Heteroscedasticity

1. Nature of Heteroscedasticity

Heteroscedasticity refers to unequal variances of the error ε_i for different observations. It may be visually revealed by a "funnel shape" in the plot of the residuals e_i against the estimates $\hat{Y}_i$ or against one of the independent variables X_k. Effects of heteroscedasticity are the following

- heteroscedasticity *does not* bias OLS coefficient estimates
- heteroscedasticity means that OLS standard errors of the estimates are incorrect (often underestimated); therefore statistical inference is invalid
- heteroscedasticity means that OLS is not the best (= most efficient, minimum variance) estimator of β

2. Formal Diagnostic Tests for Heteroscedasticity

There are many diagnostic tests for heteroscedasticity. Tests vary with respect to the statistical assumptions required and their sensitivity to departure from these assumptions (robustness).

1. Brown-Forsythe Test

Properties

This test is robust against even serious departures from normality of the errors.

Principle

Find out whether the error variance σ_i^2 increases or decreases with values of an independent variable X_k (or with values of the estimates $\hat{Y}$) by the following procedure:

1. split the observations into 2 groups: one group with low values of X_k (or low values of $\hat{Y}$) and another group with high values of X_k (or high values of $\hat{Y}$)
2. calculate the median value of the residuals within each group, and the absolute deviations of the residuals from their group median
3. then do a t-test of the difference in the means of these absolute deviations between the two groups; the test statistic is distributed as t with (n – 2) df where n is the total number of cases

2. Breusch-Pagan *aka* Cook-Weisberg Test

Properties

This is a large sample test; it assumes normality of errors; it assumes σ_i^2 is a specific function of one or several X_k.

References

This test was developed independently by Breusch and Pagan (1979) and Cook and Weisberg (1983).

- Cook, R. D. and S. Weisberg. 1983. "Diagnostics for Heteroscedasticity in Regression." *Biometrika* 70:1-10.
- Breusch, T. S. and A. R. Pagan. 1979. "A Simple Test for Heteroscedasticity and Random Coefficient Variation." *Econometrica* 47:1287-1294.

3. Remedial Approach I: Transforming Y

If heteroscedasticity is found the first strategy is to try finding a transformation of Y that stabilizes the error variance. One can try various transformations along the ladder of powers or estimate the optimal transformation using the Box-Cox procedure. One variant of the Box-Cox procedure automatically finds the optimal transformation of Y given a multiple regression model with p independent variables. Note that transforming Y can change the regression relationship with the independent variables X_k.

4. Remedial Approach II: Weighted Least Squares (WLS)

1. Principle of WLS

Unequal error variance implies that the variance-covariance matrix of the errors ε_i,

$$\boldsymbol{\sigma}^2\{\boldsymbol{\varepsilon}\} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

is such that the variance σ_i^2 of ε_i may be different for each observation. Errors are still assumed uncorrelated across observations. Hence the off-diagonal entries of $\boldsymbol{\sigma}^2\{\boldsymbol{\varepsilon}\}$ are zeroes and the matrix is diagonal.

Assume that the σ_i^2 are known.
Then the weighted least squares (WLS) criterion is to minimize "

$$Q_w = \Sigma_{i=1 \text{ to } n}\, w_i(Y_i - \beta_0 - \beta_1 X_{i1} - \ldots - \beta_{p-1} X_{i,p-1})^2$$

where the weights $w_i=1/\sigma_i^2$ are inversely proportional to the σ_i^2; thus WLS gives *less weight* to observations with *large error variance*, and vice-versa.

2. WLS in Practice

1. Estimating the σ_i^2

In practice the σ_i^2 (and the weights w_i) are not known and must be estimated. The general strategy for estimating the σ_i^2 (and w_i) is

- estimate the regression of Y on the X_k with OLS and obtain the residuals e_i; then
  - e_i^2 is an estimator of σ_i^2
  - $|e_i|$ (the absolute value of e_i) is an estimator of σ_i
- on the basis of visual evidence (residual plots), regress either e_i^2 (to estimate the *variance function*) or $|e_i|$ (to estimate the *standard deviation function*) on
  - one X_k, or
  - several X_k, or
  - $\hat{Y}$ (from the OLS regression), or
  - a polynomial function of any of the above
- the fitted value (estimate) from the regression is an estimate
  - $\hat{v}_i$ of the variance σ_i^2 (if dependent variable is e_i^2), or
  - $\hat{s}_i$ of the standard deviation σ_i (if dependent variable is $|e_i|$)
- calculate the weights w_i as either
  - $w_i = 1/(\hat{s}_i)^2$ (if $\hat{s}_i$ was estimated), or
  - $w_i = 1/\hat{v}_i$ (if $\hat{v}_i$ was estimated)

2. Estimating the WLS Regression

Having estimated the w_i, the WLS regression can be done either

- using a WLS-capable program, by simply providing the program with a variable containing the weights, say w; the program automatically minimizes Q_w
- using OLS, by multiplying each variable (both dependent and independent, including the constant) by *the square root of the w_i* corresponding to a given observation and running an OLS regression without a constant with the transformed data

3. Recommendations on WLS

The WLS approach to heteroscedasticity has at least two drawbacks.

1. WLS usually necessitates strong assumptions about the nature of the error variance, e.g. that it is a function of particular X variable or of $\hat{Y}$. Sometimes the assumption appears reasonable (e.g., error variance is proportional to population size, when the units are real units); other times it is not.
2. WLS produces an alternative unbiased estimate of $\boldsymbol{\beta}$; but the OLS estimate is also unbiased. When $\mathbf{b}_{OLS}$ and $\mathbf{b}_{WLS}$ differ, which one should one choose?

5. Conclusion: Dealing with Heteroscedasticity

Provisional guidelines for dealing with the possibility of heteroscedasticity are

1. look at the plot of OLS residuals against estimates; if there is a suggestion of a funnel shape use a test of heteroscedasticity; you can use the Breusch-Pagan a.k.a. Cook-Weisberg test or another test (modified Levene or Goldfeld-Quandt) if you have a reason to, such as a small sample or doubts about normality of errors
2. if there is heteroscedasticity look first for a reasonable transformation that might stabilize the variances of the errors, but without introducing problems of interpretation or upsetting the functional relationship of Y with the independent variables; if such a transformation is found it is a desirable solution
3. if a suitable transformation cannot be found, investigate the possibility of WLS; try estimating the variance function or the standard deviation function; if a convincing function is found (one that has substantial R^2 and/or one that makes substantive sense) then try WLS; otherwise, use the robust standard error approach instead
4. if the transformation approach and the WLS approach do not seem promising, then use the robust standard errors approach
    - The alternative strategy can be used even when the form of the heteroscedasticity is unknown. It consists of
        1. estimating **b** using OLS as usual
        2. use a *heteroscedasticity consistent covariance matrix* (HCCM) to estimate the standard errors of the estimates; these standard errors are then called *robust standard errors*