

# Erros heteroscedástico-consistentes

Teoria e simulações

*Luiz Fernando Palin Droubi\**

*Norberto Hochheim†*

*Willian Zonato‡*

10/06/2019

## 1 INTRODUÇÃO

A homoscedasticidade é uma das hipóteses que devem ser necessariamente verificadas na inferência clássica. A ausência da homoscedasticidade, ou seja, a heteroscedasticidade, no entanto, não invalida completamente um modelo de regressão linear, pois a heteroscedasticidade se relaciona apenas aos erros do modelo, o que afeta o cálculo dos intervalos de confiança das previsões e dos coeficientes do modelo, mas não invalida os valores centrais dos mesmos, ou seja, mesmo que um modelo contenha erros heteroscedásticos, este modelo pode ser utilizado para fazer previsões, desde que não se esteja interessado nos seus intervalos de confiança. Na Engenharia de Avaliações, contudo, frequentemente ou quase sempre, existe interesse em avaliar o grau de precisão do modelo, ou mesmo o grau de fundamentação de um laudo. Dado que estes enquadramentos normativos são feitos com a utilização dos intervalos de confiança das estimativas, ou com os p-valores dos testes de hipótese realizado pelos regressores, a utilização de modelos com erros heteroscedásticos nesta área torna-se complicada, o que não significa que isto seja impossível. Infelizmente a NBR14.653-2 não aborda este assunto, mas acredita-se que, mesmo assim, é possível se utilizar de modelos heteroscedásticos e ainda estar de acordo com a norma. Demonstrar como isso pode ser feito é o objetivo deste artigo.

## 2 DESENVOLVIMENTO E FUNDAMENTAÇÃO

### 2.1 A Hipótese de homoscedasticidade

Segundo Long (2000, p. 5), na regressão linear, a variância do estimador MQO  $\hat{\beta}$  é:

$$\text{var}(\hat{\beta}) = (X'X)^{-1}X'\Omega X(X'X)^{-1}$$

---

\*SPU/SC, [luiz.droubi@planejamento.gov.br](mailto:luiz.droubi@planejamento.gov.br)

†UFSC, [hochheim@gmail.com](mailto:hochheim@gmail.com)

‡SPU/SC, [willian.zonato@planejamento.gov.br](mailto:willian.zonato@planejamento.gov.br)



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

A adoção da hipótese da homoscedasticidade na inferência clássica, em conjunto com a hipótese da independência dos erros, pode ser resumida na equação abaixo:

$$\Omega_{MQO} = \sigma^2 I$$

onde  $\sigma^2$  é a variância dos erros do modelo, suposta constante, e  $I$  é a matriz identidade  $n \times n$ , onde  $n$  é o número de dados do modelo. A matriz  $\Omega$  assim obtida, pode-se demonstrar, simplifica demasiado o cálculo da matriz de variância-covariância dos coeficientes de estimação, da maneira a seguir:

$$\text{var}(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

## 2.2 Detecção da heteroscedasticidade

Diversos testes foram desenvolvidos visando a detecção da heteroscedasticidade, dentro os quais citamos o teste de Goldfeld-Quandt (1965), o teste de Park (1966), o teste de Glejser (1969), o teste de Breusch-Pagan (1979) e o teste de White (1980).

Deve-se salientar que nenhum dos testes mencionados pode ser considerado melhor do que o outro, porque, conforme será visto, cada um deles testa uma hipótese em particular para a estrutura do erro. O teste de Breusch-Pagan, por exemplo, assume que a estrutura da variância do termo de erro seja linear, sendo muito eficiente para detectar este tipo de heteroscedasticidade. Se, no entanto, se estiver lidando com uma estrutura não-linear do termo de erro, este teste não será eficaz, sendo o teste de White melhor para estes casos.

### 2.2.1 Teste de Goldfeld-Quandt

O teste de Goldfeld-Quandt (1965) é aplicável a grandes amostras e consiste em omitir algumas amostras do modelo e comparar a estrutura da variância do termo de erro em diferentes subgrupos.

### 2.2.2 Teste de Park

O teste de Park (1966) assume que a variância do termo de erro é proporcional a alguma potência da variável independente ( $\sigma_{u_i}^2 = \sigma^2 X_i^\beta e^v$ ). Desta maneira, o teste de Park consiste em estimar os resíduos da equação de regressão e então fazer uso de uma regressão destes resíduos em relação à variável dependente, da seguinte maneira:

$$\ln(u_i^2) = \ln(\sigma^2) + \beta \ln(X_i) + v_i$$

Se o coeficiente  $\beta$  for significativo, é indicativo da presença de heteroscedasticidade.



