste de Levene (LV). Assim, as simulações a seguir têm o intuito de avaliar a robustez dos testes que supõem normalidade em situações em que essa suposição não é satisfeita.

Observa-se que todos os testes que supõem normalidade foram altamente anticonservativos nas situações apresentadas. Isso ocorreu inclusive com testes que apresentaram, nas simulações anteriores, comportamento conservativo. Um ponto importante a ser observado é que grande parte dos testes foram consideravelmente anticonservativos no caso em que os erros foram gerados a partir da distribuição t_{10} . No caso do teste escore (S), por exemplo, ao nível de significância nominal $\alpha = 5\%$, observou-se uma taxa de rejeição de 9%, portanto quase igual ao dobro do esperado. Conforme se observa na Figura 3.1.a, as caudas dessa distribuição são razoavelmente próximas das caudas da distribuição normal. Além disso, a Figura 3.1.b traz um gráfico de resíduos externamente studentizados³ (ver Neter, 1996) obtidos a partir de uma regressão com erros t_{10} , o qual apresenta evidências muito fracas de não-normalidade, com a ocorrência de dois pontos um pouco superiores a 2. Isso indica que, em geral, os testes são muito sensíveis à suposição de normalidade de erros. Portanto, ao menor indício de não-normalidade, deve-se avaliar com cuidado os resultados dos testes de heteroscedasticidade, dando maior atenção aos testes robustos mencionados.

Assim, as simulações a respeito dos tamanhos empíricos dos testes trouxeram várias conclusões interessantes. Observou-se, por exemplo, que o teste da razão de verossimilhanças (RV) foi anticonservativo em todas as situações consideradas. Outros testes também se mostraram anticonservativos em diversas situações, embora em menor intensidade. Foram eles o teste da razão de verossimilhanças com correção de Bartlett (RVB), o teste da razão de verossimilhanças condicional (RC) e o teste de Ramsey (RS). Assim, o uso desses testes não é recomendável em amostras pequenas, uma vez que os mesmos conduzem à rejeição da hipótese nula de homoscedasticidade erroneamente com uma probabilidade superior à desejada. Por outro lado, os testes de White (W_1 e W_2) mostraram-se, quase sempre, excessivamente conservativos para os tamanhos amostrais observados. Logo, apesar de esses testes terem características bastante desejáveis, tais como a não-necessidade de que se especifique a distribuição dos erros nem a forma das variâncias, os mesmos não parecem adequados para aplicação em tamanhos amostrais menores ou iguais a 100, o que deverá ser confirmado a partir dos resultados das análises sobre o poder. Observou-se também que alguns testes são bastante sensíveis à presença de pontos de alavanca, tornando-se

²A variável foi padronizada subtraindo 5 (média da distribuição χ_5^2) dos valores gerados e dividindo-os por $\sqrt{10}$ (desvio padrão da distribuição χ_5^2), de forma a obter uma variável com média zero e variância unitária

³Esses resíduos têm distribuição t -Student com $n - p - 2$ graus de liberdade sob homoscedasticidade, sendo amplamente utilizados no diagnóstico de modelos de regressão linear

Tabela 3.5: Tamanhos observados dos testes para diversas distribuições dos erros, com covariáveis uniformes, $n = 50$, $p = 5$ e $k = 2$

Estatística	t_3			t_{10}			χ^2_3 (padronizada)					
	10%	5%	1%	10%	5%	1%	10%	5%	1%			
<i>RV</i>	48,6	40,0	25,9	21,4	22,9	14,8	5,2	3,6	33,2	24,0	11,3	8,2
<i>RVB</i>	43,3	34,4	20,7	16,9	17,2	10,6	3,4	1,9	27,3	18,5	7,5	4,9
<i>RVM</i>	39,7	30,9	18,0	14,8	14,6	8,3	2,3	1,3	23,9	15,6	5,6	3,6
<i>RVM₁</i>	40,2	31,4	18,3	15,1	15,0	8,5	2,5	1,4	24,4	16,0	5,9	3,7
<i>RVMB₁</i>	41,0	32,4	18,9	15,6	15,8	9,1	2,7	1,5	25,3	16,8	6,5	4,0
<i>RVC</i>	38,9	30,9	19,1	15,5	15,8	9,7	3,3	2,0	24,7	16,9	7,1	4,8
<i>RVR</i>	41,3	32,7	19,2	15,8	16,0	9,3	2,9	1,6	25,6	17,0	6,7	4,3
<i>S</i>	39,3	31,0	18,6	14,9	15,2	9,0	2,7	1,6	24,3	16,1	6,2	4,1
<i>SB</i>	40,0	31,3	18,7	15,0	15,7	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>SB₁</i>	40,1	31,5	18,9	15,1	15,7	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>SB₂</i>	40,1	31,5	18,9	15,1	15,7	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>SB₃</i>	40,1	31,5	18,9	15,2	15,8	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>S*</i>	40,1	31,5	18,9	15,1	15,8	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>S*₁</i>	40,1	31,5	18,9	15,1	15,8	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>S*₂</i>	40,1	31,5	18,9	15,1	15,8	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>S*₃</i>	40,1	31,5	18,9	15,1	15,7	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>S*₄</i>	40,1	31,5	18,9	15,2	15,8	9,3	2,8	1,7	24,9	16,5	6,4	4,2
<i>S*_{BO}</i>	15,9	8,1	2,3	1,3	11,3	5,7	0,9	0,4	14,8	7,9	1,7	0,9
<i>SM</i>	39,3	30,7	18,6	15,0	15,3	9,0	2,8	1,6	24,2	16,0	6,2	4,2
<i>SS</i>	9,4	4,0	0,4	0,2	10,0	4,6	0,5	0,2	12,7	6,0	0,8	0,3
<i>SSM</i>	9,8	4,2	0,5	0,2	10,1	4,7	0,5	0,2	12,8	6,2	0,9	0,3
<i>SR</i>	38,9	30,3	18,3	14,7	14,9	8,8	2,7	1,6	23,7	15,6	5,9	4,0
<i>W₁</i>	0,3	0,0	0,0	0,0	0,5	0,1	0,0	0,0	1,0	0,2	0,0	0,0
<i>W₂</i>	11,9	6,6	1,6	0,9	10,0	4,8	0,7	0,4	12,1	6,3	1,2	0,4
<i>GQ</i>	21,2	15,3	7,3	5,5	12,4	6,7	1,9	1,2	15,4	9,7	3,1	1,9
<i>LV</i>	10,7	5,0	0,8	0,3	10,2	4,9	0,8	0,4	10,9	5,3	0,9	0,4
<i>SZ</i>	21,7	16,0	8,0	6,3	13,3	7,5	1,9	1,1	16,8	10,5	3,7	2,4
<i>RS</i>	46,2	36,9	22,3	18,2	17,7	10,3	3,5	2,1	27,1	18,2	6,7	4,5

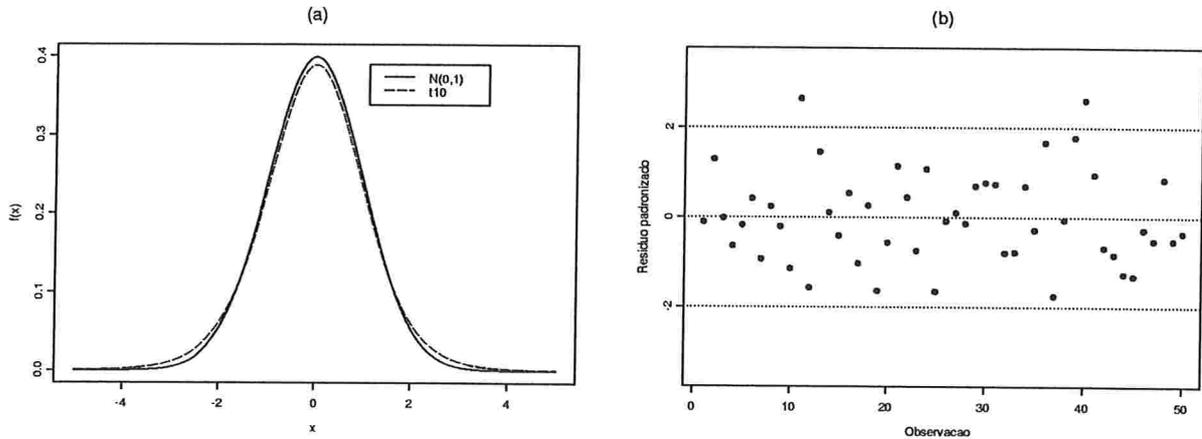


Figura 3.1: (a) Funções densidade de probabilidade das distribuição $N(0, 1)$ e t_{10} ; (b) Resíduos ordinários obtidos a partir de uma regressão com $n = 50$, $p = 5$ e erros homoscedásticos com distribuição t_{10}

nesse caso muito conservativos. Foi o caso, por exemplo, dos testes RVM , RVM_1 , RVR , S e SS . Um outro ponto importante é o fato de que os testes que supõem normalidade dos erros são altamente sensíveis a essa suposição, mostrando-se anticonservativos até mesmo quando o desvio é pequeno, como no caso da distribuição t_{10} . Somente foram robustos os testes escore studentizado (SS), sua versão modificada (SSM), os testes de White (W_1 e W_2) e o teste de Levene (LV).

3.3 Poder dos testes

Nesta seção, será avaliado o poder dos testes em estudo, sob diversas situações. Serão desconsiderados os testes que, nas simulações de tamanho, mostraram-se anticonservativos em grande parte das situações. São eles o teste da razão de verossimilhanças (RV), sua versão com correção de Bartlett (RVB), o teste da razão de verossimilhanças condicional (RVC) e o teste de Ramsey (RS).

Será introduzida uma medida do grau de heteroscedasticidade exibido por um determinado conjunto de dados simulados. Essa medida é dada por $\lambda = \max(\sigma_t^2)/\min(\sigma_t^2)$. Sob homoscedasticidade, tem-se que $\lambda = 1$.

Em primeiro lugar, apresentaremos os resultados relativos ao poder dos testes em uma situação muito simples: um modelo de análise de variância balanceado com um fator e dois níveis⁴. Nesse caso, tem-se que $p = 2$ e $k = 1$, de forma que $\sigma_1^2 = \sigma^2$ e $\sigma_2^2 = \sigma^2 \exp(\delta)$, sendo σ_1^2 e σ_2^2 as variâncias das observações dos grupos 1 e 2, respectivamente. Assim, tomou-se $\delta = \log(b)$, com b variando entre 1 e 7, com incrementos iguais a 0,5, de forma a obter valores de λ variando entre 1 e 7. O tamanho amostral foi fixado em $n = 30$, e os dados foram gerados a partir da distribuição normal. Os resultados obtidos para $\alpha = 5\%$ são exibidos na Tabela 3.6. Foram

⁴Considerou-se a parametrização do tipo casela de referência, sendo o grupo 1 o de referência.

omitidos os resultados do teste escore com valor crítico corrigido S_4^* . Nessa situação, os valores críticos corrigidos assumiram valores negativos, de forma que o teste rejeita a hipótese nula de homoscedasticidade 100% das vezes para qualquer λ . O teste de Goldfeld–Quandt também foi omitido, devido aos problemas que o mesmo apresenta no caso de modelos de análise de variância, conforme mencionado anteriormente.

Na Tabela 3.6, observa-se que na situação considerada, grande parte dos testes apresentaram desempenhos próximos. Foi o caso dos testes baseados no teste da razão de verossimilhanças, do teste escore, e de suas versões com correções tipo–Bartlett e correções no valor crítico, e do teste de Levene (LV). Dentre esses, os testes que se mostraram ligeiramente melhores foram RVB_1 , RVR , testes escore com correções tipo–Bartlett e LV . A Figura 3.3 traz as curvas de poder desses quatro testes⁵. Por outro lado, os testes que se destacaram com poder razoavelmente mais baixo que os demais foram o teste escore com correção por bootstrap (S_{BO}^*), teste escore studentizado e sua versão modificada (SS e SSM) e os testes não–construtivos de White (W_1 e W_2) e de Szroeter (SZ). No caso do teste escore com correção por bootstrap, portanto, há evidências de que a maior complexidade computacional requerida para sua aplicação não reflete em ganho de poder em relação aos demais testes.

Os resultados dessa tabela revelam que os testes, em geral, só apresentam poder maior que 50% para valores de λ superiores a 3,0, isto é, quando a variância do grupo 2 é mais de 3 vezes superior à variância do grupo 1. Além disso, o poder dos testes só excede 80% para valores de λ próximos a 5,0. Também foram obtidos os resultados para $\alpha = 10\%$, os quais serão aqui omitidos. Observou-se, nesse caso, que a maioria dos testes têm poder superior a 50% para valores de λ iguais ou superiores a 2,5. Além disso, o poder supera 80% para valores de λ maiores ou iguais a 4,0, em grande parte dos testes.

Para visualizar melhor o que representam esses graus de heteroscedasticidade, foram construídos gráficos de resíduos externamente studentizados, simulados a partir de valores de λ iguais a 2,0, 3,0, 4,0 e 5,0 (Figura A.1 do Apêndice A). Observa-se que nos casos em que $\lambda \geq 3,0$, os resíduos apresentaram variabilidades com diferenças consideráveis.

Para avaliar a partir de qual valor de n os testes possuem poder razoável no caso em que $\lambda = 2,0$, isto é, no caso em que a variância do grupo 2 é igual ao dobro da variância do grupo 1, foram feitas simulações considerando valores de n entre 20 e 200. Os resultados são exibidos na Tabela 3.7. Observa-se que os testes, em geral, apresentam poder superior a 50% somente para tamanhos amostrais maiores ou iguais a 80, e superior a 80% somente para tamanhos amostrais maiores ou iguais a 140. Também foram obtidos os resultados tomando $\alpha = 10\%$ (aqui omitidos). Obteve-se, nesse caso, que o poder dos testes, em geral, é maior que 50% apenas para tamanhos amostrais maiores ou iguais a 60, e maior que 80% para tamanhos amostrais maiores ou iguais a 120. Em ambos os casos, não se observaram diferenças muito acentuadas entre os valores do poder dos diversos testes, ocorrendo apenas uma ligeira superioridade do teste de Levene (LV) em relação aos demais. Portanto, para tamanhos amostrais pequenos, menores que 80, digamos, é baixa a probabilidade de se detectar heteroscedasticidade no caso em que a variância do segundo grupo é igual ao dobro da do primeiro, sendo que isso vale para todos os testes de heteroscedasticidade considerados neste estudo.

