Correlação Serial e Heterocedasticidade em Regressões de Séries Temporais

Wooldridge, Cap. 12

Porto Alegre, 11 de novembro de 2010

CORRELAÇÃO SERIAL

- Ocorrência
- Conseqüência
- Análise gráfica
- Autocorrelação
- Exemplo
- Testes
- Correlação serial devido a dinâmicas mal especificadas

Ocorrência

Ocorre principalmente quando os dados são observados ao longo do tempo. No caso de uma *cross-section*, pode ocorrer, por exemplo, no caso de omissão de variável relevante.

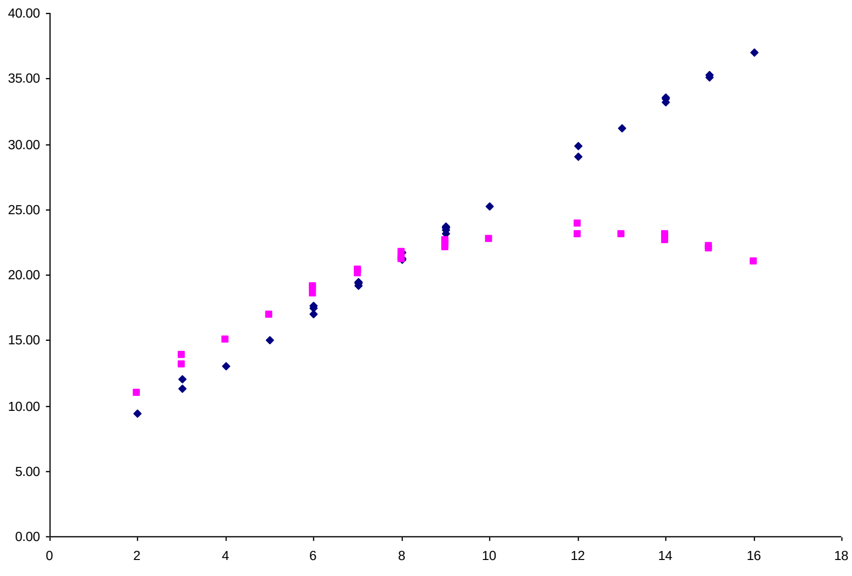
Exemplo:

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{1t}^2 + u_t$$

(Modelo verdadeiro)

$$y_t = \beta_0^* + \beta_1^* x_{1t} + u_t^*$$

(Modelo especificado)



Seqüência de resíduos + e -

Conseqüências

- **Os estimadores de MQO continuam lineares, não viesados, consistentes e assintoticamente normais, mas não são mais eficientes (BLUE).**

- **As variâncias estimadas serão viesadas (em geral, subestimadas).**

Conseqüências

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki} + e_i$$

Suposições:

- $Var(e_i) = \sigma_e^2$: **constante, não varia com i (homocedasticidade).**
- $e_i \sim N(0; \sigma_e^2)$
- **Os erros são independentes, ou seja,**

$Corr(e_i, e_j) = 0$ **para i ≠ j.**

Conseqüências

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_k x_{ki} + e_i$$

No caso da presença de correlação serial teremos que,

$$Cov\left(e_i, e_j\right) = E\left[e_i e_j\right] = \sigma_{ij} \neq 0$$

Análise Gráfica

- Gráfico de e_i *vs* ordem que as observações foram coletadas: verifica a presença de dependência;
- Para que não haja dependência entre os resíduos, o gráfico obtido não deve conter seqüências muito longas de resíduos de mesmo sinal.

Contexto

- **Dados extraídos numa ordem conhecida.**
- **Objetivo**: **identificar se o erro relativo à observação i sofre a influência dos erros relativos às observações anteriores.**

Série de erros: $e_1, e_2, \ldots, e_i, \ldots$

Autocorrelação

Autocorrelação (correlação serial) **de primeira ordem**: correlação existente entre uma observação i qualquer e a observação imediatamente anterior (i-1).

Autocorrelação (correlação serial) **de ordem q**: correlação existente entre uma observação i qualquer e a observação anterior (i-q).

Modelagem da autocorrelação

$$e_i = \rho e_{i-1} + u_i$$

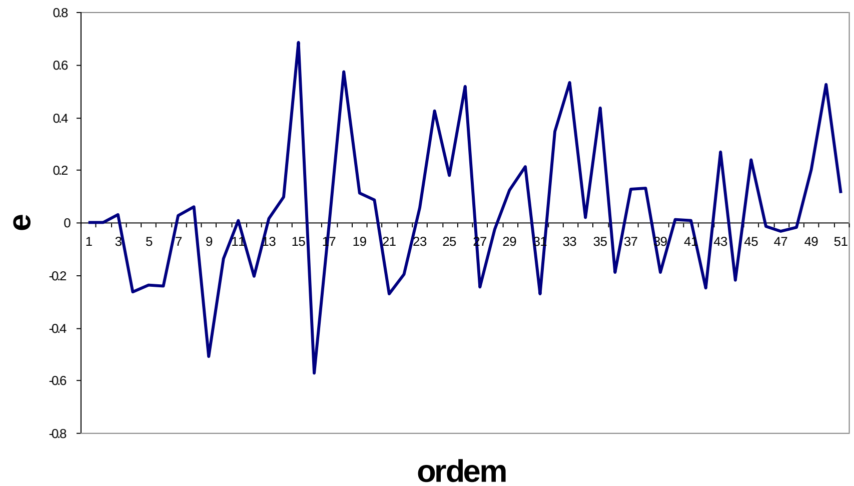
Suposições:

- $u_i \sim N(0; \sigma_u^2)$
- u_i independentes entre si
- u_i independente de e_{i-1}
- ρ: número entre -1 e 1.
- $\rho = 0 \Rightarrow e_i$ independentes e com variância σ_e^2

