LINEAR REGRESSION ANALYSIS

MODULE – XI

Lecture - 34

Autocorrelation

Dr. Shalabh
Department of Mathematics and Statistics
Indian Institute of Technology Kanpur

Tests of autocorrelation

Durbin-Watson (D-W) test

The Durbin-Watson (D-W) test is used for testing the hypothesis of lack of first order autocorrelation in the disturbance term. The null hypothesis is

$$H_0 : \rho = 0$$

Use OLS to estimate β in $y = X\beta + u$ and obtain residual vector

$$e = y - Xb = \bar{H}y$$

where

$$b = (X'X)^{-1}X'y,\ \bar{H} = I - X(X'X)^{-1}X'.$$

The D - W test statistic is .

$$d = \frac{\sum_{t=2}^{n}\left(e_t - e_{t-1}\right)^2}{\sum_{t=1}^{n} e_t^2}$$

$$= \frac{\sum_{t=2}^{n} e_t^2}{\sum_{t=1}^{n} e_t^2} + \frac{\sum_{t=2}^{n} e_{t-1}}{\sum_{t=1}^{n} e_t^2} - 2\frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=1}^{n} e_t^2}.$$

For large *n*,

$$d \approx 1 + 1 - 2r$$

$$d \simeq 2(1 - r).$$

where r is the sample autocorrelation coefficient from residuals based on OLSE and can be regarded as the regression coefficient of e_t on e_{t-1}. Here

positive autocorrelation of e_t's $\Rightarrow d < 2$

negative autocorrelation of e_t's $\Rightarrow d > 2$

zero autocorrelation of e_t's $\Rightarrow d \approx 2.$

As $-1 < r < 1$, so

if $-1 < r < 0$, then $2 < d < 4$ and

if $0 < r < 1$, then $0 < d < 2$.

So d lies between 0 and 4.

Since e depends on X, so for different data sets, different values of d are obtained. So the sampling distribution of d depends on X. Consequently exact critical values of d cannot be tabulated owing to their dependence on X. Durbin and Watson therefore obtained two statistics $\underline{d}$ and $\overline{d}$ such that

$$\underline{d} < d < \overline{d}$$

and their sampling distributions do not depend upon X.

Considering the distribution of $\underline{d}$ and $\overline{d}$, they tabulated the critical values as d_L and d_U respectively. They prepared the tables of critical values for $15 < n < 100$ and $k \leq 5$. Now tables are available for $6 < n < 200$ and $k \leq 10$.

The test procedure is as follows:

$H_0 : \rho = 0$			
Nature of H_1	Reject H_0 when	Retain H_0 when	The test is inconclusive when
$H_1 : \rho > 0$	$d < d_L$	$d > d_U$	$d_L < d < d_U$
$H_1 : \rho < 0$	$d > (4 - d_L)$	$d < (4 - d_U)$	$(4 - d_U) < d < (4 - d_L)$
$H_1 : \rho \neq 0$	$d < d_L$ or $d > (4 - d_L)$	$d_U < d < (4 - d_U)$	$d_L < d < d_U$ or $(4 - d_U) < d < (4 - d_L)$
Values of d_L and d_U are obtained from tables.			

Limitations of D-W test

1. If d falls in the inconclusive zone, then no conclusive inference can be drawn. This zone becomes fairly larger for low degrees of freedom. One solution is to reject H_0 if the test is inconclusive. A better solutions is to modify the test as
   - Reject H_0 when $d < d_U$.
   - Accept H_0 when $d \geq d_U$.

   This test gives satisfactory solution when values of x_i' s change slowly, e.g., price, expenditure etc.

2. The D-W test is not applicable when intercept term is absent in the model. In such a case, one can use another critical values, say d_M in place of d_L. The tables for critical values d_M are available.

3. The test is not valid when lagged dependent variables appear as explanatory variables. For example,

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \ldots. + \beta_r y_{t-r} + \beta_{r+1} x_{t1} + \ldots + \beta_k x_{t,k-r} + u_t$$
$$u_t = \rho u_{t-1} + \varepsilon_t.$$

In such case, **Durbin's *h* test** is used which is given as follows.

Durbin's *h*-test

Apply OLS to

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + ... + \beta_r y_{t-r} + \beta_{r+1} x_{t1} + ... + \beta_k x_{t,k-r} + u_t$$

$$u_t = \rho u_{t-1} + \varepsilon_t$$

and find OLSE b_1 of β_1. Let its variance be $Var(b_1)$ and its estimator is $\widehat{Var}(b_1)$. Then the Dubin's *h*-statistic is

$$h = r\sqrt{\frac{n}{1 - n\ \widehat{Var}(b_1)}}$$

which is asymptotically distributed as $N(0,1)$ and

$$r = \frac{\sum_{t=2}^{n} e_t e_{t-1}}{\sum_{t=2}^{n} e_t^2}.$$

This test is applicable when *n* is large. When $\left[1 - n\widehat{Var}(b_1)\right] < 0,$ then test breaks down. In such cases, the following test procedure can be adopted.

Introduce a new variable ε_{t-1} to $u_t = \rho u_{t-1} + \varepsilon_t$. Then

$$e_t = \delta\rho_{t-1} + y_t.$$

Now apply OLS to this model and test $H_{0A} : \delta = 0$ versus $H_{1A} : \delta \neq 0$ using *t* - test . It H_{0A} is accepted then accept $H_0 : \rho = 0.$

If $H_{0A} : \delta = 0$ is rejected, then reject $H_0 : \rho = 0.$

4. If $H_0 : \rho = 0$ is rejected by D-W test, it does not necessarily mean the presence of first order autocorrelation in the disturbances. It could happen because of other reasons also, e.g.,

- distribution may follows higher order *AR* process.
- some important variables are omitted.
- dynamics of model is misspecified.
- functional term of model is incorrect.