

**UNIVERSIDADE FEDERAL DE VIÇOSA**

**Desempenhos de Testes de Heterocedasticidade na Regressão Linear Simples**

Jonas Firmiano da Silva  
*Magister Scientiae*

**VIÇOSA - MINAS GERAIS**  
**2024**

**JONAS FIRMIANO DA SILVA**

**Desempenhos de Testes de Heterocedasticidade na Regressão Linear Simples**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

Orientador: Paulo Cesar Emiliano

**VIÇOSA - MINAS GERAIS  
2024**

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade  
Federal de Viçosa - Campus Viçosa**

T

S586d  
2024  
Silva, Jonas Firmiano da, 1998-  
Desempenho de testes de heterocedasticidade na regressão  
linear simples / Jonas Firmiano da Silva. – Viçosa, MG, 2024.  
1 dissertação eletrônica (53 f.): il. (algumas color.).

Inclui apêndices.

Orientador: Paulo César Emiliano.

Dissertação (mestrado) - Universidade Federal de Viçosa,  
Departamento de Estatística, 2024.

Referências bibliográficas: f. 36-37.

DOI: <https://doi.org/10.47328/ufvbbt.2025.146>

Modo de acesso: World Wide Web.

1. Análise de regressão - Testes. 2. Mínimos quadrados.  
I. Emiliano, Paulo César, 1982-. II. Universidade Federal de  
Viçosa. Departamento de Estatística. Programa de  
Pós-Graduação em Estatística Aplicada e Biometria. III. Título.

CDD 22. ed. 519.536

**JONAS FIRMIANO DA SILVA**

**Desempenhos de Testes de Heterocedasticidade na Regressão Linear Simples**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Estatística Aplicada e Biometria, para obtenção do título de *Magister Scientiae*.

APROVADA: 27 de agosto de 2024.

Assentimento:

---

Jonas Firmiano da Silva  
Autor

---

Paulo Cesar Emiliano  
Orientador

Essa dissertação foi assinada digitalmente pelo autor em 24/03/2025 às 17:16:52 e pelo orientador em 25/03/2025 às 17:27:22. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **YYN2.CZSX.UB5A** e clique no botão 'Validar documento'.

## AGRADECIMENTOS

Primeiramente, quero expressar minha profunda gratidão a Deus, cuja orientação e força foram essenciais para a realização deste trabalho. A fé e a confiança em Sua sabedoria me deram coragem e persistência para superar os desafios ao longo desta jornada acadêmica.

A minha família, pelo amor incondicional e pelo apoio constante durante todo o percurso. Um agradecimento especial ao meu pai, João Batista da Silva, cuja dedicação e valores sempre foram uma inspiração para mim; à minha mãe, Tânia Firmiano da Silva, cuja paciência e compreensão foram fundamentais para minha motivação; e ao meu irmão, Elias Batista da Silva, pelo incentivo e pelo apoio incondicional. Vocês são o meu alicerce e a minha fonte de força, e sem o apoio de vocês, este trabalho não teria sido possível.

A meu orientador, Paulo César Emiliano, expresso minha sincera gratidão pela orientação, paciência e apoio contínuo. Sua experiência e conhecimento foram cruciais para o desenvolvimento deste trabalho, e suas sugestões e críticas construtivas ajudaram a aprimorar minha pesquisa. Agradeço também pela confiança depositada em mim e pelo incentivo constante ao longo desta jornada acadêmica.

Aos membros da banca examinadora, expresso minha profunda gratidão pela leitura atenta e pelas valiosas contribuições feitas a este trabalho. Suas observações e sugestões foram fundamentais para o aprimoramento desta dissertação, e sua dedicação em avaliar meu trabalho com rigor e cuidado é muito apreciada. Agradeço pela oportunidade de compartilhar minhas ideias e por todo o conhecimento que pude adquirir a partir de nossas discussões.

Quero expressar minha gratidão aos meus amigos da Matemática, como Sheila e Letícia, que estiveram ao meu lado desde o início da minha jornada acadêmica, e também aos novos amigos, representados pela Carol, que têm sido uma parte importante da minha vida. Cada um de vocês desempenhou um papel significativo, proporcionando momentos de descontração e troca de ideias que foram verdadeiros refúgios durante os períodos intensos de trabalho.

Além disso, gostaria de reconhecer as amizades mais antigas que ainda mantenho, como Isabela e Ana Júlia, que sempre foram essenciais para me

apoiar e motivar ao longo dessa trajetória. Agradeço a todos pelos momentos compartilhados e por tornar essa jornada mais leve.

Aos meus companheiros de mestrado em Estatística, com quem compartilhei não apenas os desafios acadêmicos, mas também momentos muito especiais que vão além do ambiente acadêmico. A camaradagem e o suporte mútuo foram fundamentais para o sucesso deste projeto e para o enriquecimento da minha experiência acadêmica.

À Universidade Federal de Viçosa, expresso meu agradecimento pela oportunidade de realizar a pós-graduação e pelo ambiente acadêmico de excelência proporcionado. A infraestrutura e o apoio da universidade foram essenciais para o desenvolvimento deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. A quem expresso minha sincera gratidão pela concessão da bolsa de estudos, que possibilitou a dedicação integral a este projeto e foi crucial para o progresso da minha pesquisa.

A todos que, de alguma forma, contribuíram para a realização deste trabalho, meu sincero agradecimento. Cada contribuição, grande ou pequena, foi valiosa e teve um impacto significativo na conclusão deste projeto.

## RESUMO

SILVA, Jonas Firmiano da, M.Sc., Universidade Federal de Viçosa, agosto de 2024.  
**Desempenhos de Testes de Heterocedasticidade na Regressão Linear Simples.**  
Orientador: Paulo Cesar Emiliano.

A regressão linear é uma ferramenta fundamental em diversas áreas de pesquisa, sendo o método dos mínimos quadrados ordinários (MQO) um dos mais utilizados para estimar os parâmetros da regressão. No entanto, para que o MQO seja considerado o melhor estimador linear não viesado, algumas hipóteses precisam ser atendidas, entre elas a homocedasticidade dos resíduos. A violação dessa hipótese pode levar a problemas na inferência dos parâmetros estimados, tornando crucial a detecção de heterocedasticidade nos modelos lineares. Este estudo tem como objetivo avaliar a eficácia e o poder de diferentes testes estatísticos na identificação da heterocedasticidade. Utilizando simulação de dados, foram avaliados os testes de White, Breusch-Pagan e Glejser, analisando seu desempenho quando há o aumento do coeficiente de variação das variâncias (CVV) e o aumento das amostras de 15 a 150 observações. Os resultados indicam que o poder dos testes aumenta com o crescimento do CVV, tornando-os mais eficazes na detecção de heterocedasticidade à medida que a variabilidade da variância dos resíduos se intensifica. Além disso, o aumento do tamanho da amostra melhora a confiabilidade dos testes. O teste de White apresentou um poder ligeiramente superior em comparação aos demais, mas todos demonstraram poder crescente na medida em que o CVV e o tamanho da amostra aumentam. Observou-se que o poder dos testes é mais estável em amostras maiores, mantendo-se elevado para diferentes valores de CVV.

Palavras-chave: Teste de White; Teste de Breusch-Pagan; Teste de Glejser; Mínimos quadrados; Poder do teste

## ABSTRACT

SILVA, Jonas Firmiano da, M.Sc., Universidade Federal de Viçosa, August, 2024. **Performance of Heteroscedasticity Tests in Simple Linear Regression.** Adviser: Paulo Cesar Emiliano.

Linear regression is a fundamental tool in various research fields, with the ordinary least squares (OLS) method being one of the most commonly used approaches for estimating regression parameters. However, for OLS to be considered the best linear unbiased estimator, certain assumptions must be met, including the homoskedasticity of residuals. Violating this assumption can lead to issues in parameter inference, making the detection of heteroskedasticity in linear models crucial. This study aims to evaluate the effectiveness and power of different statistical tests in identifying heteroskedasticity. Using data simulation, the White, Breusch-Pagan, and Glejser tests were assessed, analyzing their performance as the coefficient of variation of variance (CVV) increased and as sample sizes ranged from 15 to 150 observations. The results indicate that the power of the tests increases with a higher CVV, making them more effective in detecting heteroskedasticity as the variability of residual variance intensifies. Additionally, increasing the sample size improves the reliability of the tests. The White test showed slightly higher power compared to the others, but all demonstrated increasing effectiveness as CVV and sample size grew. It was observed that the power of the tests is more stable in larger samples, remaining high across different CVV values.

Keywords: White Test; Breusch-Pagan Test; Glejser Test; Least Squares; Test power

## LISTA DE FIGURAS

Figura 1: Fluxograma da simulação para a regressão linear e aplicação dos testes. . .	23
Figura 2: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150) para o cenário 1. . . . .	28
Figura 3: Poder dos testes de Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o cenário 2. . . . .	29
Figura 4: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o cenário 3. . . . .	31
Figura 5: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o cenário 4. . . . .	33

## LISTA DE ABREVIATURAS E SIGLAS

BLUE	<i>Best linear unbiased estimator</i>
CVV	Coefficiente de variação da variância
MQO	Mínimos quadrados ordinários
MQP	Mínimos quadrados ponderados
OLS	<i>Ordinary least squares</i>
RLS	Regressão linear simples
RLM	Regressão linear múltipla

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>10</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>12</b>
2.1	Regressão linear simples (RLS) . . . . .	12
2.2	Regressão linear múltipla (RLM) . . . . .	12
2.3	Mínimos quadrados ponderados na RLS . . . . .	13
2.4	Mínimos quadrados ordinários . . . . .	14
2.5	Heterocedasticidade . . . . .	15
2.6	Testes de hipóteses . . . . .	16
2.6.1	Teste de White . . . . .	16
2.6.2	Teste de Glejser . . . . .	17
2.6.3	Teste de Breusch–Pagan . . . . .	17
2.6.4	Poder de um teste . . . . .	19
2.7	Coeficiente de variação das variâncias (CVV) . . . . .	19
<b>3</b>	<b>METODOLOGIA</b>	<b>22</b>
3.1	Simulação . . . . .	22
3.2	Testes de heterocedasticidade . . . . .	23
3.2.1	Teste de White . . . . .	23
3.2.2	Teste de Glejser . . . . .	24
3.2.3	Teste de Breusch–Pagan . . . . .	25
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>27</b>
<b>5</b>	<b>CONCLUSÕES</b>	<b>35</b>
	<b>REFERÊNCIAS</b>	<b>36</b>
	<b>APÊNDICE</b>	<b>38</b>

# 1 INTRODUÇÃO

Quando alguém se propõe a estudar a relação entre variáveis, supõe-se que estas apresentam algum tipo de dependência que possa ser estimada. Ao longo do tempo, uma abordagem utilizada para analisar essa relação é a regressão linear. Essa técnica procura explicar a associação entre uma variável dependente e as variáveis independentes, visando estimar valores para os parâmetros da regressão, a fim de representar da forma mais precisa possível a relação linear entre essas variáveis. Entretanto, para entender a importância da regressão linear, é fundamental conhecer suas origens e os métodos desenvolvidos para sua aplicação.

O Método dos Mínimos Quadrados Ordinários (MQO) foi desenvolvido de forma independente por Karl Friedrich Gauss (1777-1855) e Adrien-Marie Legendre (1752-1855), deixando um grande impacto em diversas áreas do conhecimento. No início, sua aplicação estava muito ligada à astronomia, mas com o avanço da tecnologia e o surgimento de novos desafios, o método passou a ser utilizado em diferentes contextos. No entanto, nem sempre as condições ideais para sua aplicação eram atendidas, especialmente no que diz respeito à igualdade das variâncias dos erros, ou seja, à homocedasticidade. Quando isso não ocorre, os parâmetros ainda podem ser estimados, mas há riscos, como a subestimação ou superestimação da variância dos estimadores (Hoffmann e Vieira, 1998). Por isso, entender os efeitos da heterocedasticidade é fundamental para um uso mais preciso da regressão linear.

