

Lecture 28: The Breusch–Pagan test

Econometrics ■ Unit II: Multicollinearity and heteroskedasticity

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

B.Sc. (Hons) Statistics, Semester V
Academic Session 2026–2027

Previous lecture

Lecture 27: Bartlett's test.

Today

- Breusch–Pagan test
 - Setting
 - Procedure
 - Why the auxiliary response and why divide by 2

This test was proposed by Breusch and Pagan (1979). Suppose the variance may depend on some variables $z_2, \dots, z_p$ (often some or all of the regressors):

$$\sigma_i^2 = h(\alpha_1 + \alpha_2 z_{2i} + \dots + \alpha_p z_{pi}),$$

where h is a positive function, for example $h(t) = e^t$.

If $\alpha_2 = \dots = \alpha_p = 0$, then $\sigma_i^2 = h(\alpha_1)$ is constant. So the hypotheses are

$$H_0 : \alpha_2 = \alpha_3 = \dots = \alpha_p = 0 \quad \text{against} \quad H_1 : \text{not all are zero.}$$

The test assumes normal disturbances.

1. Estimate $y = X\beta + \varepsilon$ by OLS and obtain the residuals $e_1, \dots, e_n$.
2. Compute $\tilde{\sigma}^2 = \frac{1}{n} \sum_i e_i^2$ (the maximum likelihood estimate under H_0).
3. Construct $p_i = e_i^2 / \tilde{\sigma}^2$.
4. Regress p_i on a constant and $z_{2i}, \dots, z_{pi}$, and obtain the explained sum of squares $\text{ESS} = \sum_i (\hat{p}_i - \bar{p})^2$.
5. Compute

$$\Theta = \frac{\text{ESS}}{2} \underset{\sim}{\text{approx.}} \chi_{p-1}^2 \quad \text{under } H_0.$$

Reject H_0 if $\Theta > \chi_{\alpha, p-1}^2$.

Note that $\bar{p} = 1$, because $\sum_i p_i = \sum_i e_i^2 / \tilde{\sigma}^2 = n\tilde{\sigma}^2 / \tilde{\sigma}^2 = n$.

Why the auxiliary response and why divide by 2

- Since $E(\varepsilon_i) = 0$, $E(\varepsilon_i^2) = \sigma^2$. If the variance depends on z , the squared disturbances, and hence the squared residuals, will be related to z . The regression in Step 4 measures this relation.
- Under H_0 , p_i is close to ε_i^2/σ^2 , which has mean 1.
- Its variance is 2. Indeed, for a normal variable $E(\varepsilon_i^4) = 3\sigma^4$, so

$$\text{Var}(\varepsilon_i^2) = E(\varepsilon_i^4) - \{E(\varepsilon_i^2)\}^2 = 3\sigma^4 - \sigma^4 = 2\sigma^4, \quad \text{Var}\left(\frac{\varepsilon_i^2}{\sigma^2}\right) = \frac{2\sigma^4}{\sigma^4} = 2.$$

- In an ordinary regression, explained sum of squares divided by the error variance has a chi-square distribution. Here the “error variance” of p_i under H_0 is 2, which explains the division by 2.
- The $p - 1$ degrees of freedom are the number of slope coefficients $\alpha_2, \dots, \alpha_p$ being tested.

Example: variance different in two groups

Example 4.2 (Variance different in two groups)

An intercept-only model is fitted to 40 observations. In the first group of 20, ten outcomes are -1 and ten are $+1$; in the second group of 20, ten are -3 and ten are $+3$. The fitted mean is 0, so the residuals equal the outcomes.

Step 2. $\sum e_i^2 = 20(1) + 20(9) = 200$, so $\tilde{\sigma}^2 = 200/40 = 5$.

Step 3. $p_i = 1/5 = 0.2$ in the first group and $9/5 = 1.8$ in the second.

Example: variance different in two groups (cont.)

Example 4.2 (cont.)

Step 4. Let $z_i = 0$ in group 1 and 1 in group 2. The regression of p_i on $(1, z_i)$ fits the group means exactly: $\hat{p}_i = 0.2 + 1.6z_i$. With $\bar{p} = 1$,

$$\text{ESS} = 20(0.2 - 1)^2 + 20(1.8 - 1)^2 = 20(0.64) + 20(0.64) = 25.6.$$

Step 5. $\Theta = 25.6/2 = 12.8$ with $p - 1 = 1$ degree of freedom. Since $\chi^2_{0.05,1} = 3.8415$ and $12.8 > 3.8415$, reject homoskedasticity (p -value ≈ 0.00035).

Remarks

- The test detects only variance patterns related to the chosen z variables.
- Its chi-square distribution is a large-sample approximation.

Practice

- Exercise 4.4 (Exercises and Solutions booklet).

Next lecture

Lecture 29: The Goldfeld–Quandt test; comparison of the three tests.