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
**SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA**  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

### 2.2.3 Teste de Glejser

O teste de Glejser (1969; *apud* NWAKUYA; NWABUEZE, 2018) consiste em testar se os valores absolutos dos resíduos ( $|\hat{u}_i|$ ) da regressão original tem correlação com alguma variável independente do modelo ( $X_i$ ), conforme formas abaixo:

$$|\hat{u}_i| = \beta_0 + \beta_1 X_i + v_i$$

$$|\hat{u}_i| = \beta_0 + \beta_1 \sqrt{X_i} + v_i$$

$$|\hat{u}_i| = \beta_0 + \beta_1 \frac{1}{X_i} + v_i$$

$$|\hat{u}_i| = \beta_0 + \beta_1 \frac{1}{\sqrt{X_i}} + v_i$$

$$|\hat{u}_i| = \sqrt{\beta_0 + \beta_1 X_i} + v_i$$

$$|\hat{u}_i| = \sqrt{\beta_0 + \beta_1 X_i^2} + v_i$$

### 2.2.4 Teste de Breusch-Pagan

O teste de Breusch-Pagan (1979) consiste em assumir que a estrutura da variância do termo de erro seja linear em função das variáveis independentes do modelo. Assim, o método consiste em estimar a variância do erro a partir do quadrado dos resíduos da regressão linear e efetuar então uma regressão auxiliar deste termo em função dos regressores do mesmo modelo. Calcula-se o valor do coeficiente de determinação deste modelo auxiliar e aplica-se o teste do  $\chi^2$  para obter a sua significância. KOENKER (1981) propõe um ajuste ao teste original, fazendo uso dos resíduos studentizados.

### 2.2.5 Teste de White

O teste de White (1980) consiste numa aplicação do teste de Breusch-Pagan, porém com a adição dos termos quadráticos e de interação do modelo original. Obviamente, o teste de White necessita de um número maior de graus de liberdade no modelo para ter eficácia, haja vista que a regressão auxiliar necessitará de diversos graus de liberdade.

## 2.3 Alternativas para contornar a heteroscedasticidade

Ocorre que apenas raramente os dados do mundo real podem ser considerados homoscedásticos e independentes, o que faz com que procedimentos especiais tenham que ser tomados para a utilização do estimador MQO. A tentativa de se contornar a heteroscedasticidade pode ser feita de diversas formas, algumas mais trabalhosas, outras mais complexas, outras causando maior distorção ao modelo.



## 2.4 Transformação da variável dependente

Na Engenharia de Avaliações, a primeira e mais usual forma de se contornar a heteroscedasticidade é através da aplicação de transformações à variável dependente. No entanto, esta não é a única alternativa e a transformação dos dados pode causar distorções significativas, o que pode ser um problema, especialmente considerando o problema da retransformação das variáveis transformadas.

Segundo MATLOFF (2017), se  $\varphi(x)$  é uma função convexa, então a desigualdade de Jensen se exprime na seguinte desigualdade:

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)]$$

## 2.5 Outros estimadores

Outra forma de contornar a heteroscedasticidade é através da utilização de outros estimadores, como o estimador dos mínimos quadrados ponderados. Este método consiste na adoção de um vetor de pesos conveniente escolhido e aplicado a cada uma das observações, de maneira que se elimine a heteroscedasticidade do modelo. O inconveniente deste método está na determinação do vetor de pesos a ser utilizado, o que não é trivial.

## 2.6 Estimação heteroscedástico-consistente da matriz de variância-covariância

Uma maneira mais interessante de se detectar a heteroscedasticidade é através da estimação da matriz de variância-covariância dos coeficientes heteroscedástico-consistente, desfazendo-se assim, da hipótese da homoscedasticidade.

Existem diversas maneiras de se estimar a matriz heteroscedástico-consistente, sendo que este método foi originalmente concebido por White (1980; *apud* LONG; ERVIN, 2000, p. 6), da seguinte forma:

$$HC0 = (X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1} = (X'X)^{-1}X'diag[e_i^2]X(X'X)^{-1}$$

Onde  $diag[e_i^2]$  é a matriz diagonal em que os elementos da diagonal são compostos pelo vetor do erros amostrais elevado ao quadrado.

