Autocorrelation

Pravat Uprety

Central Department of Statistics

Tribhuvan University

Autocorrelation

- One of the assumptions of the Classical Linear Regression Model (CLRM) is that the disturbance term of the model is independent.

$$\textbf{Cov}\ (\epsilon_t, \epsilon_s) = \textbf{E}\ (\epsilon_t, \epsilon_s) = 0 \text{ for } t \neq s.$$

This feature of regression disturbance is known as serial independence or non autocorrelation. It implies that the value of disturbance term in one period is not correlated with its value in another period.

In time series the disturbance term at period t may be related with the disturbance term at t-1, t-2, …..and t+1, t+2,…….and so on. In that case $cov(\epsilon_t, \epsilon_s) \neq 0$ for $t \neq s$ and we and we say that the disturbances are autocorrelated.

Specification of Autocorrelation Relationship

$\epsilon_t = \rho\ \epsilon_{t-1} + u_t$ ---------- (1)

$E(u_t) = 0$

$Var\ (u_t) = E(u_t^2) = \sigma_u^2$ for all t.

u_t is normally distributed

$E(u_t\ u_{t-1}) = 0$

Equation 1 is known as first order autoregression scheme and it is denoted by AR (1).

By successive substitution for ϵ_t, ϵ_{t-1}, …..in (1) we get,

$\epsilon_t = \rho(\rho\, \epsilon_{t-2} + u_{t-1}) + u_t$

$= \rho^2\, \epsilon_{t-2} + \rho\, u_{t-1} + u_t$

$= \rho^2\, (\rho\, \epsilon_{t-3} + u_{t-2}) + \rho\, u_{t-1} + u_t$

$= \rho^3 \epsilon_{t-3} + \rho^2 u_{t-2} + \rho\, u_{t-1} + u_t$

$= u_t + \rho\, u_{t-1} + \rho^2 u_{t-2} + \rho^3 u_{t-3} + \ldots\ldots\ldots\ldots$ ---------------(2)

This shows that under the first-order autoregressive scheme, the effect of past disturbances wears off gradually as $|\rho| < 1$.

- ρ represents the correlation coefficient between ϵ_t and ϵ_{t-1} (first order autocorrelation), ρ^2 is the correlation coefficient between ϵ_t and ϵ_{t-2} (second order autocorrelation) and ρ^s is the correlation coefficient between ϵ_t and ϵ_{t-s} (S order autocorrelation).

- When $\rho = 0$ there is no autocorrelation because ϵ_t becomes u_t and which does not suffer from autocorrelation.
- The strength of autocorrelation becomes high as ρ approaches to unity (+1).
- If ρ approaches to -1, the strength of autocorrelation becomes high again.

Mean

$E(\epsilon_t) = 0$

Variance of ϵ_t

$$Var(\epsilon_t) = \frac{\sigma_u^2}{1-\rho^2} = \sigma_\epsilon^2$$

Covariance of ϵ_t and ϵ_{t-1}

$Cov(\epsilon_t \epsilon_{t-1}) = \rho\, \sigma_\epsilon^2$

Consequences of autocorrelation are

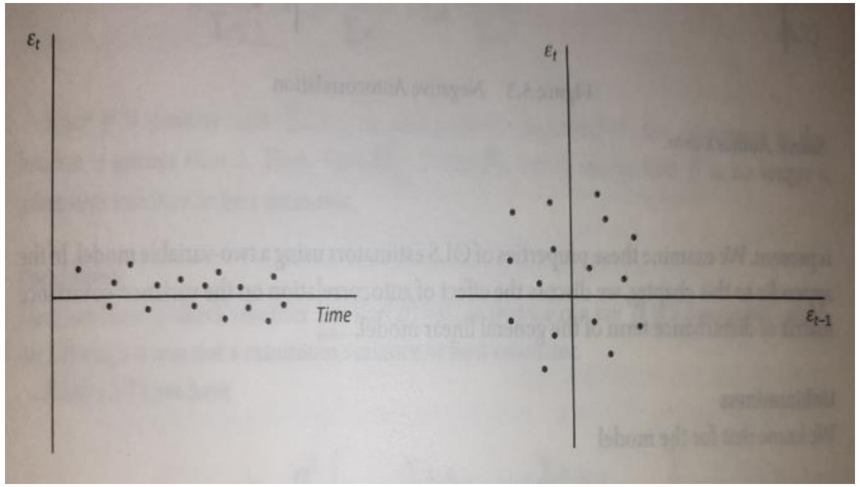
1) The OLS estimators are still unbiased and consistent.
2) The OLS estimators are no longer minimum variance or best estimators. Hence, they are not efficient and BLUE.
3) If we disregard the problem of autocorrelation and believe that all assumptions are valid, following problems will arise