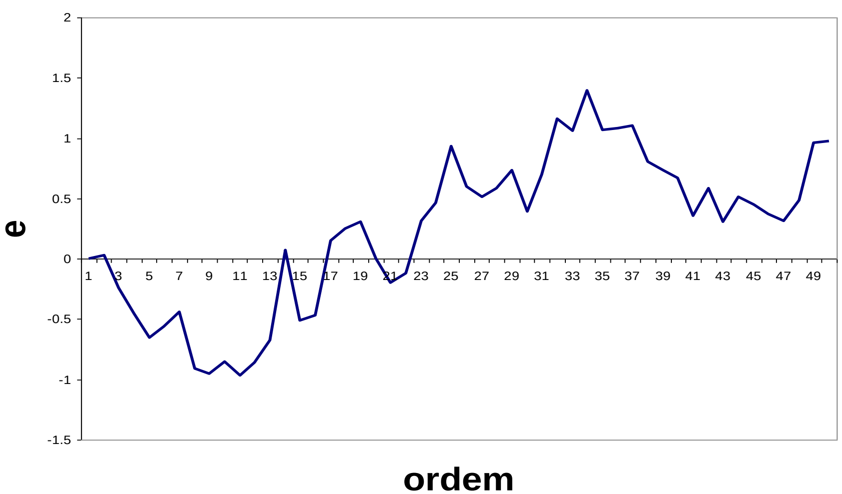
Exemplos

Simulações de séries de erros com autocorrelação de 1ª ordem

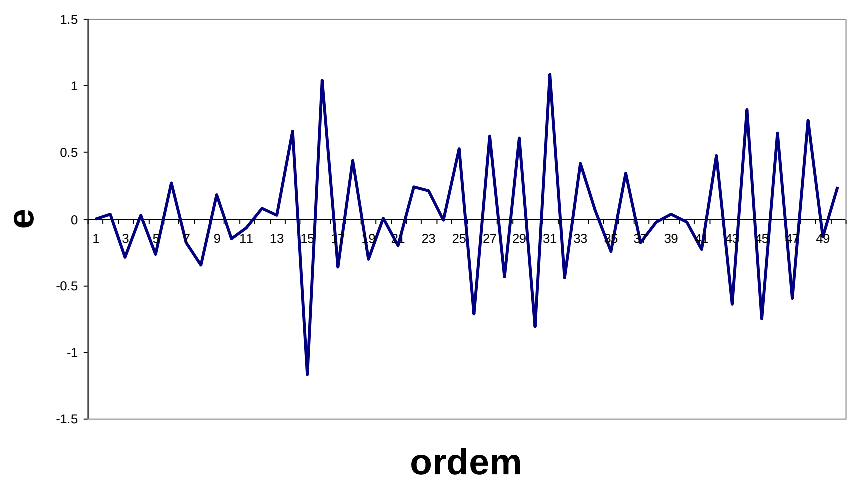
$\rho = 0$



$\rho = 0,9$

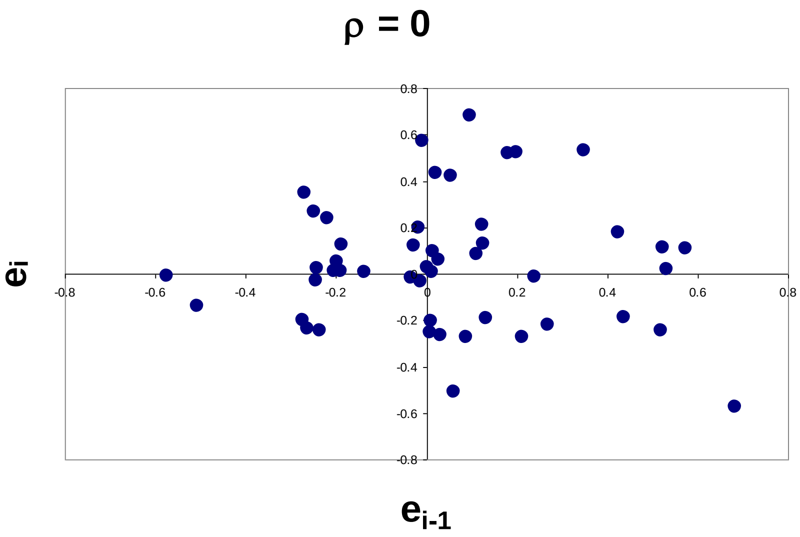


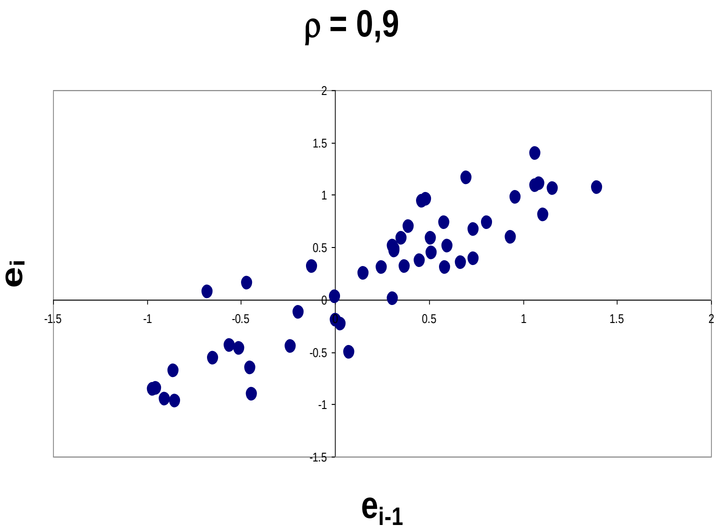
$\rho = -0,9$

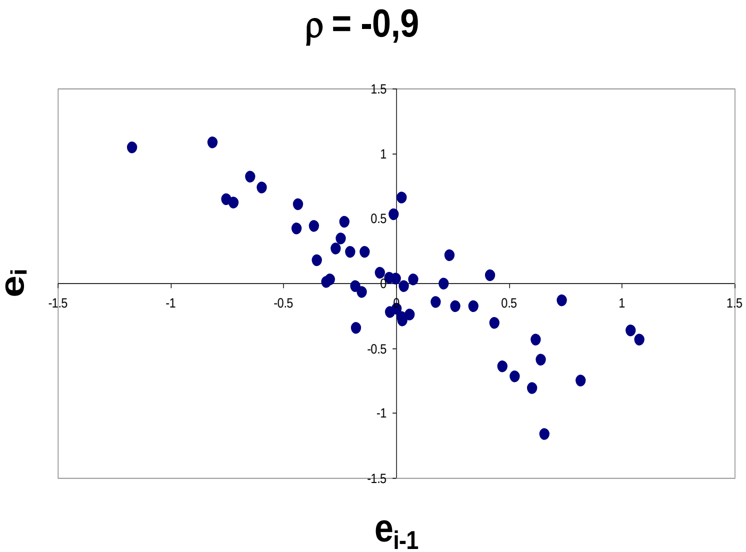


Diagramas de dispersão

$$e_i \quad vs \quad e_{i-1}$$







Estrutura AR(1) para a Matriz de Variâncias e Covariâncias

$$Var\left(\underset{\sim}{e}\right)=\sigma_e^2\,\underset{\sim}{\Omega}=\sigma_e^2\begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \dots & 1 \end{pmatrix}$$

A variância do erro é constante.

Teste para Verificação de Ausência de Correlação Serial (Teste de Durbin-Watson)

$$H_0 : \rho = 0$$

Estatística do teste:

$$d = \frac{\sum_{i=2}^{n} \left(\hat{e}_i - \hat{e}_{i-1} \right)^2}{\sum_{i=1}^{n} \hat{e}_i^2}$$

$$0 \le d \le 4$$

Estatística *d*

Note que

$$d = \frac{\sum \hat{e}_i^2 + \sum \hat{e}_{i-1}^2 - 2\sum \hat{e}_i \hat{e}_{i-1}}{\sum \hat{e}_i^2}$$

Como

$$\sum \hat{e}_i^2 \approx \sum \hat{e}_{i-1}^2$$

E, ainda, a correlação entre $\hat{e}_i$ **e** $\hat{e}_{i-1}$ **é**

$$\hat{\rho} = \frac{\sum \hat{e}_i \hat{e}_{i-1}}{\sum \hat{e}_i^2}$$

Estatística *d*

Então,

$$d \cong 2\left(1-\frac{\sum \hat{e}_i \hat{e}_{i-1}}{\sum \hat{e}_i^2}\right)=2(1-\hat{\rho})$$

$$-1 \leq \hat{\rho} \leq 1 \Rightarrow 0 < d < 4$$

Estatística de Durbin-Watson

- **Se a autocorrelação for positiva, o valor de *d* será baixo.**

- **Se a autocorrelação for negativa, o valor de *d* será alto.**

- **Valores próximos a 2 indicam autocorrelação próxima de zero.**

Teste de Durbin-Watson

$$H_0 : \rho = 0 \qquad\qquad d = \frac{\sum_{i=2}^{n}\left(\hat{e}_i - \hat{e}_{i-1}\right)^2}{\sum_{i=1}^{n}\hat{e}_i^2} \cong 2\left(1 - \hat{\rho}\right)$$

H_0 deve ser rejeitada para valores distantes de 2.

A distribuição de *d* depende do tamanho amostral (*n*) e do número de variáveis independentes (*k*).

Teste de Durbin-Watson

$$H_0 : \rho = 0$$
$$H_A : \rho > 0$$

$$d = \frac{\sum_{i=2}^{n} \left(\hat{e}_i - \hat{e}_{i-1} \right)^2}{\sum_{i=1}^{n} \hat{e}_i^2} \cong 2\left(1 - \hat{\rho}\right)$$

n = 20, k = 3 e nível de significância de 5%:

Tabela (Gujarati): d_L = 0,998 e d_U = 1,676

se $d < d_L$ então rejeita-se H_0 (correlação positiva)

se $d > d_U$ então não há evidências para rejeitar H_0

e se $d_L < d < d_U$ então teste inconclusivo.

Teste de Durbin-Watson

$$H_0 : \rho = 0$$
$$H_A : \rho < 0$$

$$d = \frac{\sum_{i=2}^{n} \left(\hat{e}_i - \hat{e}_{i-1}\right)^2}{\sum_{i=1}^{n} \hat{e}_i^2} \cong 2\left(1 - \hat{\rho}\right)$$

n = 20, *k* = 3 e nível de significância de 5%:

Tabela (Gujarati): d_L= 0,998 e d_U= 1,676

se $d > 4 - d_L$ então rejeita-se H_0

se $d < 4 - d_U$ então não há evidências para rejeitar H_0

se $4 - d_U < d < 4 - d_L$ então teste inconclusivo.

Exemplo

Base de dados de Wooldridge (2006) com preços de imóveis americanos (*hprice1*) para estimar o modelo a seguir:

$$preco = \beta_0 + \beta_1 terreno + \beta_2 area + \beta_3 quartos + e$$

preco: **preço da casa (em milhares de dólares)**

terreno: **tamanho do terreno**

area: **tamanho da casa**

quartos: **número de quartos**

Supondo que os dados estejam ordenados de acordo com a seqüência em que foram coletados. Verificar a existência de correlação serial entre os erros do modelo.