Posteriormente diversas pequenas modificações foram propostas ao método, como a correção dos resíduos pelos graus de liberdade ( $n/(n - k)$ ) e outras.

## 3 ESTUDO DE CASO

Para este estudo de caso foi utilizado software R 3.6.0. Os funções utilizadas para a execução dos testes são do pacote lmttest (ZEILEIS; HOTHORN, 2002). Os gráficos foram confeccionados com o

auxílio do pacote mosaic (PRUIM et al., 2017).

### 3.1 Exemplo 1

#### 3.1.1 Criação dos dados

Neste primeiro exemplo foram gerados randomicamente para a variável dependente ( $Y$ ). Para a variável independente ( $X$ ), foram criados 100 dados em sequência de 1 a 100. Para a variável dependente, com a utilização da função `rnorm`, foram criados dados de distribuição normal com média igual a média do vetor da variável independente ( $\bar{Y} = \bar{X}$ ) e desvio-padrão variável em função de  $X$ , da seguinte maneira:

$$\text{desvio}(Y) = 0,4 \cdot X$$

Ou seja, foi criada, randômicamente, um tipo linear de heteroscedasticidade. Os gráfico dos dados assim criados pode ser vistos na figura 1.

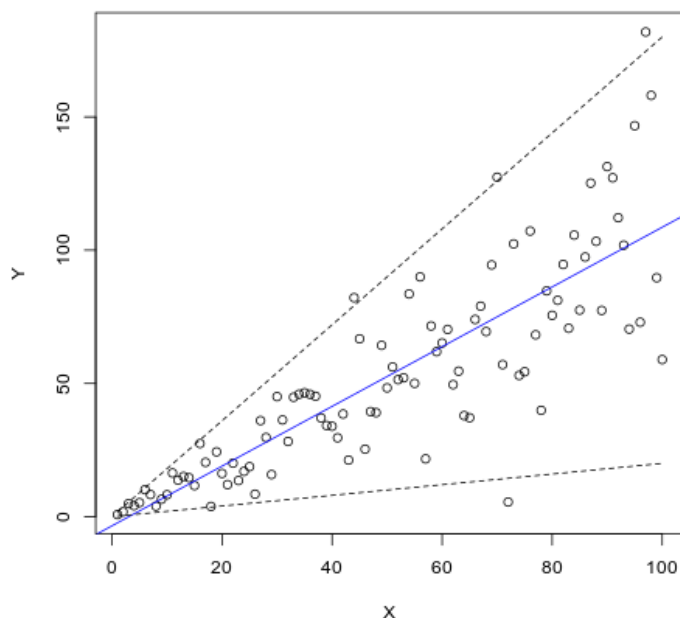


Figura 1: Forma linear de heteroscedasticidade.

##### 3.1.1.1 Testes



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

Os resultados das estatísticas dos testes e suas significâncias estão resumidos na tabela 1. Pode-se verificar que, para este caso, todos os testes foram eficazes em não rejeitar a hipótese nula da homoscedasticidade, ou seja, todos os testes detectaram a presença da heteroscedasticidade.

#### **3.1.1.2 Park**

Apesar da forma linear da heteroscedasticidade presente, como o valor da estatística do teste  $t$  da variável  $\ln(X)$  resultou alto (5,48) e significativo ( $p\text{-valor} = 3.335e-07$ ), conclui-se que não se pode descartar a hipótese da heteroscedasticidade, ou seja, o teste de Park funcionou, mesmo na presença de uma forma de heteroscedasticidade diferente da pressuposta.

#### **3.1.1.3 Glejser**

Foi utilizada a primeira forma (linear) do teste de Glejser. A significância do teste foi de  $3.55e-09$ . Conclui-se que o teste funciona. Deve-se atentar, porém, que a forma de heteroscedasticidade, neste caso, era previamente conhecida, o que possibilitou a aplicação da forma “correta” do teste, o que não ocorre na prática.

#### **3.1.1.4 Breusch-Pagan**

A forma original do teste de BP foi aplicada, através da função `bptest`:

O teste de BP, neste caso, não logrou a rejeição da hipótese nula, ou seja, a da homoscedasticidade. Desta maneira, pode-se afirmar que o teste funcionou a contento.