A heterocedasticidade nos resíduos de um modelo de regressão linear é um desafio comum em diversas áreas, como a agricultura e as ciências ambientais. Segundo Gujarati (2000), em muitos casos, é mais fácil analisar as consequências da heterocedasticidade do que identificá-la diretamente. Pesquisas recentes reforçam essa ideia, como o estudo de Dutta Roy, Das e Debnath (2022) sobre a estimativa de biomassa florestal em Tripura, na Índia, o trabalho de Sari et al. (2019) na modelagem do crescimento de tomates e a pesquisa de Yu, Clark e Tian (2022) sobre o rendimento de arroz na China. Esses estudos aplicam testes estatísticos para verificar a homocedasticidade nos resíduos, evidenciando a importância dessa verificação para a confiabilidade dos modelos estatísticos. Diante disso, é essencial conhecer e utilizar as ferramentas adequadas para detectar a heterocedasticidade e garantir análises mais confiáveis.

Neste trabalho, serão analisados três testes estatísticos para detectar heterocedasticidade: o teste de White (White, 1980), o teste de Glejser (Glejser, 1969) e o teste de Breusch-Pagan (Breusch e Pagan, 1979). O objetivo é avaliar o poder desses testes em cenários onde o coeficiente de variação da variância ( $CVV$ ) é ajustado de acordo com condições pré-definidas. Embora Uyanto (2022) também tenha utilizado simulação de variáveis em sua

análise, este estudo se diferencia pela abordagem adotada na geração dos dados, permitindo uma avaliação específica dos efeitos da heterocedasticidade sob diferentes condições. Dessa forma, busca-se aprofundar a compreensão da variabilidade dos erros com o intuito de fornecer uma base sólida para a interpretação e aplicação de modelos de regressão linear em diferentes contextos.

Especificamente, serão avaliados o controle do tamanho do erro de tipo I e o poder dos diferentes testes estatísticos aqui estudados para detectar heterocedasticidade nos resíduos da regressão linear, além de definir o que é heterocedasticidade e discutir suas implicações na análise de regressão, comparando como a mesma afeta os testes de White, Glejser e Breusch-Pagan utilizando simulações. O que faremos é determinar o poder estatístico dos testes mencionados, avaliando a taxa de rejeição correta da hipótese nula em diferentes cenários de heterocedasticidade; e analisar e discutir os resultados obtidos, destacando as condições sob as quais cada teste se mostra mais eficaz. Ao explorar esses objetivos, esperamos contribuir significativamente para a compreensão e aplicação dos testes de heterocedasticidade em diversos contextos, oferecendo *insights* valiosos para pesquisadores e profissionais das mais diversas áreas.

## 2 REFERENCIAL TEÓRICO

### 2.1 Regressão linear simples (RLS)

Segundo Hoffmann e Vieira (1998) a regressão linear simples é um método estatístico utilizado para explorar a relação entre duas variáveis. O modelo matemático é dado por,

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad (1)$$

em que  $Y_i$  para  $i = 1, \dots, n$  é a variável dependente,  $X_i$  é a variável explicativa,  $\beta_0$  e  $\beta_1$  os parâmetros da regressão e  $\varepsilon_i$  os erros da regressão.

O objetivo da regressão linear é ajustar a melhor reta que represente a relação entre as variáveis, mas a estimação dos parâmetros do modelo de regressão para atingir esse objetivo é um desafio.

### 2.2 Regressão linear múltipla (RLM)

Segundo Gujarati (2000), a regressão linear múltipla, tem como estrutura uma variável dependente e mais de uma variável explicativa, sendo que o modelo matemático para a regressão linear múltipla é da forma:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + \varepsilon_i, \quad i = 1, \dots, n$$

ou ainda na forma matricial:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

na qual,

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & \cdots & X_{1k} \\ 1 & X_{21} & X_{22} & \cdots & X_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & X_{n2} & \cdots & X_{nk} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

em que  $Y_i$  é a variável dependente,  $X_{ij}$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, k$  são as variáveis explicativas;  $\beta_j$ ,  $j = 0, \dots, k$  são os coeficientes da regressão e  $\varepsilon_i$ ,  $i = 1, \dots, n$  os erros da regressão.

### 2.3 Mínimos quadrados ponderados na RLS

Quando o objetivo é estimar os parâmetros da regressão linear simples, um dos métodos sugeridos por Hoffman e Vieira (1998) é o dos mínimos quadrados ponderados. Esse método visa atribuir pesos diferentes às observações, de forma que, ao invés de assumir que todas as observações têm a mesma variabilidade, considera-se que algumas podem ser mais confiáveis do que outras. Isso é especialmente útil em situações onde as observações possuem variância heterogênea (ou seja, variância não constante), o que é uma condição comum em muitas aplicações práticas.

Nesse contexto, o modelo descrito pela Equação (1) é ajustado para incorporar a ponderação  $w_i$ , e pode ser reescrito como:

$$w_i Y_i = w_i \beta_0 + w_i \beta_1 X_i + w_i \varepsilon_i, \quad (2)$$

no qual as variáveis  $Y_i$  e  $X_i$  são ponderadas pelos pesos  $w_i$ . Reescrevendo, obtemos:

$$Y_i^* = \beta_0 w_i + \beta_1 w_i X_i + \varepsilon_i^*,$$

em que  $Y_i^* = w_i Y_i$  e  $\varepsilon_i^* = w_i \varepsilon_i$ . O objetivo agora é minimizar a soma dos quadrados dos erros  $\varepsilon_i^*$ :

$$\sum_{i=1}^n \varepsilon_i^{*2} = \sum_{i=1}^n w_i^2 (Y_i - \beta_0 - \beta_1 X_i)^2.$$

Com ferramentas de cálculo diferencial e integral Hoffman e Vieira (1998) determinam os estimadores para  $\beta_0$  e  $\beta_1$ , dados por

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n w_i (X_i - \bar{X}_w)(Y_i - \bar{Y}_w)}{\sum_{i=1}^n w_i (X_i - \bar{X}_w)^2},$$

e

$$\hat{\beta}_0 = \bar{Y}_w - \hat{\beta}_1 \bar{X}_w,$$

no qual,

- $w_i$  é o peso associado a cada observação  $i$ ,

- $\bar{X}_w = \frac{\sum_{i=1}^n w_i X_i}{\sum_{i=1}^n w_i}$  é a média ponderada da variável explicativa  $X$ ,
- $\bar{Y}_w = \frac{\sum_{i=1}^n w_i Y_i}{\sum_{i=1}^n w_i}$  é a média ponderada da variável dependente  $Y$ .

## 2.4 Mínimos quadrados ordinários

O caso particular em que  $w_i$  é constante é o método dos mínimos quadrados ordinários (MQO), que tem como modelo,

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i.$$

Segundo Gujarati (2000) os estimadores de  $\beta_0$  e  $\beta_1$  de MQO para a regressão linear simples são dados por:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X}_w)(Y_i - \bar{Y}_w)}{\sum_{i=1}^n (X_i - \bar{X}_w)^2},$$

e

$$\hat{\beta}_0 = \bar{Y}_w - \hat{\beta}_1 \bar{X}_w,$$

sendo,

- $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$  é a média da variável explicativa  $X$ ,
- $\bar{Y} = \frac{\sum_{i=1}^n Y_i}{n}$  é a média da variável dependente  $Y$ .

Gujarati (2000) também mostrou uma generalização para a estimativa dos parâmetros da RLM, que resulta em,

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

sendo  $\mathbf{X}$  e  $\mathbf{Y}$  iguais ao da Seção 2.2.

## 2.5 Heterocedasticidade

Se o objetivo de um estudo com regressão linear fosse apenas encontrar os parâmetros da regressão, o método dos Mínimos Quadrados Ordinários (MQO) seria suficiente, já que se trata de uma abordagem matemática. No entanto, como também é necessário fazer inferências sobre as estimativas desses parâmetros, algumas hipóteses, formuladas por Gauss (1809), se tornam essenciais. Uma dessas hipóteses é a homocedasticidade dos erros, ou seja, a variância dos erros da regressão deve ser constante e igual.

Segundo uma interpretação das ideias de Gujarati (2000), a heterocedasticidade nos erros de uma regressão pode ser causada por alguns fatores, como omissão de variáveis relevantes, especificação incorreta do modelo, variação na escala dos dados, mudanças estruturais, presença de *outliers* ou erro de medição na variável dependente.

Esse cenário de heterocedasticidade é um obstáculo para encontrar o melhor modelo linear. Isso se dá porque afeta os estimadores de mínimos quadrados ordinários (MQO). Como mencionado por Hoffmann e Vieira (1998), as estimativas de variância dos parâmetros podem ser distorcidas quando há heterocedasticidade. A variância dos estimadores de MQO pode ser superestimada ou subestimada.

Hoffmann (2016) demonstra que, ao aplicar o método dos mínimos quadrados ordinários em um modelo com heterocedasticidade, o estimador permanece não viesado, mas pode não ser o mais eficiente, pois não minimiza a variância. Nesse contexto, e considerando que se conhece as variâncias dos erros, o método dos mínimos quadrados ponderados (MQP) se mostra superior, pois ajusta a estimação dos parâmetros considerando a variabilidade dos erros. Comparando a eficiência dos estimadores MQO e MQP em um modelo específico, Hoffmann (2016) calcula que a eficiência relativa do MQO é de apenas 53,2%, evidenciando a perda de precisão ao ignorar a heterocedasticidade.

O uso de estimadores de variância tendenciosos compromete a validade dos intervalos de confiança e dos testes de hipóteses, uma vez que esses procedimentos pressupõem uma estimativa correta da variância (Gujarati, 2000). Quando essa condição não é atendida, as conclusões podem ser equivocadas, pois os erros padrão das estimativas podem ser subestimados ou superestimados. Isso pode distorcer a análise de significância estatística dos coeficientes, resultando em decisões inadequadas na inferência estatística (Gujarati, 2000).

A verificação da heterocedasticidade nos resíduos é essencial na análise de modelos estatísticos aplicados. Os testes para sua detecção permitem avaliar a variabilidade dos erros residuais, contribuindo para uma interpretação mais fiel dos resultados. Essa abordagem melhora a confiabilidade das conclusões e auxilia na tomada de decisões em contextos agrícolas e ambientais.

## 2.6 Testes de hipóteses

Existem vários testes para detectar a heterocedasticidade, tais como o teste de White (White, 1980), o teste de Glejser (Glejser, 1969) e o teste de Breusch-Pagan (Breusch e Pagan, 1979), dentre outros. O teste de White é um teste geral que não requer especificação da forma da heterocedasticidade, o teste de Glejser é baseado na relação entre os resíduos e as variáveis independentes, e o teste de Breusch-Pagan verifica a dependência da variância dos erros em relação às variáveis explicativas.

### 2.6.1 Teste de White

White (1980) desenvolveu um teste que se destaca pela sua aplicabilidade prática e simplicidade de implementação. O teste de White é utilizado para verificar se há heterocedasticidade nos resíduos de um modelo de regressão linear. Para aplicá-lo, primeiramente estima-se os coeficientes do modelo usando o método dos mínimos quadrados ordinários.

O próximo passo é calcular os quadrados dos resíduos estimados do modelo para cada observação. Em seguida, realiza-se uma regressão auxiliar dos quadrados dos resíduos em função da variável independente  $X$  e do quadrado dessa variável. Essa regressão é representada pela equação:

$$\hat{\varepsilon}_i^2 = \gamma_0 + \gamma_1 X_i + \gamma_2 X_i^2 + u_i,$$

em que  $\gamma_i$  são os coeficientes estimados da regressão auxiliar e  $u_i$  são os erros dessa regressão.

A próxima etapa é avaliar os coeficientes da regressão auxiliar com base nas hipóteses estatísticas:

$$H_0 : \gamma_1 = \gamma_2 = 0 \text{ (Homocedasticidade),}$$

$$H_1 : \gamma_1 \neq 0 \text{ ou } \gamma_2 \neq 0 \text{ (Heterocedasticidade).}$$

Segundo White (1980), na regressão auxiliar, calcula-se o coeficiente de determinação  $R_{\text{aux}}^2$  e o utiliza para determinar a estatística do teste, dada por:

$$LM = n \cdot R_{\text{aux}}^2,$$

em que  $n$  é o número de observações utilizadas na regressão auxiliar. Sob a hipótese nula de homocedasticidade e assumindo que o modelo original está corretamente especificado, a estatística  $LM$  segue uma distribuição qui-quadrado com 2 graus de liberdade (o número de coeficientes da regressão auxiliar menos um).

