Testing for heteroskedasticity.

Breusch-Pagan test for heteroskedasticity.

$Var(u|x) = E(u^2|x)$

$$\boxed{\begin{aligned} H_0 &: E(u^2|x) = \sigma^2 \\ H_1 &: E(u^2|x) = h(x) = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k \end{aligned}}$$

$u^2 = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k + v$. $E(v|x) = 0$.

$E(u^2|x) = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k$

$\Downarrow$

$H_0: \delta_1 = \delta_2 = \cdots = \delta_k = 0$

$H_1: H_0$ not true $\Rightarrow$ heteroskedasticity.

$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k + \text{error}$.

OLS residual from the regression of y on $x_1, \ldots, x_k$

$$F = \frac{(R_{ur}^2 - \underbrace{R_r^2}_{=0})/k}{(1 - R_{ur}^2)/(n-k-1)}.$$

test method 1

$$= \frac{R_{\hat{u}^2}^2/k}{(1 - R_{\hat{u}^2}^2)/(n-k-1)} \overset{asy}{\sim} F_{k,\, n-k-1}$$

model = lm(y ~ x1 + x2 + ---- + xk)

uhat = resid(model)

what_sq = What ^2.

A = lm (what_sq ~ $X_1 + X_2 + \cdots + X_k$).

summary (A) ← get $R^2_{\hat{u}^2}$

Lagrange Multiplier test of heteroskedasticity.

$LM = n \cdot R^2_{\hat{u}^2} \overset{asy}{\sim} \chi^2_k$ test method 2.

$\boxed{\text{White test}}$

- **The White test for heteroskedasticity**

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

Regress squared residuals on all explanatory variables, their squares, and in-teractions (here: example for k=3)

$$H_0 : \delta_1 = \delta_2 = \cdots = \delta_9 = 0$$

$$LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_9$$

The White test detects more general deviations from heteroskedasticity than the Breusch-Pagan test

- Disadvantage of this form of the White test
  - Including all squares and interactions leads to a large number of estimated parameters (e.g. k=6 leads to 27 parameters to be estimated).

- **Alternative form of the White test**

special cases.

$$\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + error$$

get $\hat{y}$ by { model = lm (y ~ $x_1 + \cdots + X_k$), y-hat = fitted (model) }

This regression indirectly tests the dependence of the squared residuals on the explanatory variables, their squares, and interactions, because the predicted value of y and its square implicitly contain all of these terms.

- Example: Heteroskedasticity in (log) housing price equations

$$H_0 : \delta_1 = \delta_2 = 0, \; LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_2$$

$$R^2_{\hat{u}^2} = .0392, LM = 88(.0392) \approx 3.45, p\text{-}value_{LM} = .178$$

- **Weighted least squares estimation**
- Heteroskedasticity is known up to a multiplicative constant

$$Var(u_i|\mathbf{x}_i) = \sigma^2 h(\mathbf{x}_i),\ \ h(\mathbf{x}_i) = h_i > 0$$

The functional form of the heteroskedasticity is known

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + u_i$$

$$\Rightarrow \left[\frac{y_i}{\sqrt{h_i}}\right] = \beta_0 \left[\frac{1}{\sqrt{h_i}}\right] + \beta_1 \left[\frac{x_{i1}}{\sqrt{h_i}}\right] + \cdots + \beta_k \left[\frac{x_{ik}}{\sqrt{h_i}}\right] + \left[\frac{u_i}{\sqrt{h_i}}\right]$$

$$\Leftrightarrow \ y_i^* = \beta_0 x_{i0}^* + \beta_1 x_{i1}^* + \cdots + \beta_k x_{ik}^* + u_i^*$$

Transformed model

$$Var\left(\frac{u_i}{h_i} \,\Big|\, x_i\right) = \frac{1}{h_i^2} Var(u_i | x_i)$$

$$= \frac{1}{h_i^2} \sigma^2 h_i = \frac{\sigma^2}{h_i}$$

$$Var\left(\frac{u_i}{\sqrt{h_i}} \,\Big|\, x_i\right) = \sigma^2$$

$\Longrightarrow$

$$y_i^* = \beta_0 x_{i0}^* + \beta_1 x_{i1}^* + \cdots + \beta_k x_{ik}^* + u_i^*$$

error is homoskedasticity.

- **OLS in the transformed model is weighted least squares (WLS)**

$$\min \sum_{i=1}^{n} \left(\left[\frac{y_i}{\sqrt{h_i}}\right] - b_0 \left[\frac{1}{\sqrt{h_i}}\right] - b_1 \left[\frac{x_{i1}}{\sqrt{h_i}}\right] - \cdots - b_k \left[\frac{x_{ik}}{\sqrt{h_i}}\right] \right)^2$$

$$\Leftrightarrow \ \min \sum_{i=1}^{n} (y_i - b_0 - b_1 x_{i1} - \cdots - b_k x_{ik})^2 / h_i$$

Observations with a large variance get a smaller weight in the optimization problem

- Why is WLS more efficient than OLS in the original model?
  - Observations with a large variance are less informative than observations with small variance and therefore should get less weight.
- WLS is a special case of generalized least squares (GLS)

- **Unknown heteroskedasticity function (feasible GLS)**

$$Var(u|\mathbf{x}) = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k) = \sigma^2 h(\mathbf{x})$$

Assumed general form of heteroskedasticity; exp-function is used to ensure positivity

$$u^2 = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k) \cdot v$$

Multiplicative error (assumption: independent of the explanatory variables)

$$\Rightarrow \log(u^2) = \alpha_0 + \delta_1 x_1 + \cdots + \delta_k x_k + e$$

$$\log(\hat{u}^2) = \hat{\alpha}_0 + \hat{\delta}_1 x_1 + \cdots + \hat{\delta}_k x_k + error$$

$$\Rightarrow \hat{h}_i = \exp(\hat{\alpha}_0 + \hat{\delta}_1 x_1 + \cdots + \hat{\delta}_k x_k)$$

Use inverse values of the estimated heteroskedasticity funtion as weights in WLS

Δ. Logarithmic transformation of the dependent variable often mitigates heteroskedasticity.

减轻.