#### **3.1.1.5 Breusch-Pagan studentizado (Koenker)**

O mesmo ocorre para o teste de BP com a utilização dos resíduos studentizados:

#### **3.1.1.6 Goldfeld-Quandt**

O teste de GQ neste caso foi eficaz em detectar a modificação na variância nas partições testadas.

#### **3.1.1.7 White**

Por fim, o teste de White mostrou-se adequado para a detecção da forma linear da heteroscedasticidade. No entanto, deve-se salientar, que para isto foi inutilmente utilizado um grau de liberdade a mais do que o teste de Breusch-Pagan, o que pode fazer a diferença no caso de pequenas amostras.

### **3.1.2 Quadro-Resumo**

Tabela 1: Estatísticas e p-valores para os dados do Exemplo 1.



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

Teste	Estatística	Significância
Park	5,48	3.3e-07
Glejser	6,49	3.5e-09
GQ	15,95	2,5e-13
BP	40,28	2,2e-10
Koenker	18,86	1,4e-05
White	21,66	1,98e-05

## 3.2 Exemplo 2

### 3.2.1 Criação dos dados

Neste exemplo foram gerados randomicamente dados de distribuição normal para a variável dependente ( $Y$ ), com a utilização da função `rnorm`, com média igual a média do vetor da variável independente ( $\bar{Y} = \bar{X}$ ) e desvio-padrão variável em função de  $X$ , da seguinte maneira:

$$\text{desvio}(Y) = 0,001 \cdot X^2$$

Ou seja, foi criada, randômicamente, um tipo suavemente não-linear de heteroscedasticidade. Os gráfico dos dados assim criados pode ser vistos na figura 2.

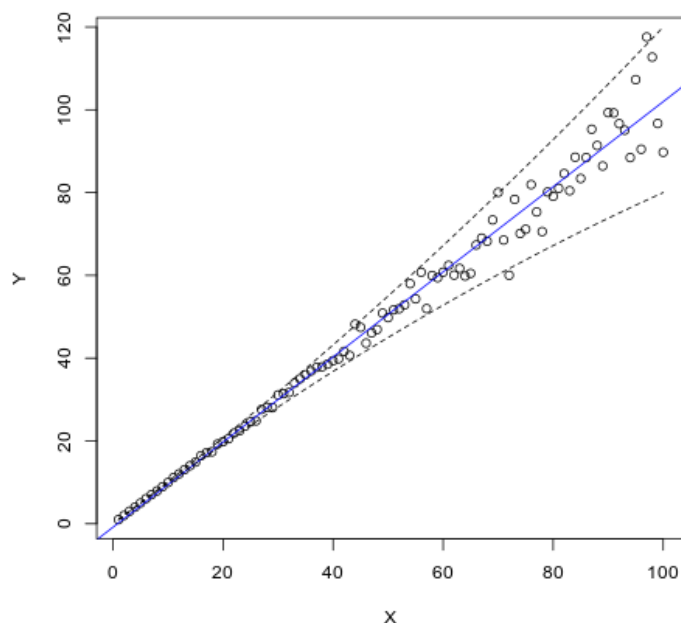


Figura 2: Forma não-linear de heteroscedasticidade.



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
**SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA**  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

### **3.2.2 Testes**

Os resultados das estatísticas dos testes e suas significâncias estão resumidos na tabela 2. Pode-se verificar que, para este caso, todos os testes foram eficazes em não rejeitar a hipótese nula da homoscedasticidade, ou seja, todos os testes detectaram a presença da heteroscedasticidade.

#### **3.2.2.1 Park**

Apesar da forma linear da heteroscedasticidade presente, como o valor da estatística do teste  $t$  da variável  $\ln(X)$  resultou alto (5,64) e significativo ( $p$ -valor =  $1.652e-07$ ), conclui-se que não se pode descartar a hipótese da heteroscedasticidade, ou seja, o teste de Park funcionou, mesmo na presença de uma forma de heteroscedasticidade diferente da pressuposta.

#### **3.2.2.2 Glejser**

Foi utilizada a primeira forma (linear) do teste de Glejser. A significância do teste foi de  $3.55e-09$ . Conclui-se que o teste funciona. Deve-se atentar, porém, que a forma de heteroscedasticidade, neste caso, era previamente conhecida, o que possibilitou a aplicação da forma “correta” do teste, o que não ocorre na prática.