Por fim, compara-se o valor de  $LM$  ao valor crítico da distribuição qui-quadrado com 2 graus de liberdade e o nível de significância escolhido. Se  $LM$  for maior que o valor crítico,

rejeitamos  $H_0$ , indicando a presença de heterocedasticidade nos resíduos da regressão. Caso contrário, existem indícios que os coeficientes da regressão auxiliar são iguais a zero, o que indica que não há relação entre o quadrado dos erros da regressão e a variável dependente, e, assim, os erros podem ser considerados homocedásticos.

### 2.6.2 Teste de Glejser

O teste desenvolvido por Glejser (1969) procura verificar a relação entre os resíduos da regressão e as variáveis independentes do modelo. Ele desenvolveu um teste que tem como hipóteses estatísticas,

$H_0$  : As variâncias são homocedásticas ( $\sigma_i^2 = \sigma^2, \forall i = 1, \dots, n$ ),

$H_1$  : As variâncias são heterocedásticas ( $\sigma_i^2 \neq \sigma^2$ , para algum,  $i = 1, \dots, n$ ).

Em seu trabalho, Glejser, propôs ajustar uma regressão para avaliar a relação dos resíduos com as variáveis independentes, mas como os valores dos resíduos reais são desconhecidos para aplicar o teste é necessário estimar por MQO os parâmetros da regressão linear e, em seguida, estimar os resíduos da forma  $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$ , sendo que, com essa estimativa ajusta-se a regressão,

$$|\hat{\varepsilon}_i| = a + bZ_i + \alpha_i,$$

em que  $|\hat{\varepsilon}_i|$  são os valores absolutos dos resíduos,  $Z_i$  é a variável de interesse,  $a$  e  $b$  são coeficientes de regressão a serem estimados, e  $\alpha_i$  são os termos de erro da nova regressão.

Após realizar essa regressão, Glejser (1969) propõe testar a significância dos coeficientes, isso pode ser feito através de algum teste como o teste  $t$  de Student. Se os coeficientes forem estatisticamente significativos, isso sugere a presença de heterocedasticidade. Com base nos resultados do teste, pode-se inferir se há evidências de heterocedasticidade nos dados.

### 2.6.3 Teste de Breusch–Pagan

Para verificar se há heterocedasticidade nos resíduos de um modelo de regressão linear, Breusch e Pagan (1979) desenvolveram um teste cuja hipótese nula é de homocedasticidade das variâncias dos resíduos da regressão linear, enquanto a hipótese alternativa indica heterocedasticidade,

$H_0$  : As variâncias são homocedásticas ( $\sigma_i^2 = \sigma^2, \forall i = 1, \dots, n$ )

$H_1$  : As variâncias são heterocedásticas ( $\sigma_i^2 \neq \sigma^2$ , para algum  $i = 1, \dots, n$ ).

A aplicação do teste se dá através da análise dos erros da regressão por MQO. Primeiramente, ajusta-se o modelo da forma:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i,$$

e, em seguida, calcula-se os resíduos, representados por  $\hat{\varepsilon}_i$ , que são a diferença entre os valores reais e os valores previstos pelo modelo:

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i.$$

O próximo passo é realizar uma regressão dos resíduos ao quadrado ( $\hat{\varepsilon}_i^2$ ) em relação à variável independente  $X$ , que é chamada de regressão auxiliar. Esta regressão pode ser expressa como:

$$\hat{\varepsilon}_i^2 = \lambda_0 + \lambda_1 X_i + u_i,$$

em que  $\lambda_0$  e  $\lambda_1$  são os coeficientes estimados na regressão auxiliar, e  $u_i$  são os termos de erro da regressão.

Após a regressão dos resíduos ao quadrado, testamos a significância conjunta dos coeficientes  $\lambda_0$  e  $\lambda_1$  com base nas hipóteses estatísticas:

$$H_0 : \lambda_1 = 0 \text{ (Homocedasticidade),}$$

$$H_1 : \lambda_1 \neq 0 \text{ (Heterocedasticidade).}$$

Da regressão auxiliar, calcula-se o coeficiente de determinação ( $R_{\text{aux}}^2$ ) e o utiliza para calcular a estatística do teste,

$$LM = n \cdot R_{\text{aux}}^2,$$

em que  $n$  é o número de observações usadas para a regressão auxiliar. Sob a hipótese nula, a estatística  $LM$  segue uma distribuição qui-quadrado com 1 grau de liberdade (número de variáveis explicativas na regressão auxiliar).

Por fim, compara-se o valor de  $LM$  com o valor crítico da distribuição qui-quadrado com 1 grau de liberdade e o nível de significância desejado. Se  $LM$  for maior que o valor crítico, rejeita-se  $H_0$  e conclui-se que há indícios de heterocedasticidade nos resíduos da regressão. Caso contrário, existem indícios que os coeficientes da regressão auxiliar são iguais a zero, o que indica que não há relação entre o quadrado dos erros da regressão e a variável dependente, e, assim, os erros podem ser considerados homocedásticos.

#### 2.6.4 Poder de um teste

Quando testes de hipóteses são aplicados, geralmente, partimos de uma afirmação inicial chamada de hipótese nula ( $H_0$ ). Esta hipótese em geral representa uma condição de não efeito ou não diferença, e muitas vezes, ao propor o teste, o objetivo é rejeitar  $H_0$  em favor de uma hipótese alternativa ( $H_1$ ), que sugere a presença de um efeito ou diferença.

Morettin e Bussab (2017) afirmam que durante a aplicação do teste de hipóteses, a decisão tomada está sujeita a erros, os quais podem ser de dois tipos:

Erro tipo I: Rejeitar a hipótese  $H_0$ , dado que ela é verdadeira,

Erro tipo II: Não rejeitar a hipótese  $H_0$ , dado que ela não é verdadeira.

Eles ainda asseguram que as probabilidades de cometer esses erros são definidas como:

$$\alpha = P(\text{erro de tipo I}) = P(\text{rejeitar } H_0 \mid H_0 \text{ é verdadeira}),$$

e

$$\beta = P(\text{erro de tipo II}) = P(\text{não rejeitar } H_0 \mid H_0 \text{ é falsa}).$$

Segundo Morettin e Bussab (2017), o nível de significância  $\alpha$  é a probabilidade de rejeitar a hipótese nula  $H_0$  quando ela é verdadeira, dito de outra forma a probabilidade  $\alpha$  de cometer um erro de tipo I é conhecida como nível de significância do teste. Este valor é frequentemente pré-definido e comumente adotado em níveis como 0,01, 0,05 ou 0,10, representando uma confiança de 99%, 95% ou 90%, respectivamente, de que não se está cometendo um erro tipo I ao rejeitar  $H_0$ .

Por outro lado, a probabilidade  $\beta$  de cometer um erro de tipo II está relacionada com a capacidade do teste de detectar um efeito quando ele realmente existe. A partir de  $\beta$ , podemos calcular o poder do teste, que é dado por  $1 - \beta$ . O poder do teste é, portanto, a probabilidade de corretamente rejeitar  $H_0$  quando  $H_1$  é verdadeira, e um poder alto (próximo de 1) é desejável para garantir que o teste é eficaz em identificar efeitos reais. Como mencionado por Morettin e Bussab (2017), o poder de um teste é a probabilidade de rejeitar a hipótese nula  $H_0$  quando ela é falsa.

## 2.7 Coeficiente de variação das variâncias (CVV)

Desenvolvido por Box (1954) e sugerido como métrica padronizada por Kennedy (1985), o Coeficiente de Variação das Variâncias (CVV) oferece uma maneira sólida e acessível de compreender como a heterocedasticidade pode afetar modelos estatísticos, independentemente do tamanho da amostra ou da escala dos dados.

O CVV é uma ferramenta útil para avaliar o efeito da heterocedasticidade em modelos estatísticos. Esse coeficiente é calculado dividindo o desvio padrão pela média das variâncias dos erros. A fórmula é expressa da seguinte maneira:

$$CVV = \frac{\sqrt{\frac{1}{n-1} \sum_{i=1}^n \left( \sigma_i^2 - \frac{\sum_{i=1}^n \sigma_i^2}{n} \right)^2}}{\frac{1}{n} \sum_{i=1}^n \sigma_i^2}, \quad (3)$$

em que  $\sigma_i^2$  representa a variância de cada erro, e  $n$  é o número de observações do modelo. Essa fórmula quantifica o grau de dispersão das variâncias dos erros em relação à sua média.

Kennedy (1985) argumenta que estudos de simulação Monte Carlo podem ser significativamente aprimorados utilizando o CVV como uma medida padronizada. Uma das principais vantagens do CVV é sua simplicidade de cálculo. Além disso, ele é capaz de capturar todas as variações relevantes nas variâncias dos erros, proporcionando uma visão mais clara da heterocedasticidade.

Para mensurar o CVV na prática, é possível estimá-lo a partir dos resíduos de um modelo de regressão. Primeiramente, calcula-se a variância dos resíduos para cada observação ou grupo de observações. Em seguida, aplica-se a fórmula do CVV, permitindo avaliar a dispersão dessas variâncias. Esse procedimento pode ser implementado em *softwares* estatísticos como R ou Python, facilitando sua aplicação em estudos empíricos.

O uso do CVV permite que os pesquisadores avaliem a heterocedasticidade em diversos contextos de maneira mais consistente e frequente. Song e Taamouti (2021) afirmam que mensurar a heterocedasticidade pode ajudar a entender melhor a relação entre as variáveis, mas destacam que não há uma medida universalmente aceita na literatura. Eles apontam, entretanto, que o CVV, conforme sugerido por Kennedy (1985), é uma alternativa viável. Kennedy e Sudmant (1990) utilizaram o CVV como medida de heterocedasticidade e observaram que ele auxiliou na compreensão de aspectos da inferência estatística na presença de variâncias não constantes, especialmente em contextos sem observações replicadas. Essa abordagem melhora a interpretação dos resultados estatísticos e facilita comparações entre diferentes estudos.

Kennedy (1985) também destaca que o CVV é particularmente útil em modelos de regressão linear, em que a heterocedasticidade pode impactar tanto as estimativas dos coeficientes quanto a precisão das previsões. Ao incorporar o CVV na análise, os pesquisadores

podem ajustar seus modelos com maior precisão, considerando a variabilidade das variâncias dos erros e melhorando a robustez dos resultados.

### 3 METODOLOGIA

Nesta seção, é descrito o processo utilizado para simular os dados e aplicar testes de heterocedasticidade em modelos de regressão linear. A metodologia foi desenvolvida com o objetivo de avaliar o poder dos testes em diferentes cenários, que são:

- Cenário 1:  $\sigma_i^2 = |x_i|\sigma^2$ ,
- Cenário 2:  $\sigma_i^2 = x_i^2\sigma^2$ ,
- Cenário 3:  $\sigma_i^2 = |E(y_i)|\sigma^2$ ,
- Cenário 4:  $\sigma_i^2 = [E(y_i)]^2\sigma^2$ .

A simulação envolveu o ajuste de um modelo com erros não homocedásticos, considerando os valores de  $\sigma_i^2$  de cada cenário. Em seguida, foram aplicados os testes de White, Glejser e Breusch-Pagan, repetindo o processo 1000 vezes para calcular o poder de cada teste. Posteriormente, todo o procedimento foi repetido 20 vezes, sendo calculada a média do poder para cada cenário.

#### 3.1 Simulação

A simulação seguiu uma série de etapas. Inicialmente, foram gerados  $\frac{n}{3} - 1$  valores a partir de uma distribuição normal com média  $k$  e variância 1. Em seguida, adicionou-se um valor  $h$  para ajustar o Coeficiente de Variação da Variância (CVV). Esses  $\frac{n}{3}$  valores foram replicados três vezes, gerando um total de  $n$  observações (Figura 1).

Utilizando os valores de  $x$ , calculou-se  $\sigma_i^2$  para os cenários 1 e 2. A partir do valor esperado de  $y_i$ , dado por  $E(y_i) = a + bx$ , com  $a = 2$  e  $b = 3$ , determinou-se  $\sigma_i^2$  para os cenários 3 e 4. Após essa definição de  $\sigma_i^2$ , calculou-se o CVV correspondente a cada cenário. Se o CVV não alcançasse o valor desejado, ajustava-se o valor de  $h$  e repetiu-se o processo até atingir o CVV especificado. Em cada cenário, foram simulados valores para a variável independente para que o CVV assumisse os valores 0; 0,4; 0,6; 0,9; 1,1; 1,3; 1,5 e 1,7.