Exemplo

Dependent Variable: PRECO
Method: Least Squares
Date: 07/30/10 Time: 13:55
Sample: 1 88
Included observations: 88

	Coefficient	Std. Error	t-Statistic	Prob.
C	-21.77031	29.47504	-0.738601	0.4622
TERRENO	0.002068	0.000642	3.220096	0.0018
AREA	0.122778	0.013237	9.275093	0.0000
QUARTOS	13.85252	9.010145	1.537436	0.1279

R-squared	0.672362	Mean dependent var	293.5460
Adjusted R-squared	0.660661	S.D. dependent var	102.7134
S.E. of regression	59.83348	Akaike info criterion	11.06540
Sum squared resid	300723.8	Schwarz criterion	11.17800
Log likelihood	-482.8775	Hannan-Quinn criter.	11.11076
F-statistic	57.46023	Durbin-Watson stat	2.109796
Prob(F-statistic)	0.000000		

n = 88, k = 3, α = 5% , d = 2,1098 d_L = 1,589 e d_U = 1,726

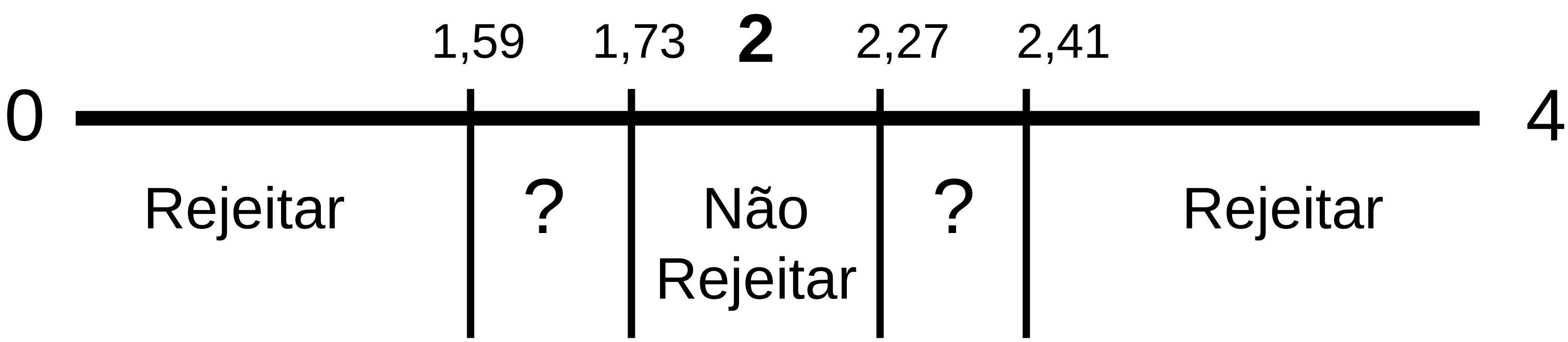
Exemplo

$$H_0 : \rho = 0 \qquad H_A : \rho \neq 0$$

n = 88, k = 3, α = 5% , d = 2,1098

d_L = 1,589 e d_U = 1,726

4 - d_L = 2,41 4 - d_U= 2,27



Observações

Se H_0 for rejeitada, então:

- **Os estimadores de MQO são ineficientes;**
- **A correção depende do conhecimento que temos sobre a natureza da interdependência dos termos de erro, isto é, do conhecimento da estrutura de correlação.**

Observações

Se H_0 for rejeitada, então:

- **Se, por exemplo, houver uma estrutura AR(1) com ρ conhecido, pode-se estimar os parâmetros do modelo original, de maneira eficiente, a partir de uma transformação neste modelo (conhecida como quase-diferença), que é equivalente a empregar o método dos mínimos quadrados generalizados (*GLS*).**

Correlação Serial devido a dinâmicas mal especificadas

- Uma estatística DW significante não necessariamente implica que tenhamos um problema de correlação serial. Este ponto foi argumentado por Sargan (1964) e também por Hendry e Mizon (1978). Consideramos que,

$$y_t = \beta x_t + u_t \quad \text{com} \quad u_t = \rho u_t + e_t \qquad (1)$$

- E e_t são independentes com uma variância comum σ^2.
- Podemos escrever esse modelo como

$$y_t = \rho y_{t-1} + \beta x_t - \beta \rho x_{t-1} + e_t \qquad (2)$$

Correlação Serial devido a dinâmicas mal especificadas

- E considerando um modelo dinâmico estável alternativo,

$$y_t = \beta_1 y_{t-1} + \beta_2 x_t - \beta_3 x_{t-1} + e_t \quad \text{com} \quad |\beta_1| < 1 \qquad (3)$$

- A primeira equação (2) é a mesma que a segunda (3) com a restrição:

$$\beta_1 \beta_2 + \beta_3 = 0 \qquad (4)$$

Correlação Serial devido a dinâmicas mal especificadas

Um teste para $\rho = 0$ é um teste para $\beta_1 = 0$ (e $\beta_3 = 0$). Mas antes de testarmos isso, deveríamos primeiramente testar a restrição (4) e testar para $\rho = 0$ apenas se a hipótese nula $\beta_1 \beta_2 + \beta_3 = 0$ não for rejeitada.

Se esta hipótese for rejeitada, não teremos um modelo de correlação serial e a correlação serial nos erros em (1) deve-se a um problema de especificação dinâmica, ou seja, neste caso, à omissão das variáveis y_{t-1} e x_{t-1} da equação.

Correlação Serial devido a dinâmicas mal especificadas

Alternativamente, Hendry & Mizon desenvolvem o argumento da seguinte maneira: a partir de uma equação na forma de (3), aplicando o operador de diferenças $L^n y_t = y_{t-n}$ encontramos a equação

$$y_t = \beta_1 L y_t + \beta_2 x_t - \beta_3 L x_t + e_t$$

$$(1 - \beta_1 L) y_t = (\beta_2 - \beta_3 L) x_t + e_t$$

E, se a restrição (4) vale, então,

$$(1 - \beta_1 L) y_t = \beta_2 (1 - \beta_1 L) x_t + e_t \qquad (5)$$

Que é a equação (1), ao dividirmos ambos os lados pelo fator comum, tornando o termo de erro, $u_t = e_t / (1 - \beta_1 L)$

Correlação Serial devido a dinâmicas mal especificadas

- Logo, se o modelo da forma em (3) com variáveis defasadas em um período satisfazem a restrição acima, então, os polinômios no operador de diferença têm uma raiz comum e, esta raiz é o coeficiente de correlação serial de primeira ordem do termo de erro quando o modelo é definido na sua forma estática, como em (1).

- Neste caso haverá uma correspondência um para um entre o fator dinâmico comum e os erros auto regressivos.

- Contudo, nem sempre é possível proceder desta forma, uma vez que existe uma infinidade de situações nas quais o aspecto dinâmico não pode ser adequadamente captado pelo erro auto regressivo, ou seja, o fator comum não necessariamente será encontrado.

Correlação Serial devido a dinâmicas mal especificadas

- A restrição (4) é não linear nos parâmetros e, portanto, é necessário utilizar os testes Wald ou ML. Se o teste DW for significante, uma abordagem apropriada é testar a restrição (4) para estarmos seguros de que possuímos um modelo de correlação serial antes de aplicar qualquer transformação autoregressiva nas variáveis.

- A sugestão de Sargan é iniciar com o modelo geral (3) e testar a restrição (4) antes de se aplicar qualquer teste de correlação serial.

- No caso explicitado, não haverá um teste t exato como no caso das restrições lineares. Um procedimento seria linearizar a restrição por meio de uma série de expansão de Taylor e utilizar o teste Wald, que é assintótico, ou mesmo o ML.

HETEROCEDASTICIDADE

- Definição
- Conseqüências para os estimadores de Mínimos Quadrados (OLS)
- Testes
- Exemplo
- Estimação via Mínimos Quadrados Generalizados
- Inferência

HETEROCEDASTICIDADE

"A variância do termo erro, dadas as variáveis explicativas, não é constante"

Suposições do modelo de Mínimos Quadrados Ordinários (OLS)

- Regressores fixos e matriz **X** de posto completo;
- Erro aleatório (de média zero);
- **Homocedasticidade**;
- Ausência de correlação;
- Parâmetros constantes;
- Modelo Linear;
- Normalidade.