- The estimated variance of disturbance term will be under estimate of its true variance.
- The standard error of the estimated slope coefficient will be much smaller if it is computed usual OLS formula. Which will provide spurious (wrong) impression about the statistical significance.
- The usual t and F test will become invalid.

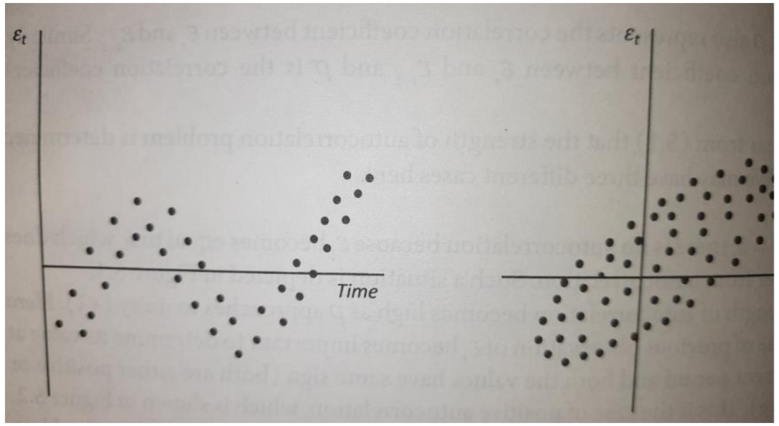
Detection

- **Graphical Method (Residual Plot)**
  - Graph of ϵ_t against time
  - Graph of ϵ_t against ϵ_{t-1}

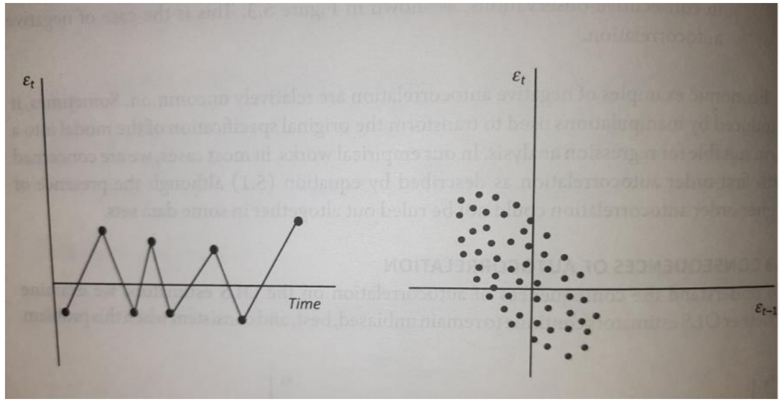
1. When $\rho = 0$ there is no autocorrelation because ϵ_t becomes u_t and which does not suffer from autocorrelation.



2) The strength of autocorrelation becomes high as ρ approaches to unity (+1). Here the value of previous observation of ϵ_t becomes important to determine its value at the current period and both the values have same sign (both are either positive or negative). This is the case of +ve autocorrelation.



3) If ρ approaches to -1, the strength of autocorrelation becomes high again. Here also the past values of the disturbance term become important in determining its value in the current period but the signs of the disturbance term switch in consecutive observations. This is the case of –ve autocorrelation.