A seguir, são apresentados resultados relativos à aplicação dos testes de heteroscedasticidade

⁵No caso dos testes escore com correções tipo–Bartlett, escolheu-se SB_3 , que apresentou resultados superiores aos demais em algumas situações consideradas nas simulações de tamanho

Tabela 3.6: Poder dos testes no modelo de análise de variância com um fator e dois níveis, com $\alpha = 5\%$, $n = 30$ e erros normais

Estatística	λ												
	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0	5,5	6,0	6,5	7,0
<i>RVM</i>	4,6	10,5	21,9	35,8	48,9	60,0	68,8	76,6	81,8	86,0	89,0	91,2	93,0
<i>RVM</i> ₁	4,6	10,5	21,9	35,8	48,9	60,0	68,8	76,6	81,8	86,0	89,0	91,2	93,0
<i>RVMB</i> ₁	5,0	11,1	23,1	37,3	50,1	61,6	70,0	77,5	82,9	86,7	89,6	91,8	93,5
<i>RVR</i>	5,5	11,9	24,3	38,8	51,9	63,2	71,3	78,7	83,9	87,5	90,4	92,5	94,0
<i>S</i>	4,6	10,5	21,8	35,8	48,8	60,0	68,7	76,5	81,8	85,9	89,0	91,2	93,0
<i>SB</i>	5,0	11,0	22,8	37,1	49,9	61,4	69,9	77,3	82,7	86,6	89,5	91,6	93,4
<i>SB</i> ₁	5,0	11,0	22,8	37,1	49,9	61,4	69,9	77,3	82,7	86,6	89,5	91,6	93,4
<i>SB</i> ₂	5,0	11,1	22,8	37,1	49,9	61,4	69,9	77,3	82,7	86,6	89,5	91,7	93,4
<i>SB</i> ₃	5,0	11,1	22,9	37,1	50,0	61,5	70,0	77,3	82,7	86,6	89,6	91,7	93,5
<i>S*</i>	5,0	11,1	23,0	37,2	50,0	61,5	70,0	77,4	82,8	86,6	89,6	91,7	93,5
<i>S</i> ₁ *	5,0	11,1	22,9	37,1	50,0	61,5	70,0	77,3	82,7	86,6	89,6	91,7	93,5
<i>S</i> ₂ *	5,0	11,1	22,8	37,1	49,9	61,4	69,9	77,3	82,7	86,6	89,5	91,7	93,4
<i>S</i> ₃ *	4,9	11,0	22,7	36,9	49,8	61,3	69,8	77,3	82,6	86,5	89,5	91,6	93,4
<i>S</i> _{BO} *	4,6	9,7	19,7	31,3	42,8	53,0	61,4	68,8	74,1	78,9	82,4	85,0	87,1
<i>SM</i>	4,6	10,5	21,8	35,8	48,8	60,0	68,7	76,5	81,8	85,9	89,0	91,2	93,0
<i>SS</i>	5,2	10,3	19,9	30,8	40,5	49,5	56,7	63,3	68,3	72,5	75,8	77,8	79,9
<i>SSM</i>	5,2	10,3	19,9	30,8	40,5	49,5	56,7	63,3	68,3	72,5	75,8	77,8	79,9
<i>SR</i>	3,9	9,1	19,4	32,7	45,4	56,4	65,7	73,8	79,2	84,0	87,2	89,7	92,0
<i>W</i> ₁	5,2	10,3	19,9	30,8	40,5	49,5	56,7	63,3	68,3	72,5	75,8	77,8	79,9
<i>W</i> ₂	5,2	10,3	19,9	30,8	40,5	49,5	56,7	63,3	68,3	72,5	75,8	77,8	79,9
<i>LV</i>	4,1	14,3	27,6	41,5	53,0	62,6	69,4	76,6	80,7	85,1	88,2	89,5	91,6
<i>SZ</i>	4,8	14,4	26,5	38,1	48,9	58,3	64,4	70,2	75,5	78,7	82,2	84,4	85,8

Nota: $\lambda = \max(\sigma_t^2) / \min(\sigma_t^2)$

Tabela 3.7: Poder dos testes no modelo de análise de variância com um fator e dois níveis, com $\alpha = 5\%$, erros normais e $\lambda = 2,0$

Estatística	n									
	20	40	60	80	100	120	140	160	180	200
<i>RVM</i>	13,6	30,0	44,5	56,3	67,0	75,2	81,7	87,1	90,1	92,9
<i>RVM₁</i>	13,6	30,0	44,5	56,3	67,0	75,2	81,7	87,1	90,1	92,9
<i>RVMB₁</i>	15,0	30,9	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>RVR</i>	16,6	32,0	45,9	57,3	67,8	75,8	82,0	87,3	90,2	93,0
<i>S</i>	13,2	30,0	44,5	56,3	67,0	75,3	81,7	87,1	90,1	92,9
<i>SB</i>	14,6	30,7	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>SB₁</i>	14,6	30,7	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>SB₂</i>	14,7	30,8	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>SB₃</i>	14,7	30,8	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>S*</i>	14,8	30,8	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>S₁*</i>	14,7	30,8	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>S₂*</i>	14,7	30,8	45,1	56,7	67,4	75,5	81,8	87,2	90,1	93,0
<i>S₃*</i>	14,5	30,7	45,0	56,6	67,4	75,5	81,8	87,2	90,1	93,0
<i>S_{BO}*</i>	11,3	28,1	41,9	54,3	64,9	73,4	80,6	86,0	89,6	92,5
<i>SM</i>	13,2	30,0	44,5	56,3	67,0	75,3	81,7	87,1	90,1	92,9
<i>SS</i>	12,7	27,4	41,1	53,3	64,0	72,8	80,3	85,7	89,2	92,2
<i>SSM</i>	12,7	27,4	41,1	53,3	64,0	72,8	80,3	85,7	89,2	92,2
<i>SR</i>	10,0	28,0	42,8	55,1	66,2	74,7	81,3	86,7	89,9	92,8
<i>W₁</i>	12,7	27,4	41,1	53,3	64,0	72,8	80,3	85,7	89,2	92,2
<i>W₂</i>	12,7	27,4	41,1	53,3	64,0	72,8	80,3	85,7	89,2	92,2
<i>LV</i>	20,0	37,1	50,7	61,8	71,2	78,3	84,0	88,9	91,7	94,1
<i>SZ</i>	19,2	33,9	45,9	56,5	65,4	73,2	78,6	83,9	87,4	90,1

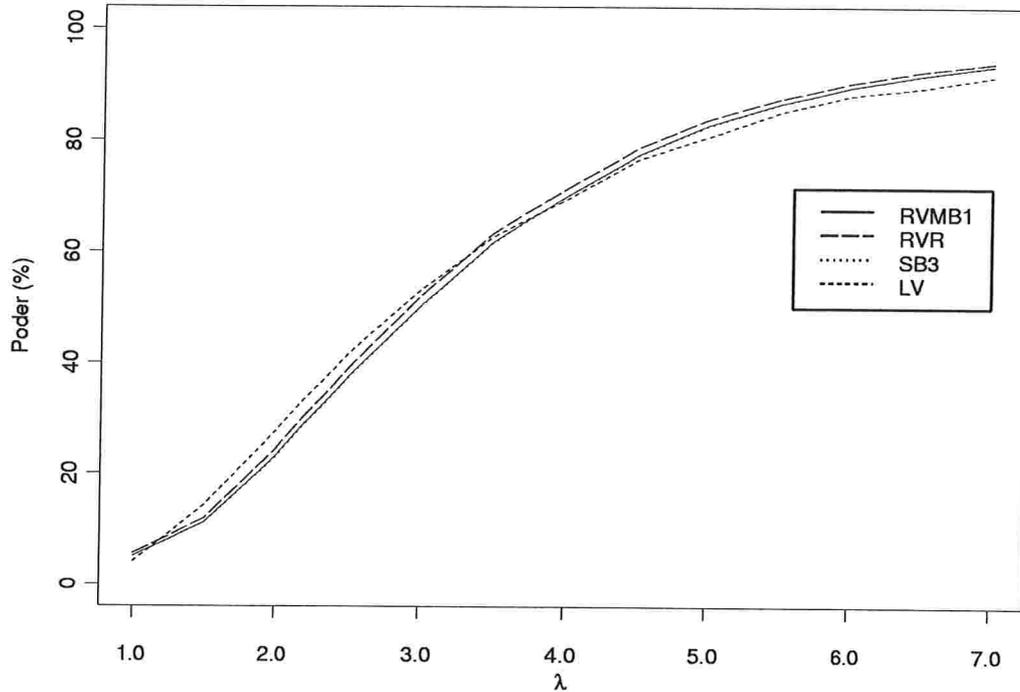


Figura 3.2: Poder dos testes $RVMB_1$, RVR , SB_3 e LV no modelo de análise de variância com 1 fator e 2 grupos, com $\alpha = 5\%$, $n = 30$ e erros normais

na ausência de réplicas, sendo as covariáveis geradas a partir da distribuição $\mathcal{U}(0,1)$. Foram consideradas duas situações. Na primeira, tomou-se $k = 1$, e na segunda, $k = 4$. Em ambas, fixou-se $n = 30$, $p = 5$ e $\alpha = 5\%$. É importante lembrar que na primeira situação os testes não-constructivos que exigem a ordenação das observações segundo suas variâncias podem ser aplicados adequadamente, enquanto que na segunda situação, as observações foram ordenadas somente de acordo com os valores de z_{t1} . Os resultados são exibidos nas Tabelas 3.8 e 3.9.

Observa-se na Tabela 3.8 que, em geral, os testes só apresentam poder superior a 50% para valores de δ iguais ou superiores a 2,0, o que corresponde a valores de λ iguais ou superiores a 7,0. Além disso, o poder em geral supera 80% para valores de δ iguais ou superiores a 3,0, o que corresponde a valores de λ iguais ou superiores a 18,3. Situações típicas de resíduos externamente studentizados gerados sob heteroscedasticidade, para alguns valores de δ considerados, são ilustradas na Figura A.2 (Apêndice A).

Destaca-se, na Tabela 3.8, a superioridade do poder do teste de Szroeter (SZ), para quase todos os valores de δ . Outros testes com poder razoavelmente alto, embora menor que o exibido pelo teste de Szroeter, foram o teste da razão de verossimilhanças modificado com correção de Bartlett ($RVMB_1$), teste da razão de verossimilhanças residual (RVR) e testes escore com

Tabela 3.8: Poder dos testes, com $\alpha = 5\%$, $n = 30$, $p = 5$, $k = 1$, covariáveis uniformes e erros normais

Estatística	δ										
	0,0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
<i>RVM</i>	3,7	6,2	14,3	28,2	46,8	65,9	80,8	90,5	95,8	98,2	99,2
<i>RVM</i> ₁	4,2	6,9	15,3	29,7	48,5	67,5	82,4	91,4	96,3	98,5	99,4
<i>RVMB</i> ₁	4,7	7,7	16,4	31,2	50,4	69,1	83,6	92,1	96,7	98,7	99,5
<i>RVR</i>	5,0	7,9	16,8	31,8	51,0	69,7	83,8	92,2	96,7	98,6	99,4
<i>S</i>	4,3	7,8	17,0	31,6	49,0	66,0	79,3	88,0	93,7	96,6	98,2
<i>SB</i>	4,5	8,2	17,4	32,1	50,1	66,7	79,9	88,4	93,9	96,8	98,3
<i>SB</i> ₁	4,5	8,2	17,4	32,1	50,1	66,7	79,9	88,4	93,9	96,8	98,3
<i>SB</i> ₂	4,5	8,2	17,4	32,2	50,1	66,7	79,9	88,4	93,9	96,8	98,3
<i>SB</i> ₃	4,5	8,2	17,4	32,2	50,1	66,7	79,9	88,4	93,9	96,8	98,3
<i>S*</i>	4,5	8,2	17,4	32,2	50,2	66,8	79,9	88,4	93,9	96,8	98,3
<i>S*</i> ₁	4,5	8,2	17,4	32,2	50,2	66,8	79,9	88,4	93,9	96,8	98,3
<i>S*</i> ₂	4,5	8,2	17,4	32,2	50,1	66,7	79,9	88,4	93,9	96,8	98,3
<i>S*</i> ₃	4,5	8,2	17,4	32,1	50,1	66,6	79,9	88,4	93,9	96,8	98,3
<i>S*</i> ₄	4,5	8,2	17,4	32,2	50,1	66,7	79,9	88,4	93,9	96,8	98,3
<i>S*</i> _{BO}	4,8	8,3	17,2	30,7	47,2	63,6	76,1	85,4	91,5	94,8	96,9
<i>SM</i>	4,4	7,3	15,9	29,7	47,1	64,1	77,9	87,1	93,1	96,2	97,9
<i>SS</i>	5,4	8,3	16,7	29,1	43,3	58,3	70,0	79,1	85,2	88,6	91,2
<i>SSM</i>	5,5	7,8	15,4	27,2	40,5	55,1	66,7	76,1	82,0	85,8	88,4
<i>SR</i>	3,9	6,7	14,8	28,1	45,3	62,6	76,4	86,2	92,5	95,7	97,7
<i>W</i> ₁	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<i>W</i> ₂	1,4	1,4	1,7	2,1	2,7	3,2	4,0	5,1	5,7	6,4	7,1
<i>GQ</i>	5,0	10,2	18,6	30,4	45,4	60,7	74,2	84,6	91,4	95,3	97,7
<i>LV</i>	5,0	10,4	18,7	29,6	43,0	56,5	68,7	78,0	84,5	89,3	92,4
<i>SZ</i>	5,3	13,4	26,4	43,3	61,8	76,6	86,6	93,0	96,4	98,2	99,0

Nota: Entre parênteses é dado $\lambda = \max(\sigma_t^2)/\min(\sigma_t^2)$

correção tipo-Bartlett. A Figura 3.3 traz as curvas de poder desses quatro testes. Destacaram-se, devido ao baixo poder, os testes robustos à não-normalidade dos erros, identificados no estudo sobre os tamanhos dos testes (SS , SSM , W_1 , W_2 e LV).

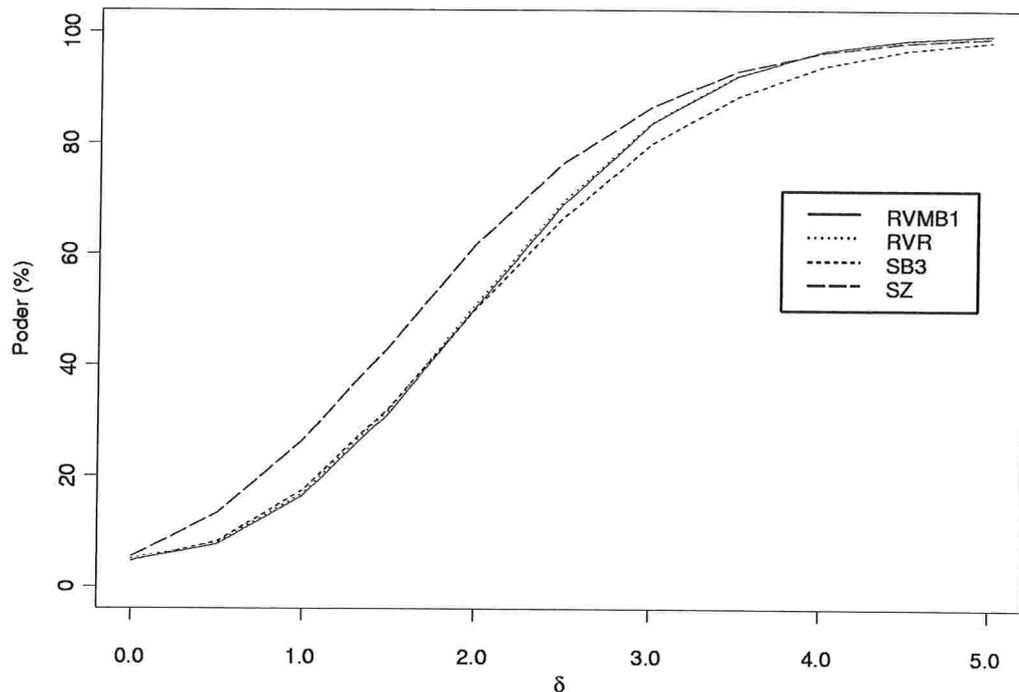


Figura 3.3: Poder dos testes $RVMB_1$, RVR , SB_3 e SZ , com $n = 30$, $p = 5$, $k = 1$, covariáveis uniformes e erros normais

Na Tabela 3.9, por outro lado, observa-se que sob ordenação incorreta das observações, o teste de Szroeter (SZ) não manteve os resultados favoráveis observados no caso em que $k = 1$. Neste caso, mereceram destaque o teste da razão de verossimilhanças modificado (RVM_1), sua versão com correção de Bartlett ($RVMB_1$) e o teste da razão de verossimilhanças residual (RVR), tendo apresentado valores de poder superiores aos demais testes. O teste escore e todas as suas versões corrigidas e modificadas apresentaram resultados consideravelmente inferiores aos anteriores. Dentre os testes não-constructivos, o que em geral apresentou maior poder foi o de Goldfeld-Quandt (GQ). A Figura 3.4 traz as curvas de poder dos testes RVM_1 , $RVMB_1$, RVR e GQ . Assim como na situação anterior, com $k = 1$, observou-se que os testes robustos à não-normalidade dos erros apresentaram poder bem mais baixo que os demais. Nessa tabela, observa-se que o poder dos testes mencionados como mais poderosos, apenas excede 50% para valores de $\delta = \delta_1 = \dots = \delta_4$ superiores a 2,0, o que corresponde a λ maior ou igual a 67,5. Portanto, mais uma vez houve evidências de que todos os testes têm poder baixo, até mesmo

para níveis razoavelmente altos de heteroscedasticidade, para o tamanho amostral considerado ($n = 30$).