Homocedasticidade

- **A hipótese de homocedasticidade para a regressão múltipla significa que a variância do erro não observável *u*, condicional nas variáveis explicativas, é constante, ou seja, não se mantém quando a variância dos fatores não-observáveis muda ao longo de diferentes segmentos da população.**
- **Por exemplo, a heterocedasticidade está presente se a variância de *u* que afeta *y* aumenta com *x*.**

Homocedasticidade: Suposição de que

$$Var(\underset{\sim}{\mathbf{e}}) = E(\underset{\sim}{\mathbf{e}}\,\underset{\sim}{\mathbf{e}}') = \sigma^2 \underset{\sim}{I_n} = \begin{pmatrix} \sigma^2 & & & & \\ & \sigma^2 & & & \\ & & \sigma^2 & & \\ & & & \ddots & \\ & & & & \sigma^2 \end{pmatrix}$$

(matriz de covariância)

implica em:

$$Var(e_i) = E(e_i^2) = \sigma^2, \ \forall i$$

Esta hipótese não é correta quando *Var(e)* varia de um grupo para outro da população, ou seja, varia com os valores das variáveis explicativas

Heterocedasticidade

Aqui, consideraremos mantidas todas as outras suposições de OLS, exceto a de homocedasticidade. Assim, teremos algo como, por exemplo

$$Var\left(\underset{\sim}{\mathbf{e}}\right) = E\left(\underset{\sim}{\mathbf{e}}\,\underset{\sim}{\mathbf{e}}'\right) = \underset{\sim}{\Omega} = \begin{pmatrix} \sigma_1^2 & & & & \\ & \sigma_2^2 & & & \\ & & \sigma_3^2 & & \\ & & & \ddots & \\ & & & & \sigma_n^2 \end{pmatrix}$$

Propriedades dos estimadores

$$\underset{\sim}{\hat{\beta}}^{(OLS)} = \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{y} = \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\left(\underset{\sim}{X}\,\underset{\sim}{\beta}+\underset{\sim}{e}\right) = \underset{\sim}{\beta}+\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}$$

$$E\left(\underset{\sim}{\hat{\beta}}^{(OLS)}\right) = E\left[\underset{\sim}{\beta}+\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\right] = \underset{\sim}{\beta}+\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'E\left[\underset{\sim}{e}\right] = \underset{\sim}{\beta}$$

$$p\lim\left(\underset{\sim}{\hat{\beta}}^{(OLS)}\right) = p\lim\left(\underset{\sim}{\beta}+\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\right) = \underset{\sim}{\beta}$$

Conseqüências para os estimadores de Mínimos Quadrados (OLS)

Note que a suposição de homocedasticidade não desempenha papel algum na demonstração de que os estimadores de mínimos quadrados dos parâmetros do modelo de regressão são não viesados e consistentes;

Propriedades

$$\hat{\underset{\sim}{\beta}}^{(OLS)} = \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{y} = \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\left(\underset{\sim}{X}\,\underset{\sim}{\beta}+\underset{\sim}{e}\right) = \underset{\sim}{\beta}+\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}$$

$$Var\left(\hat{\underset{\sim}{\beta}}^{(OLS)}\right) = E\left[\left(\hat{\underset{\sim}{\beta}}^{(OLS)}-\underset{\sim}{\beta}\right)\left(\hat{\underset{\sim}{\beta}}^{(OLS)}-\underset{\sim}{\beta}\right)'\right] =$$

$$= E\left\{\left[\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\right]\left[\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\right]'\right\} =$$

$$= E\left\{\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\,\underset{\sim}{e}'\underset{\sim}{X}\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\right\} =$$

$$= \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'E\left\{\underset{\sim}{e}\,\underset{\sim}{e}'\right\}\underset{\sim}{X}\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1} =$$

$$= \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{\Omega}\,\underset{\sim}{X}\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}$$

Conseqüências para os estimadores de Mínimos Quadrados (OLS)

Aqui, verifica-se que a expressão usual de cálculo da variância dos estimadores, quando a suposição de homocedasticidade é válida,

$$Var\left(\hat{\underset{\sim}{\beta}}^{(OLS)}\right)=E\left[\left(\hat{\underset{\sim}{\beta}}^{(OLS)}-\underset{\sim}{\beta}\right)\left(\hat{\underset{\sim}{\beta}}^{(OLS)}-\underset{\sim}{\beta}\right)'\right]=\sigma^2\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}$$

não se aplica mais.

Conseqüências para os estimadores de Mínimos Quadrados (OLS)

- A partir das expressões anteriores, pode-se demonstrar que os estimadores das variâncias dos estimadores dos parâmetros do modelo de regressão são viesados, se não for válida a suposição de homocedasticidade, o que afeta o erro-padrão dos estimadores de mínimos quadrados;
- Isso significa que os intervalos de confiança e os testes *t* e *F* são prejudicados;
- Além disso, os estimadores de mínimos quadrados não são mais BLUE e nem assintoticamente eficientes.

Conseqüências para os estimadores de Mínimos Quadrados (OLS)

Quando existe heterocedasticidade, os estimadores usuais por mínimos quadrados dão mais peso para os resíduos com maior variância, já que a soma de quadrados dos resíduos (SSR) associados com os termos de maior variância tende a ser maior que aquela associada aos termos de menor variância.

A suposição de homocedasticidade é necessária para a determinação das distribuições das somas de quadrados e testes de hipóteses.

Conseqüências para os estimadores de Mínimos Quadrados (OLS)

Observação:

A suposição de homocedasticidade entra fundamentalmente na derivação das distribuições das variáveis presentes nos testes, toda a análise neles baseada não é válida (a falha na suposição de homocedasticidade é mais grave que a falha na suposição de normalidade).

Testes de Heterocedasticidade

A primeira forma de detectar a existência de heterocedasticidade é através da análise de resíduos (construir gráficos dos resíduos ao quadrado *versus* cada uma das variáveis explicativas e *versus* os valores ajustados da variável resposta).

Exemplo

Base de dados de Wooldridge (2006) com preços de imóveis americanos (*hprice1*) para estimar o modelo a seguir:

$$preco = \beta_0 + \beta_1 terreno + \beta_2 area + \beta_3 quartos + e$$

preco: **preço da casa (em milhares de dólares)**

terreno: **tamanho do terreno**

area: **tamanho da casa**

quartos: **número de quartos**

Exemplo

Estimativas de mínimos quadrados:

Dependent Variable: PRECO
Method: Least Squares
Date: 07/30/10 Time: 13:55
Sample: 1 88
Included observations: 88

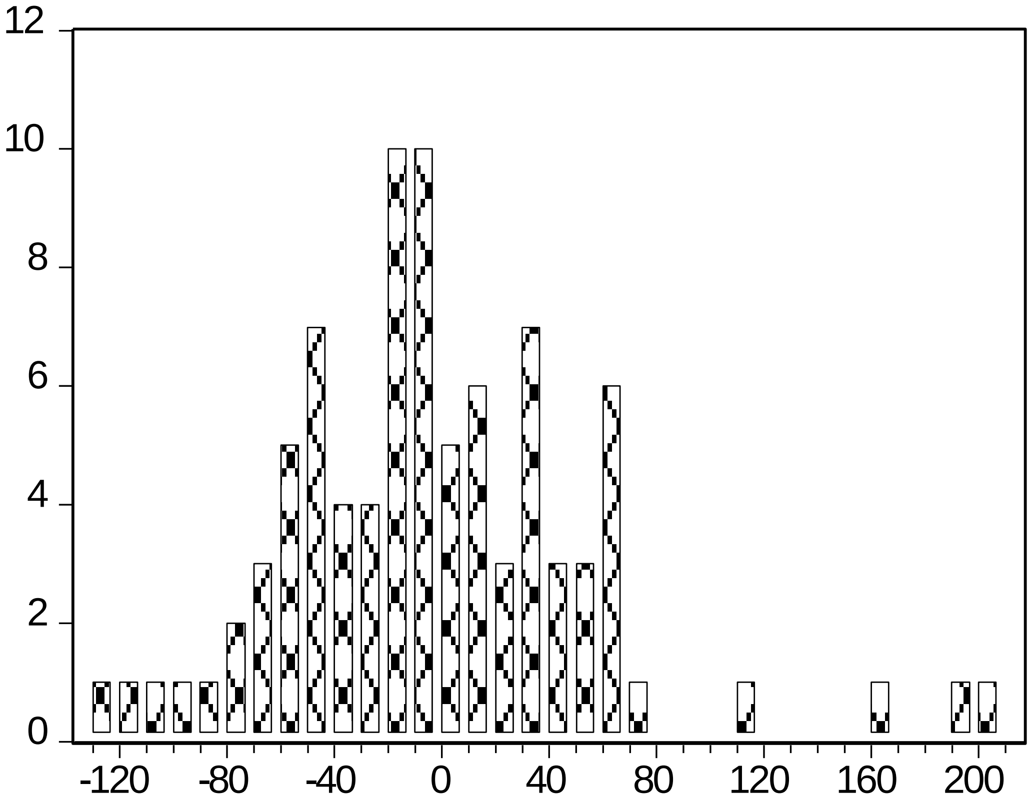
	Coefficient	Std. Error	t-Statistic	Prob.
C	-21.77031	29.47504	-0.738601	0.4622
TERRENO	0.002068	0.000642	3.220096	0.0018
AREA	0.122778	0.013237	9.275093	0.0000
QUARTOS	13.85252	9.010145	1.537436	0.1279