Durbin Watson Test (D-W test)

This is the simplest and most widely used test for autocorrelation. It is based on following assumptions:

a) The regression model includes a constant or intercept term.

b) We are examining presence of first order autocorrelation.

c) The regression model does not include a lagged dependent variable as an explanatory variable.

Now, to understand how the test is performed, consider the model

$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots\ldots\ldots\ldots\ldots\ldots + \beta_k X_{kt} + \epsilon_t$

Where $\epsilon_t = \rho\, \epsilon_{t-1} + u_t \qquad |\rho| < 1$

Null hypothesis (H_0): $\rho = 0$

Or $H_0 : \rho \geq 0$

Or $H_0 : \rho \leq 0$

Alternative hypothesis (H_1): $\rho \neq 0$

Or $H_1 : \rho < 0$

Or $H_1 : \rho > 0$

The Durbin-Watson (D-W) statistic is

$$d = \frac{\sum_{t=2}^{n}(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{n}\hat{\varepsilon}_t^{\,2}} = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n}e_t^{\,2}} = \frac{\sum_{t=2}^{n}e_t^2 + \sum_{t=2}^{n}e_{t-1}^2 - 2\sum_{t=2}^{n}e_t e_{t-1}}{\sum_{t=1}^{n}e_t^{\,2}}$$

For large sample

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

Decision Rule

From D-W table

We can get the value of $\mathbf{d_L}$ and d_u for given value of n, k and α.

If	$0 < d < d_L$	reject H_0 (+ve autocorrelation)
If	$d_L \leq d \leq d_u$	no conclusion
If	$d_u < d < 4 - d_u$	do not reject H_0 (no significant autocorrelation)
If	$4 - d_u \leq d \leq 4 - d_L$	no conclusion
If	$4 - d_L < d \leq 4$	reject H_0 (-ve autocorrelation)

Limitations

1) It can not be used for testing higher order autocorrelation
2) This test is biased towards non-rejection of null hypothesis when a lagged dependent variable is included in the model as an explanatory variable.
3) It becomes inapplicable if the model does not contain intercept term.
4) Sometime, the test produces no conclusion result.
5) This test is not robust for small sample.

Example

Suppose that the residuals for a set of data collected over 12 consecutive time periods were as follows:

Time Period	1	2	3	4	5	6	7	8	9	10	11	12
Residual	+5	+6	+3	-4	-4	+2	0	-5	-3	+4	-2	-2

Compute the Durbin-Watson statistic. At the 0.05 level of significance, is there evidence of positive autocorrelation among the residuals?

Breusch-Godfrey Lagrange Multiplier Test (LM test)

Given the problems of Durbin-Watson test, it is always good to use some other test for autocorrelation. Although there are many alternative here, the most widely used test is the Breusch- Godfrey Lagrange Multiplier test (BG test) developed by Breusch (1978) and Godfrey (1978). This test can pick up higher order autocorrelation and can be performed using many econometrics software packages, including Eviews/Stata/R/python. It is also a more powerful test as it is not biased towards non rejection of the null hypothesis when lagged dependent variables is included in the model as an explanatory variable.

To understand the BG test procedure, consider the model,

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots\dots\dots\dots\dots\dots + \beta_k X_{kt} + \epsilon_t \quad \text{---------------------(1)}$$

Where,

$$\epsilon_t = \rho_1 \epsilon_{t-1} + \rho_2 \epsilon_{t-2} + \dots\dots\dots\dots\dots\dots + \rho_p \epsilon_{t-p} + u_t \quad \text{----------------------(2)}$$

Null hypothesis (H_0): $\rho_1 = \rho_2 = \rho_3 = \dots\dots\dots\dots = \rho_p = 0$ (no auto correlation)

Alternative hypothesis (H_1): at least one ρ is not zero (autocorrelation)

Steps

1) Estimate equation (1) by OLS and estimate $\hat{\varepsilon}_t = e_t$
2) Run the auxiliary regression of e_t on X_{1t}, X_{2t},........., X_{kt}, e_{t-1}, e_{t-2}, ...e_p.
3) To test the validity of above null hypothesis compute LM statistic as follows

$$LM = (n-p)\, R^2$$

Where R^2 is coefficient of determination of auxiliary equation

n is no of observations

p is order of autocorrelation

This LM statistic follows a chi square distribution with degrees of freedom P.