Com as variáveis independentes simuladas, gerou-se os erros seguindo uma distribuição normal com média 0 e variância  $\sigma_i^2$ , assegurando que os erros fossem heterocedásticos em cada cenário. Em seguida, a variável dependente  $y$  foi simulada de acordo com a expressão  $y = a + bx + erro$ , utilizando os valores de  $a$ ,  $b$ ,  $x$  (variável independente) e os erros.

Por fim, com as variáveis dependente ( $y$ ) e independente ( $x$ ), o modelo foi ajustado usando a função  $lm$  do pacote  $lmtest$ . Este modelo foi ajustado de forma a garantir que os erros fossem heterocedásticos, permitindo, assim, avaliar o poder dos testes.

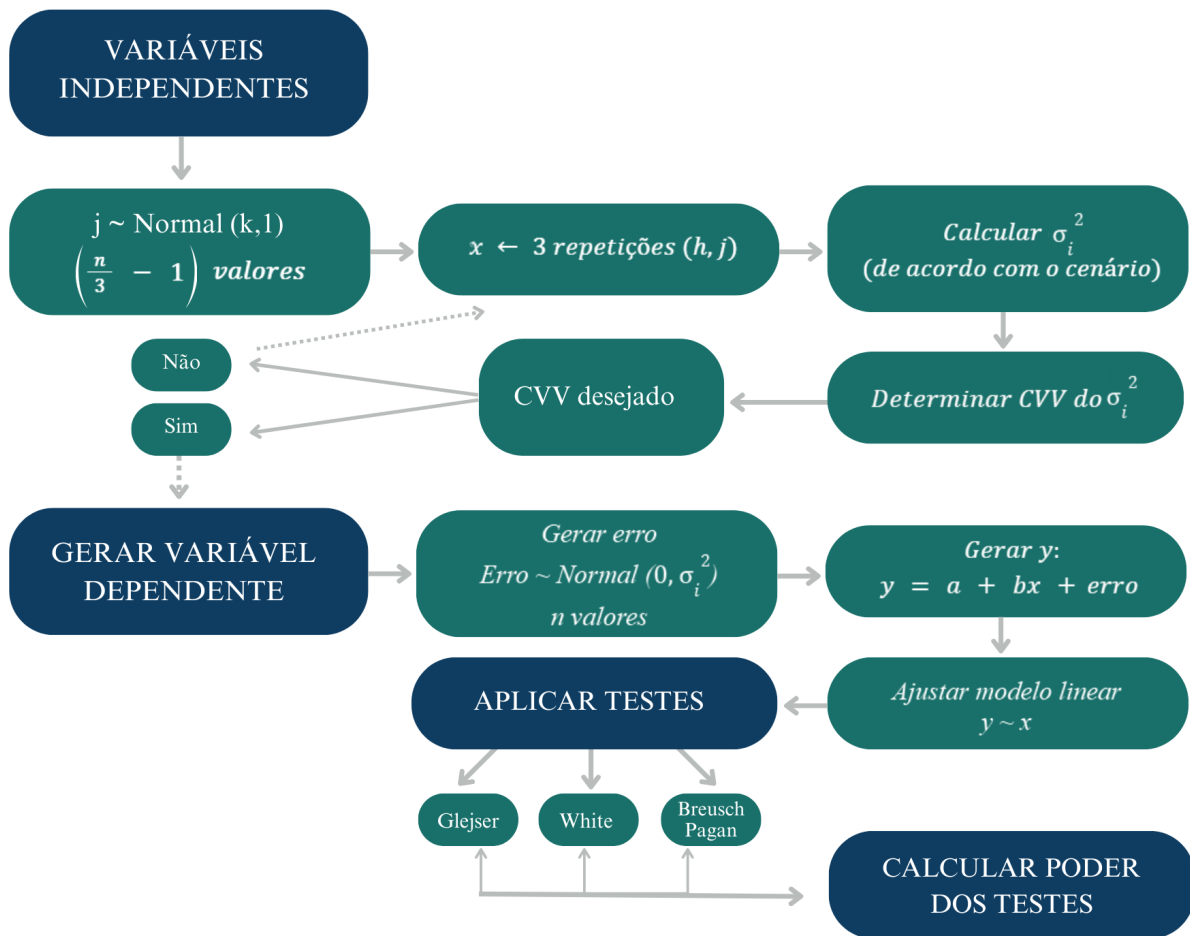


Figura 1: Fluxograma da simulação para a regressão linear e aplicação dos testes.

## 3.2 Testes de heterocedasticidade

### 3.2.1 Teste de White

O teste de White foi utilizado para verificar se há heterocedasticidade nos resíduos de um modelo de regressão linear. Para aplicá-lo, primeiramente estimou-se os coeficientes do modelo usando o método dos mínimos quadrados ordinários.

O próximo passo foi calcular os quadrados dos resíduos estimados do modelo para cada observação. Em seguida, realizou-se uma regressão auxiliar dos quadrados dos resíduos em função dos valores da variável independente  $x$  e do quadrado dessa variável. Essa regressão é representada pela equação:

$$\hat{\varepsilon}_i^2 = \gamma_0 + \gamma_1 x_i + \gamma_2 x_i^2 + u_i,$$

em que  $\gamma_i$  são os coeficientes estimados da regressão auxiliar e  $u_i$  são os erros dessa regressão.

A próxima etapa foi avaliar os coeficientes da regressão auxiliar com base nas hipóteses estatísticas:

$$H_0 : \gamma_1 = \gamma_2 = 0 \text{ (Homocedasticidade),}$$

$$H_1 : \gamma_1 \neq 0 \text{ ou } \gamma_2 \neq 0 \text{ (Heterocedasticidade).}$$

Depois, foi calculado o coeficiente de determinação  $R_{\text{aux}}^2$  para determinar a estatística do teste, dada por:

$$LM = n \cdot R_{\text{aux}}^2,$$

em que  $n$  é o número de observações utilizadas na regressão auxiliar. Sob a hipótese nula de homocedasticidade e assumindo que o modelo original está corretamente especificado, a estatística  $LM$  segue uma distribuição qui-quadrado com 2 graus de liberdade (o número de coeficientes da regressão auxiliar menos um).

Por fim, comparamos o valor de  $LM$  ao valor crítico da distribuição qui-quadrado com 2 graus de liberdade e o nível de significância escolhido. Se  $LM$  for maior que o valor crítico, rejeitamos  $H_0$ , indicando a presença de heterocedasticidade nos resíduos da regressão. Esse processo foi realizado utilizando a função `bptest` do R, especificando que a regressão auxiliar seria estimada com  $x_i$  e  $x_i^2$ .

### 3.2.2 Teste de Glejser

O teste de Glejser procura verificar a relação entre os resíduos da regressão e as variáveis independentes do modelo. O teste tem como hipóteses estatísticas,

$$H_0 : \text{As variâncias são homocedásticas } (\sigma_i^2 = \sigma^2, \forall i = 1, \dots, n),$$

$$H_1 : \text{As variâncias são heterocedásticas } (\sigma_i^2 \neq \sigma^2, \text{ para algum, } i = 1, \dots, n).$$

Ajustou-se uma regressão para avaliar a relação dos resíduos com as variáveis independentes, mas como os valores dos resíduos reais são desconhecidos para aplicar o teste foi necessário estimar por MQO os parâmetros da regressão linear e, em seguida, estimar os resíduos da forma  $\hat{\varepsilon}_i = y_i - \hat{y}_i$ , sendo que, com essa estimativa ajusta-se a regressão,

$$|\hat{\varepsilon}_i| = a + bx_i + \alpha_i,$$

em que  $|\hat{\varepsilon}_i|$  são os valores absolutos dos resíduos,  $x_i$  é o valor da variável independente do modelo original,  $a$  e  $b$  são coeficientes de regressão que foram estimados, e  $u_i$  são os termos de erro da nova regressão.

Após realizar essa regressão, foi testado a significância dos coeficientes, através do

teste  $t$  de Student. Foi analisado se os coeficientes foram estatisticamente significativos, o que indica a presença de heterocedasticidade.

### 3.2.3 Teste de Breusch–Pagan

Para verificar se há heterocedasticidade nos resíduos de um modelo de regressão linear, foi utilizado o teste de Breusch-Pagan (1979) cuja hipótese nula é de homocedasticidade das variâncias dos resíduos da regressão linear, enquanto a hipótese alternativa indica heterocedasticidade,

$H_0$  : As variâncias são homocedásticas ( $\sigma_i^2 = \sigma^2, \forall i = 1, \dots, n$ )

$H_1$  : As variâncias são heterocedásticas ( $\sigma_i^2 \neq \sigma^2$ , para algum  $i = 1, \dots, n$ ).

A aplicação do teste se deu através da análise dos erros da regressão por MQO. Primeiramente, ajustou-se o modelo da forma:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i,$$

e, em seguida, calculou-se os resíduos, representados por  $\hat{\varepsilon}_i$ , que são a diferença entre os valores reais e os valores preditos pelo modelo:

$$\hat{\varepsilon}_i = y_i - \hat{y}_i.$$

O próximo passo foi realizar uma regressão dos resíduos ao quadrado ( $\hat{\varepsilon}_i^2$ ) em relação aos valores da variável independente  $x_i$ , que é chamada de regressão auxiliar. Esta regressão pode ser expressa como:

$$\hat{\varepsilon}_i^2 = \lambda_0 + \lambda_1 x_i + u_i,$$

em que  $\lambda_0$  e  $\lambda_1$  são os coeficientes estimados na regressão auxiliar, e  $u_i$  são os termos de erro da regressão.

Após a regressão dos resíduos ao quadrado, testou-se a significância dos coeficientes  $\lambda_0$  e  $\lambda_1$  com base nas hipóteses estatísticas:

$H_0$  :  $\lambda_1 = 0$  (Homocedasticidade),

$H_1$  :  $\lambda_1 \neq 0$  (Heterocedasticidade).

Da regressão auxiliar, calculou-se o coeficiente de determinação ( $R_{\text{aux}}^2$ ) o que foi

utilizado para calcular a estatística do teste,

$$LM = n \cdot R_{\text{aux}}^2,$$

em que  $n$  é o número de observações usadas para a regressão auxiliar. Sob a hipótese nula, a estatística  $LM$  segue uma distribuição qui-quadrado com 1 grau de liberdade (número de variáveis explicativas na regressão auxiliar).

Por fim, comparou-se o valor de  $LM$  com o valor crítico da distribuição qui-quadrado com 1 grau de liberdade e o nível de significância desejado. Se  $LM$  for maior que o valor crítico, rejeitamos  $H_0$  e conclui-se que há indícios de heterocedasticidade nos resíduos da regressão. Esse teste foi aplicado usando a função `bptest` do pacote `lmtest` (Ilmarinen, 2021) do *software* R (R, 2025)

## 4 RESULTADOS E DISCUSSÃO

A metodologia aplicada permitiu avaliar o desempenho dos testes de detecção de heterocedasticidade em diferentes cenários, fornecendo uma visão confiável sobre o poder de cada teste nas situações simuladas. Ao variar os valores de CVV e as formas de heterocedasticidade, foi possível compreender como esses fatores influenciam a capacidade dos testes em identificar corretamente a presença de heterocedasticidade.

Nas figuras apresentadas (Figuras 2 a 5) há uma linha pontilhada indicando o nível de significância de 0,05. Os pontos para o CVV= 0, podem ser interpretados como sendo a variância constante (homocedasticidade), sendo assim nesse ponto o poder representa a probabilidade de, quando é verdadeira, rejeitar a hipótese nula de homocedasticidade, ou seja, nesse ponto está representado a probabilidade do erro de tipo I de cada teste. Em todos os testes aplicados foi observado o a probabilidade do erro de tipo I sendo próximo de 0,05 (Figuras 2 a 5). O poder dos testes aumenta com o CVV indicando uma maior capacidade de detectar heterocedasticidade quando ela realmente existe. Portanto, a linha pontilhada é um ponto de referência para a probabilidade do erro de tipo I.

