LECTURE - APPLIED Regression

Time Series - 1

Lecture 10 - Part 1

Many applications of regression involve both predictor and response variables that are time series, that is, the variables are time-oriented. Regression models using time series data occur relatively often in economics, business, stock market etc. The usual assumptions are not usually appropriate for time series data. Usually the errors in time series data exhibit some type of autocorrelated structure. By autocorrelation we mean that the errors are correlated with themselves at different time periods.

The presence of autocorrelation in the errors has several effects on the ordinary least-squares regression procedure.

1. The ordinary least squares (OLS) regression coefficients are still unbiased, but they are no longer minimum-variance estimates.

2. When the errors are positively autocorrelated, the residual mean square may seriously underestimate the error variance σ^2. Consequently, many of the inferences based on confidence interval and testing are not correct. Generally, underestimation of the error variance σ^2 gives the results a false impression of precision of estimation and potential forecast accuracy.

3. The confidence intervals, prediction intervals, and tests of hypotheses based on the t and F distributions are, strictly speaking, no longer valid.

DETECTING AUTOCORRELATION: THE DURBINWATSON TEST

Residual plots can be useful for the detection of autocorrelation. The most useful display is the plot of residuals versus time.

If there is positive autocorrelation, residuals of identical sign occur in clusters. That is, there are not enough changes of sign in the pattern of residuals. On the other hand, if there is negative autocorrelation, the residuals will alternate signs too rapidly.

The test developed by Durbin and Watson is a very widely used procedure. This test is based on the assumption that the errors in the regression model are generated by a first-order autoregressive process observed at equally spaced time periods, that is,

$$\epsilon_t = \phi\, \epsilon_{t-1} + a_t$$

where ϵ_t is the error term in the model at time period t, and $a_t \sim i.i.d\ N(0, \sigma_a^2)$ and ϕ is a parameter that defines the relationship between successive values of the model errors ϵ_t & ϵ_{t-1} and the time index is t = 1, 2, ..., T. For technical reason we require that $|\phi| < 1$. In time series regression models ϕ is called the autocorrelation parameter.

Thus, a simple linear regression model with first-order autoregressive errors would be

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t, \quad \text{where} \quad \epsilon_t = \phi\, \epsilon_{t-1} + a_t$$

where y_t & x_t are the observations on the response and predictor variables at time period t.

By successively substituting for ϵ_t, $\epsilon_{t-1}, \ldots$ on the right-hand side of the above equation we obtain

$$\epsilon_t = \sum_{j=0}^{\infty} \phi^j\, \epsilon_{t-j}$$

Furthermore we can show that,

$$E(\epsilon_t) = 0,\;\; Var(\epsilon_t) = \sigma^2 = \sigma_a^2 \left(\frac{1}{1-\phi^2}\right) \text{ and } Cov(\epsilon_t, \epsilon_{t\pm j}) = \phi^j\, \sigma_a^2 \left(\frac{1}{1-\phi^2}\right)$$

The autocorrelation between two errors that are one period apart is called **lag one autocorrelation**, and it is

$$\rho_1 = \frac{Cov(\epsilon_t, \epsilon_{t+1})}{\sqrt{Var(\epsilon_t)\, Var(\epsilon_{t+1})}}$$

$$= \frac{\phi\, \sigma_a^2 \left(\frac{1}{1-\phi^2}\right)}{\sqrt{\sigma_a^2 \left(\frac{1}{1-\phi^2}\right) \sigma_a^2 \left(\frac{1}{1-\phi^2}\right)}} = \phi$$

Similarly, the autocorrelation between two errors that are k periods apart is $\rho_k = \phi^k$.

The DurbinWatson test is a statistical test for the presence of positive autocorrelation in regression model errors. Specifically, the hypotheses considered in the DurbinWatson test are $H_0 : \phi = 0 \; vs \; H_1 : \phi > 0$.

The DurbinWatson test statistic is

$$
\begin{aligned}
d &= \frac{\sum_{i=2}^{T} (e_t - e_{t-1})^2}{\sum_{i=1}^{T} e_t^2} \\
&= \frac{\sum_{i=2}^{T} e_t^2 + \sum_{i=2}^{T} e_{t-1}^2 - 2 \sum_{i=2}^{T} e_t \, e_{t-1}}{\sum_{i=1}^{T} e_t^2} \approx 2(1 - r_1)
\end{aligned}
$$

For uncorrelated errors (i.e $r_1 \approx 0$) the value of the DurbinWatson statistic should be approximately 2. Statistical testing is necessary to determine just how far away from 2 the statistic must fall in order for us to conclude that the assumption of uncorrelated errors is violated. Unfortunately, the distribution of the DurbinWatson test statistic d depends on the X matrix, and this makes critical values for a statistical test difficult to obtain. However, Durbin and Watson showed that d lies between lower and upper bounds, say d_L and d_U. If d is outside these limits, a conclusion regarding the hypotheses can be drawn.