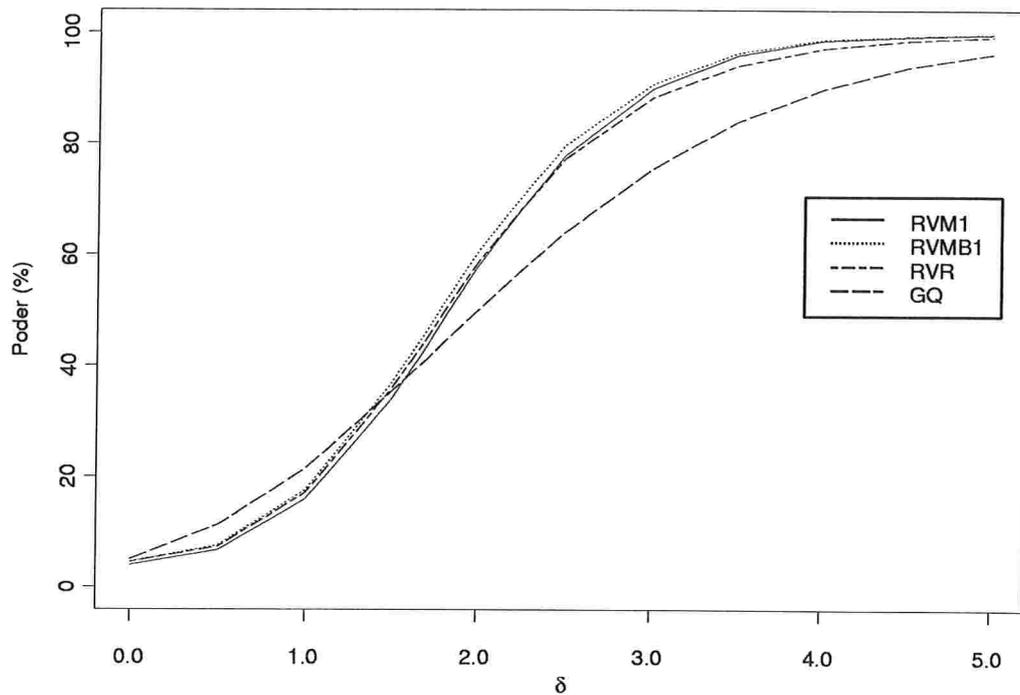


Figura 3.4: Poder dos testes RVM_1 , $RVMB_1$, RVR e GQ , com $n = 30$, $p = 5$, $k = 4$, covariáveis uniformes e erros normais

A seguir, será avaliado o poder dos testes sob uma das condições que mais afetaram os tamanhos dos mesmos, conforme observado na seção anterior: a presença de pontos de alavanca. Para tanto, considerou-se o caso em que as covariáveis têm distribuição t_3 . Fixou-se $n = 30$, $p = 5$, $k = 1$ e erros normais. Os valores do poder para δ variando entre 0,0 e 0,4 são exibidos na Tabela 3.10. Observou-se que, dentre os testes considerados, o que apresentou maior poder para a maioria dos valores de δ foi o teste da razão de verossimilhanças modificado com correção de Bartlett, seguido pelos testes da razão de verossimilhanças modificado (RVM_1) e pelo teste da razão de verossimilhanças residual (RVR). O teste escore e suas versões corrigidas e modificadas apresentaram valores de poder bem inferiores aos primeiros, com exceção do teste escore residual (SR), para o qual a inferioridade não é tão acentuada. Os testes não-construtivos também apresentaram valores baixos de poder. É possível observar que o poder dos testes, em geral, não ultrapassa 20% para valores de δ inferiores a 0,12, o que corresponde a um valor de λ igual a 35,5. Essa medida, no entanto, deve ser vista com uma certa cautela, uma vez que foi influenciada por dois pontos com alavanca muito alta existentes na particular matriz Z simulada. Para avaliar melhor o grau

Tabela 3.9: Poder dos testes, com $\alpha = 5\%$, $n = 30$, $p = 5$, $k = 4$, covariáveis uniformes e erros normais

Estatística	$\delta = \delta_1 = \delta_2 = \delta_3 = \delta_4$										
	0,0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
	(1,0)	(2,9)	(8,2)	(23,6)	(67,5)	(193,6)	(555,1)	(1.591,2)	(4.561,5)	(13.077,0)	(37.487,0)
<i>RVM</i>	2,8	4,7	12,3	28,1	50,4	70,4	83,8	91,2	95,3	97,3	98,6
<i>RVM</i> ₁	4,0	6,7	15,8	33,8	57,9	77,8	89,8	95,9	98,6	99,3	99,7
<i>RVMB</i> ₁	4,5	7,6	17,5	36,6	60,5	79,7	90,7	96,3	98,8	99,3	99,7
<i>RVR</i>	4,5	7,3	17,0	35,7	58,6	77,2	88,2	94,0	97,1	98,5	99,2
<i>S</i>	4,3	6,0	13,3	25,8	41,2	57,2	69,9	79,0	86,4	90,8	93,8
<i>SB</i>	4,3	6,0	13,3	25,7	41,1	56,9	69,4	77,9	83,9	86,5	86,6
<i>SB</i> ₁	4,3	6,0	13,3	25,7	41,1	57,0	69,8	78,9	86,2	90,8	93,8
<i>SB</i> ₂	4,3	6,0	13,3	25,7	41,2	57,0	69,8	78,9	86,2	90,8	93,8
<i>SB</i> ₃	4,3	6,0	13,3	25,8	41,2	57,1	69,9	79,0	86,3	90,8	93,8
<i>S*</i>	4,3	6,0	13,3	25,7	41,1	57,0	69,8	78,9	86,2	90,8	93,8
<i>S*</i> ₁	4,2	6,0	13,3	25,7	41,1	57,0	69,8	78,9	86,2	90,7	93,8
<i>S*</i> ₂	4,2	6,0	13,3	25,7	41,1	56,9	69,7	78,9	86,1	90,7	93,8
<i>S*</i> ₃	4,3	6,0	13,3	25,7	41,1	57,0	69,8	78,9	86,2	90,8	93,8
<i>S*</i> ₄	4,3	6,0	13,3	25,8	41,2	57,1	69,9	79,0	86,3	90,8	93,8
<i>S*BO</i>	3,6	4,4	8,3	15,6	25,3	35,2	43,6	50,5	55,5	58,9	60,6
<i>SM</i>	4,3	5,7	12,4	24,0	38,7	54,5	67,3	76,5	84,5	89,8	93,1
<i>SS</i>	4,4	4,7	7,4	10,9	14,4	17,0	18,3	18,1	16,7	15,2	13,3
<i>SSM</i>	4,4	4,3	6,3	9,1	11,7	13,4	14,5	13,7	12,6	11,0	9,6
<i>SR</i>	4,0	5,0	11,4	22,1	36,2	51,6	64,8	74,6	82,7	88,4	92,0
<i>W</i> ₁	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<i>W</i> ₂	1,4	1,7	2,2	2,8	3,7	5,1	6,7	8,2	9,8	10,9	11,6
<i>GQ</i>	5,0	11,3	21,3	35,2	49,9	63,9	75,4	83,9	89,8	93,8	96,2
<i>LV</i>	5,0	10,9	18,2	26,6	34,8	41,4	46,3	49,3	50,8	50,9	49,9
<i>SZ</i>	5,3	12,4	23,2	35,6	48,5	59,8	69,3	77,0	83,1	87,7	90,9

Nota: Entre parênteses é dado $\lambda = \max(\sigma_t^2)/\min(\sigma_t^2)$

de heteroscedasticidade representado pelos valores de δ considerados, foram construídos gráficos de resíduos externamente studentizados, gerados a partir de dados heteroscedásticos (Figura A.3). Esses gráficos ilustram a dificuldade de se diagnosticar a heteroscedasticidade caso ocorram pontos com alavancas altas, uma vez que a maioria dos resíduos encontram-se distribuídos em uma faixa razoavelmente horizontal, com apenas os resíduos relativos aos pontos de alavanca fugindo a esse padrão.

Por último, avaliamos o poder dos testes robustos à não-normalidade (SS , SSM , W_1 , W_2 e LV), na situação em que os erros têm distribuição t_{10} . Os resultados são exibidos na Tabela 3.11. Fixou-se $n = 50$, $p = 5$, $\alpha = 5\%$, as covariáveis foram geradas a partir da distribuição $\mathcal{U}(0, 1)$, considerando-se $k = 1$ e $k = 4$. Observa-se que no caso em que $k = 1$, o teste de Levene apresentou poder mais alto para todos os valores de δ considerados. O teste escore studentizado (SS) apresentou resultados ligeiramente inferiores, e sua versão modificada (SSM) produziu resultados piores que a versão original. Os testes de White apresentaram poder muito mais baixo que os demais testes. No caso em que $k = 4$, o teste de Levene não pôde ter as observações ordenadas corretamente, conforme mencionado anteriormente. Apesar desta limitação, esse teste manteve seu desempenho superior aos demais. É interessante observar que nesse caso, só foram obtidos valores do poder superiores a 50% para valores de δ superiores a 2,0, o que corresponde a valores de λ iguais ou superiores a 85,2.

3.4 Impacto da heteroscedasticidade sobre os testes t usuais envolvendo β

Um dos principais objetivos na aplicação de modelos de regressão é fazer inferências a respeito do vetor paramétrico β . Conforme discutido anteriormente, inferências usuais sobre tal vetor paramétrico baseadas no método de MQO podem ser pouco precisas na presença de heteroscedasticidade dos erros, o que leva a concluir pela importância da detecção da mesma. Uma pergunta a ser feita neste ponto é a seguinte: os testes de heteroscedasticidade apresentados possuem um poder razoável em situações nas quais a heteroscedasticidade ocorre em grau capaz de prejudicar consideravelmente as inferências baseadas em MQO? Essa seção apresenta alguns resultados que visam a responder tal questão.

Recordemo-nos da situação ilustrada na Tabela 3.8, em que os dados foram gerados com heteroscedasticidade multiplicativa, sendo $n = 30$, $p = 5$, $k = 1$, covariáveis uniformes e erros normais. Conforme observado anteriormente, o poder da maioria dos testes é próximo a 50% quando $\delta = 2,0$, e só se aproxima de 80% para $\delta = 3,0$. Considere que haja interesse em testar a hipótese nula $H_0 : \beta_1 = 0$ contra a alternativa $H_1 : \beta_1 \neq 0$. A estatística do teste t usual para tal hipótese é dada por

$$t = \frac{\widehat{\beta}_1}{\sqrt{s^2(\widehat{\beta}_1)}},$$

em que $\widehat{\beta}_1$ é o estimador de MQO de β , correspondendo ao segundo elemento em (2.1), e $s^2(\widehat{\beta}_1)$ é a estimativa da variância de $\widehat{\beta}_1$ sob homoscedasticidade, dada pelo elemento (2, 2) da matriz (1.2). Sob H_0 , t tem distribuição t -Student com $n - p = 25$ graus de liberdade. A Tabela

Tabela 3.10: Poder dos testes, com $\alpha = 5\%$, $n = 30$, $p = 5$, $k = 1$, covariáveis t_3 e erros normais

Estatística	δ												
	0,00	0,02	0,04	0,06	0,08	0,10	0,12	0,14	0,16	0,18	0,20	0,30	0,40
	(1,0)	(1,8)	(3,3)	(6,0)	(10,8)	(19,6)	(35,5)	(64,4)	(116,8)	(211,8)	(384,0)	(7,523,7)	(147,430,0)
<i>RVM</i>	1,2	1,3	1,7	2,9	4,6	8,3	12,1	19,1	25,9	34,1	42,6	76,9	92,1
<i>RVM</i> ₁	2,8	2,6	3,1	4,7	6,8	11,0	15,7	23,3	30,1	38,6	47,5	80,1	93,9
<i>RVMB</i> ₁	3,8	3,7	4,2	6,2	8,8	13,5	18,2	26,3	33,8	42,6	50,7	82,1	94,7
<i>RVR</i>	2,0	1,9	2,6	4,1	6,1	10,3	14,6	22,4	29,2	37,7	46,2	79,1	92,9
<i>S</i>	2,0	1,5	1,7	2,0	2,9	4,1	5,8	8,4	11,0	15,2	19,0	46,1	70,1
<i>SB</i>	3,3	2,9	2,7	3,2	4,3	5,6	6,9	9,4	12,1	14,7	17,3	27,6	30,0
<i>SB</i> ₁	3,5	3,0	3,0	3,4	4,6	6,1	7,9	11,2	14,5	18,7	22,8	47,0	65,3
<i>SB</i> ₂	4,2	3,6	3,5	4,1	5,5	7,3	9,6	13,4	17,2	22,1	26,9	55,8	78,7
<i>SB</i> ₃	4,8	4,1	3,9	4,6	6,3	8,2	10,6	14,7	18,6	23,5	28,5	57,9	80,2
<i>S*</i>	2,9	2,2	2,2	2,7	3,9	5,3	7,2	10,4	13,4	18,2	22,5	50,2	73,8
<i>S</i> ₁ *	2,6	2,0	2,1	2,5	3,6	4,9	6,8	9,9	12,7	17,5	21,5	48,9	72,9
<i>S</i> ₂ *	2,2	1,6	1,9	2,2	3,2	4,4	6,3	9,0	11,9	16,1	20,2	47,5	71,4
<i>S</i> ₃ *	2,5	1,9	2,0	2,4	3,5	4,7	6,6	9,6	12,5	17,1	21,0	48,6	72,5
<i>S</i> ₄ *	4,8	4,1	3,9	4,6	6,3	8,2	10,6	14,7	18,6	23,5	28,5	57,9	80,2
<i>S</i> ^{*BO}	5,4	4,8	4,5	5,3	6,6	8,7	11,0	14,6	17,9	22,1	25,4	40,3	42,2
<i>SM</i>	1,9	1,5	1,8	2,4	3,4	5,1	7,1	10,3	13,5	18,4	22,8	50,6	74,3
<i>SS</i>	2,3	1,9	1,6	2,1	2,7	3,9	5,1	6,6	8,4	10,8	11,9	17,1	13,2
<i>SSM</i>	2,0	2,0	1,8	2,5	3,6	5,2	6,5	8,9	11,3	14,6	16,6	26,1	24,1
<i>SR</i>	5,4	5,5	6,4	8,6	11,6	15,5	19,3	25,1	31,0	37,0	42,4	70,0	86,8
<i>W</i> ₁	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,0	0,1	0,1	0,2	0,2
<i>W</i> ₂	1,4	1,4	1,3	1,3	1,3	2,2	3,0	4,6	7,0	10,7	15,8	47,2	71,3
<i>GQ</i>	5,3	5,7	5,9	7,4	8,7	8,9	10,8	12,4	13,7	16,8	19,6	40,1	67,4
<i>LV</i>	5,5	6,4	7,0	8,2	9,3	9,6	10,8	11,1	11,9	12,2	11,3	9,7	5,8
<i>SZ</i>	4,9	6,5	7,8	10,3	12,6	14,5	17,3	19,7	23,0	26,1	27,6	35,1	33,9

Nota: Entre parênteses é dado $\lambda = \max(\sigma_t^2) / \min(\sigma_t^2)$

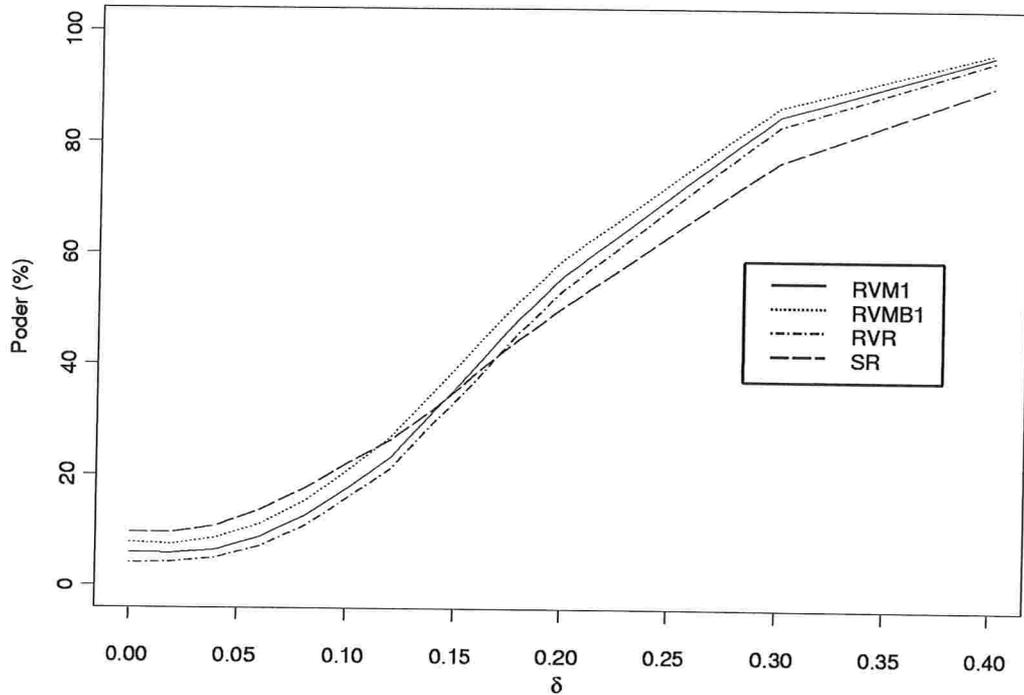


Figura 3.5: Poder dos testes RVM_1 , $RVMB_1$, RVR e SR , com $n = 30$, $p = 5$, $k = 1$, covariáveis t_3 e erros normais

3.12 traz os tamanhos observados desse teste sob homoscedasticidade e sob diversos graus de heteroscedasticidade dos erros, considerando-se níveis de significância nominais de 10%, 5%, 1% e 0,5%.

Observa-se que sob heteroscedasticidade o teste é anticonservativo. Considerando um nível de significância de 5%, tem-se que o tamanho do teste é igual a 9,3% e 10,9%, para $\delta = 2,0$ e $\delta = 2,5$, respectivamente. Nesses casos, o tamanho observado é aproximadamente o dobro do nominal. Esse exemplo simples mostra que há casos em que os testes de heteroscedasticidade tratados nesse trabalho não são altamente poderosos contra alternativas que prejudicam substancialmente as inferências usuais sobre β . Logo, conclui-se que em pequenas amostras os resultados dos testes devem ser vistos com cuidado, e complementados por outros métodos de diagnóstico, como os gráficos de resíduos, por exemplo.