No cenário 1, o poder dos testes mostrou-se mais estável à medida que o número de observações aumentou. À medida que o CVV aumentou, observou-se também um incremento no poder dos três testes (Figura 2). Com mais observações, o teste de White destacou-se, apresentando um poder mais alto e comportamento mais estável, aproximando-se de 1. Embora os demais testes também tenham registrado aumento no poder com o aumento das observações, seus valores ainda se mantiveram ligeiramente mais baixos em comparação ao teste de White.

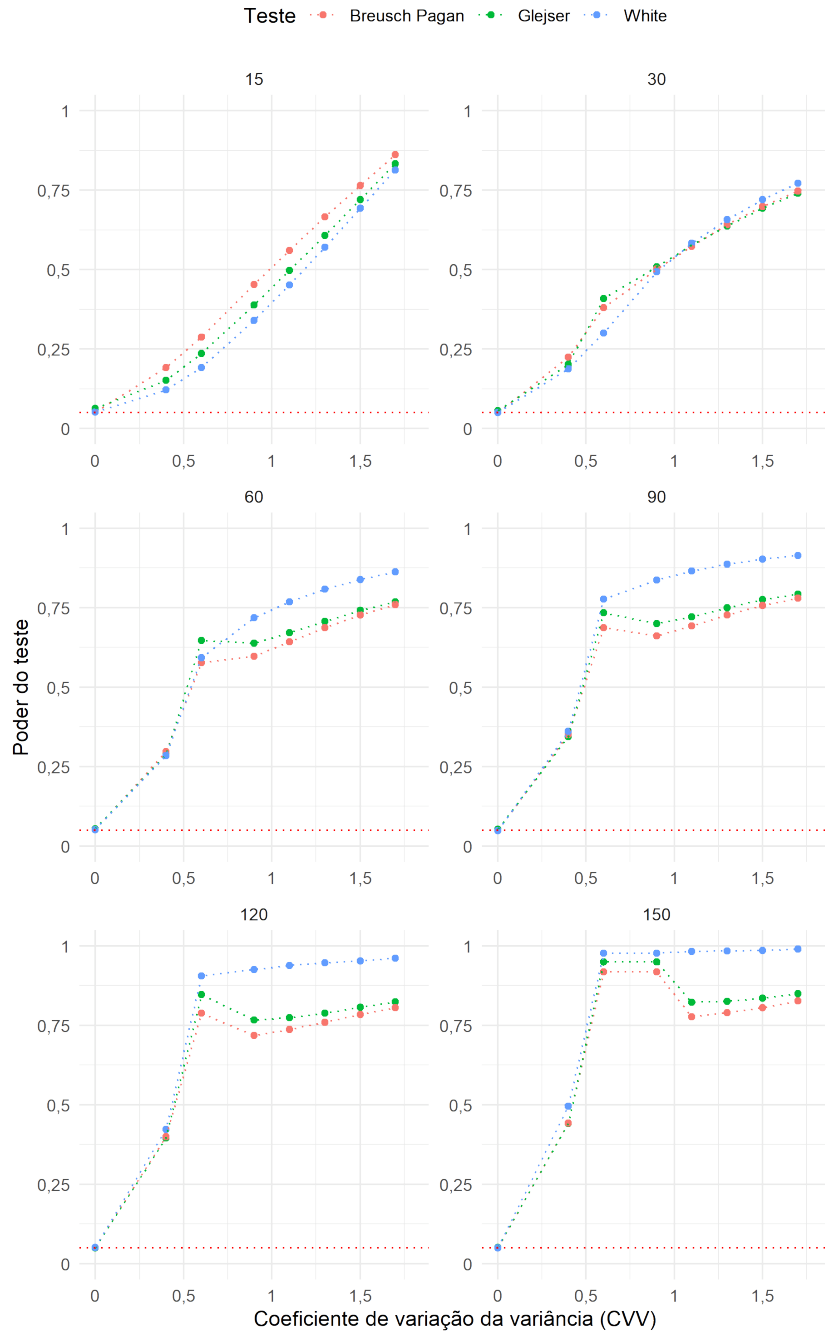


Figura 2: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150) para o cenário 1.

A Figura 3 apresenta o poder em relação ao CVV, para o cenário 2. O poder dos testes variou de acordo com o tamanho amostral. Para  $n = 15$ ,  $n = 30$  e  $n = 60$  o poder não atingiu estabilidade, mas quanto maior o CVV maior o poder. Já para os valores mais altos ( $n = 90$ ,  $n = 120$  e  $n = 150$ ) a medida que o CVV aumenta o poder tende a se estabilizar se

aproximando de 1. O poder dos três testes é muito próximo sendo que para valores maiores de  $n$  o teste de White tem poder ligeiramente maior.

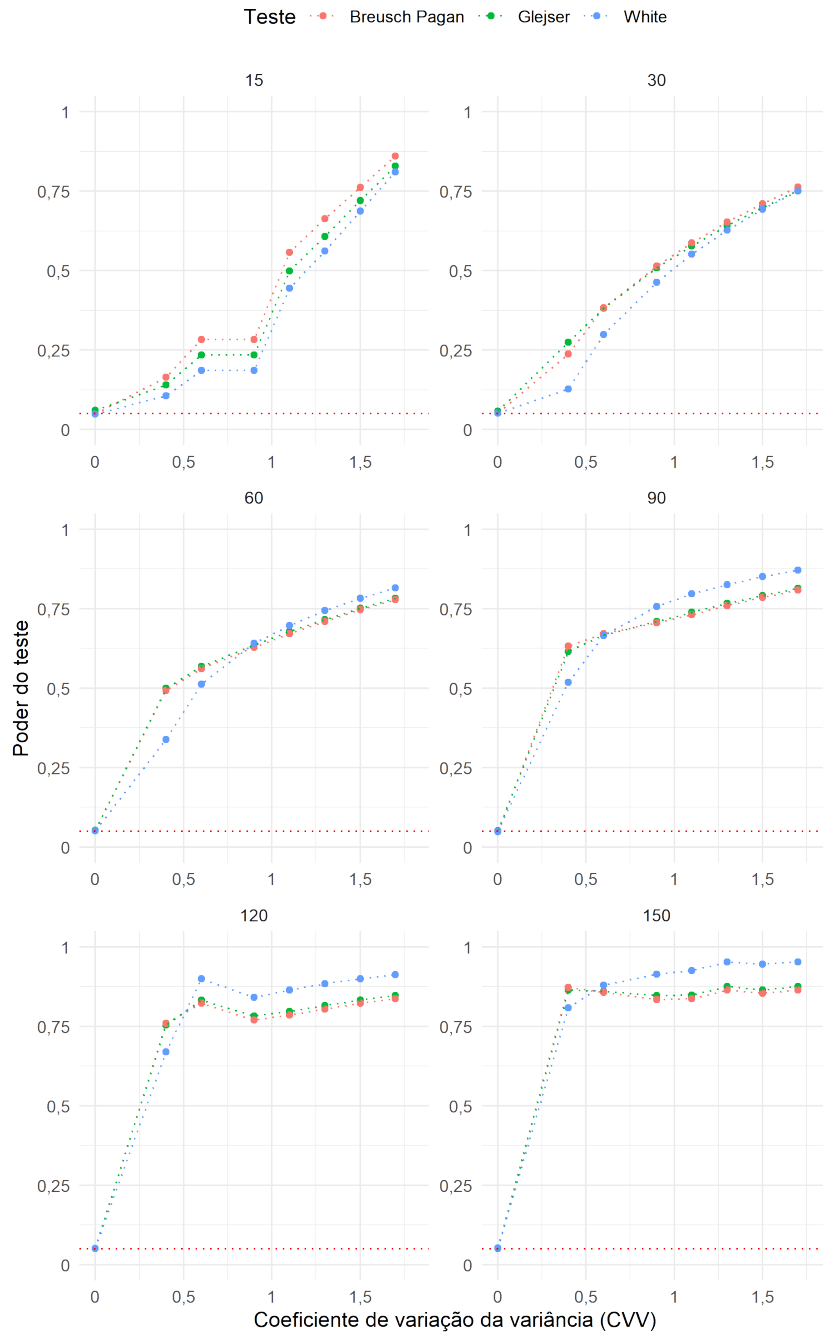


Figura 3: Poder dos testes de Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o cenário 2.

A Figura 4 apresenta os resultados do poder dos testes Glejser, Breusch-Pagan e

White para o cenário 3. Nesse cenário o poder dos três testes tem comportamento similar. A medida que o CVV aumenta o poder também aumentou, para todo o valor de  $n$ , mas para  $n = 15$  e  $n = 30$  a tendência no aumento foi aparentemente linear, para os outros valores teve uma tendência a estabilizar quando o CVV aumentou.

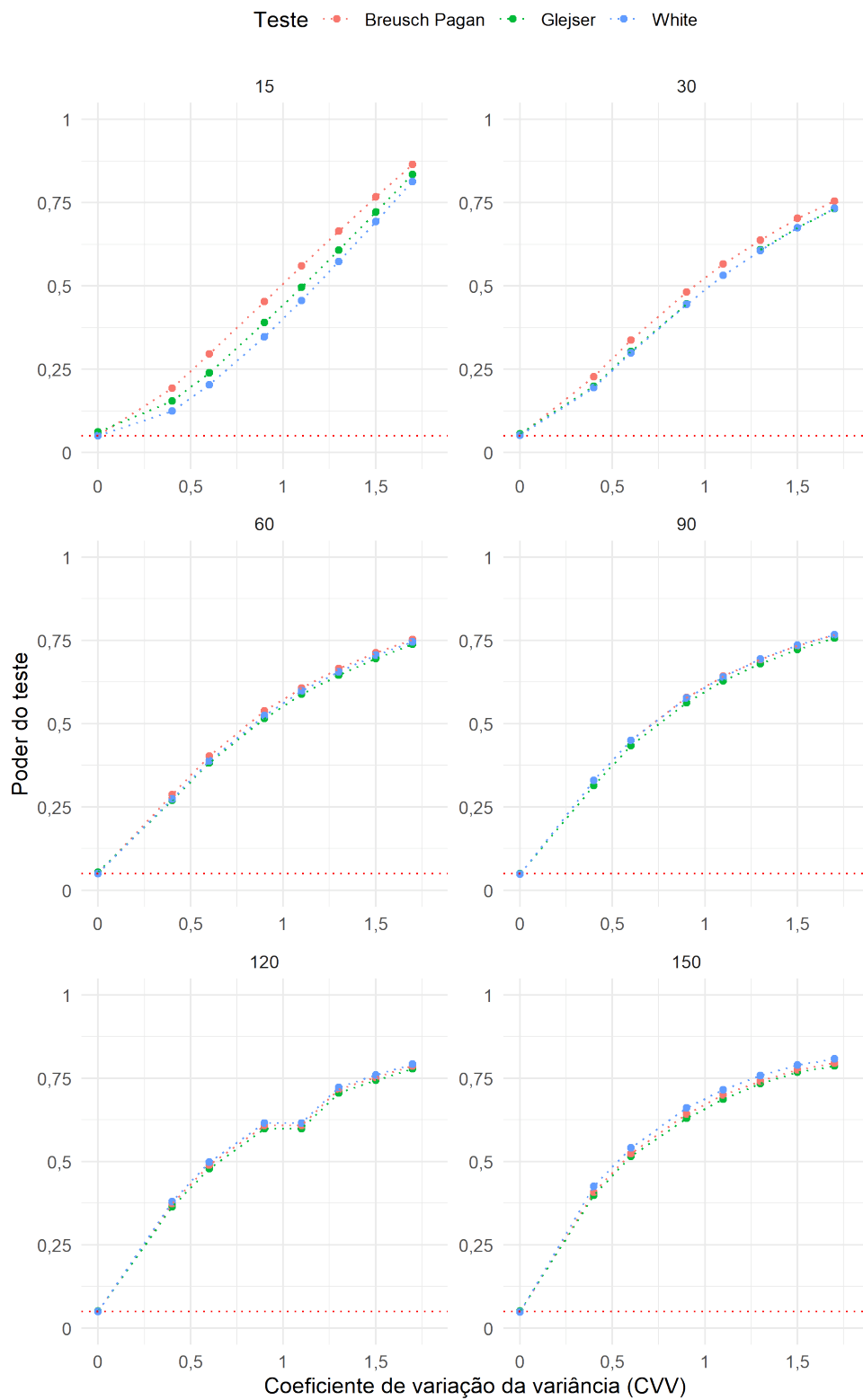


Figura 4: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o cenário 3.