#### **3.2.2.3 Breusch-Pagan**

O teste de BP, apesar de assumir a forma linear de heteroscedasticidade, logrou a rejeição da hipótese da homoscedasticidade. Isto ocorreu, talvez, pela não-linearidade da heteroscedasticidade, neste caso, ser singela.

#### **3.2.2.4 Breusch-Pagan studentizado (Koenker)**

O mesmo ocorre para o teste de BP studentizado.

#### **3.2.2.5 Goldfeld-Quandt**

O teste de Goldfeld-Quandt foi eficaz também neste caso, detectando a presença da heteroscedasticidade.

#### **3.2.2.6 White**

Por fim, o teste de White, concebido para encontrar as formas não-lineares de heteroscedasticidade, também foi eficaz neste caso, rejeitando a hipótese nula da homoscedasticidade.

### **3.2.3 Quadro-Resumo**

Tabela 2: Estatísticas e  $p$ -valores para os dados do Exemplo 2.





**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

Teste	Estatística	Significância
Park	5,64	1.7e-07
Glejser	6,49	3.5e-09
GQ	220,9	5,4e-29
BP	71,84	2,3e-17
Koenker	20,96	4,68e-06
White	30,68	2,18e-07

### 3.3 Exemplo 3

#### 3.3.1 Criação dos dados

Os dados foram criados da mesma maneira do exemplo anterior, porém com uma nova função para a variância, da seguinte maneira:

$$\text{desvio}(Y) = 0,5 \cdot (X - 1)^2$$

Ou seja, foi criada, randômicamente, um tipo de heteroscedasticidade não-linear mais drástico, onde a variância é menor no centro da amostra, aumentando nas extremidades. Os gráfico dos dados assim criados pode ser vistos na figura 3.

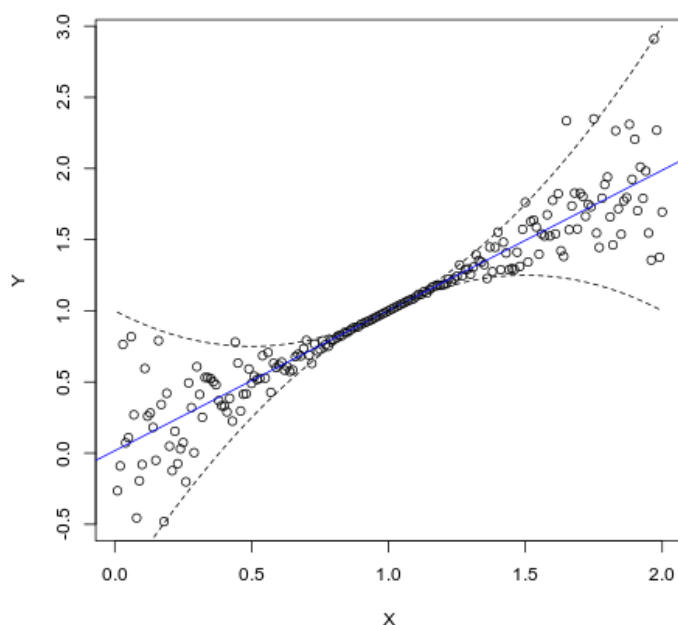


Figura 3: Forma drasticamente não-linear de heteroscedasticidade.



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
**SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA**  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

### 3.3.2 Testes

#### 3.3.2.1 Park

Como o valor da estatística do teste t da variável  $\ln(X)$  resultou alto (-3,09) e significativo (p-valor = 0.002313), conclui-se que não se pode descartar a hipótese da heteroscedasticidade.

#### 3.3.2.2 Glejser

Neste exemplo, o teste de Glejser detecta a heteroscedasticidade, mas apenas na equação da forma quadrática, falhando em detectá-la na função de teste linear.

Quando utilizada a forma quadrática do teste de Glejser. A significância do teste foi de 0.0168.

#### 3.3.2.3 Breusch-Pagan

Na ocorrência da heteroscedasticidade não-linear extrema deste caso, o teste de BP não logra a rejeição da hipótese nula, ou seja, a da homoscedasticidade.

#### 3.3.2.4 Breusch-Pagan studentizado (Koenker)

Também o teste de BP studentizado falha neste caso.

#### 3.3.2.5 Goldfeld-Quandt

Neste caso, o teste de Goldfeld=Quandt não logrou a rejeição da hipótese nula, ou seja, a da homoscedasticidade, apesar de ser sabido que há heteroscedasticidade no modelo. Ou seja, o teste falha em detectar a heteroscedasticidade pela própria característica do mesmo, que divide em a amostra em duas partições, excluindo-se uma fração dos dados centrais (aqui 25%) e compara suas variâncias. Como no presente caso a variância muda igualmente em ambas as extremidades da amostra, o teste não a detecta.