Tabela 3.11: Poder dos testes, com $\alpha = 5\%$, $n = 50$, $p = 5$, covariáveis uniformes e erros t_{10}

k	Estat.	δ										
		0,0	0,5	1,0	1,5	2,0	2,5	3,0	3,5	4,0	4,5	5,0
1	SS	(1,0)	(1,6)	(2,6)	(4,3)	(7,0)	(11,3)	(18,3)	(29,8)	(48,4)	(78,6)	(127,6)
	SSM	5,3	9,3	21,0	39,7	59,5	76,0	86,4	93,0	95,3	96,8	97,9
	W_1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1
	W_2	5,0	4,6	5,4	7,0	9,5	13,2	17,5	21,3	24,9	29,4	33,7
	LV	5,2	13,5	26,3	43,5	61,8	77,1	87,8	94,3	97,0	98,6	99,5
	4	SS	(1,0)	(3,0)	(9,2)	(28,1)	(85,2)	(259,0)	(787,0)	(2.391,0)	(7.265,1)	(22.075,0)
SSM		4,3	7,0	16,5	32,6	48,9	58,1	63,6	65,1	63,8	62,3	59,9
W_1		0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0
W_2		5,0	5,2	7,7	13,1	21,4	28,2	36,3	42,0	46,0	48,6	51,4
LV		5,7	13,9	25,1	39,8	52,9	61,6	66,9	69,2	68,8	66,2	62,8

Nota: Entre parênteses é dado $\lambda = \max(\sigma_t^2)/\min(\sigma_t^2)$

3.5 Robustez dos testes à má especificação da função cedástica

A escolha da função cedástica $w_t = w(z_t, \delta)$, em geral, não é uma tarefa simples. Em grande parte dos casos, não há réplicas das covariáveis em X , de forma que não é possível verificar se a particular função cedástica assumida é compatível com as variâncias observadas nos grupos determinados pelas réplicas. Assim, faz-se necessário avaliar a robustez dos testes de heteroscedasticidade à escolha inadequada de tal função. Nessa seção são apresentados alguns resultados obtidos com essa finalidade. Foram considerados três possíveis tipos de inadequação da função cedástica: escolha incorreta da função, porém com as covariáveis corretas em Z , escolha correta da função, porém com omissão de covariáveis em Z e escolha correta da função, porém com escolha incorreta das covariáveis em Z .

No primeiro caso, foram gerados dados a partir do modelo aditivo dado em (1.5), com $n = 50$, $p = 3$, $k = 1$, covariáveis uniformes e erros normais. A Tabela 3.12 traz o poder dos testes para alguns valores de δ , assumindo o modelo correto e assumindo o modelo multiplicativo. É importante lembrar que os testes não-constructivos não são afetados pela escolha da função cedástica, o mesmo ocorrendo com o teste score e suas versões modificadas no caso em que as variâncias dos erros são da forma (2.5). Portanto, os únicos testes que poderiam ter seus desempenhos afetados pela escolha incorreta são o TRV e suas versões corrigidas e modificadas. Observa-se que, no caso desses testes, houve uma redução muito pequena do poder ao assumir

Tabela 3.12: Tamanho do teste t de $H_0 : \beta_1 = 0$ contra $H_1 : \beta_1 \neq 0$, sob homoscedasticidade e diversos graus de heteroscedasticidade, assumindo o modelo multiplicativo com $n = 30$, $p = 5$, $k = 1$ e $\sigma_t^2 = \exp(\delta x_{t1})$

δ	10%	5%	1%	0,5%
0,0	9,7	4,9	0,9	0,4
0,5	10,6	5,5	1,3	0,6
1,0	13,3	6,8	1,6	0,9
1,5	14,0	7,7	1,9	1,0
2,0	16,3	9,3	2,7	1,6
2,5	18,7	10,9	3,2	1,9
3,0	19,5	12,1	3,8	2,1
3,5	21,8	14,0	4,9	3,0
4,0	25,2	16,6	5,9	3,7
4,5	25,6	17,4	6,4	4,3
5,0	27,0	18,3	7,4	4,8
5,5	28,6	19,6	8,2	5,4
6,0	29,7	21,1	9,1	6,2

o modelo incorreto, sugerindo que a escolha correta da função cedástica não é um ponto crítico para o desempenho dos testes.

No segundo caso, foram geradas observações a partir do modelo multiplicativo com $n = 50$, $p = 3$, $k = 2$, erros normais e covariáveis uniformes, sendo $\sigma_t^2 = \exp(\delta_1 z_{t1} + \delta_2 z_{t1}^2)$. A Tabela 3.13 traz o poder dos testes assumindo o modelo correto, e assumindo o modelo $\sigma^2 = \exp(\delta_1 z_{t1})$, em que ocorre a omissão de uma covariável em Z . Os resultados revelam que, curiosamente, o poder dos testes tende a ser maior no caso em que se omite uma covariável. Portanto, esse exemplo simples não trouxe evidências de que a omissão de covariáveis em Z possa ser prejudicial ao desempenho dos testes de heteroscedasticidade.

Por último, foram gerados dados a partir do modelo multiplicativo com $n = 50$, $p = 3$, $k = 1$, erros normais e covariáveis uniformes, com $\sigma_t^2 = \exp(\delta \log z_{t1})$. Na Tabela 3.14 são apresentados os resultados obtidos ao assumir o modelo correto, e ao assumir o modelo dado por $\sigma^2 = \exp(\delta z_{t1})$, em que a covariável em Z é mal especificada. Nesse caso, ao contrário dos demais, observa-se que ocorre, sobretudo no caso do TRV e de suas versões corrigidas e modificadas, uma razoável diminuição do poder sob a má especificação do modelo. No caso do teste score e de suas versões modificadas, observa-se que a redução ocorre, em geral, apenas para os níveis mais altos de heteroscedasticidade.

Logo, dentre os tipos de má especificação considerados, o que se mostrou mais prejudicial ao desempenho dos testes foi a escolha inadequada da covariável em Z .

Tabela 3.13: Poder dos testes para dados gerados sob o modelo aditivo, com $\sigma_t^2 = (1 + \delta z_{t1})^2$, $n = 50$, $p = 3$, $k = 1$, erros normais e covariáveis uniformes, assumindo-se o modelo correto e o modelo multiplicativo com $\sigma_t^2 = \exp(\delta z_{t1})$

Estatística	Modelo aditivo (correto)				Modelo multiplicativo (incorreto)			
	$\delta = 0,0$	$\delta = 1,0$	$\delta = 2,0$	$\delta = 3,0$	$\delta = 0,0$	$\delta = 1,0$	$\delta = 2,0$	$\delta = 3,0$
<i>RV</i>	6,6	53,7	88,1	96,5	6,3	52,1	86,5	95,3
<i>RVB</i>	5,3	50,0	86,0	95,8	4,9	48,2	84,3	94,3
<i>RVM</i>	4,6	48,3	85,0	95,3	4,4	46,9	83,3	93,8
<i>RVM₁</i>	4,6	48,4	85,1	95,3	4,4	46,9	83,3	93,8
<i>RVMB₁</i>	4,9	49,4	85,5	95,5	4,7	47,9	84,0	94,2
<i>RVC</i>	5,6	30,7	70,3	87,4	5,7	33,1	71,5	87,7
<i>RVR</i>	5,1	50,0	85,9	95,7	4,9	48,6	84,5	94,3
<i>S</i>	4,5	46,4	80,4	91,5	4,5	46,4	80,4	91,5
<i>SB</i>	4,7	47,3	81,1	91,9	4,7	47,3	81,1	91,9
<i>SB₁</i>	4,7	47,3	81,1	91,9	4,7	47,3	81,1	91,9
<i>SB₂</i>	4,7	47,3	81,1	91,9	4,7	47,3	81,1	91,9
<i>SB₃</i>	4,7	47,3	81,1	91,9	4,7	47,3	81,1	91,9
<i>S*</i>	4,7	47,4	81,2	92,0	4,7	47,4	81,2	92,0
<i>S₁*</i>	4,7	47,4	81,2	92,0	4,7	47,4	81,2	92,0
<i>S₂*</i>	4,7	47,3	81,2	92,0	4,7	47,3	81,2	92,0
<i>S₃*</i>	4,7	47,3	81,1	91,9	4,7	47,3	81,1	91,9
<i>S₄*</i>	4,7	47,3	81,1	91,9	4,7	47,3	81,1	91,9
<i>S_{BO}*</i>	5,6	45,4	77,6	89,0	5,6	45,4	77,6	89,0
<i>SM</i>	4,5	46,8	80,7	91,7	4,5	46,8	80,7	91,7
<i>SS</i>	4,9	44,1	75,4	87,4	4,9	44,1	75,4	87,4
<i>SSM</i>	4,9	44,6	75,9	87,7	4,9	44,6	75,9	87,7
<i>SR</i>	4,2	46,0	80,2	91,3	4,2	46,0	80,2	91,3
<i>W₁</i>	1,4	4,9	13,0	19,7	1,4	4,9	13,0	19,7
<i>W₂</i>	4,2	17,7	35,5	47,5	4,2	17,7	35,5	47,5
<i>GQ</i>	5,2	54,1	85,6	94,9	5,2	54,1	85,6	94,9
<i>LV</i>	4,2	45,1	75,1	88,1	4,2	45,1	75,1	88,1
<i>SZ</i>	4,6	58,8	88,6	95,9	4,6	58,8	88,6	95,9
<i>RS</i>	5,4	37,2	73,4	89,1	5,4	37,2	73,4	89,1

Tabela 3.14: Poder dos testes para dados gerados sob o modelo multiplicativo, com $\sigma_t^2 = \exp(\delta_1 z_{t1} + \delta_2 z_{t1}^2)$, $n = 50$, $p = 3$, $k = 2$, erros normais e covariáveis uniformes, assumindo-se o modelo correto e o modelo $\sigma_t^2 = \exp(\delta_1 z_{t1})$

Estatística	Modelo correto				Modelo incorreto			
	$\delta = 0,0$	$\delta = 0,5$	$\delta = 1,0$	$\delta = 1,5$	$\delta = 0,0$	$\delta = 0,5$	$\delta = 1,0$	$\delta = 1,5$
<i>RV</i>	7,4	26,8	76,5	97,8	6,3	32,8	82,9	98,7
<i>RVB</i>	4,9	22,3	71,7	96,7	4,9	29,5	80,6	98,4
<i>RVM</i>	4,2	21,5	71,1	96,6	4,4	28,4	79,8	98,3
<i>RVM₁</i>	4,2	21,6	71,3	96,6	4,4	28,4	79,8	98,3
<i>RVMB₁</i>	4,4	22,2	71,8	96,8	4,7	29,3	80,5	98,4
<i>RVC</i>	5,1	14,0	58,4	92,2	5,7	17,9	68,5	96,2
<i>RVR</i>	4,8	23,1	72,8	97,0	4,9	30,0	81,0	98,4
<i>S</i>	4,0	22,7	70,6	95,7	4,5	30,2	80,5	97,9
<i>SB</i>	4,6	24,1	66,0	63,2	4,7	31,1	81,0	98,0
<i>SB₁</i>	4,6	24,5	72,7	96,2	4,7	31,1	81,0	98,0
<i>SB₂</i>	4,6	24,5	72,7	96,2	4,7	31,1	81,0	98,0
<i>SB₃</i>	4,6	24,5	72,8	96,2	4,7	31,1	81,0	98,0
<i>S*</i>	4,6	24,6	72,8	96,2	4,7	31,1	81,0	98,0
<i>S₁*</i>	4,6	24,5	72,8	96,2	4,7	31,1	81,0	98,0
<i>S₂*</i>	4,6	24,5	72,7	96,2	4,7	31,1	81,0	98,0
<i>S₃*</i>	4,5	24,0	72,3	96,1	4,7	31,1	81,0	98,0
<i>S₄*</i>	4,6	24,5	72,8	96,2	4,7	31,1	81,0	98,0
<i>S_{BO}*</i>	4,9	22,0	66,5	92,9	5,6	29,1	77,9	97,0
<i>SM</i>	4,0	23,2	71,2	95,7	4,5	30,6	80,8	97,9
<i>SS</i>	4,2	20,9	62,3	89,2	4,9	28,6	75,7	96,1
<i>SSM</i>	4,3	21,4	63,3	89,9	4,9	29,1	76,1	96,3
<i>SR</i>	4,0	22,7	70,5	95,6	4,2	29,9	80,3	97,9
<i>W₁</i>	1,4	2,8	7,9	15,2	1,4	2,8	7,9	15,2
<i>W₂</i>	4,2	12,5	40,2	69,6	4,2	12,5	40,2	69,6
<i>GQ</i>	5,2	34,0	81,0	98,1	5,2	34,0	81,0	98,1
<i>LV</i>	4,2	28,4	69,6	94,4	4,2	28,4	69,6	94,4
<i>SZ</i>	4,6	40,1	86,9	99,1	4,6	40,1	86,9	99,1
<i>RS</i>	5,4	21,6	69,3	95,7	5,4	21,6	69,3	95,7

Nota: $\delta_1 = \delta_2 = \delta$

Tabela 3.15: Poder dos testes para dados gerados sob o modelo multiplicativo, com $\sigma_t^2 = \exp(\delta \log z_{t1})$, $n = 50$, $p = 3$, $k = 1$, erros normais e covariáveis uniformes, assumindo-se o modelo correto e o modelo $\sigma_t^2 = \exp(\delta z_{t1})$

Estatística	Modelo correto				Modelo incorreto			
	$\delta = 0,0$	$\delta = 0,3$	$\delta = 0,6$	$\delta = 0,9$	$\delta = 0,0$	$\delta = 0,3$	$\delta = 0,6$	$\delta = 0,9$
<i>RV</i>	6,6	40,7	91,0	99,7	6,3	27,5	70,3	93,2
<i>RVB</i>	5,1	35,7	88,5	99,5	4,9	24,4	66,9	91,8
<i>RVM</i>	4,5	31,4	85,4	99,1	4,4	23,3	65,4	91,1
<i>RVM₁</i>	4,6	31,8	85,8	99,2	4,4	23,4	65,4	91,2
<i>RVMB₁</i>	4,8	32,6	86,5	99,3	4,7	24,2	66,5	91,6
<i>RVC</i>	5,7	43,0	92,1	99,7	5,7	13,3	49,0	82,2
<i>RVR</i>	5,0	33,3	86,9	99,3	4,9	24,7	67,1	91,8
<i>S</i>	4,0	15,1	60,7	91,2	4,5	21,9	59,2	85,8
<i>SB</i>	4,4	18,7	66,0	89,8	4,7	22,7	60,1	86,4
<i>SB₁</i>	4,8	18,7	66,2	93,4	4,7	22,7	60,1	86,4
<i>SB₂</i>	4,8	18,8	66,4	93,4	4,7	22,7	60,1	86,4
<i>SB₃</i>	4,8	18,9	66,5	93,4	4,7	22,7	60,1	86,4
<i>S*</i>	4,8	18,6	66,1	93,3	4,7	22,7	60,1	86,5
<i>S₁*</i>	4,7	18,5	66,1	93,3	4,7	22,7	60,1	86,5
<i>S₂*</i>	4,7	18,5	66,0	93,2	4,7	22,7	60,1	86,4
<i>S₃*</i>	4,6	17,6	64,4	92,8	4,7	22,7	60,1	86,4
<i>S₄*</i>	4,8	18,9	66,5	93,4	4,7	22,7	60,1	86,4
<i>S_{BO}*</i>	5,4	16,5	57,4	84,6	5,6	22,2	56,5	81,9
<i>SM</i>	4,1	14,4	59,4	90,6	4,5	22,3	59,5	86,0
<i>SS</i>	4,6	13,6	50,7	77,5	4,9	21,0	55,0	80,0
<i>SSM</i>	4,7	12,8	48,9	75,7	4,9	21,3	55,5	80,4
<i>SR</i>	4,1	14,4	59,4	90,6	4,2	21,7	58,9	85,6
<i>W₁</i>	1,4	3,2	10,9	22,3	1,4	3,2	10,9	22,3
<i>W₂</i>	4,2	8,1	21,7	38,5	4,2	8,1	21,7	38,5
<i>GQ</i>	5,2	31,4	70,9	92,4	5,2	31,4	70,9	92,4
<i>LV</i>	4,2	25,0	59,1	83,8	4,2	25,0	59,1	83,8
<i>SZ</i>	4,6	33,5	73,5	93,3	4,6	33,5	73,5	93,3
<i>RS</i>	5,4	19,8	57,2	85,8	5,4	19,8	57,2	85,8

3.6 Conclusões

As simulações feitas conduziram a diversas conclusões importantes a respeito do comportamento, em pequenas amostras, dos testes de heteroscedasticidade em estudo.