A Figura 5 apresenta os resultados do poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra ( $n = 15, 30, 60, 90, 120, 150$ ) no cenário 4.

Para todos os testes o poder aumenta com o CVV para todos os tamanhos de amostra, sendo que quanto maior o  $n$  mais perto da estabilização próximo de 1 o poder do teste se encontrava.

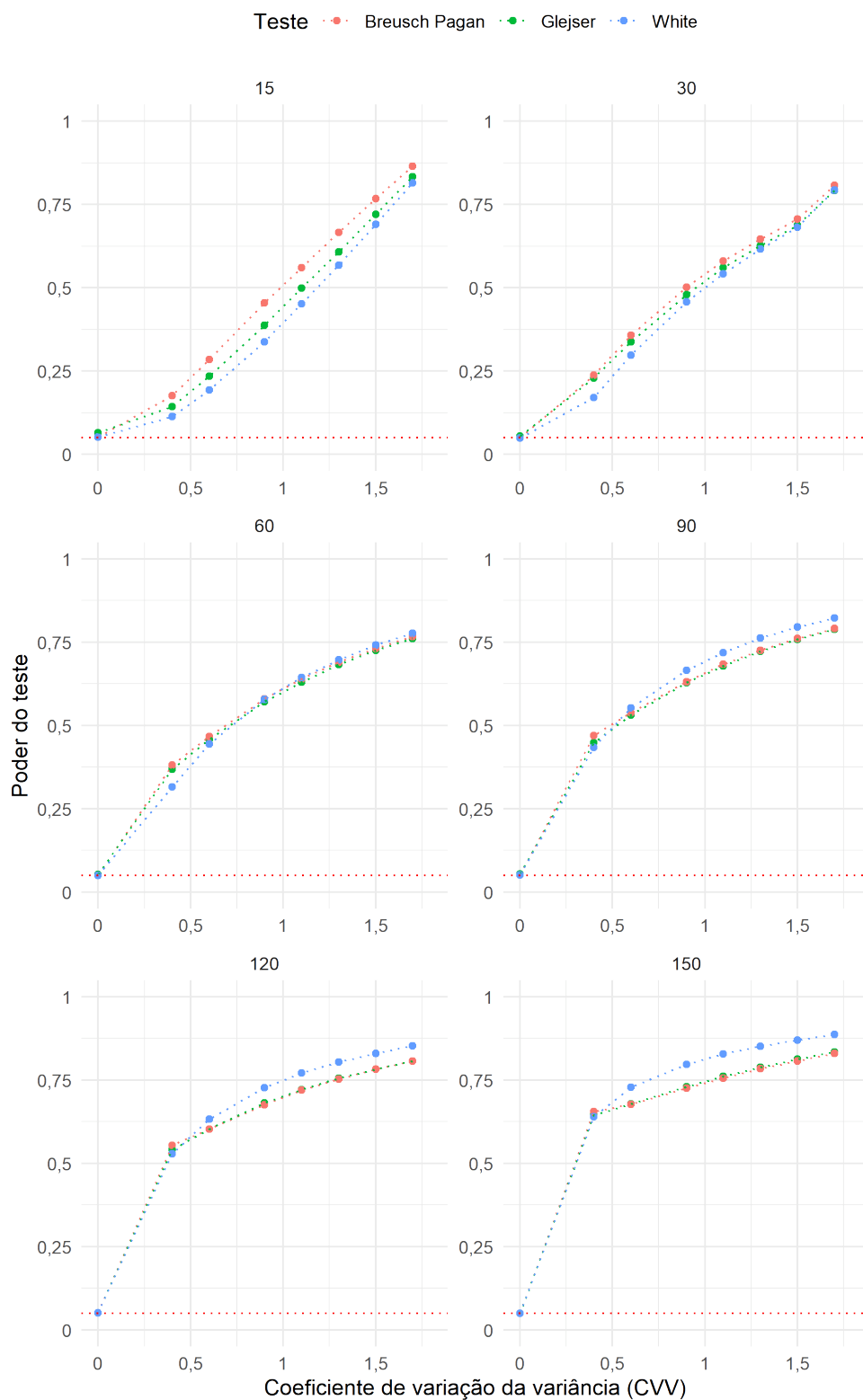


Figura 5: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o cenário 4.

Para todos os testes, o poder aumenta com o aumento do CVV, indicando que os testes se tornam mais eficazes na detecção de heterocedasticidade à medida que a variabilidade dos resíduos aumenta (Figura 2, Figura 3, Figura 4, Figura 5). O aumento do tamanho da amostra também incrementa o poder dos testes, permitindo uma detecção mais confiável de heterocedasticidade. O teste de White tende a mostrar um poder ligeiramente mais alto em comparação com os testes de Glejser e Breusch-Pagan. No entanto, todos os testes são eficazes na detecção de heterocedasticidade conforme o CVV e o tamanho da amostra aumentam. O poder dos testes é mais estável para amostras maiores e permanece alto para diferentes valores de CVV.

## 5 CONCLUSÕES

O estudo realizado demonstrou a eficácia dos testes de Glejser, Breusch-Pagan e White na detecção de heterocedasticidade em resíduos de regressão linear. Percebe-se que todos os testes exibiram boa capacidade de identificar heterocedasticidade conforme o coeficiente de variação das variâncias aumenta. Isso é observado para todos os cenários estudados.

O teste de Glejser, é uma ferramenta que pode ter uma gama de variações interessantes para a detecção da heterocedasticidade, já que é possível analisar diversas formas de relação dos resíduos e as variáveis independentes. Nesse estudo observou-se um aumento no poder do teste à medida que o CVV aumentava.

O teste de Breusch-Pagan apresentou um comportamento similar ao do teste de Glejser, com uma tendência de aumento no poder conforme o CVV cresce. Este teste mostrou-se consistente em sua capacidade de detectar heterocedasticidade através de diferentes tamanhos de amostra, indicando que ele é uma ferramenta sólida para análises estatísticas em diferentes contextos de pesquisa.

O teste de White mostrou um poder levemente superior em comparação aos outros testes, especialmente em CVV's elevados. Esse resultado indica que o teste de White pode ser mais sensível à presença de heterocedasticidade em situações onde a variabilidade dos resíduos é substancial. Uyanto (2022) encontrou resultados diferentes, sugerindo que o teste de White apresentava menor poder para alguns cenários análogos. Este aspecto destaca a importância de considerar as características específicas dos dados e do contexto ao escolher o teste mais apropriado para a análise.

Em termos gerais, todos os testes estudados exibiram um aumento no poder à medida que o CVV cresce, o que indica que eles se tornam mais eficazes na detecção de heterocedasticidade com o aumento da variabilidade dos resíduos. Além disso, o aumento do tamanho da amostra também melhora o poder dos testes, proporcionando uma detecção mais confiável de heterocedasticidade.

## REFERÊNCIAS

- BOX, G. E. P. Some theorems on quadratic forms applied in the study of analysis of variance problems, I. Effect of inequality of variance in the one-way classification. **The annals of mathematical statistics**, v. 25, n. 2, p. 290–302, 1954.
- BREUSCH, T. S.; PAGAN, A. R. A simple test for heteroscedasticity and random coefficient variation. **Econometrica**, v. 47, n. 5, p. 1287–1294, 1979.
- DUTTA, A. R.; DAS, S.; DEBNATH, B. A non-destructive approach to develop tree-level allometric equations for estimating aboveground biomass in the forests of Tripura, Northeast India. **Tropical Ecology**, v. 64, n. 3, p. 532-542 2022.
- GAUSS, C. F. **Theoria motus corporum coelestium in sectionibus conicis solem ambientium**. Boston, 326 p. 1809.
- GLEJSER, H. A new test for heteroskedasticity. **Journal of the american statistical association**, v. 64, n. 325, p. 316–323, 1969.
- GUJARATI, D. N. **Econometria básica**. 3<sup>a</sup> ed. São Paulo: Makron Books. 2000.
- HOFFMANN, R.; VIEIRA, S. **Análise de regressão: uma introdução à econometria**. 3<sup>a</sup> ed. São Paulo: Editora Hucitec. 1998.
- HOFFMANN, R. **Análise de regressão: uma introdução à econometria**, 2016.
- ILMARINEN, V. J. (2021). **multid**: multivariate difference between two groups. <https://CRAN.R-project.org/package=multid>
- KENNEDY, P. E. A suggestion for measuring heteroskedasticity. **Communications in statistics - Simulation and computation**, v. 14, n. 4, p. 845–851, 1985.
- MORETTIN, P. A ; BUSSAB W. O. **Estatística básica**. [s.l.]: Saraiva, 2017.
- R Core Team (2025). **R**: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. Disponível em: <https://www.R-project.org> (Acessado em 18 de março de 2024).
- SARI, B. G.; LÚCIO, A. D.; SANTANA, C. S.; OLIVOTO, T.; DIEL, M.; KRYSCZUN, D. K. Nonlinear growth models: an alternative to ANOVA in tomato trials evaluation. **European journal of agronomy**, v. 104, p. 21–36, 2019.
- SONG, X.; TAAMOUTI, A. A nonparametric measure of heteroskedasticity. **Journal of statistical planning and inference**, v. 212, p. 45–68, 2021. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378375820301105?via%3Dihub>. Acesso em: 3 jun.2022.
- SUDMANT, W. ; KENNEDY, P. On inference in the presence of heteroskedasticity without replicated observations. **Communications in Statistics - Simulation and Computation**, v. 19, n. 2, p. 491–504, 1990.

UYANTO, S. S. Monte Carlo power comparison of seven most commonly used heteroscedasticity tests. **Communications in Statistics - Simulation and Computation**, v. 51, n. 4, p. 2065–2082, 2019.

WHITE, H. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. **Econometrica**. v. 48. n. 4, p. 817, 1980.

YU, Y.; CLARK, J. S.; TIAN, Q.; YAN, F. Rice yield response to climate and price policy in high-latitude regions of China. **Food Security**, v. 14, n. 5, p. 1143-1157, 2022.

## APÊNDICE

	<b>Página</b>
<b>APÊNDICE A:</b> Script para determinar o poder dos testes de Breusch-Pagan, Glejser e White para $CVV = 0,6$ .....	39
<b>APÊNDICE B:</b> <i>Script</i> para determinar os valores de $h$ para $n = 150$ .....	48
<b>APÊNDICE C:</b> Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 1 ...	50
<b>APÊNDICE D:</b> Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 2 ...	51
<b>APÊNDICE E:</b> Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 3 ...	52
<b>APÊNDICE F:</b> Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 4 ...	53