R-squared	0.672362	Mean dependent var	293.5460
Adjusted R-squared	0.660661	S.D. dependent var	102.7134
S.E. of regression	59.83348	Akaike info criterion	11.06540
Sum squared resid	300723.8	Schwarz criterion	11.17800
Log likelihood	-482.8775	Hannan-Quinn criter.	11.11076
F-statistic	57.46023	Durbin-Watson stat	2.109796
Prob(F-statistic)	0.000000		

Exemplo

Análise de resíduos:

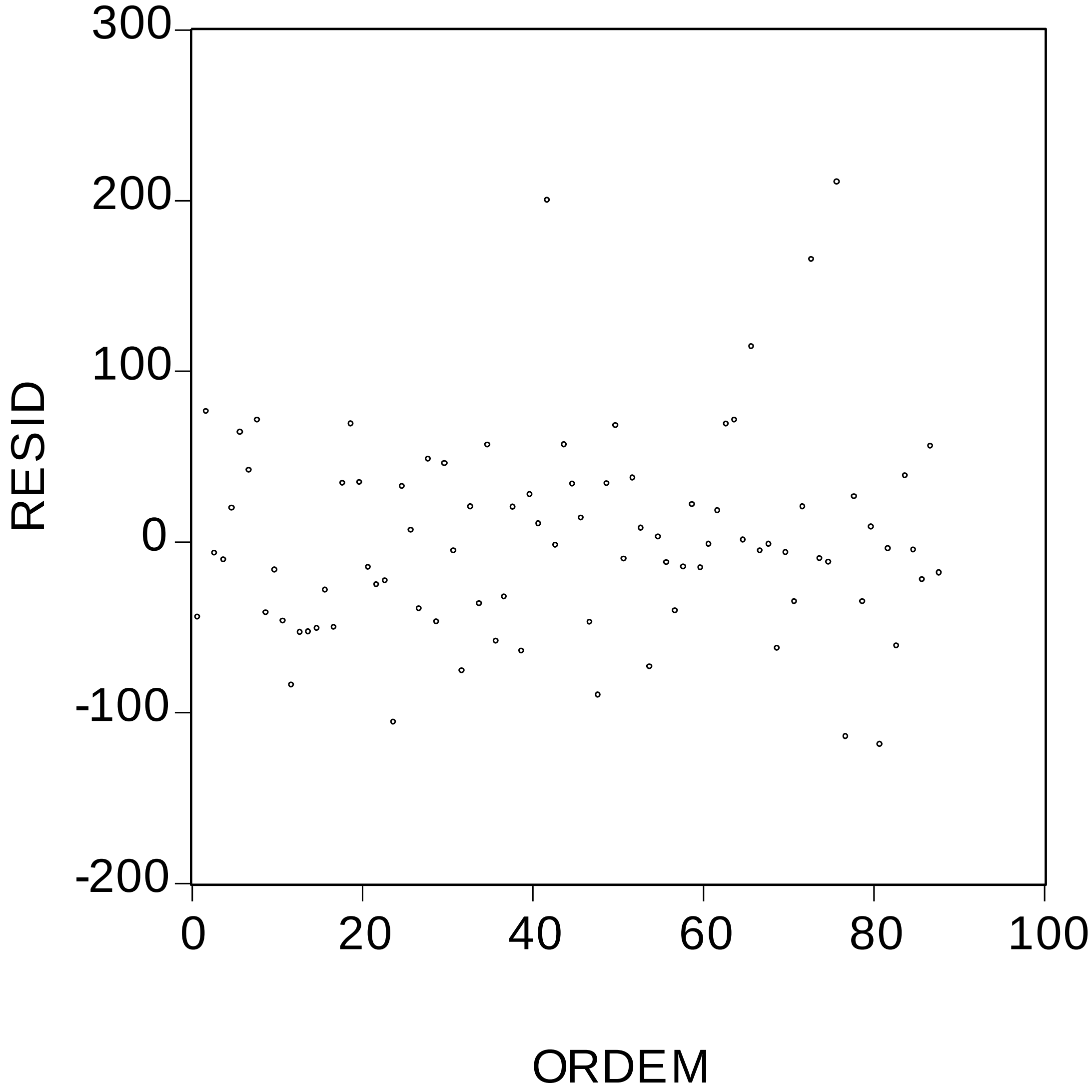
H_0: as observações são provenientes de uma população normalmente distribuída



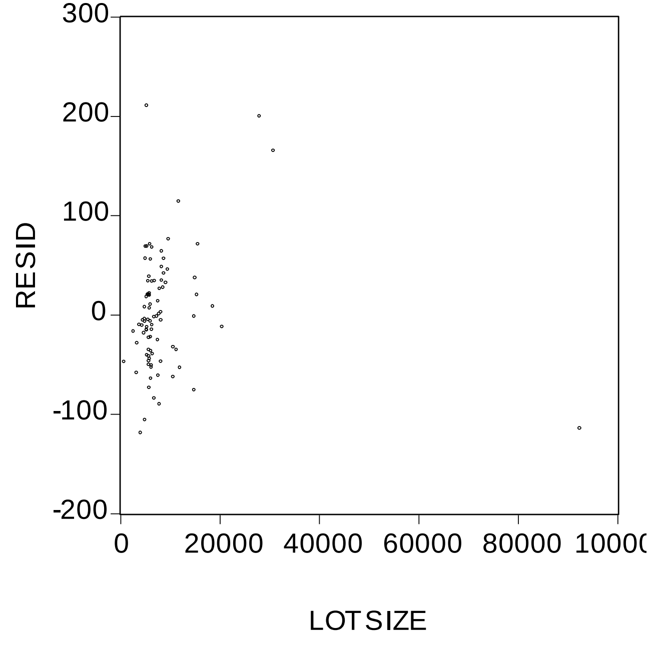
Series: Residuals	
Sample 1 88	
Observations 88	
Mean	8.07E-14
Median	-6.554850
Maximum	209.3758
Minimum	-120.0264
Std. Dev.	58.79282
Skewness	0.960683
Kurtosis	5.260844
Jarque-Bera	32.27791
Probability	0.000000

$$JB = \frac{n}{6} A\hat{s}s^2 + \frac{n}{24}\left(C\hat{u}rt - 3\right)^2 \sim \chi^2_2$$