Em primeiro lugar, o estudo dos tamanhos dos testes levou a concluir que os testes da razão de verossimilhanças (RV), sua versão com correção de Bartlett (RVB), sua versão que utiliza a função de verossimilhança condicional (RVC) e o teste de Ramsey (RS) são, em geral, anti-conservativos, conduzindo à rejeição da hipótese de homoscedasticidade erroneamente com uma probabilidade maior que o nível de significância nominal dos testes. Logo, desaconselha-se o uso de tais testes em aplicações práticas com pequenos tamanhos amostrais. Por outro lado, os testes de White (W_1 e W_2) mostraram-se exageradamente conservativos para os tamanhos amostrais considerados. Isso reflete em baixo poder dos testes, de forma que a probabilidade de que os mesmos detectem heteroscedasticidade, até mesmo quando esta ocorre em grau acentuado, é muito baixa. Portanto, esses testes também não são recomendados para aplicações com pequenos tamanhos amostrais.

As características amostrais que mais afetaram os tamanhos dos testes foram a distribuição das covariáveis e dos erros. No primeiro caso, na presença de pontos de alavanca, grande parte dos testes mostraram-se consideravelmente conservativos, de forma a apresentar baixo poder. No segundo caso, sob erros com distribuições assimétricas ou com caudas pesadas, grande parte dos testes foram bastante anticonservativos, com exceção dos testes formulados para serem robustos a essa condição.

As simulações envolvendo o poder dos testes não permitiram identificar um teste que fosse mais poderoso em todas as situações. De qualquer forma, houve alguns que tiveram desempenhos razoáveis em certas situações. No caso em que a variância dos erros é função de somente uma variável ($k = 1$) e não ocorrem pontos de alavanca, o teste de Szroeter apresentou poder razoavelmente maior que os demais, confirmando o resultado obtido por Griffiths e Surekha (1986). No entanto, este teste é sensível à ordenação das observações, de forma que sua aplicabilidade é limitada nos casos em que $k > 1$, em que na prática não é possível ordenar adequadamente as observações. No caso em que $k > 1$, destacaram-se os testes da razão de verossimilhanças modificado RVM_1 , sua versão com correção de Bartlett ($RVMB_1$) e o teste da razão de verossimilhanças residual (RVR). No caso desse último teste, confirmou-se o resultado de Lyon e Tsai (1996), em que o mesmo destacou-se por seu alto poder. Esses três testes também apresentaram bons resultados na presença de pontos de alavanca, sobretudo $RVMB_1$. O teste score e suas versões modificadas, em geral, apresentaram baixo poder, o que está de acordo com o observado por diversos autores. É interessante observar que o teste score com correção de bootstrap (S_{BO}^*), embora tenha tido melhor desempenho que o teste original, sobretudo sob a presença de pontos de alavanca nas covariáveis, o que confirma o observado por Cribari-Neto e Zarkos (1999), não apresentou comportamento melhor que algumas versões do teste da razão de verossimilhanças. Logo, a maior complexidade existente na aplicação do teste não reflete em vantagens com relação ao poder, de forma que seu uso não é aconselhado em pequenas amostras.

Os testes robustos à normalidade dos dados, isto é, testes de White (W_1 e W_2), teste score studentizado (SS) e sua versão modificada (SSM) e teste de Levene, em geral também apresentaram baixo poder. Logo, só são recomendados quando há evidências, mesmo que pequenas, de não-normalidade dos erros. Dentre esses, o mais poderoso sob erros t_{10} foi o teste de Levene,

mesmo sob má-ordenação das observações, quando $k = 4$.

Portanto, há evidências de que nas situações “bem-comportadas”, em que as variâncias são afetadas por somente uma variável, não ocorrendo pontos de alavanca, o teste de Szroeter é o preferido. Nas demais situações, sob normalidade, os três testes recomendáveis são o teste da razão de verossimilhanças modificado em sua versão alternativa à de Simonoff e Tsai (RVM_1), a versão desse teste com correção de Bartlett ($RVMB_1$) e o teste da razão de verossimilhanças residual (RVR).

É importante ressaltar que, mesmo os testes com melhores desempenhos mencionados acima, só apresentam valores razoáveis de poder para altos graus de heteroscedasticidade, sob os tamanhos amostrais considerados neste trabalho. Logo, é possível que os mesmos não sejam capazes de identificar, com alta probabilidade, graus de heteroscedasticidade capazes de prejudicar consideravelmente as inferências usuais baseadas no método de MQO. Portanto, em pequenas amostras os resultados de tais testes devem ser vistos com cautela, e complementados com outros métodos de diagnóstico, tais como os gráficos de resíduos.

Capítulo 4

Aplicação

Em estudos que relacionam gastos com alimentação e renda, é comum observar a existência de heteroscedasticidade nos erros. A explicação para esse fato é simples: “domicílios com baixas rendas não têm a opção de ter hábitos alimentares extravagantes; comparativamente, eles têm poucas escolhas e são quase que forçados a gastar uma certa porção de sua renda em alimentação. Domicílios com rendas altas, por outro lado, podem ter hábitos alimentares simples ou extravagantes. A renda, sozinha, é provavelmente menos importante como variável explicativa para os gastos com alimentos em famílias com altas rendas” (Griffiths et al, 1993).

Nesse capítulo, apresentar-se-á a aplicação dos testes de heteroscedasticidade tratados nesse trabalho a dados de um estudo sobre gastos com alimentação. O objetivo desse estudo, cujos dados são apresentados na Tabela 15.4 de Griffiths et al (1993), é relacionar o gasto total anual com alimentação dos domicílios (y_t , em milhares de dólares) com a renda anual per capita (x_t , em milhares de dólares) e o número de pessoas do domicílio (n_t). A amostra é composta por 38 domicílios, e os dados são exibidos na Tabela 4.1.

Inicialmente, ajustou-se o modelo homoscedástico da forma

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 n_t + u_t. \quad (4.1)$$

A Figura 4.1 traz os gráficos de resíduos externamente studentizados plotados versus as variáveis explicativas x_t e n_t (gráficos a e b). No gráfico (a), observa-se uma tendência de aumento da variabilidade dos resíduos com o aumento da renda familiar (x_t), sugerindo, conforme esperado, a existência de heteroscedasticidade dos erros. No gráfico (b), no entanto, não há um padrão nítido de aumento da variabilidade dos resíduos com o aumento do número de pessoas no domicílio (n_t). Em (c) e (d) são apresentados os gráficos de variáveis adicionadas descritos na Seção 2.3, úteis para a escolha das covariáveis em Z . No gráfico (c), é apresentada a situação em que é eliminado o efeito de n_t sobre x_t e r_t (2.22). Esse gráfico apresenta uma certa curvatura, sugerindo que se deva incluir na matriz de covariáveis Z a variável x_t^2 . No gráfico (d), é apresentada a situação em que é eliminado o efeito de x_t sobre n_t e r_t . Nesse gráfico, os pontos parecem dispersos em uma faixa horizontal, o que reforça as evidências, fornecidas pelo gráfico (b), de inexistência de efeito do número de pessoas sobre as variâncias dos erros.

Com base nesses gráficos, assumir-se-á para as variâncias dos erros um modelo multiplicativo

Tabela 4.1: Dados sobre gastos com alimentação

t	y_t	x_t	n_t	t	y_t	x_t	n_t	t	y_t	x_t	n_t
1	15,998	62,476	1	14	21,141	82,251	2	27	18,561	82,469	1
2	16,652	82,304	5	15	17,446	59,862	3	28	11,629	44,208	2
3	21,741	74,679	3	16	9,629	26,563	3	29	18,067	49,467	5
4	7,431	39,151	3	17	14,005	61,818	2	30	14,539	25,905	5
5	10,481	64,724	5	18	9,160	29,682	1	31	19,192	79,178	5
6	13,548	36,786	3	19	18,831	50,825	5	32	25,918	75,811	3
7	23,256	83,052	4	20	7,641	71,062	4	33	28,833	82,718	6
8	17,976	86,935	1	21	13,882	41,990	4	34	15,869	48,311	4
9	14,161	88,233	2	22	9,670	37,324	3	35	14,910	42,494	5
10	8,825	38,695	2	23	21,604	86,352	5	36	9,550	40,573	4
11	14,184	73,831	7	24	10,866	45,506	2	37	23,066	44,872	6
12	19,604	77,122	3	25	28,980	69,929	6	38	14,751	27,167	7
13	13,728	45,519	2	26	10,882	61,041	2				

Fonte: Tabela 15.4 de Griffiths et al (1993), p. 512.

da seguinte forma:

$$\sigma_t^2 = \sigma^2 \exp(\delta_1 x_t^* + \delta_2 x_t^{*2}),$$

em que $x_t^* = x_t/10$. Essa transformação foi feita para evitar problemas de singularidade aparente no processo de maximização da função de verossimilhança, os quais ocorrem ao se considerar a variável não transformada, devido aos altos valores assumidos por x_t^2 . Em (4.1), será utilizada a variável transformada x_t^* no lugar de x_t . A escolha do modelo multiplicativo foi arbitrária. Porém, conforme se observou na Seção 3.5, a escolha da função cedástica não parece ter grande influência sobre o desempenho dos testes de heteroscedasticidade. Os resultados desses testes são apresentados na Tabela 4.2.

Observa-se que a maioria dos testes conduz à rejeição da hipótese nula de homoscedasticidade ao nível de significância de 5%. As únicas exceções foram o teste da razão de verossimilhanças condicional (*RVC*), cuja estatística assumiu um valor negativo, o teste score studentizado (*SS*), sua versão modificada (*SSM*), o teste de White (W_1) e o teste de Goldfeld-Quandt (*GQ*). Ao nível de 10%, o único teste que não conduz à rejeição da hipótese nula de homoscedasticidade é o teste de White (W_1), que conforme se verificou anteriormente, apresenta baixo poder para pequenos tamanhos amostrais. Portanto, há evidências de que os erros do modelo proposto são heteroscedásticos, o que está de acordo com o comportamento observado nos gráficos de resíduos, e com o que se espera desse tipo de estudo.

Na Tabela 4.3 são apresentadas as estimativas de MQO dos parâmetros do modelo homoscedástico (com x_t^*) e as estimativas de máxima verossimilhança dos parâmetros do modelo heteroscedástico em estudo. São apresentados ainda as estimativas dos erros padrão dos estimadores dos componentes de β , as estatísticas t e z , e os *p-values* dos testes de significância

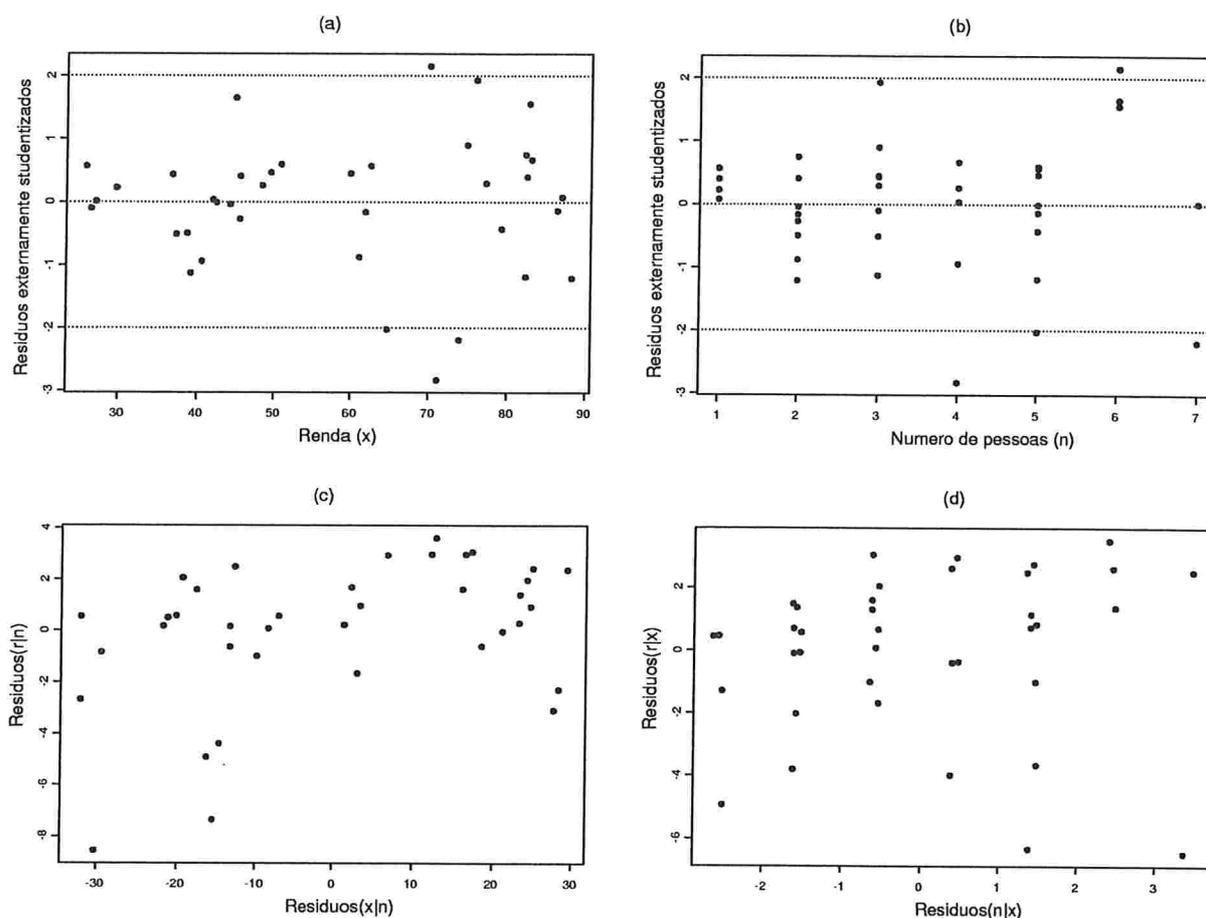


Figura 4.1: (a) e (b) Gráficos de resíduos externamente studentizados versus variáveis explicativas; (c) e (d) Gráficos de variável adicionada, apresentados na Seção 2.3

desses parâmetros. No caso do modelo heteroscedástico, as estimativas dos erros padrão são dadas pela raiz quadrada dos elementos da diagonal de (1.12).

Nota-se que as estimativas dos parâmetros do modelo heteroscedástico não diferiram muito das obtidas assumindo-se o modelo homoscedástico. No entanto, as estimativas dos erros padrão foram consideravelmente menores no caso do modelo heteroscedástico. Os testes de significância de β_1 e β_2 tiveram resultados semelhantes, com a rejeição da hipótese de nulidade em ambos os casos ao nível de significância de 1%. Portanto, há evidências de que os gastos com alimentação aumentam, em média, com a renda e com o número de pessoas no domicílio. Além disso, conforme esperado nesse tipo de estudo, há evidências da existência de heteroscedasticidade dos erros, sendo que a variância desses aumenta com o aumento da renda dos domicílios.

Tabela 4.2: Resultados dos testes de heteroscedasticidade para os dados sobre gastos com alimentação, assumindo o modelo $\sigma_t^2 = \sigma^2 \exp(\delta_1 x_t^* + \delta_2 x_t^{*2})$

Teste	Estatística	v.c. 10%	v.c. 5%	v.c. 1%	v.c. 0,5%	<i>p-value</i>
<i>RV</i>	12,677	4,605	5,991	9,210	10,597	0,002
<i>RVB</i>	10,574	4,605	5,991	9,210	10,597	0,005
<i>RVM</i>	9,225	4,605	5,991	9,210	10,597	0,010
<i>RVM₁</i>	9,395	4,605	5,991	9,210	10,597	0,009
<i>RVMB₁</i>	9,592	4,605	5,991	9,210	10,597	0,008
<i>RVC</i>	-6,559	4,605	5,991	9,210	10,597	1,000
<i>RVR</i>	9,892	4,605	5,991	9,210	10,597	0,007
<i>S</i>	6,801	4,605	5,991	9,210	10,597	0,033
<i>SB</i>	7,340	4,605	5,991	9,210	10,597	0,025
<i>SB₁</i>	7,340	4,605	5,991	9,210	10,597	0,025
<i>SB₂</i>	7,354	4,605	5,991	9,210	10,597	0,025
<i>SB₃</i>	7,369	4,605	5,991	9,210	10,597	0,025
<i>S*</i>	6,801	4,196	5,467	9,004	10,940	< 0,050
<i>S₁[*]</i>	6,801	4,206	5,480	9,029	11,050	< 0,050
<i>S₂[*]</i>	6,801	4,218	5,494	9,063	11,129	< 0,050
<i>S₃[*]</i>	6,801	4,311	5,643	9,095	10,774	< 0,050
<i>S₄[*]</i>	6,801	4,220	5,472	8,866	10,919	< 0,050
<i>S_{BO}[*]</i>	6,801	5,437	6,826	9,707	13,161	< 0,050
<i>SM</i>	6,455	4,605	5,991	9,210	10,597	0,040
<i>SS</i>	5,636	4,605	5,991	9,210	10,597	0,060
<i>SSM</i>	5,290	4,605	5,991	9,210	10,597	0,071
<i>SR</i>	6,656	4,605	5,991	9,210	10,597	0,036
<i>W₁</i>	8,484	9,236	11,070	15,086	16,750	0,131
<i>W₂</i>	12,009	9,236	11,070	15,086	16,750	0,035
<i>GQ</i>	2,496	2,323	2,978	4,849	5,847	0,083
<i>LV</i>	-2,936	-1,306	-1,688	-2,434	-2,719	0,003
<i>SZ</i>	1,760	1,282	1,645	2,326	2,576	0,039
<i>RS</i>	8,232	4,605	5,991	9,210	10,597	0,016

Nota: v.c. 10%, v.c. 5%, v.c. 1% e v.c. 0,5% são os valores críticos assintóticos para níveis de significância de 10%, 5%, 1% e 0,5%.