## APÊNDICE A: *Script* para determinar o poder dos testes de Breusch-Pagan, Glejser e White para $CVV = 0,6$

```
rm(list = ls())
library(lmtest)
library(dplyr)
library(openxlsx)
# Define o diretório de trabalho
setwd("C:/Users/jonas/Dropbox/R/Script/Heterocedasticidade")
set.seed(8)

getwd()

tempoi=proc.time()
#cvv = 0.6
# Definir valores de n, j, k e alpha
valores_n <- c(seq(15, 30, 15), seq(60, 150, 30))
j <- 1000 # Número de repetições
alpha <- 0.05 # Nível de significância
k <- 20 # Número de vezes que você deseja repetir o processo
c <- 2
b <- 3
a <- -2
# Inicializar tabela para armazenar resultados
resultados1 <- data.frame(
  M = character(),
  N = integer(),
  J = integer(),
  glejser_poder = double(),
  bp_poder = double(),
  White_poder = double()
)

# Definindo os valores dos vetores x
sigma2 <- 2 # Definindo sigma2
```

```

xe15 <- rep(c(10.351256, 2.187070, 3.996029, 4.299370, 3.127458)
            /(sigma2), 3)

xe30 <- rep(c(7.2216480, 2.1870696, 3.9960295, 4.2993699, 3.1274577,
              3.0811140, 2.5691663, 1.0928952, 2.0298989, 0.4409717)
            /(sigma2), 3)

xe60 <- rep(c(7.452527, 2.1870696, 3.9960295, 4.2993699, 3.1274577,
              3.0811140, 2.5691663, 1.0928952, 2.0298989, 0.4409717,
              2.2974868, 0.3911907, 2.0275031, 0.7517841, 2.4503220,
              2.4927427, 2.5491113, 3.4418546, 2.3631123, 1.7651367)
            /(sigma2),3)

xe90 <- rep(c(7.9667480, 2.1870696, 3.9960295, 4.2993699, 3.1274577,
              3.0811140, 2.5691663, 1.0928952, 2.0298989, 0.4409717,
              2.2974868, 0.3911907, 2.0275031, 0.7517841, 2.4503220,
              2.4927427, 2.5491113, 3.4418546, 2.3631123, 1.7651367,
              2.6601284, 1.1501214, 2.4462922, 1.2575124, 1.6720574,
              1.4028931, 1.4763645, 2.4069560, 1.5756891, 2.2414025)
            /(sigma2), 3)

xe120 <- rep(c(8.7581330, 2.1870696, 3.9960295, 4.2993699, 3.1274577,
               3.0811140, 2.5691663, 1.0928952, 2.0298989, 0.4409717,
               2.2974868, 0.3911907, 2.0275031, 0.7517841, 2.4503220,
               2.4927427, 2.5491113, 3.4418546, 2.3631123, 1.7651367,
               2.6601284, 1.1501214, 2.4462922, 1.2575124, 1.6720574,
               1.4028931, 1.4763645, 2.4069560, 1.5756891, 2.2414025,
               3.0981025, 1.1944513, 0.2197650, 3.6317085, 2.8502559,
               2.5206710, 1.5001600, 2.1468230, 2.8973434, 2.1052041)
            /(sigma2), 3)

xe150 <- rep(c(8.4740450, 2.1870696, 3.9960295, 4.2993699, 3.1274577,
               3.0811140, 2.5691663, 1.0928952, 2.0298989, 0.4409717,
               2.2974868, 0.3911907, 2.0275031, 0.7517841, 2.4503220,
               2.4927427, 2.5491113, 3.4418546, 2.3631123, 1.7651367,
               2.6601284, 1.1501214, 2.4462922, 1.2575124, 1.6720574,

```

```

1.4028931, 1.4763645, 2.4069560, 1.5756891, 2.2414025,
3.0981025, 1.1944513, 0.2197650, 3.6317085, 2.8502559,
2.5206710, 1.5001600, 2.1468230, 2.8973434, 2.1052041,
2.3961597, 1.4864547, 1.5891875, 1.2621406, 1.0316954,
1.6742230, 0.3235675, 2.7837089, 3.2076640, 0.5703912)
/(sigma2), 3)

xf15 <- rep(x = c(11.25142, 5.187070, 6.996029, 7.299370, 6.127458)
/sqrt(sigma2), 3)
xf30 <- rep(c(9.877336, 5.187070, 6.996029, 7.299370, 6.127458,
6.081114, 5.569166, 4.092895, 5.029899, 3.440972)
/sqrt(sigma2), 3)
xf60 <- rep(c(10.21733, 5.187070, 6.996029, 7.299370, 6.127458,
6.081114, 5.569166, 4.092895, 5.029899, 3.440972,
5.297487, 3.391191, 5.027503, 3.751784, 5.450322,
5.492743, 5.549111, 6.441855, 5.363112, 4.765137)
/sqrt(sigma2), 3)
xf90 <- rep(c(10.564826, 5.187070, 6.996029, 7.299370, 6.127458,
6.081114, 5.569166, 4.092895, 5.029899, 3.440972,
5.297487, 3.391191, 5.027503, 3.751784, 5.450322,
5.492743, 5.549111, 6.441855, 5.363112, 4.765137,
5.660128, 4.150121, 5.446292, 4.257512, 4.672057,
4.402893, 4.476364, 5.406956, 4.575689, 5.241402)
/sqrt(sigma2), 3)
xf120 <- rep(c(19.1439, 5.187070, 6.996029, 7.299370, 6.127458,
6.081114, 5.569166, 4.092895, 5.029899, 3.440972,
5.297487, 3.391191, 5.027503, 3.751784, 5.450322,
5.492743, 5.549111, 6.441855, 5.363112, 4.765137,
5.660128, 4.150121, 5.446292, 4.257512, 4.672057,
4.402893, 4.476364, 5.406956, 4.575689, 5.241402,
6.098103, 4.194451, 3.219765, 6.631708, 5.850256,
5.520671, 4.500160, 5.146823, 5.897343, 5.105204)
/sqrt(sigma2), 3)

xf150 <- rep(c(11.245232, 5.187070, 6.996029, 7.299370, 6.127458,
6.081114, 5.569166, 4.092895, 5.029899, 3.440972,

```

```

5.297487, 3.391191, 5.027503, 3.751784, 5.450322,
5.492743, 5.549111, 6.441855, 5.363112, 4.765137,
5.660128, 4.150121, 5.446292, 4.257512, 4.672057,
4.402893, 4.476364, 5.406956, 4.575689, 5.241402,
6.098103, 4.194451, 3.219765, 6.631708, 5.850256,
5.520671, 4.500160, 5.146823, 5.897343, 5.105204,
5.396160, 4.486455, 4.589187, 4.262141, 4.031695,
4.674223, 3.323568, 5.783709, 6.207664, 3.570391)
/sqrt(sigma2), 3)

xg15 <- rep(c(26.272780, 7.187070, 8.996029, 9.299370, 8.127458)
/(b*(sigma2)) - (c/b), 3)
xg30 <- rep(c(25.89919, 7.187070, 8.996029, 9.299370, 8.127458,
8.081114, 7.569166, 6.092895, 7.029899, 5.440972)
/(b*(sigma2)) - (c/b), 3)
xg60 <- rep(c(29.967670, 7.187070, 8.996029, 9.299370, 8.127458,
8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137)
/(b*(sigma2)) - (c/b), 3)
xg90 <- rep(c(33.376780, 7.187070, 8.996029, 9.299370, 8.127458,
8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137,
7.660128, 6.150121, 7.446292, 6.257512, 6.672057,
6.402893 , 6.476364, 7.406956, 6.575689, 7.241402)
/(b*(sigma2)) - (c/b), 3)
xg120 <- rep(c(36.890810, 7.187070, 8.996029, 9.299370, 8.127458,
8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137,
7.660128, 6.150121, 7.446292, 6.257512, 6.672057 ,
6.402893 , 6.476364, 7.406956, 6.575689, 7.241402,
8.098103, 6.194451, 5.219765, 8.631708 , 7.850256,
7.520671, 6.500160, 7.146823, 7.897343, 7.105204)
/(b*(sigma2)) - (c/b), 3)

```

```

xg150 <- rep(c(39.300119, 7.187070, 8.996029, 9.299370, 8.127458,
              8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
              7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
              7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137,
              7.660128, 6.150121, 7.446292, 6.257512, 6.672057 ,
              6.402893 , 6.476364, 7.406956, 6.575689, 7.241402,
              8.098103, 6.194451, 5.219765, 8.631708 , 7.850256,
              7.520671, 6.500160, 7.146823, 7.897343, 7.105204,
              7.396160, 6.486455 , 6.589187, 6.262141, 6.031695,
              6.674223, 5.323568, 7.783709, 8.207664, 5.570391)
          /(b*(sigma2)) - (c/b), 3)

xh15 <- rep(c(14.820002, 7.187070, 8.996029, 9.299370, 8.127458)
          /(b*sqrt(sigma2)) - (c/b), 3)
xh30 <- rep(c(13.76482, 7.187070, 8.996029, 9.299370, 8.127458,
              8.081114, 7.569166, 6.092895, 7.029899, 5.440972)
          /(b*sqrt(sigma2)) - (c/b), 3)
xh60 <- rep(c(14.50489, 7.187070, 8.996029, 9.299370, 8.127458,
              8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
              7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
              7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137)
          /(b*sqrt(sigma2)) - (c/b), 3)
xh90 <- rep(c(15.14005, 7.187070, 8.996029, 9.299370, 8.127458,
              8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
              7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
              7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137,
              7.660128, 6.150121, 7.446292, 6.257512, 6.672057 ,
              6.402893 , 6.476364, 7.406956, 6.575689, 7.241402)
          /(b*sqrt(sigma2)) - (c/b), 3)
xh120 <- rep(c(15.93338, 7.187070, 8.996029, 9.299370, 8.127458,
              8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
              7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
              7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137,
              7.660128, 6.150121, 7.446292, 6.257512, 6.672057 ,
              6.402893 , 6.476364, 7.406956, 6.575689, 7.241402,
              8.098103, 6.194451, 5.219765, 8.631708 , 7.850256,

```

```

7.520671, 6.500160, 7.146823, 7.897343, 7.105204)
/(b*sqrt(sigma2)) - (c/b), 3)
xh150 <- rep(c(16.281121, 7.187070, 8.996029, 9.299370, 8.127458,
8.081114, 7.569166, 6.092895, 7.029899, 5.440972,
7.297487, 5.391191, 7.027503, 5.751784, 7.450322,
7.492743, 7.549111 ,8.441855 , 7.363112, 6.765137,
7.660128, 6.150121, 7.446292, 6.257512, 6.672057 ,
6.402893 , 6.476364, 7.406956, 6.575689, 7.241402,
8.098103, 6.194451, 5.219765, 8.631708 , 7.850256,
7.520671, 6.500160, 7.146823, 7.897343, 7.105204,
7.396160, 6.486455 , 6.589187, 6.262141, 6.031695,
6.674223, 5.323568, 7.783709, 8.207664, 5.570391)
/(b*sqrt(sigma2)) - (c/b), 3)

# Listas de vetores x para cada m
x_list <- list(
  e = list(xe15, xe30, xe60, xe90, xe120, xe150),
  f = list(xf15, xf30, xf60, xf90, xf120, xf150),
  g = list(xg15, xg30, xg60, xg90, xg120, xg150),
  h = list(xh15, xh30, xh60, xh90, xh120, xh150)
)

# Loop para diferentes valores de m e n
for (m in names(x_list)) {
  for (index in seq_along(valores_n)) {
    n <- valores_n[index]
    x <- x_list[[m]][[index]]

    if (length(x) == 0) next # Pula se o vetor x estiver vazio

    media_contador1 <- 0
    media_contador2 <- 0
    media_contador3 <- 0
    media_contador4 <- 0 # Para gqtest
    media_contador5 <- 0 # Para hmctest
  }
}

```

```

for (rep in 1:k) {
  contador1 <- 0
  contador2 <- 0
  contador3 <- 0
  contador4 <- 0 # Para gqtest
  contador5 <- 0 # Para hmctest

  for (i in 1:j) {

    if (m == "e") {
      erro <- rnorm(n, 0, sd = sqrt(2*abs(x)))
    } else if (m == "f") {
      erro <- rnorm(n, 0, sd = sqrt(2*x^2))
    } else if (m == "g") {
      y_esp <- a + b * x
      erro <- rnorm(n, 0, sd = sqrt(abs(2*y_esp)))
    } else if (m == "h") {
      y_esp <- a + b * x
      erro <- rnorm(n, 0, sd = sqrt(2*y_esp^2))
    }

    y <- a + b * x + erro
    modelo <- lm(y ~ x)

    # Teste Glejser
    uhat <- modelo$residuals
    uhat.abs <- abs(uhat)
    reg.glejser <- lm(uhat.abs ~ x)
    p_value <- summary(reg.glejser)$coefficients["x", "Pr(>|t|)"]

    # Teste Breusch-Pagan
    bp.het <- bptest(modelo, studentize = TRUE)
    bp_pvalue <- bp.het$p.value

    # Teste White
    white <- bptest(modelo, ~ x + I(x^2))
  }
}

```

```

white_pvalue <- white$p.value

# Atualizar contadores
if (p_value > alpha) {
  contador1 <- contador1 + 1
}
if (bp_pvalue > alpha) {
  contador2 <- contador2 + 1
}
if (white_pvalue > alpha) {
  contador3 <- contador3 + 1
}

}

media_contador1 <- media_contador1 + contador1
media_contador2 <- media_contador2 + contador2
media_contador3 <- media_contador3 + contador3

}

# Calcular médias dos contadores
media_contador1 <- media_contador1 / k
media_contador2 <- media_contador2 / k
media_contador3 <- media_contador3 / k

# Calcular o poder do teste
glejser_poder <- 1 - media_contador1 / j
bp_poder <- 1 - media_contador2 / j
White_poder <- 1 - media_contador3 / j

# Armazenar resultados na tabela
resultados1 <- rbind(resultados1, data.frame(