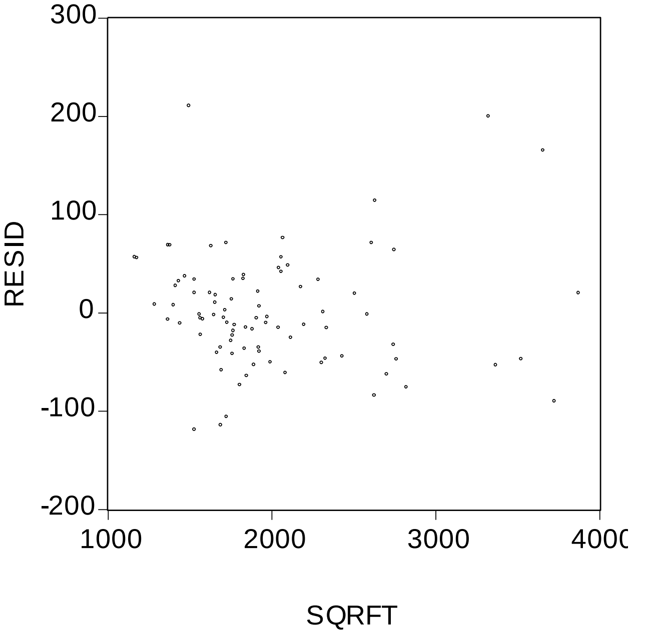
Exemplo



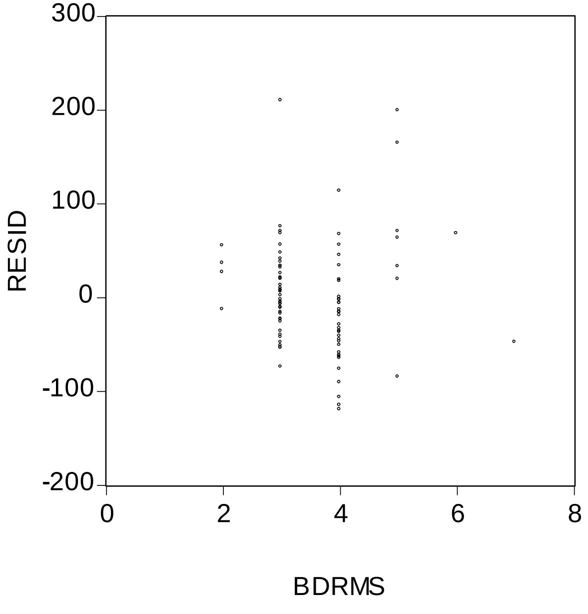
Exemplo



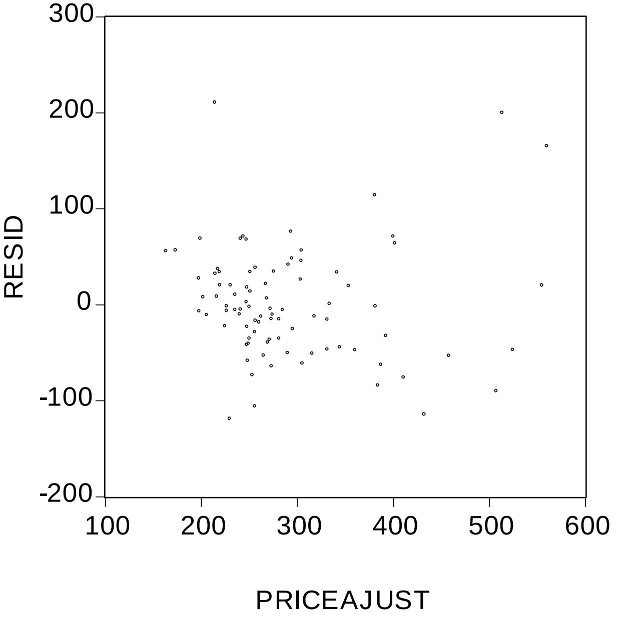
300
200
100
0
-100
-200
RESID
0
20000
40000
60000
80000
LOTSIZE



300
200
100
0
-100
-200
RESID
1000
2000
3000
SQRFT

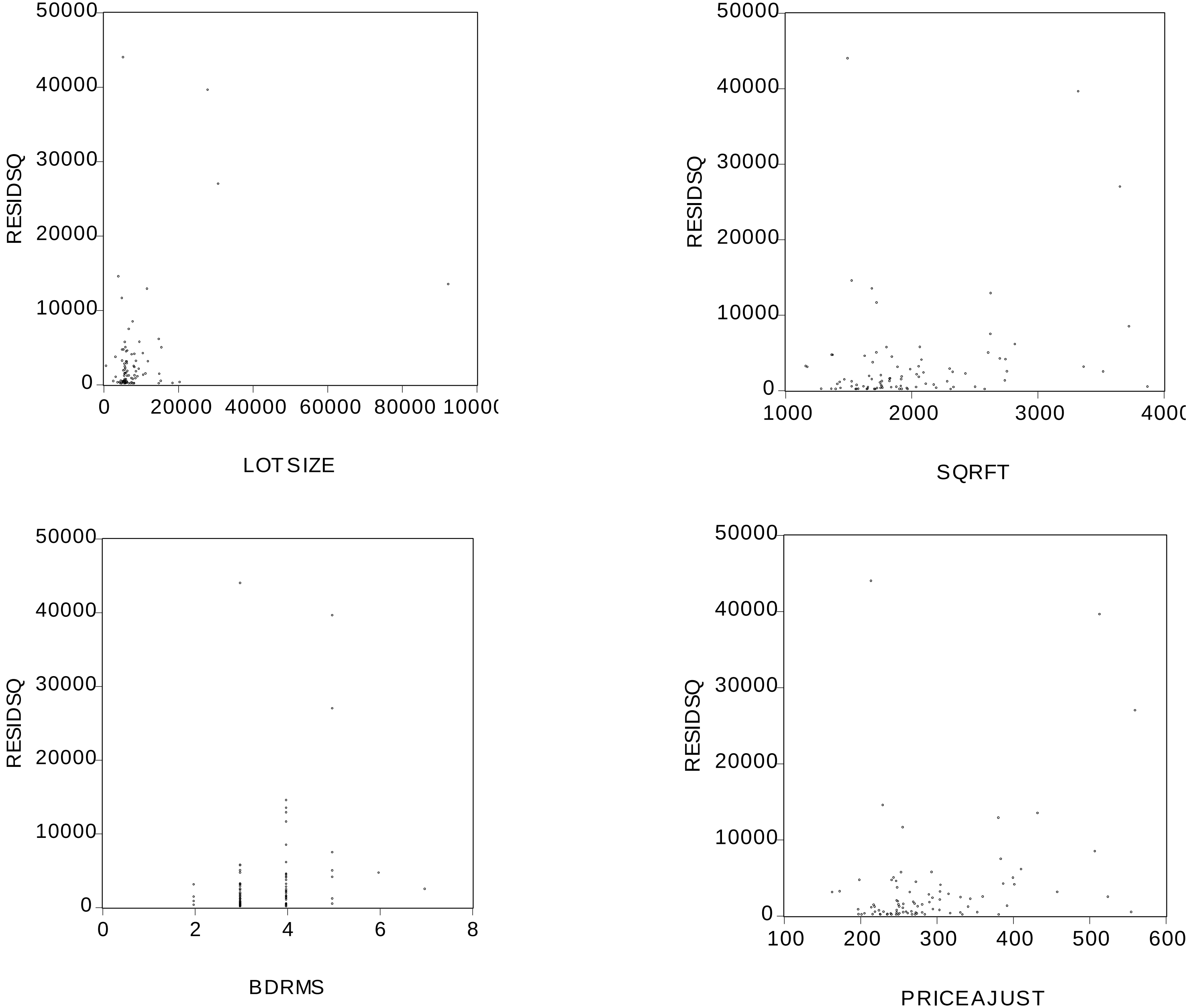


300
200
100
0
-100
-200
RESID
0
2
4
6
8
BDRMS



300
200
100
0
-100
-200
RESID
100
200
300
400
500
600
PRICEAJUST

Exemplo



Testes de Heterocedasticidade

Considerando $$H_0: \quad Var(e \mid x_1, x_2 \ldots, x_k) = \sigma^2$$

Homocedasticidade →

$$H_0 : Var\left(e \mid x_1, x_2, ..., x_k\right) =$$

$$= E\left(e^2 \mid x_1, x_2, ..., x_k\right) - \left[E\left(e \mid x_1, x_2, ..., x_k\right)\right]^2 =$$

$$= E\left(e^2 \mid x_1, x_2, ..., x_k\right) = \sigma^2$$

Para H_0 ser rejeitada, precisamos encontrar relação entre o e^2 e variáveis explicativas; por exemplo:

$e^2 = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + \nu$**, onde ν é variável erro.**

Testes de Heterocedasticidade

Neste caso,

H_0 (hipótese de homocedasticidade): $\delta_1 = ... = \delta_k = 0$,

a qual pode ser testada com uma estatística F ou LM.