Tabela 4.3: Estimativas e testes de significância dos parâmetros dos modelos homoscedástico e heteroscedástico

Parâmetro	Modelo homoscedástico				Modelo heteroscedástico			
	Estimativa de MQO	Erro padrão	t	p-value	Estimativa de MV	Erro padrão	z	p-value
β_0	2,243	2,669	0,841	0,406	2,347	1,457	1,610	0,107
β_1	1,645	0,354	4,645	0,000	1,527	0,233	6,561	0,000
β_2	1,145	0,414	2,763	0,009	1,304	0,231	5,637	0,000
σ^2	16,957	—	—	—	0,005	—	—	—
δ_1	—	—	—	—	2,613	—	—	—
δ_2	—	—	—	—	-0,194	—	—	—

Apêndice A

Gráficos de resíduos heteroscedásticos

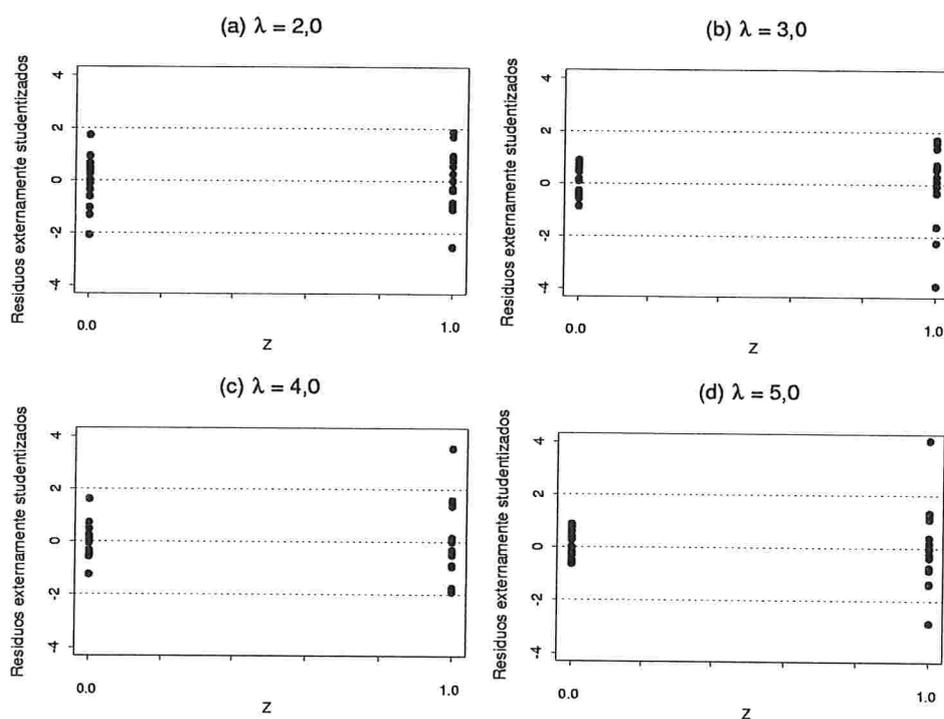


Figura A.1: Gráficos de resíduos externamente studentizados para o modelo de análise de variância com 1 fator e dois grupos (dados simulados, $n = 30$)

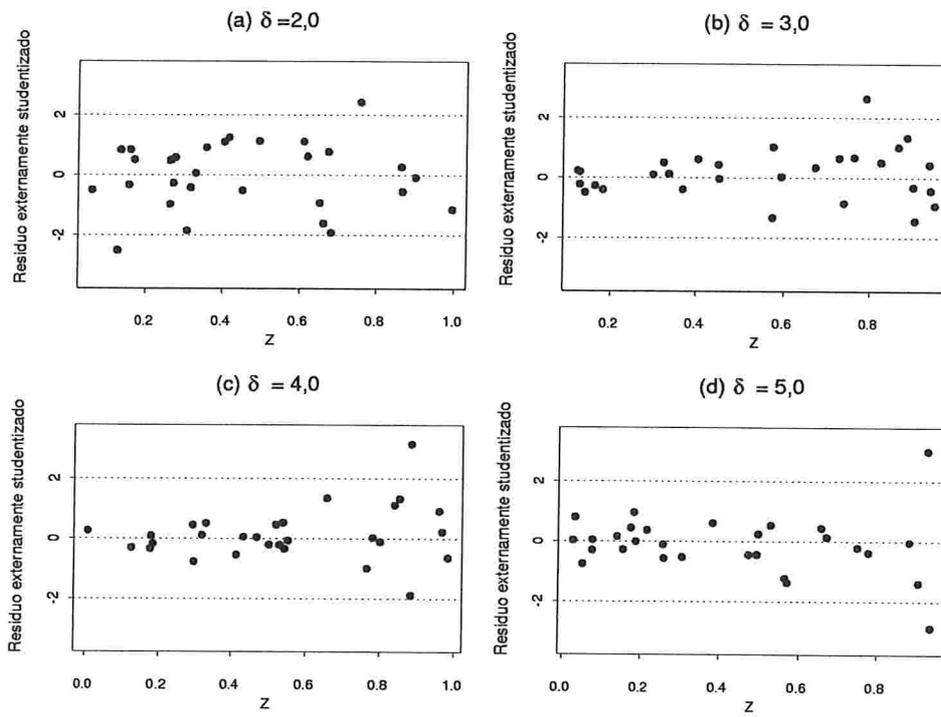


Figura A.2: Gráficos de resíduos externamente studentizados para dados heteroscedásticos simulados com $n = 30$, $p = 5$, $k = 1$, covariáveis uniformes e erros normais

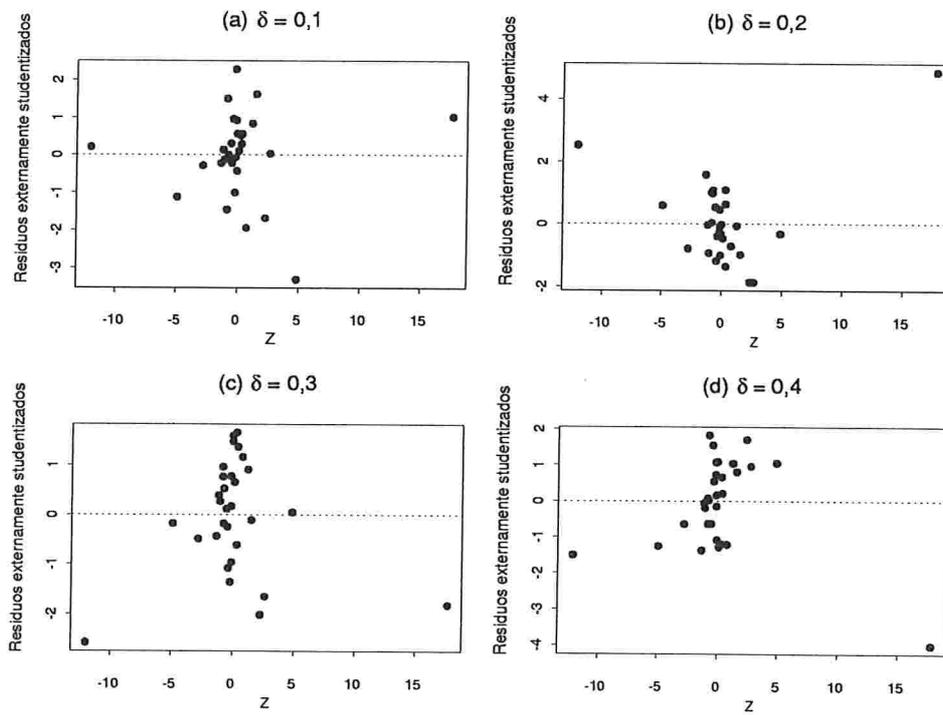


Figura A.3: Gráficos de resíduos externamente studentizados para dados heteroscedásticos simulados com $n = 30$, $p = 5$, $k = 1$, covariáveis com distribuição t_3 e erros normais

Apêndice B

Programa

Este apêndice traz um programa desenvolvido na linguagem de programação R para o cálculo das estatísticas dos 28 testes de heteroscedasticidade tratados neste trabalho. Esse programa está disponível no site <http://sites.uol.com.br/biancamatos>.

```
#####
# PROGRAMA: heteros_tests.src
# USO: Calcula as estatísticas de 28 testes de heteroscedasticidade no modelo de regressão linear.
# AUTORAS: Bianca de Carla Hernandez Matos (biancamatos@uol.com.br)
#         Silvia L. de Paula Ferrari (sferrari@ime.usp.br)
#####

tr <- function(mX) # função para calcular o traço de uma matriz
{sum(diag(mX))}

print.linha<-function(x=1) # função para imprimir x linhas em branco
{for (i in 1:x) {cat("\n")}}

# Função para calcular estatísticas dos testes baseados no teste da razão de verossimilhanças
fTRVbased<-function(dbartlettTRV, dlogdet1, vzbar,dbartlettTRVM, dsigma2hat0, fMultip, ik, IN, ip, vy, mX,
                    mZ, stepmax, steptol, vP)
{
  # Logaritmo da função de verossimilhança perfilada para o modelo multiplicativo
  fLogLikProf <- function(vP)
  {
    mWinv <- diag(1 / (exp(mZ%*%vP))[,1])
    vw <- exp(mZ%*%vP)
    vbetahatdelta <- solve(t(mX)%*%mWinv%*%mX, tol=1e-15)%*%t(mX)%*%mWinv%*%vy
    dsigma2hatdelta <- (1/IN)*t(vy - mX%*%vbetahatdelta)%*%mWinv%*%(vy - mX%*%vbetahatdelta)
    adFunc <- -(-(IN/2)*log(dsigma2hatdelta)-0.5*sum(log(vw)))
  }

  # Logaritmo da função de verossimilhança perfilada para o modelo aditivo
  fLogLikProfAddit<-function(vP)
  {
    mWinv <- diag((1 / (1+mZ%*%vP)^2)[,1])
    vw <- 1/diag(mWinv)
    vbetahatdelta <- solve(t(mX)%*%mWinv%*%mX, tol=1e-15)%*%t(mX)%*%mWinv%*%vy
    dsigma2hatdelta <- (1/IN)*t(vy - mX%*%vbetahatdelta)%*%mWinv%*%(vy - mX%*%vbetahatdelta)
    adFunc <- -(-(IN/2)*log(dsigma2hatdelta)-0.5*sum(log(vw)))
  }
}
```

```

# Logaritmo da função de verossimilhança perfilada sob homoscedasticidade:
fLogLikProfNull <- function(vP)
{
  vbetahat <- solve(t(mX)%*%mX, tol=1e-15)%*%t(mX)%*%vy
  dsigma2hat <- (1/IN)*t(vy - mX)%*%vbetahat)%*%(vy - mX)%*%vbetahat)
  adFunc <- -(-(IN/2)*log(dsigma2hat))
}

# Logaritmo da função de verossimilhança perfilada modificada para o modelo multiplicativo
fLogLikProfModif <- function(vP)
{
  mWinv <- diag((1 / exp(mZ)%*%vP))[,1])
  vw <- exp(mZ)%*%vP
  vbetahatdelta <- solve(t(mX)%*%mWinv)%*%mX, tol=1e-15)%*%t(mX)%*%mWinv)%*%vy
  dsigma2hatdelta <- (1/IN)*t(vy - mX)%*%vbetahatdelta)%*%mWinv)%*%(vy - mX)%*%vbetahatdelta)
  aux.det<-log(det(t(mX)%*%mWinv)%*%mX))
  adFunc <- -(-(IN-ip-2)/2)*log(dsigma2hatdelta)-0.5*sum(log(vw))-0.5*aux.det+(1/IN)*sum(log(vw)))
}

# Logaritmo da função de verossimilhança perfilada modificada para o modelo aditivo
fLogLikProfModifAddit <- function(vP)
{
  mWinv <- diag((1 / (1+mZ)%*%vP)^2)[,1])
  vw <- 1/diag(mWinv)
  vbetahatdelta <- solve(t(mX)%*%mWinv)%*%mX, tol=1e-15)%*%t(mX)%*%mWinv)%*%vy
  dsigma2hatdelta <- (1/IN)*t(vy - mX)%*%vbetahatdelta)%*%mWinv)%*%(vy - mX)%*%vbetahatdelta)
  aux.det<-log(det(t(mX)%*%mWinv)%*%mX))
  adFunc <- -(-(IN-ip-2)/2)*log(dsigma2hatdelta)-0.5*sum(log(vw))-0.5*aux.det+(1/IN)*sum(log(vw)))
}

# Logaritmo da função de verossimilhança modificada sob homoscedasticidade
fLogLikProfNullModif <- function(vP)
{
  vbetahat <- solve(t(mX)%*%mX, tol=1e-15)%*%t(mX)%*%vy
  dsigma2hat <- (1/IN)*t(vy - mX)%*%vbetahat)%*%(vy - mX)%*%vbetahat)
  aux.det <- log(det(t(mX)%*%mX))
  adFunc <- -(-(IN-ip-2)/2)*log(dsigma2hat)-0.5*aux.det)
}

# Valores iniciais para delta
vunit <- rep(1,ik)

# Maximização da função de verossimilhança
if (fMultip)
{
  vp1 <- vP
  vp2 <- vP
  ir1 <- nlm(fLogLikProf, vp1, steptol=steptol, stepmax = stepmax, ndigit=400, iterlim=100,
            gradtol=1e-7)$code
  ir2 <- nlm(fLogLikProfModif, vp2, steptol=steptol, stepmax = stepmax, ndigit=400, iterlim=100,
            gradtol=1e-7)$code
  vp1 <- nlm(fLogLikProf, vp1, steptol=steptol, stepmax = stepmax, ndigit=400, iterlim=100,
            gradtol=1e-7)$estimate
  vp2 <- nlm(fLogLikProfModif, vp2, steptol=steptol, stepmax = stepmax, ndigit=400, iterlim=100,
            gradtol=1e-7)$estimate
}
else
{
  vp1 <- vP
  vp2 <- vP
  ir1 <- nlm(fLogLikProfAddit, vp1, steptol=steptol, stepmax = stepmax, ndigit=100, iterlim=100000,
            gradtol=1e-6)$code
  ir2 <- nlm(fLogLikProfModifAddit, vp2, steptol=steptol, stepmax = stepmax, ndigit=100, iterlim=100000,

```

```

        gradtol=1e-6)$code
vp1 <- nlm(fLogLikProfAddit, vp1, steptol=steptol, stepmax = stepmax, ndigit=100, iterlim=100000,
        gradtol=1e-6)$estimate
vp2 <- nlm(fLogLikProfModifAddit, vp2, steptol=steptol, stepmax = stepmax, ndigit=100, iterlim=100000,
        gradtol=1e-6)$estimate
}
if ((ir1==1 || ir1==2) && (ir2==1 || ir2==2))
{
  vdeltahat <- vp1
  if (fMultip)
  {
    mWinv <- diag((1 / exp(mZ**vdeltahat))[,1])
    dMaxLogLikProf <- -fLogLikProf(vp1)
    dMaxLogLikProfModif <- -fLogLikProfModif(vp2)
  }
  else
  {
    mWinv <- diag((1 / (1+mZ**vdeltahat)^2)[,1])
    dMaxLogLikProf <- -fLogLikProfAddit(vp1)
    dMaxLogLikProfModif <- -fLogLikProfModifAddit(vp2)
  }
  vw <- 1/diag(mWinv)
  vbetahatdelta <- solve(t(mX)**mWinv**mX, tol=1e-15)**t(mX)**mWinv**vy
  dsigma2hatdelta <- (1/IN)*t(vy - mX**vbetahatdelta)**mWinv**t(vy - mX**vbetahatdelta)
  dMaxLogLikNull <- -fLogLikProfNull(vp1)
  dMaxLogLikProfNullModif <- -fLogLikProfNullModif(vp2)
  dconst <- (IN-ip-2)/IN
  mGhatinv <- solve(diag(vw/((prod(vw))^(1/IN))))

  # Cálculo das estatísticas
  dRV <- 2*(dMaxLogLikProf - dMaxLogLikNull)
  dRVB <- dRV/(1+dbartlettTRV)
  dRVM <- dconst*dRV + dlogdet1 - log(det(t(mX)**mGhatinv**mX))
  dRVM2 <- 2*(dMaxLogLikProfModif - dMaxLogLikProfNullModif)
  dRVMB2 <- dRVM2/(1+dbartlettTRVM)
  dRVC <- dRVM -(2/IN)*sum(mZ ** vdeltahat)
  dRVR <- dRVC + 2*(log(dsigma2hat0)-log(dsigma2hatdelta))
  voutput <- c(dRV , dRVB , dRVM , dRVM2 , dRVMB2 , dRVC , dRVR, ir1, ir2)
  return(voutput)
}
else
{
  return(c(rep(NA,7),ir1,ir2))
}
}

# Função para calcular estatísticas dos testes baseados no teste escore
fScorebased <- function(vbetahat0, dsigma2hat0, mIdent, vunit, mDbar,dc1, dc2, dc3, vh, mAux, mV, IN,
  ip, ik, vy, mX, mZ)
{
  vu0 <- vy - mX**vbetahat0
  vu02 <- as.matrix(vu0^2)
  dS <- (t(vu02)**mDbar**solve(t(mDbar)**mDbar, tol=1e-15)**t(mDbar)**vu02)/(2*dsigma2hat0^2)
  dpsihat <- sum((vu02-rep(dsigma2hat0, IN))^2)/IN
  dsigma2til<-IN*dsigma2hat0/(IN-ip)
  va<-rep((1/dsigma2til), IN)*vu02-(vunit-vh)
  vtau <- solve(t(mDbar)**mDbar, tol=1e-15)**t(mDbar)**(vu02*(1/as.double(dsigma2hat0)))
  dSB <- dS*(1-(dc1+dc2*dS+dc3*dS^2))
  dts <- qchisq(0.995, ik)
  if (dS<dts)
  {
    dSB1 <- dSB
  }
}
}