The decision procedure is as follows:

$$\begin{aligned} \text{if } d < d_L, &\ \text{then Reject } H_0 : \phi = 0 \\ \text{if } d > d_U, &\ \text{then do not Reject } H_0 : \phi = 0 \\ \text{if } d_L \leq d \leq d_U, &\ \text{then the test is inconclusive} \end{aligned}$$

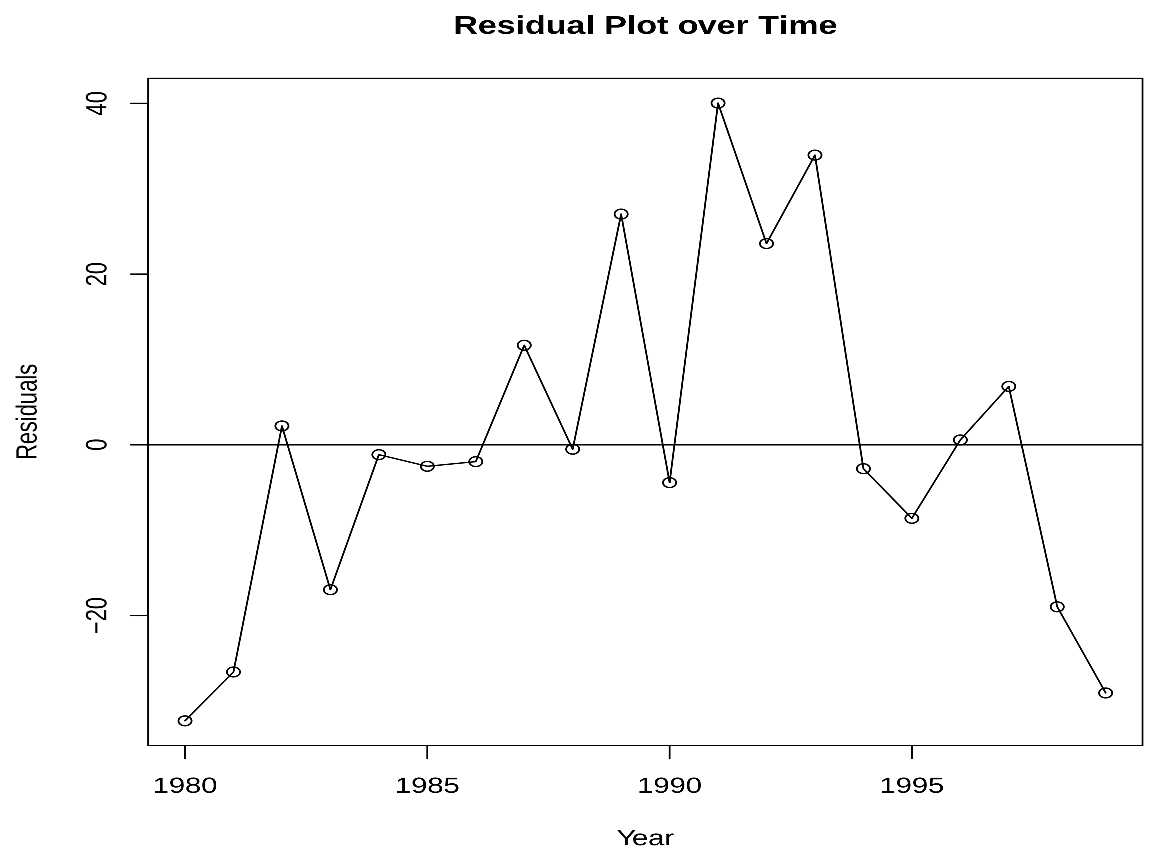
Table (at the back of the book) gives the bounds d_L and d_U for a range of sample sizes, various numbers of predictors, and three α values (0.05, 0.01). However, if a test for negative autocorrelation is desired, one can use the statistic (4 - d). Then the decision rules for testing the hypotheses $H_0 : \phi = 0 \ vs \ H_1 : \phi < 0$ are the same as those used in testing for positive autocorrelation. It is also possible to test a two-sided alternative hypothesis $H_0 : \phi = 0 \ vs \ H_1 : \phi \neq 0$ by using both of the one-sided tests simultaneously. If this is done, the two-sided procedure has type I error 2α.