Como não se conhece os erros no modelo populacional, deve estima-los, e, $\hat{e}_i$ é uma estimativa do erro e_i. Então, pode-se estimar a equação

$$\hat{e}^2 = \delta_0 + \delta_1 x_1 + ... + \delta_k x_k + \nu,$$

e calcular a estatística F ou LM para verificar a significância conjunta de de x_1, ..., x_k, como segue:

Testes de Heterocedasticidade

$$F_{obs} = \frac{R^2_{\hat{e}^2} / k}{(1 - R^2_{\hat{e}^2}) / (n - k - 1)} \sim F_{[k;\ n-k-1]}$$

ou

$$LM = n \cdot R^2_{\hat{e}^2} \sim \chi^2_k$$

Testes de Heterocedasticidade

- **Rejeitamos H_0 quando o valor observado for superior ao crítico;**
- **A versão LM deste teste é conhecida como TESTE DE BREUSCH-PAGAN (Teste BP).**

Observações:

- **Podemos considerar apenas um sub-conjunto das variáveis explicativas;**
- **Se H_0 for rejeitada, então, precisaremos recorrer a algum método de estimação que leve em conta a violação da suposição de homocedasticiadade.**

Exemplo

Voltando ao modelo,

$$preco = \beta_0 + \beta_1 terreno + \beta_2 area + \beta_3 quartos + e$$

a fim de se fazer um teste BP para verificar se os erros são homocedásticos. Após a estimação do modelo, o quadrado dos resíduos são salvos e regredidos em função das variáves explicativas:

Modelo estimado para os resíduos ao quadrado:

Dependent Variable: RESIDSQ
Method: Least Squares
Date: 07/30/10 Time: 14:43
Sample: 1 88
Included observations: 88

	Coefficient	Std. Error	t-Statistic	Prob.
C	-5522.795	3259.478	-1.694380	0.0939
TERRENO	0.201521	0.071009	2.837961	0.0057
AREA	1.691037	1.463850	1.155198	0.2513
QUARTOS	1041.760	996.3810	1.045544	0.2988

R-squared	0.160141	Mean dependent var	3417.316
Adjusted R-squared	0.130146	S.D. dependent var	7094.384
S.E. of regression	6616.646	Akaike info criterion	20.47695
Sum squared resid	3.68E+09	Schwarz criterion	20.58956
Log likelihood	-896.9860	Hannan-Quinn criter.	20.52232
F-statistic	5.338919	Durbin-Watson stat	2.351111
Prob(F-statistic)	0.002048		

Exemplo

Solução:

$\chi^2_{obs} = 88 * 0,16 = 14,08$

$\chi^2(0,05; 3) = 7,81$

p-valor = 0,0028

$F_{obs} = 5,34$

$F(0,05; 3; 84) = 2,71$

p-valor = 0,002048

Testes de Heterocedasticidade

TESTE DE WHITE

Este teste consiste em recorrer a uma suposição menos rigorosa do que a de homocedasticidade:

H_0: e^2 não é correlacionado com as variáveis explicativas, seus quadrados e seus produtos cruzados (interações).

Testes de Heterocedasticidade

TESTE DE WHITE

Por exemplo, quando o modelo contém k = 3 variáveis independentes, o teste de White fica baseado na estimação do modelo

$$\hat{e}^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3 +$$
$$+ \delta_4 x_1^2 + \delta_5 x_2^2 + \delta_6 x_3^2 +$$
$$+ \delta_7 x_1 x_2 + \delta_8 x_1 x_3 + \delta_9 x_2 x_3 + erro$$

Testes de Heterocedasticidade

TESTE DE WHITE

- **A utilidade deste teste consiste em identificar a forma de heterocedasticidade, ou de erro de especificação, ou de ambos.**
- **Comparado ao teste BP, este modelo tem 6 parâmetros a mais para ser estimado, logo, há uma perda no número de graus de liberdade.**
- **Para minimizar a perda de graus de liberdade, este teste pode ser feito com a regressão:**

$$\hat{e}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + erro$$

Exemplo

Voltando ao modelo para realizar o Teste de White a fim de verificar se os erros são homocedásticos:

$$preco = \beta_0 + \beta_1 terreno + \beta_2 area + \beta_3 quartos + e$$

Heteroskedasticity Test: White

F-statistic	5.386953	Prob. F(9,78)	0.0000
Obs*R-squared	33.73166	Prob. Chi-Square(9)	0.0001
Scaled explained SS	65.47818	Prob. Chi-Square(9)	0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 07/30/10 Time: 15:01
Sample: 1 88
Included observations: 88

	Coefficient	Std. Error	t-Statistic	Prob.
C	15626.24	11369.41	1.374411	0.1733
TERRENO	-1.859507	0.637097	-2.918719	0.0046
TERRENO^2	-4.98E-07	4.63E-06	-0.107498	0.9147
TERRENO*AREA	0.000457	0.000277	1.649673	0.1030
TERRENO*QUARTOS	0.314647	0.252094	1.248135	0.2157
AREA	-2.673918	8.662183	-0.308689	0.7584
AREA^2	0.000352	0.001840	0.191484	0.8486
AREA*QUARTOS	-1.020860	1.667154	-0.612337	0.5421
QUARTOS	-1982.841	5438.483	-0.364595	0.7164
QUARTOS^2	289.7541	758.8303	0.381843	0.7036

R-squared	0.383314	Mean dependent var	3417.316
Adjusted R-squared	0.312158	S.D. dependent var	7094.384
S.E. of regression	5883.814	Akaike info criterion	20.30444
Sum squared resid	2.70E+09	Schwarz criterion	20.58596
Log likelihood	-883.3955	Hannan-Quinn criter.	20.41786
F-statistic	5.386953	Durbin-Watson stat	2.052712
Prob(F-statistic)	0.000010		

Exemplo

Para minimizar a perda de graus de liberdade, será refeito o exemplo anterior utilizando o modelo

$$\hat{e}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + erro$$

para testar H_0: o erro é homocedástico ($\delta_1 = \delta_2 = 0$).

Exemplo

Dependent Variable: RESIDSQ
Method: Least Squares
Date: 07/30/10 Time: 15:13
Sample: 1 88
Included observations: 88

	Coefficient	Std. Error	t-Statistic	Prob.
C	19071.59	8876.227	2.148615	0.0345
PRECO_AJUST	-119.6554	53.31721	-2.244217	0.0274
PRECO_AJUST^2	0.208947	0.074596	2.801037	0.0063

R-squared	0.184868	Mean dependent var	3417.316
Adjusted R-squared	0.165689	S.D. dependent var	7094.384
S.E. of regression	6480.055	Akaike info criterion	20.42434
Sum squared resid	3.57E+09	Schwarz criterion	20.50880
Log likelihood	-895.6710	Hannan-Quinn criter.	20.45837
F-statistic	9.638819	Durbin-Watson stat	2.031774
Prob(F-statistic)	0.000169		

Exemplo

Solução:

$\chi^2_{obs} = 88 * 0{,}1848 = 16{,}26$

$\chi^2(0{,}05; 2) = 5{,}99$

p-valor = 0,000295

MÍNIMOS QUADRADOS GENERALIZADOS

Estimação via Mínimos Quadrados Generalizados

As técnicas inferenciais são componentes importantes em muitas análises de dados. Na presença de heterocedasticidade, como foi verificado anteriormente, toda a análise baseada em testes de hipóteses se torna inválida.

Será apresentado o ajuste dos erros-padrão e as estatísticas t e F, obtidos por mínimos quadrados, na presença de heterocedasticidade cuja forma é conhecida.

Estimação via Mínimos Quadrados Generalizados

$$\hat{Y}^* = X^* \hat{\beta} + \varepsilon^* \quad \text{onde}: \; Y^* = PY, \quad X^* = PX, \quad \varepsilon^* = P\varepsilon$$

$$\mathrm{E}\left[\varepsilon^* \varepsilon^* ' | \mathrm{X}\right] = \sigma^2 I$$

$$= P\mathrm{E}\left[\varepsilon\varepsilon'\right] P' = \sigma^2 I$$

$$= \mathrm{E}\left[P\varepsilon\varepsilon' P \;\; | \mathrm{X}\right] = \sigma^2 I$$

$$= P\Omega P' = \sigma^2 I$$

$$= P\sigma^2 \Phi P' = \sigma^2 I$$

$$= \sigma^2 P\Phi P' = \sigma^2 I$$

$$\Rightarrow \quad P\Phi P' = I \qquad\qquad \Phi^{-1} = P' P$$

Estimação via Mínimos Quadrados Generalizados

Pode-se demonstrar que no caso em que a matriz Φ conhecida, os parâmetros do modelo de regressão podem ser estimados a partir de um método chamado de Mínimos Quadrados Generalizados (GLS).