```

```

else
{
  dSB1 <- dS
}
dSB2 <- dSB+0.25*(dc1^2*dS+2*dc1*dc2*(dS^2)+(2*dc1*dc3+(4/3)*dc2^2)*(dS^3)+3*dc2*dc3*(dS^4)+
(9/5)*dc3^2*(dS^5))
if (dc2==0 && dc3==0)
{
  dSB3 <- dSB
}
else
{
  if (dc3>0)
  {
    dSB3 <- sqrt(pi/(3*dc3))*exp((dc2^2)/(3*dc3)-dc1)*(pnorm(sqrt(6*dc3)*dS+sqrt(2/(3*dc3))*dc2)-
pnorm(sqrt(2/(3*dc3))*dc2))
  }
  else
  {
    dSB3<-(1/(2*dc2))*exp(-dc1)*(1-exp(-2*dc2*dS))
  }
}
dSM <- dS+t(vh)%*%mV%*%vtau
dSS <- 2*dsigma2hat0^2*dS/dpsihat
dSSM <- dSS+t(vh)%*%mV%*%vtau
dSR <- 0.5*t(va)%*%mZ%*%solve(t(mZ)%*%mAux%*%mZ, tol=1e-15)%*%t(mZ)%*%va
voutput <- c(dS , dSB , dSB1 , dSB2 , dSB3 , dSM , dSS , dSSM ,dSR)
return(voutput)
}

# Função para calcular a estatística do teste escore para a correção do valor crítico por bootstrap
fScoreBoot <- function(vuboothat, mDbar, dsigma2hat0boot)
{
  vuboothat2 <- as.matrix(vuboothat^2)
  dS <- (t(vuboothat2)%*%mDbar%*%solve(t(mDbar)%*%mDbar, tol=1e-15)%*%t(mDbar)%*%vuboothat2)/
(2*dsigma2hat0boot^2)
  return(dS)
}

# Função para calcular estatísticas dos testes não-constitutivos
fNonConstruct <- function(vbetahat0, dsigma2hat0, mDtil, mDast, IN, ip, ik, vy, mX, mZ, icGQ,inLV)
{
  vu0 <- vy - mX%*%vbetahat0
  vu02 <- as.matrix(vu0^2)
  vu02bar <- mean(vu02)
  vunit <- rep(1, IN)
  mS <- diag(as.vector(vu02-vunit*dsigma2hat0)^2)
  mXW2 <- cbind(vunit,mDast)
  vbetahatW2 <- solve(t(mXW2)%*%mXW2, tol=1e-15)%*%t(mXW2)%*%vu02
  vfitW2 <- mXW2 %*% vbetahatW2
  dSQTW2 <- t(vu02-vu02bar)%*%(vu02-vu02bar)
  dSQRW2 <- t(vu02-vfitW2)%*%(vu02-vfitW2)
  dR2 <- 1 - dSQRW2/dSQTW2
  dW1 <- t(vu02)%*%mDtil%*%solve(t(mDtil)%*%mS%*%mDtil, tol=1e-15)%*%t(mDtil)%*%vu02
  dW2 <- IN*dR2
  # Sort according to first column of matrix Z
  vy.ord <- vy[sort.list(mZ[,1])]
  mX.ord <- mX[sort.list(mZ[,1]),]
  ve.ord <- vu0[sort.list(mZ[,1])]
  vy1 <- as.matrix(vy.ord[1:(IN-icGQ)/2])
  mX1 <- as.matrix(mX.ord[1:(IN-icGQ)/2],)
  vbetahat1 <- solve(t(mX1)%*%mX1, tol=1e-15)%*%t(mX1)%*%(vy1)
  dS1 <- sum((vy1 - mX1%*%vbetahat1)^2)
}

```

```

vy2 <- as.matrix(vy.ord[((IN-icGQ)/2+icGQ+1):IN])
mX2 <- as.matrix(mX.ord[((IN-icGQ)/2+icGQ+1):IN,])
vbetahat2 <- solve(t(mX2)%*%mX2, tol=1e-15)%*%t(mX2)%*%vy2
dS2 <- sum((vy2 - mX2)%*%vbetahat2)^2
dGQ <- dS2/dS1
v.e1 <- ve.ord[1:inLV]
v.e2 <- ve.ord[(inLV+1):IN]
in2 <- length(v.e2)
vd1 <- abs(v.e1 - median(v.e1))
vd2 <- abs(v.e2 - median(v.e2))
ds <- sqrt(((inLV-1)*var(vd1)+(in2-1)*var(vd2))/(IN-2))
dLV <- (mean(vd1)-mean(vd2))/(ds*sqrt(1/inLV+1/in2))
dhtil <- sum(t(seq(1,IN,1))*(ve.ord^2))/sum(ve.ord^2)
dSZ <- ((6*IN)/(IN^2-1))^0.5*(dhtil-(IN+1)/2)
iN1 <- trunc(IN/3)
iN2 <- trunc((IN - iN1)/2)
iN3 <- IN-iN1-iN2
dsigma2hat01 <- sum((ve.ord^2)[1:iN1])/iN1
dsigma2hat02 <- sum((ve.ord^2)[(iN1+1):(iN1+iN2)])/iN2
dsigma2hat03 <- sum((ve.ord^2)[(iN1+iN2+1):IN])/iN3
dRS <- IN*log(dsigma2hat0)-(iN1*log(dsigma2hat01)+iN2*log(dsigma2hat02)+iN3*log(dsigma2hat03))
voutput <- c(dW1 , dW2 , dGQ , dLV , dSZ , dRS)
return(voutput)
}

# Função para impressão de resultados
fPrintResults <- function(mTestStats, vcvBoot, dglWhite, dc1, dc2, dc3, fMultip, fBoot,
  INESTAT, IN, ip, ik, icGQ, inLV, INBOOT, vy, mX, mZ, fFailed.MLE, ir1, ir2, stepmax,steptol, vP)
{
  # Initialization of matrices
  mResult <- matrix(nrow=INESTAT+6,ncol=6) # matriz de resultados (estatísticas, valores críticos
    e p-values)
  vcvModif2 <- numeric(4)
  dglGQ <- (IN-icGQ-2*ip)/2

  # Valores críticos assintóticos e modificados
  vcv <- qchisq(c(0.90,0.95,0.99,0.995),ik)
  vcvWhite <- qchisq(c(0.90,0.95,0.99,0.995),dglWhite)
  vcvModif <- vcv*(1+(dc1+dc2*vcv+dc3*(vcv^2)))
  vcvModif1 <- vcvModif+0.25*(dc1^2*vcvModif+2*dc1*dc2*(vcvModif^2)+(2*dc1*dc3+(4/3)*dc2^2)*(vcvModif^3)+
    3*dc2*dc3*(vcvModif^4)+(9/5)*dc3^2*(vcvModif^5))
  if (dc2==0 && dc3==0)
  {
    vcvModif2 <- vcvModif
    vcvModif4 <- vcvModif
  }
  else
  {
    if (dc3!=0)
    {
      FFalhou <- T
      vcvModif4 <- (1/sqrt(6*dc3))*(qnorm(vcv*(sqrt(pi/(3*dc3))*exp(dc2^2/(3*dc3)-dc1))^(-1) +
        pnorm(dc2*sqrt(2/(3*dc3))))-sqrt(2/(3*dc3))*dc2)
    }
    else
    {
      vcvModif2 <- (exp(dc1)*(exp(2*dc2*vcv)-1))/(2*dc2)
      vcvModif4 <- (-1/(2*dc2))*log(vcv*(exp(-dc1/(2*dc2))+1)^(-1))
    }
  }
}
if (dc1>=0 && dc2>=0)
{
  vcvModif3 <- vcvModif
}

```

```

}
else
{
  if (dc1<0 && dc2>=0)
  {
    vcvModif3 <- (1/(1-dc1))*(vcv+(dc2*vcv^2+dc3*vcv^3))
  }
  else
  {
    if (dc1>=0 && dc2<0)
    {
      vcvModif3 <- (1/(2*dc2))*(1-sqrt(1-4*dc2*(vcv+dc1*vcv+dc3*vcv^3)))
    }
    else
    {
      vcvModif3 <- (1/(2*dc2))*(1-dc1-sqrt((1-dc1)^2-4*dc2*(vcv+dc3*vcv^3)))
    }
  }
}
vcvGQ <- qf(c(0.90, 0.95, 0.99, 0.995), dglGQ, dglGQ)
vcvLV <- qt(c(0.10, 0.05, 0.01, 0.005), (IN-2))
vcvSZ <- qnorm(c(0.90, 0.95, 0.99, 0.995))
vcvRS <- qchisq(c(0.90, 0.95, 0.99, 0.995),2)

# Composição da matriz de resultados

# Primeira coluna: Estatísticas dos testes
mResult[1:12,1] <- mTestStats[1:12,1]
for (ci in 13:18)
{
  mResult[ci,1] <- mTestStats[8,1] # teste score com valores críticos corrigidos
}
mResult[19:28,1] <- mTestStats[13:22,1]

# Segunda à quinta colunas: valores críticos
for (ci in 1:12)
{
  mResult[ci,2:5] <- vcv
}
mResult[13,2:5] <- vcvModif
mResult[14,2:5] <- vcvModif1
mResult[15,2:5] <- vcvModif2
mResult[16,2:5] <- vcvModif3
mResult[17,2:5] <- vcvModif4
mResult[18,2:5] <- vcvBoot
for (ci in 19:22)
{
  mResult[ci,2:5] <- vcv
}
for (ci in 23:24)
{
  mResult[ci,2:5] <- vcvWhite
}
mResult[25,2:5] <- vcvGQ
mResult[26,2:5] <- vcvLV
mResult[27,2:5] <- vcvSZ
mResult[28,2:5] <- vcvRS

# Sexta coluna: p-values
for (ci in 1:12)
{
  mResult[ci,6] <- 1- pchisq(mResult[ci,1], ik)
}

```

```

mResult[13,6] <- NA
mResult[14,6] <- NA
mResult[15,6] <- NA
mResult[16,6] <- NA
mResult[17,6] <- NA
mResult[18,6] <- NA
for (ci in 19:22)
{
  mResult[ci,6] <- 1-pchisq(mResult[ci,1], ik)
}
for (ci in 23:24)
{
  mResult[ci,6] <- 1-pchisq(mResult[ci,1], dglWhite)
}
mResult[25,6] <- 1-pf(mResult[25,1], dglGQ, dglGQ)
mResult[26,6] <- pt(mResult[26,1], (IN-2))
mResult[27,6] <- 1-pnorm(mResult[27,1])
mResult[28,6] <- 1-pchisq(mResult[28,1], 2)

vnomes.testes <- c("RV", "RVB", "RVM", "RVM2", "RVMB2", "RVC", "RVR", "S", "SB", "SB1", "SB2",
                  "SB3", "S*", "S1*", "S2*", "S3*", "S4*", "SBoot", "SM", "SS", "SSM", "SR", "W1",
                  "W2", "GQ", "LV", "SZ", "RS")

if (!fBoot)
{
  mResult <- mResult[-18,]
  vnames.testes <- vnames.testes[-18]
}
if (FFalhou)
{
  mResult <- mResult[-15,]
  vnames.testes <- vnames.testes[-15]
}
if (fFailed.MLE)
{
  mResult <- mResult[-(1:7),]
  vnames.testes <- vnames.testes[-(1:7)]
}
print.linha(3)
cat("ESTATÍSTICAS DOS TESTES DE HETEROSCEDASTICIDADE E VALORES CRÍTICOS", fill=T)
if (fMultip) {cat("MODELO MULTIPLICATIVO", fill=T)}
else {cat("MODELO ADITIVO", fill=T)}
if (!fBoot) {cat("SOMENTE VALORES CRÍTICOS ASSINTÓTICOS", fill=T)}
else {cat("VALORES CRÍTICOS ASSINTÓTICOS E CORRIGIDOS POR BOOTSTRAP", fill=T)}
cat("-----", fill=T)
if (fBoot) {cat("NÚMERO DE AMOSTRAS DE BOOTSTRAP: ", INBOOT, fill=T)}
cat("NÚMERO DE OBSERVAÇÕES (n): ", IN, fill=T)
cat("NÚMERO DE COVARIÁVEIS PARA O MODELO DA MÉDIA (p): ", ip, fill=T)
cat("NÚMERO DE COVARIÁVEIS PARA O MODELO DA VARIÂNCIA (k): ", ik, fill=T)
cat("NÚMERO DE OBSERVAÇÕES ELIMINADAS NO TESTE DE GOLDFELD-QUANDT: ", icGQ, fill=T)
cat("NÚMERO DE OBSERVAÇÕES NO PRIMEIRO GRUPO NO TESTE DE LEVENE: ", inLV, fill=T)
cat("DATA: ", date(), fill=T)
cat(" ", fill=T)
dimnames(mResult) <- list(vnames.testes, c("Estat", "vc 10%", "vc 5%", "vc 1%", "vc 0,5%", "p-value"))
print(round(mResult,3))
print.linha()
cat("Nota: vc 10%, vc 5%, vc 1% and vc 0,5% são os valores críticos dos testes")
print.linha(2)
if (fFailed.MLE)
{
  cat("AVISO: Não foi possível calcular as estimativas de máxima verossimilhança. As estatísticas
      baseadas no teste da razão de verossimilhanças foram omitidas.", fill=T)
  cat("DICA: Você pode tentar alterar stepmax ou o valor inicial vP.", fill=T)
}

```

```

else
{
  cat("A maximização da função de verossimilhança foi feita com sucesso.", fill=T)
}
cat("Código resultante para o logaritmo da função de verossimilhança: ", ir1, fill=T)
cat("Código resultante para o logaritmo da função de verossimilhança modificada: ", ir2, fill=T)
if (FFalhou)
{
  print.linha()
  cat("AVISO: Não foi possível calcular os valores críticos modificados para S2* (c3!=0).
      A função de erro complexa não está disponível no R.")
}
print.linha()
cat("----- ", fill=T)
}
#####
# FUNÇÃO PRINCIPAL
# Argumentos: vy é o vetor de valores da variável resposta, mX é a matriz de covariáveis para o modelo
# das médias e mZ é a matriz de covariáveis para o modelo das variâncias.
# fBoot e fMultip são variáveis indicadoras para realização da correção por bootstrap e modelo
# multiplicativo. Se fMultip=1, assume-se o modelo multiplicativo e se fMultip=0, assume-se o modelo
# aditivo. Stepmax e steptol são parâmetros para a maximização do logaritmo da verossimilhança perfilada.
# vP é o valor inicial para a maximização do logaritmo da função de verossimilhança e de sua versão
# modificada.
#####
heteros.tests<-function(vy, mX, mZ, fBoot=1, fMultip=1, stepmax=2, steptol=1e-7, vP=0*rep(1,ncol(mZ)))
{
  cat("Wait...")
  vy <- as.matrix(vy)
  mX <- as.matrix(mX)
  mZ <- as.matrix(mZ)
  INBOOT <- 500
  INESTAT <- 22
  IN <- nrow(mX)
  ip <- ncol(mX)
  ik <- ncol(mZ)
  if (length(vP) != ncol(mZ))
  {
    cat("\n Dimensão inadequada do valor inicial vP. Este deve ter ",ik, "elemento(s).", fill=T)
    return(cat(fill=T))
  }
  inLV <- trunc(IN/2)
  icGQ <- trunc(IN/3)
  if ((IN-icGQ)%2!=0) {icGQ <- icGQ-1}
  mIdent <- diag(IN)
  vunit <- rep(1,IN)
  dlogdet1 <- log(det(t(mX)%*%mX))
  vzbar <- apply(mZ, 2, mean)
  mR <- cbind(1, mZ)
  mA <- mR%*%solve(t(mR)%*%mR, tol=1e-15)%*%t(mR)
  mAd <- diag(diag(mA))
  mH <- mX%*%solve(t(mX)%*%mX, tol=1e-15)%*%t(mX)
  mHd <- diag(diag(mH))
  mH2 <- mH * mH
  mH3 <- mH2 * mH
  mV <- matrix(nrow=IN, ncol=ik)
  for (ci in 1:IN)
  {
    mV[ci,] <- mZ[ci,] - vzbar
  }
  mJ <- mV%*%solve(t(mV)%*%mV, tol=1e-15)%*%t(mV)
  mJd <- diag(diag(mJ))
}

