

Lecture 27: Bartlett's 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 26: Heteroskedasticity: concept, sources and consequences for OLS.

Today

- Bartlett's test
 - Setting
 - Derivation by the likelihood ratio
 - The test

- Bartlett's test (Bartlett, 1937) is used when the observations fall into $m \geq 2$ independent groups, for example regions or firm-size classes. We want to test whether the error variance is the same in all groups.
- In group j ($j = 1, \dots, m$), fit the model separately. Let RSS_j be its residual sum of squares and ν_j its residual degrees of freedom ($\nu_j = n_j - 1$ if only a group mean is fitted, $\nu_j = n_j - k$ if a regression with k coefficients is fitted). Define

$$s_j^2 = \frac{\text{RSS}_j}{\nu_j}, \quad \nu = \sum_{j=1}^m \nu_j, \quad s_p^2 = \frac{\sum_j \nu_j s_j^2}{\nu} \text{ (pooled variance).}$$

- Assume the errors are normal. Then $\nu_j s_j^2 / \sigma_j^2 \sim \chi_{\nu_j}^2$, independently for different groups. The hypotheses are

$$H_0 : \sigma_1^2 = \dots = \sigma_m^2 \quad \text{against} \quad H_1 : \text{at least two variances differ.}$$

The residual information in group j is equivalent to ν_j independent $N(0, \sigma_j^2)$ observations whose sum of squares is $\nu_j s_j^2$. So the likelihood is

$$L = \prod_{j=1}^m (2\pi\sigma_j^2)^{-\nu_j/2} \exp \left\{ -\frac{\nu_j s_j^2}{2\sigma_j^2} \right\},$$

and, writing c for terms free of the variances,

$$\log L = c - \sum_{j=1}^m \frac{\nu_j}{2} \log \sigma_j^2 - \sum_{j=1}^m \frac{\nu_j s_j^2}{2\sigma_j^2}.$$

Step 1 (maximum without restriction). Differentiate with respect to σ_j^2 :

$$\frac{\partial \log L}{\partial \sigma_j^2} = -\frac{\nu_j}{2\sigma_j^2} + \frac{\nu_j s_j^2}{2\sigma_j^4} = 0 \implies \hat{\sigma}_j^2 = s_j^2.$$

Substituting back, the maximum is

$$\log L_1 = c - \sum_j \frac{\nu_j}{2} \log s_j^2 - \sum_j \frac{\nu_j}{2} = c - \sum_j \frac{\nu_j}{2} \log s_j^2 - \frac{\nu}{2}.$$

Step 2 (maximum under H_0). Put every $\sigma_j^2 = \sigma^2$:

$$\begin{aligned} \log L &= c - \frac{\nu}{2} \log \sigma^2 - \frac{\sum_j \nu_j s_j^2}{2\sigma^2}, \\ \frac{\partial \log L}{\partial \sigma^2} &= -\frac{\nu}{2\sigma^2} + \frac{\sum_j \nu_j s_j^2}{2\sigma^4} = 0 \implies \hat{\sigma}^2 = s_p^2. \end{aligned}$$

The maximum is $\log L_0 = c - \frac{\nu}{2} \log s_p^2 - \frac{\nu}{2}$.

Step 3 (the ratio). With $\Lambda = L_0/L_1$,

$$-2 \log \Lambda = 2(\log L_1 - \log L_0) = \nu \log s_p^2 - \sum_{j=1}^m \nu_j \log s_j^2.$$

Call this T . Under H_0 , T is approximately χ^2 with $m - 1$ degrees of freedom (m variances reduced to one).

Step 4 (Bartlett's correction). Bartlett improved the chi-square approximation by dividing T by

$$C = 1 + \frac{1}{3(m-1)} \left(\sum_{j=1}^m \frac{1}{\nu_j} - \frac{1}{\nu} \right).$$

$$B = \frac{\nu \log s_p^2 - \sum_j \nu_j \log s_j^2}{C} \underset{\text{approx.}}{\sim} \chi_{m-1}^2 \quad \text{under } H_0. \quad (4.1)$$

- Reject H_0 at level α if $B > \chi_{\alpha, m-1}^2$. Natural logarithms are used.
- $T \geq 0$ always, and $T = 0$ only when all s_j^2 are equal.
- This is because the log of a weighted average is at least the weighted average of the logs (the logarithm is concave): with weights ν_j/ν ,

$$\log s_p^2 \geq \sum_j (\nu_j/\nu) \log s_j^2.$$

- So a large B means the group variances are far apart.

Example 4.1 (Three groups)

Three independent normal groups each have 11 observations, and their sample variances are $s_1^2 = 4$, $s_2^2 = 9$, $s_3^2 = 16$. Then $\nu_1 = \nu_2 = \nu_3 = 10$ and $\nu = 30$.

Pooled variance.

$$s_p^2 = \frac{10(4 + 9 + 16)}{30} = \frac{29}{3} = 9.6667.$$

Numerator.

$$\begin{aligned} 30 \log(29/3) &= 30(2.26868) = 68.0605 \quad \text{and} \\ 10(\log 4 + \log 9 + \log 16) &= 10(1.38629 + 2.19722 + 2.77259) \\ &= 63.5611. \end{aligned}$$

So $T = 68.0605 - 63.5611 = 4.4994$.

Example 4.1 (cont.)

Correction.

$$C = 1 + \frac{1}{3(2)} \left(\frac{3}{10} - \frac{1}{30} \right) = 1 + \frac{0.26667}{6} = 1.04444.$$

Statistic. $B = 4.4994/1.04444 = 4.3080$. Since $\chi^2_{0.05,2} = 5.9915$ and $4.3080 < 5.9915$, H_0 is not rejected at the 5% level (p -value ≈ 0.116).

Remarks

- Bartlett's test is sensitive to non-normality: a rejection may be due to non-normal errors rather than unequal variances.
- It needs independent groups, each with its own fitted model; the residuals of one pooled regression must not simply be divided into groups.

Practice

- Exercise 4.3 (Exercises and Solutions booklet).

Next lecture

Lecture 28: The Breusch–Pagan test.