```

```
M = m,  
N = n,  
J = j,  
glejser_poder = glejser_poder,  
bp_poder = bp_poder,  
White_poder = White_poder  
))  
}  
}  
  
# Exibir tabela de resultados  
print(resultados1)  
  
write.xlsx(resultados1, file = "resultados1.xlsx")  
  
tempof=proc.time()  
tempof - tempoi
```

## APÊNDICE B: *Script* para determinar os valores de $h$ $n = 150$

```
rlibrary("pacman")
p_load("multid")
set.seed(123)

# Definir os valores iniciais
n = 150
sigma2 = 2
a = 50
c = 2
b = 3
targets_cvv <- c( 0.4) # CVVs alvo
tol <- 0.0000001
j <- c(7.187070, 8.996029, 9.299370, 8.127458, 8.081114,
       7.569166, 6.092895, 7.029899, 5.440972, 7.297487,
       5.391191, 7.027503, 5.751784, 7.450322, 7.492743,
       7.549111 ,8.441855 , 7.363112, 6.765137, 7.660128,
       6.150121, 7.446292, 6.257512, 6.672057 , 6.402893 ,
       6.476364, 7.406956, 6.575689, 7.241402, 8.098103,
       6.194451, 5.219765, 8.631708 , 7.850256, 7.520671,
       6.500160, 7.146823, 7.897343, 7.105204, 7.396160,
       6.486455 , 6.589187, 6.262141, 6.031695, 6.674223,
       5.323568, 7.783709, 8.207664, 5.570391)

# Lista para armazenar os resultados
results <- list()

# Iterar sobre os valores de CVV alvo
for (target_cvv in targets_cvv) {
  cvv <- 0
  i <- 0

  # Iterar sobre os valores de x1
```

```

for (x1 in seq(8,50, by = 0.000001)) {

  x = c(x1, j) / (b*sqrt(sigma2)) - (c/b)
  y_esp <- c + b * x

  cvv=cvv_manual(sample_sizes=c(n/a,n/a,n/a,n/a,n/a,n/a,
  n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,
  n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,
  n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,n/a,
  n/a,n/a),
                variances=c(sigma2*(y_esp)^2))[[2]]

  if (abs(cvv - target_cvv) < tol) {
    results[[as.character(target_cvv)]] <- list(cvv =
    cvv,xa = c(x1, j), x1 = x1)
    break # Sai do loop se a diferença for menor que a tolerância
  }
}
}

# Exibir os resultados
results

```

## APÊNDICE C: Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 1

Tabela 1: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150) para o Cenário 1.

		Cenário 1 ( $\sigma_i^2 = \sigma^2 x_i $ )								
		Coeficiente de variação da variância (CVV)								
		$n$	0	0,4	0,6	0,9	1,1	1,3	1,5	1,7
Teste de Glejser	15	0,0627	0,1509	0,2360	0,3894	0,4977	0,6077	0,7200	0,8331	
	30	0,0563	0,2010	0,4088	0,5087	0,5773	0,6378	0,6933	0,7404	
	60	0,0536	0,2866	0,6467	0,6388	0,6720	0,7072	0,7414	0,7686	
	90	0,0527	0,3446	0,7348	0,7000	0,7220	0,7496	0,7755	0,7925	
	120	0,0490	0,3950	0,8462	0,7665	0,7737	0,7887	0,8072	0,8245	
	150	0,0504	0,4415	0,9493	0,9493	0,8231	0,8260	0,8350	0,8500	
Teste de Breusch-Pagan	15	0,0516	0,1914	0,2874	0,4529	0,5600	0,6664	0,7649	0,8619	
	30	0,0497	0,2248	0,3799	0,5010	0,5734	0,6418	0,6995	0,7475	
	60	0,0520	0,2968	0,5774	0,5966	0,6431	0,6864	0,7266	0,7595	
	90	0,0478	0,3512	0,6877	0,6618	0,6929	0,7270	0,7574	0,7803	
	120	0,0513	0,3990	0,7888	0,7179	0,7375	0,7601	0,7834	0,8054	
	150	0,0500	0,4423	0,9182	0,9182	0,7763	0,7897	0,8057	0,8273	
Teste de White	15	0,0516	0,1223	0,1910	0,3403	0,4521	0,5709	0,6929	0,8138	
	30	0,0497	0,1869	0,2996	0,4924	0,5825	0,6573	0,7197	0,7717	
	60	0,0520	0,2837	0,5921	0,7189	0,7682	0,8089	0,8384	0,8633	
	90	0,0478	0,3620	0,7768	0,8372	0,8661	0,8865	0,9032	0,9138	
	120	0,0513	0,4220	0,9057	0,9259	0,9379	0,9467	0,9531	0,9610	
	150	0,0500	0,4949	0,9773	0,9773	0,9821	0,9842	0,9856	0,9892	

## APÊNDICE D: Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 2

Tabela 2: Poder dos testes de Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o Cenário 2.

		Cenário 2 ( $\sigma_i^2 = \sigma^2 x_i^2$ )							
		Coeficiente de variação da variância (CVV)							
	$n$	0	0,4	0,6	0,9	1,1	1,3	1,5	1,7
Teste de Glejser	15	0,0599	0,1408	0,2350	0,2350	0,4983	0,6067	0,7205	0,8288
	30	0,0566	0,2738	0,3812	0,5081	0,5771	0,6411	0,6990	0,7520
	60	0,0533	0,5001	0,5692	0,6363	0,6781	0,7162	0,7510	0,7826
	90	0,0514	0,6161	0,6672	0,7101	0,7385	0,7669	0,7917	0,8142
	120	0,0504	0,7541	0,8319	0,7832	0,7972	0,8160	0,8319	0,8473
	150	0,0510	0,8640	0,8595	0,8472	0,8488	0,8747	0,8642	0,8747
Teste de Breusch-Pagan	15	0,0485	0,1641	0,2836	0,2836	0,5571	0,6630	0,7619	0,8607
	30	0,0509	0,2371	0,3830	0,5139	0,5875	0,6529	0,7107	0,7630
	60	0,0510	0,4933	0,5621	0,6285	0,6718	0,7105	0,7472	0,7788
	90	0,0482	0,6326	0,6710	0,7059	0,7315	0,7600	0,7855	0,8083
	120	0,0505	0,7602	0,8219	0,7697	0,7853	0,8043	0,8219	0,8375
	150	0,0521	0,8727	0,8563	0,8346	0,8373	0,8642	0,8540	0,8642
Teste de White	15	0,0485	0,1056	0,1855	0,1855	0,4450	0,5617	0,6871	0,8105
	30	0,0509	0,1273	0,2986	0,4627	0,5515	0,6268	0,6927	0,7506
	60	0,0510	0,3381	0,5128	0,6415	0,6978	0,7439	0,7826	0,8156
	90	0,0482	0,5180	0,6651	0,7568	0,7968	0,8262	0,8516	0,8714
	120	0,0505	0,6697	0,9002	0,8407	0,8643	0,8845	0,9002	0,9128
	150	0,0521	0,8081	0,8800	0,9144	0,9259	0,9526	0,9458	0,9526

## APÊNDICE E: Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 3

Tabela 3: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o Cenário 3.

		Cenário 3 ( $\sigma_i^2 = \sigma^2 E(y_i) $ )								
		Coeficiente de variação da variância (CVV)								
		$n$	0	0,4	0,6	0,9	1,1	1,3	1,5	1,7
Teste de Glejser	15	0,0615	0,1548	0,2393	0,3905	0,4953	0,6075	0,7212	0,8349	
	30	0,0558	0,1994	0,3037	0,4452	0,5296	0,6082	0,6744	0,7310	
	60	0,0547	0,2696	0,3822	0,5162	0,5882	0,6457	0,6962	0,7382	
	90	0,0503	0,3135	0,4346	0,5629	0,6282	0,6806	0,7224	0,7578	
	120	0,0508	0,3646	0,4785	0,5980	0,5980	0,7055	0,7444	0,7783	
	150	0,0505	0,3988	0,5150	0,6298	0,6875	0,7336	0,7676	0,7872	
Teste de Breusch-Pagan	15	0,0507	0,1937	0,2955	0,4532	0,5608	0,6652	0,7677	0,8648	
	30	0,0514	0,2271	0,3367	0,4820	0,5662	0,6374	0,7027	0,7544	
	60	0,0504	0,2877	0,4032	0,5379	0,6070	0,6655	0,7121	0,7533	
	90	0,0485	0,3306	0,4497	0,5778	0,6424	0,6935	0,7339	0,7672	
	120	0,0502	0,3744	0,4892	0,6087	0,6087	0,7164	0,7525	0,7872	
	150	0,0480	0,4075	0,5240	0,6429	0,6993	0,7411	0,7752	0,7949	
Teste de White	15	0,0507	0,1240	0,2029	0,3474	0,4556	0,5734	0,6932	0,8138	
	30	0,0514	0,1940	0,2990	0,4449	0,5316	0,6061	0,6739	0,7332	
	60	0,0504	0,2747	0,3891	0,5247	0,5984	0,6562	0,7061	0,7456	
	90	0,0485	0,3296	0,4502	0,5770	0,6404	0,6940	0,7350	0,7669	
	120	0,0502	0,3796	0,4989	0,6158	0,6158	0,7222	0,7596	0,7932	
	150	0,0480	0,4253	0,5408	0,6618	0,7158	0,7584	0,7898	0,8088	

## APÊNDICE F: Poder dos testes Glejser, Breusch-Pagan e White para o Cenário 4

Tabela 4: Poder dos testes Glejser, Breusch-Pagan e White para diferentes coeficientes de variação da variância (CVV) e tamanhos de amostra (15, 30, 60, 90, 120, 150), para o Cenário 4

		Cenário 4 ( $\sigma_i^2 = \sigma^2\{E(y_i)\}^2$ )								
		Coeficiente de variação da variância (CVV)								
		$n$	0	0,4	0,6	0,9	1,1	1,3	1,5	1,7
Teste de Glejser	15	0,0649	0,1437	0,2346	0,3876	0,4983	0,6077	0,7206	0,8332	
	30	0,0543	0,2294	0,3368	0,4793	0,5601	0,6254	0,6878	0,7916	
	60	0,0524	0,3690	0,4578	0,5711	0,6295	0,6822	0,7257	0,7620	
	90	0,0536	0,4483	0,5312	0,6291	0,6786	0,7233	0,7591	0,7885	
	120	0,0512	0,5409	0,6027	0,6816	0,7213	0,7561	0,7828	0,8067	
	150	0,0498	0,6433	0,6782	0,7299	0,7616	0,7882	0,8128	0,8335	
Teste de Breusch-Pagan	15	0,0521	0,1756	0,2846	0,4538	0,5602	0,6663	0,7669	0,8642	
	30	0,0487	0,2369	0,3575	0,5021	0,5807	0,6459	0,7059	0,8076	
	60	0,0504	0,3817	0,4669	0,5793	0,6400	0,6898	0,7312	0,7671	
	90	0,0520	0,4703	0,5391	0,6313	0,6836	0,7257	0,7613	0,7915	
	120	0,0513	0,5542	0,6029	0,6760	0,7190	0,7519	0,7826	0,8062	
	150	0,0498	0,6557	0,6766	0,7253	0,7556	0,7836	0,8075	0,8295	
Teste de White	15	0,0521	0,1136	0,1924	0,3370	0,4516	0,5677	0,6908	0,8142	
	30	0,0487	0,1706	0,2977	0,4570	0,5414	0,6157	0,6819	0,7928	
	60	0,0504	0,3158	0,4437	0,5783	0,6442	0,6975	0,7414	0,7777	
	90	0,0520	0,4336	0,5524	0,6658	0,7190	0,7625	0,7951	0,8226	
	120	0,0513	0,5287	0,6327	0,7273	0,7705	0,8041	0,8290	0,8520	
	150	0,0498	0,6398	0,7278	0,7965	0,8285	0,8518	0,8703	0,8867	