```

```

mJ2 <- mJ*mJ
mJ3 <- mJ2*mJ
mJd2 <- mJd*mJd
mDbar <- (mIdent-vunit)%*%t(vunit)/IN)%*%mZ
cat(" ")
dA1 <- (24*ik*(ip-1)- 24*IN*sum(diag(mJd)%*%mHd))+6*IN*t(vunit)%*%mHd)%*%mJ)%*%mHd)%*%vunit+
12*IN*t(vunit)%*%(mJ*mH2)%*%vunit)/IN
dA2 <- (-24*ik*(ik+2)+36*IN*sum(diag(mJd2))-24*IN*t(vunit)%*%mJd)%*%mJ)%*%mHd)%*%vunit)/IN
dA3 <- (24*IN*t(vunit)%*%mJd)%*%mJ)%*%mJd)%*%vunit+16*IN*t(vunit)%*%mJ3)%*%vunit)/IN
dc1 <- (dA1-dA2+dA3)/(12*ik)
dc2 <- (dA2-2*dA3)/(12*ik*(ik+2))
dc3 <- dA3/(12*ik*(ik+2)*(ik+4))
cat(" ")
vh <- diag(mH)
mM <- mH2
diag(mM) <- (vunit-vh)^2
mAux <- mM - ((vunit - vh)%*%t(vunit-vh))/(IN-ip)
dInd <- 1
mDast <- matrix(nrow=IN,ncol=ip*(ip+1)/2)
for (ci in 1:ip)
{
  for (cj in ci:ip)
  {
    mDast[,dInd] <- mX[,ci] * mX[,cj]
    dInd <- dInd+1
  }
}

mDtil <- (mIdent - vunit)%*%t(vunit)/IN)%*%mDast
if (all(mX[,1] == vunit))
{
  mDast <- mDast[,2:(ip*(ip+1)/2)]
  mDtil <- mDtil[,2:(ip*(ip+1)/2)]
}
dglWhite <- ncol(mDast)
# Fator de correção de Bartlett para o teste da razão de verossimilhanças:
dbartlettTRV <- (-0.5*tr(mAd * mAd) + (1/3)*sum(apply(mA*mA)*mA,2,sum)) +
0.5*sum(apply(mAd %*% mA %*% mAd,2,sum)) +tr(mAd %*% mHd)-
sum(apply(mH2 * mA,2,sum)) + 0.5*sum(apply(mHd)%*%mA)%*%mHd,2,sum)) +
sum(apply(mAd)%*%mA)%*%mHd,2,sum)) - (3*(ip+1)^2-1)/(6*IN)/ik

# Fator de correção de Bartlett para a estatística do teste da razão de verossimilhanças modificado:
dbartlettTRVM <- (-0.5*tr(mJd)%*%mJd)+(1/(2*IN))*ik^2+0.5*t(vunit)%*%mJd)%*%mJ)%*%mJd)%*%vunit+
(1/3)*t(vunit)%*%mJ3)%*%vunit-(2/IN)*ik+(1/IN)*t(vunit)%*%mJ2)%*%vunit)/ik

# Inicialização das matrizes:
mTestStats <- matrix(nrow=INESTAT,ncol=1)
mTestStatsBoot <- matrix(nrow=12,ncol=INBOOT)
vcvBoot <- matrix(nrow=4,ncol=1)

# Estimativas de máxima verossimilhança dos parâmetros sob homoscedasticidade
vbetahat0 <- solve(t(mX)%*%mX, tol=1e-15)%*%t(mX)%*%vy
dsigma2hat0 <- (1/IN)*t(vy - mX)%*%vbetahat0)%*%(vy - mX)%*%vbetahat0)
voutTRV <- fTRVbased(dbartlettTRV, dlogdet1, vzbar, dbartlettTRVM, dsigma2hat0, fMultip, ik, IN, ip,
vy, mX, mZ,stepmax, steptol, vP)
ir1 <- voutTRV[8]
ir2 <- voutTRV[9]
fFailed.MLE <- F

if (!is.na(voutTRV[1]))
{
  # Testes baseados no teste da razão de verossimilhanças
  cat(" ")
}

```

```

    mTestStats[1:7,1] <- voutTRV[1:7]
  }
  else
  {
    fFailed.MLE <- T
  }

# Testes baseados no teste escore
mTestStats[8:16,1] <- fScorebased(vbetahat0, dsigma2hat0, mIdent, vunit, mDbar, dc1, dc2, dc3, vh,
  mAux, mV, IN, ip, ik, vy, mX, mZ)

# Testes não-constitutivos
mTestStats[17:22,1] <- fNonConstruct(vbetahat0, dsigma2hat0, mDtil, mDast, IN, ip, ik, vy, mX,
  mZ, icGQ, inLV)

# Valores críticos corrigidos por bootstrap
if (fBoot)
{
  vu0 <- vy - mX%%vbetahat0
  for (cj in 1:INBOOT)
  {
    vuboot <- sample(vu0, IN, replace=T)
    vyboot <- mX%%vbetahat0 + vuboot
    # Estimativas de MQO para a amostra de bootstrap
    vbetahat0boot <- solve(t(mX)%*mX, tol=1e-15)%*t(mX)%*vyboot
    dsigma2hat0boot <- (1/IN)*t(vyboot - mX%%vbetahat0boot)%*(vyboot - mX%%vbetahat0boot)
    vuboothat <- vyboot - mX%%vbetahat0boot
    mTestStatsBoot[1,cj] <- fScoreBoot(vuboothat, mDbar, dsigma2hat0boot)
  }
  vcvBoot <- quantile(mTestStatsBoot[1,], probs=c(0.90, 0.95, 0.99, 0.995))
}

# Impressão de resultados
fPrintResults(mTestStats, vcvBoot, dglWhite, dc1, dc2, dc3, fMultip, fBoot, INESTAT, IN, ip, ik, icGQ,
  inLV, INBOOT, vy, mX, mZ, fFailed.MLE, ir1, ir2, stepmax, steptol, vP)

}# Fim do programa

```

Referências Bibliográficas

- [1] Ali, M.M. e Giaccotto, C. (1984). A study of several new and existing tests for heteroscedasticity in the general linear model. *Journal of Econometrics*, **26**, 355–373.
- [2] Atkinson, A.C. (1985). *Plots, transformations and regressions*. Oxford: Oxford Statistical Science Series.
- [3] Barndorff-Nielsen, O.E. e Hall, P. (1988). On the level-error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika*, **75**, 374–378.
- [4] Bickel, P. (1978). Using residuals robustly I: tests for heteroscedasticity, nonlinearity. *The Annals of Statistics*, **6**, 266–291.
- [5] Breusch, T.S. e Pagan, A.R. (1979). A simple test for heteroskedasticity and random coefficient variation. *Econometrica*, **47**, 1287–1294.
- [6] Carroll, R.J. e Ruppert, D. (1981). On robust tests for heteroscedasticity. *The Annals of Statistics*, **9**, 206–210.
- [7] Chesher, A. e Jewitt, I. (1987). The bias of a heteroskedasticity consistent covariance matrix estimator. *Econometrica*, **55**, 1217–1222.
- [8] Conover, W.J., Johnson, M.E. e Johnson, M.M. (1981). A comparative study of tests for homogeneity of variances with applications to the outer continental shelf bidding data. *Technometrics*, **23**, 351–361.
- [9] Cook, R.D. e Weisberg, S. (1983). Diagnostics for heteroscedasticity in regression. *Biometrika*, **70**, 1–10.
- [10] Cordeiro, G.M. (1993). Bartlett corrections and bias correction for two heteroscedastic regression models. *Communications in Statistics, Theory and Methods*, **22**, 169–188.
- [11] Cordeiro, G.M. e Ferrari, S.L.P. (1991). A modified score test statistic having chi-squared distribution to order n^{-1} . *Biometrika*, **78**, 573–582.
- [12] Cordeiro, G.M. e Ferrari, S.L.P. (1994). Matrix formulae for computing improved score tests. *Journal of Statistical Computation and Simulation*, **49**, 195–206.
- [13] Cordeiro, G.M., Ferrari, S.L.P. e Cysneiros, A.H.M.A. (1998). A formula to improve score test statistics. *Journal of Statistical Computation and Simulation*, **62**, 123–136.

- [14] Cox, D.R. e Hinkley, D.V. (1974). *Theoretical Statistics*. London: Chapman and Hall.
- [15] Cox, D.R. e Reid, N. (1987). Parameter orthogonality and approximate conditional inference. *Journal of the Royal Statistical Society B*, **49**, 1–39.
- [16] Cribari-Neto, F. e Ferrari, S.L.P. (1995). An improved Lagrange multiplier test for heteroskedasticity. *Communications in Statistics-Simulation and Computation*, **24**, 31–44.
- [17] Cribari-Neto, F. e Ferrari, S.L.P. (2001). Monotonic improved critical values for two χ^2 asymptotic criteria. *Economics Letters*, **71**, 307–316.
- [18] Cribari-Neto, F., Ferrari, S.L.P. e Cordeiro, G.M. (2000). Improved heteroscedasticity-consistent covariance matrix estimators. *Biometrika*, **87**, 907–918.
- [19] Cribari-Neto, F., Ferrari, S.L.P. e Oliveira, W.A.S.C. (2002). A numerical evaluation of tests based on different heteroskedasticity-consistent covariance matrix estimators. Relatório Técnico, RT-MAE 2002-13, IME-USP.
- [20] Cribari-Neto, F. e Galvão, N.M.S. (2002). A class of improved heteroskedasticity-consistent covariance matrix estimators. Disponível em <http://www.de.ufpe.br/~cribari/hccm.pdf>.
- [21] Cribari-Neto, F. e Zarkos, S.G. (1999). Bootstrap methods for heteroskedastic regression models: evidence on estimation and testing. *Econometric Reviews*, **1999**, 211–228.
- [22] Cysneiros, A.H.M.A., Cribari-Neto, F. e Ferrari, S.L.P. (2002). Testes escores modificados: uma comparação. *Em preparação*.
- [23] DiCiccio, T.J. e Stern, S.E. (1994). Frequentist and Bayesian Bartlett correction of test statistics based on adjusted profile likelihoods. *Journal of the Royal Statistical Society B*, **56**, 397–408.
- [24] Doornik, J.A. (2001). *Ox: An object-oriented matrix language*, 4 ed. London: Timberlake Consultants Press.
- [25] Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, **7**, 1–26.
- [26] Evans, M.A. e King, M.L. (1985). A point optimal test for heteroscedastic disturbances. *Journal of Econometrics*, **27**, 163–178.
- [27] Ferrari, S.L.P., Cysneiros, A.H.M.A e Cribari-Neto, F. (2002). An improved test for heteroskedasticity using adjusted modified profile likelihood inference. Relatório Técnico, RT-MAE 2002-17, IME-USP.
- [28] Fujisawa, H. (1997). Improvement on chi-squared approximation by monotone transformation. *Journal of Multivariate Analysis*, **60**, 84–89.
- [29] Gentleman, R. e Ihaka, R. (1997). The R language. In L. Billard e N. Fisher (eds), *Proceedings of the 28th Symposium on the Interface*, The Interface Foundation of North America.

- [30] Glejser, H. (1969). A new test for heteroskedasticity. *Journal of the American Statistical Association*, **64**, 316–323.
- [31] Godfrey, L.G. (1978). Testing for multiplicative heteroskedasticity. *Journal of Econometrics*, **8**, 227–236.
- [32] Godfrey, L.G. e Orme, C.D. (1999). The robustness, reliability and power of heteroskedasticity tests. *Econometric Reviews*, **18**, 169–194.
- [33] Goldfeld, S.M. e Quandt, R.E. (1965). Some tests for homoscedasticity. *Journal of the American Statistical Association*, **60**, 539–547.
- [34] Griffiths, W.E., Hill, R.C. e Judge, G.G. (1993). *Learning and practicing Econometrics*. New York: John Wiley and Sons.
- [35] Griffiths, W.E. e Surekha, K. (1986). A Monte Carlo evaluation of the power of some tests for heteroskedasticity. *Journal of Econometrics*, **31**, 219–231.
- [36] Harrison, M.J. e McCabe, B.P.M. (1979). A test for heteroscedasticity based on ordinary least squares residuals. *Journal of the American Statistical Association*, **74**, 494–500.
- [37] Harris, P. (1985). An asymptotic expansion for the null distribution of the efficient score statistic. *Biometrika*, **72**, 653–659.
- [38] Harvey, A.C. (1976). Estimating regression models with multiplicative heteroscedasticity. *Econometrica*, **44**, 461–466.
- [39] Honda, Y. (1988). A size correction to the Lagrange multiplier test for heteroskedasticity. *Journal of Econometrics*, **38**, 375–386.
- [40] Honda, Y. (1989). On the optimality of some tests of the error covariance matrix in the linear regression model. *Journal of the Royal Statistical Society B*, **51**, 71–79.
- [41] Jorgenson, E. (1965). *Income-expenditure relations of Danish wage and salary earners*. Copenhagen: The Statistical Department.
- [42] Kakizawa, Y. (1996). Higher order monotone Bartlett-type adjustment for some multivariate test statistics. *Biometrika*, **83**, 923–927.
- [43] Kendall, M.G. e Stuart, A. (1961). *The advanced theory of Statistics*, Vols. 1, 2. London: Griffin.
- [44] Koenker, R. (1981). A note on studentizing a test for heteroskedasticity. *Journal of Econometrics*, **17**, 107–112.
- [45] Levene, H. (1960). Robust tests for equality of variances. In *Contributions to Probability and Statistics*, ed. I, Olkin, Palo Alto, Calif. Stanford University Press, p. 278–292.
- [46] Long, J.S. e Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, **54**, 217–224.

- [47] Lyon, J.D. e Tsai, C.L. (1996). A comparison of tests for heteroscedasticity. *The Statistician*, **45**, 337–349.
- [48] MacKinnon, J.G. e White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite-sample properties. *Journal of Econometrics*, **29**, 305–325.
- [49] Meyer, J.R. e Kuh, E. (1957). *The investment decision*. Cambridge: Harvard University Press.
- [50] Neter, J., Kutner, M.H., Nachtshein, C.J. e Wasserman, W. (1996). *Applied Linear Statistical Models*. 4ed, Chicago: Irwin.
- [51] Prais, S.J. e Houthakker, H.S. (1955). *The analysis of family budgets*. Cambridge: Cambridge University Press.
- [52] Ramsey, J.B. (1969). Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society B*, **31**, 350–371.
- [53] Ramsey, J.B. e Gilbert, R. (1972). Some small sample properties of tests for specification error. *Journal of the American Statistical Association*, **67**, 180–186.
- [54] Rao, C.R. (1947). Large sample tests of statistical hypothesis concerning several parameters with applications to problems of estimation. *Proceedings of the Cambridge Philosophical Society*, **44**, 50–57.
- [55] Rao, C.R. (1965). *Linear Statistical Inference and Its Applications in Biometric Research*. New York: John Wiley and Sons, pp. 302–309.
- [56] Rutemiller, H.C. e Bowers, D.A. (1968). Estimation in heteroscedastic regression models. *Journal of the American Statistical Association*, **63**, 552–557.
- [57] Silvey, S.D. (1959). The lagrange multiplier test. *Annals of Mathematical Statistics*, **30**, 389–407.
- [58] Simonoff, J.S. e Tsai, C.H. (1994). Use of modified profile likelihood for improved tests of constancy of variance in regression. *Applied Statistics*, **43**, 357–370.
- [59] Szroeter, J. (1978). A class of parametric tests for heteroscedasticity in linear econometric models. *Econometrica*, **46**, 1311–1327.
- [60] Theil, H. (1965). The analysis of disturbances in regression analysis. *Journal of the American Statistical Association*, **60**, 1067–1079.
- [61] Verbyla, A.P. (1993). Modelling variance heterogeneity: residual maximum likelihood and diagnostics. *Journal of the Royal Statistical Society B*, **55**, 493–508.
- [62] White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, **48**, 817–838.