#### 3.3.2.6 White

Por fim, o teste de White mostrou-se adequado para a detecção da forma de heteroscedasticidade drasticamente não-linear, como era de se esperar.

### 3.3.3 Quadro-Resumo

Tabela 3: Estatísticas e p-valores para os dados do Exemplo 3.

Teste	Estatística	Significância
Park	-3,09	0.0023



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
**SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA**  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

Teste	Estatística	Significância
Glejser – forma linear	6,49	3.5e-09
Glejser – forma quadrática	2,41	3.5e-09
GQ	1,071	0,39
BP	0,1466	0,7
Koenker	0,04362	0,835
White	47,36	5,21e-11

## 4 CONCLUSÃO

## REFERÊNCIAS

BREUSCH, T. S.; PAGAN, A. R. A simple test for heteroscedasticity and random coefficient variation. **Econometrica**, v. 47, n. 5, p. 1287–1294, 1979. [Wiley, Econometric Society]. Disponível em: <<http://www.jstor.org/stable/1911963>>..

GLEJSER, H. A new test for heteroskedasticity. **Journal of the American Statistical Association**, v. 64, n. 325, p. 316–323, 1969. [American Statistical Association, Taylor & Francis, Ltd.]. Disponível em: <<http://www.jstor.org/stable/2283741>>..

GOLDFELD, S. M.; QUANDT, R. E. Some tests for homoscedasticity. **Journal of the American Statistical Association**, v. 60, n. 310, p. 539–547, 1965. [American Statistical Association, Taylor & Francis, Ltd.]. Disponível em: <<http://www.jstor.org/stable/2282689>>..

KOENKER, R. A note on studentizing a test for heteroscedasticity. **Journal of Econometrics**, v. 17, n. 1, p. 107–112, 1981. Disponível em: <<http://www.sciencedirect.com/science/article/pii/0304407681900622>>..

LONG, J. S.; ERVIN, L. H. Using heteroscedasticity consistent standard errors in the linear regression model. **The American Statistician**, v. 54, n. 3, p. 217–224, 2000. Taylor & Francis. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/00031305.2000.10474549>>..

MATLOFF, N. **Statistical regression and classification: From linear models to machine learning**. Boca Raton, Florida: Chapman & Hall, 2017.

NWAKUYA, M. T.; NWABUEZE, J. C. Application of Box-Cox transformation as a corrective measure to heteroscedasticity using an economic data. **American Journal of Mathematics and Statistics**, v. 8, n. 1, p. 8–12, 2018. Disponível em: <<http://article.sapub.org/10.5923.j.ajms.20180801.02.html#Ref>>..

PARK, R. E. Estimation with heteroscedastic error terms. **Econometrica**, v. 34, n. 4, p. 888–888, 1966. [Wiley, Econometric Society]. Disponível em: <<http://www.jstor.org/stable/1910108>>..

PRUIM, R.; KAPLAN, D. T.; HORTON, N. J. The mosaic package: Helping students to 'think with data' using r. **The R Journal**, v. 9, n. 1, p. 77–102, 2017. Disponível em: <<https://journal.r-project.org/>



**MINISTÉRIO DA ECONOMIA**  
**SECRETARIA DE COORDENAÇÃO E GOVERNANÇA DO PATRIMÔNIO DA UNIÃO**  
SUPERINTENDÊNCIA DO PATRIMÔNIO DA UNIÃO EM SANTA CATARINA  
Praça XV de Novembro nº 336, Centro, Florianópolis/SC  
(48)3224-5399

---

[archive/2017/RJ-2017-024/index.html](https://arquivo/2017/RJ-2017-024/index.html)>..

WHITE, H. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. **Econometrica**, v. 48, n. 4, p. 817–38, 1980. Disponível em: <<https://EconPapers.repec.org/RePEc:ecm:emetrp:v:48:y:1980:i:4:p:817-38>>..

ZEILEIS, A.; HOTHORN, T. Diagnostic checking in regression relationships. **R News**, v. 2, n. 3, p. 7–10, 2002. Disponível em: <<https://CRAN.R-project.org/doc/Rnews/>>..