4) If this $\chi^2 \leq \chi^2_{p,\alpha}$ (tabulated)

   We do not reject H_0 (no autocorrelation)

   If this $\chi^2 \geq \chi^2_{p,\alpha}$ (tabulated)

   We reject H_0 (autocorrelation)

Remedial Measure

- When the value of ρ is known (Transformation)

Consider the model,

$Y_t = \alpha + \beta X_t + \epsilon_t$ ---------------(1)

Where ϵ_t is assumed to follow first order autocorrelation, so that

$\epsilon_t = \rho \epsilon_{t-1} + u_t$ ------------------(2)

Now lagging equation (1) by one period and multiplying it by ρ.

$\rho Y_{t-1} = \rho\alpha + \rho\beta X_{t-1} + \rho \epsilon_{t-1}$ -------------(3)

Now, subtracting equation (3) from equation (1)

$Y_t - \rho Y_{t-1} = (1-\rho) \alpha + \beta (X_t - \rho X_{t-1}) + (\epsilon_t - \rho\epsilon_{t-1})$ --------------(4)

Therefore $Y_t^* = \alpha^* + \beta X_t^* + u_t$ [since $u_t = \epsilon_t - \rho\epsilon_{t-1}$]

- It is clear that we lost one observation with this transformation. In order to avoid this loss of observation, it is suggested that Y1 and X1 should be transformed for first observation as follows:

$Y_1^* = Y_1 \sqrt{1-\rho^2}$ $X_1^* = X_1 \sqrt{1-\rho^2}$

The transformation that generated Y_t^*, X_t^* and α^* is known as quasi- differencing or generalized differencing.

When the value of ρ is unknown

Cochrane – Orcutt (1949) Iterative Procedure

Under this model, we estimate the following equation by OLS method

$Y_t = \alpha + \beta X_t + \epsilon_t$

Then we obtain $\hat{\varepsilon}_t = e_t$

And which is used to calculate $\hat{\rho}$ as

$$\hat{\rho} = \frac{\sum e_t e_{t-1}}{\sum e_t^2}$$

Then,

$Y_t - \hat{\rho}\, Y_{t-1} = \alpha^* + \beta\,(X_t - \hat{\rho}\, X_{t-1}) + (\epsilon_t - \hat{\rho}\, \epsilon_{t-1})$

Thus a two step estimation is involved here

Suppose, α' and β' are Cochrane-Orcutt estimation of α and β obtaining new set of residuals

$\epsilon_t' = Y_t - \alpha' - \beta' X_t$

Also estimating new ρ say $\hat{\hat{\rho}}$ as

$$\hat{\hat{\rho}} = \frac{\sum e_t' e_{t-1}'}{\sum e_t'^2}$$

Again use $\hat{\hat{\rho}}$ to construct the model,

$Y_t - \hat{\hat{\rho}}\, Y_{t-1} = \alpha^* + \beta\,(X_t - \hat{\hat{\rho}}\, X_{t-1}) + (\epsilon_t - \hat{\hat{\rho}}\, \epsilon_{t-1})$

And this procedure is continued until estimated values of $\hat{\alpha}$ and $\hat{\beta}$ converge.

Note: Taking difference (in time series)

Command in R

Data File: hprice1

```
library(car)
reg_hprice<-lm(price~lotsize+sqrft+bdrms, data=hprice1)
summary(reg_hprice)
summ(reg_hprice)
```

Residual Plot

```
residualPlot(reg_hprice)
```

Durbin Watson Test

durbinWatsonTest(name of fitted model)

```
durbinWatsonTest(reg_hprice)
```

Breusch-Godfrey test

- bgtest(reg_hprice)