Estimação via Mínimos Quadrados Generalizados

Utilizando tal método obtemos:

$$\hat{\underset{\sim}{\beta}}^{(GLS)} = \left(\underset{\sim}{X}' \underset{\sim}{\Omega}^{-1} \underset{\sim}{X} \right)^{-1} \underset{\sim}{X}' \underset{\sim}{\Omega}^{-1} \underset{\sim}{y}$$

$$Var\left(\hat{\underset{\sim}{\beta}}^{(GLS)} \right) = \left(\underset{\sim}{X}' \underset{\sim}{\Omega}^{-1} \underset{\sim}{X} \right)^{-1}$$

Estimação de Mínimos Quadrados Ponderados

- **Detectando-se heterocedasticidade, é possível estimar erros padrão robustos em relação à heterocedasticidade após a estimação MQO.**

- **Antes das estatísticas robustas é possível modelar e estimar a forma específica da heterocedasticidade, calculando um estimador mais eficiente que o MQO, além de estatísticas *t* e *F* não viesadas. Porem, isso requer mais trabalho pois é preciso ser específico sobre a natureza da heterocedasticidade.**

Estimação de Mínimos Quadrados Ponderados

- **Para obter estimadores de β_j que tenham propriedades de eficiência melhores que MQO, estimamos a seguinte equação:**

$$y_i / \sqrt{h_i} = \beta_0 / \sqrt{h_i} + \beta_1 \left(x_{i1} / \sqrt{h_i} \right) + \cdots + \beta_k \left(x_{ik} / \sqrt{h_i} \right) + u_i / \sqrt{h_i}$$

- **Esta equação transformada satisfará as hipóteses do modelo linear clássico, se o modelo original também o fizer, com exceção da hipótese de homoscedasticidade.**

Estimação de Mínimos Quadrados Ponderados

- Considere a seguinte equação:

$$y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + u$$

- Assuma que *h(x)*é alguma função das variáveis explicativas que determina a heteroscedasticidade:

$$Var(u \mid x) = \sigma^2 h(x)$$

- *Com* *h(x)*>0 para todos valores possíveis das variáveis independentes.
- Supomos que a função *h(x)*é conhecida. Assim, mesmo que o parâmetro populacional σ^2 seja desconhecido, teremos condições de estimá-lo a partir de uma amostra de dados.

Estimação de Mínimos Quadrados Ponderados

- É necessário estimar os parâmetros da nova equação por mínimos quadrados ordinários. Os novos betas serão os estimadores de mínimos quadrados generalizados (GLS).

- Os erros-padrão, estatísticas *t* e estatísticas *F* podem ser obtidas de regressões que usem as variáveis transformadas.

- Por serem BLUE, os estimadores GLS são preferíveis aos estimadores OLS.

Estimação de Mínimos Quadrados Ponderados

- Os estimadores de mínimos quadrados generalizados (GLS) para correção da heteroscedasticidade são chamados de estimadores de mínimos quadrados ponderados (MQP) pois os novos betas minimizam a soma ponderada dos quadrados dos resíduos.

- A idéia é colocar menos peso nas observações com uma variância de erro mais alta enquanto o método OLS atribui pesos iguais a todas as observações, o que é melhor quando o erro é homocedastico.

Inferência após a estimação via Mínimos Quadrados (Estimadores Robustos)

Na prática é muito difícil conhecer a verdadeira forma como a heterocedasticidade se apresenta. Assim, precisamos buscar alguma metodologia que nos forneça resultados válidos na presença de heterocedasticidade cuja forma é desconhecida

Inferência após a estimação via Mínimos Quadrados (Estimadores Robustos)

Recentemente, muito se tem desenvolvido com relação ao ajuste de erros padrões, estatísticas t, F e LM para que os mesmos se tornem válidos na presença de heterocedasticidade.

Estes procedimentos são conhecidos como ROBUSTOS pois são válidos, pelo menos com amostras grandes, sendo ou não a variância do erro constante.

Inferência após a estimação via Mínimos Quadrados (Estimadores Robustos)

Considere que os erros são não correlacionados mas heterocedásticos, então a matriz Ω terá os elementos σ_1^2, σ_2^2, ..., σ_n^2 na diagonal principal. Assim, podemos escrever a expressão

$$Var\left(\underset{\sim}{\hat{\beta}}^{(OLS)}\right) = E\left[\left(\underset{\sim}{\hat{\beta}}^{(OLS)} - \underset{\sim}{\beta}\right)\left(\underset{\sim}{\hat{\beta}}^{(OLS)} - \underset{\sim}{\beta}\right)'\right] = E\left\{\left[\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\right]\left[\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\right]'\right\} =$$

$$= E\left\{\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{e}\,\underset{\sim}{e}'\underset{\sim}{X}\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\right\} = \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'E\left\{\underset{\sim}{e}\,\underset{\sim}{e}'\right\}\underset{\sim}{X}\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1} =$$

$$= \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\underset{\sim}{X}'\underset{\sim}{\Omega}\,\underset{\sim}{X}\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}$$

Inferência após a estimação via Mínimos Quadrados (Estimadores Robustos)

Como

$$Var\left(\hat{\underset{\sim}{\beta}}^{(OLS)}\right) = \left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\left(\sum_{i=1}^{n}\sigma_i^2 \underset{\sim}{x_i}\,\underset{\sim}{x_i}'\right)\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}$$

Não se conhece σ_i^2, $i = 1, 2, ..., n$. Porém, White (1980) demonstrou que uma estimativa bem simples dessas quantidades podem ser obtidas a partir do cálculo de $\hat{e}_i^2$ (quadrado do resíduo de OLS).

Inferência após a estimação via Mínimos Quadrados (Estimadores Robustos)

Assim,

$$Var\left(\underset{\sim}{\hat{\beta}}^{(OLS)}\right)=\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}\left(\sum_{i=1}^{n}\hat{e}_i^2\underset{\sim}{x_i}\underset{\sim}{x_i}'\right)\left(\underset{\sim}{X}'\underset{\sim}{X}\right)^{-1}$$

Ao se tomar a raiz quadrada dos elementos da diagonal principal desta matriz teremos o que usualmente costuma se chamar de erro-padrão devido a White (ou erro padrão robusto).

Inferência após a estimação via Mínimos Quadrados (Estimadores Robustos)

Observações:

- **O erro-padrão robusto pode ser maior ou menor do que o erro-padrão não robusto (não sabemos se o viés é para cima ou para baixo);**
- **Usando um método de estimação robusto, a estatística de teste t também será robusta.**

Exemplo

$$
\begin{array}{l}
preco = -21{,}77031 + 0{,}00207 \cdot terreno + 0{,}12277 \cdot area + 13{,}85252 \cdot quartos \\
\qquad (29{,}47504)\ (0{,}00064) \qquad (0{,}01324) \qquad (9{,}01014) \\
\qquad [37{,}13821]\ [0{,}00125] \qquad [0{,}01772] \qquad [8{,}47862] \\
\qquad \mathrm{n} = 88 \qquad \mathrm{R}^2 = 0{,}6724
\end{array}
$$

Entre colchetes encontram-se os erros padrões robustos.

Usando um método de estimação robusto, as estatísticas de testes (t, F e LM) também serão robustas.

Exemplo

Estatísticas *F* e *LM* Robustas:

- **Ajustar o modelo usando algum procedimento robusto:**

$$preco = \beta_0 + \beta_1 terreno + \beta_2 area + \beta_3 quartos + e$$

- **Testar a hipótese:**

$$H_0 : \beta_1 = \beta_3 = 0$$

$$H_0 : \beta_1 \neq 0 \quad e / ou \quad \beta_3 \neq 0$$

Exemplo

Wald Test:
Equation: Untitled

Test Statistic	Value	df	Probability
F-statistic	2.364911	(2, 84)	0.1002
Chi-square	4.729822	2	0.0940

Null Hypothesis Summary:

Normalized Restriction (= 0)	Value	Std. Err.
C(2)	0.002068	0.001251
C(4)	13.85252	8.478625

Restrictions are linear in coefficients.

$$F_{\text{crítico}} = F(0{,}05;\ 2;\ 84) = 3{,}11$$

$$\chi^2_{\text{crítico}} = \chi^2(0{,}05;\ 2) = 5{,}99$$

LM > χ_q^2 ou $F_{obs} > F_{[q;\ n\text{-}k\text{-}1]}$ rejeita a hipótese nula