Example-1 A company wants to use a regression model to relate annual regional advertising expenses to annual regional concentrate sales for a soft drink company. Table below presents 20 years of these data. We will initially assume that a straight-line relationship is appropriate and fit a simple linear regression model by ordinary least squares.

Year	t	y = Sales	x = Expenditure
1980	1	3083	75
1981	2	3149	78
1982	3	3218	80
1983	4	3239	82
1984	5	3295	84
...	...	...	...
...	...	...	...
1994	15	4318	135
1995	16	4493	144
1996	17	4683	153
1997	18	4850	161
1998	19	5005	170
1999	20	5236	182

The SAS output for this model is shown in next slide. Because these are time series data, there is a possibility that autocorrelation may be present. The plot of residuals versus time, (shown below) has a pattern indicative of potential autocorrelation. There is a definite upward trend in the plot, followed by a downward trend.

<u>Residuals Over time</u>



Parameter Estimates

Variable	DF	Estimate	SE	t-stat	p-value
Intercept	1	1608.50778	17.02234	94.49	$<.0001$
x	1	20.09096	0.14278	140.71	$<.0001$

Test for autocorrelation

To test the hypothesis $H_0 : \phi = 0 \; vs \; H_1 : \phi > 0$.

We see the Dubin-Watson statistic d = 1.08, and p-value = 0.0061, Also the table cutoff $d_L = 1.20$ at $\alpha = 0.05$

Hence we conclude that there exists a positive correlation in the data.

A significant value of the DurbinWatson statistic or a suspicious residual plot indicates a potential problem with autocorrelated model errors. This could be the result of an actual time dependence in the errors or an artificial time dependence caused by the omission of one or more important predictor variables. If the apparent autocorrelation results from missing predictors and if these missing predictors can be identified and incorporated into the model, the apparent autocorrelation problem may be eliminated.

Because it is reasonably likely that regional population affects soft drink sales, we have provided data on regional population for each of the study years (not shown). Table below is the SAS output for a regression model that includes both predictor variables, advertising expenditures and population. Both of these predictor variables are highly significant. Output also provides DurbinWatson statistic for this model as d = 3.05932, and the p-value 0.9822. Also the 5% critical values are $d_L = 1.10$ and $d_U = 1.54$, and since d is greater than d_U, we conclude that there is no evidence to reject the null hypothesis in favour of positive autocorrelation.

Parameter Estimates - Model - 2

Variable	DF	Estimate	SE	t-stat	p-value
Intercept	1	320.33956	217.32775	1.47	0.1588
x	1	18.43421	0.29153	63.23	<.0001
Population	1	0.00168	0.00028291	5.93	< .0001

Test for autocorrelation: $H_0 : \phi = 0 \ vs \ H_1 : \phi > 0$.

We see the Dubin-Watson statistic d = 3.059, and p-value = 0.9822, Also the table cutoff $d_L = 1.10$ at $\alpha = 0.05$. Hence we conclude that there is not sufficient evidence to support the fact that there exists a positive correlation in the data.

The Cochrane - Orcutt Method

When the observed autocorrelation in the model errors cannot be removed by adding one or more new predictor variables to the model, it is necessary to take explicit account of the autocorrelation structure in the model and use an appropriate parameter estimation method. A very good and widely used approach is the procedure devised by Cochrane and Orcutt..

We now describe the CochraneOrcutt method for the simple linear regression model with first-order autocorrelated errors. The procedure is based on transforming the response variable.

Note that, a simple linear regression model with first-order autoregressive errors is

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t, \quad \text{where} \quad \epsilon_t = \phi\, \epsilon_{t-1} + a_t$$

.

Let $y_t' = y_t - \phi\, y_{t-1}$. Then substituting for y_t and y_{t-1} we get

$$
\begin{aligned}
y'_t &= y_t - \phi\, y_{t-1} \\
&= \beta_0 + \beta_1 x_t + \epsilon_t - \phi\,(\beta_0 + \beta_1 x_{t-1} + \epsilon_{t-1} \\
&= \beta_0(1-\phi) + \beta_1(x_t - \phi\, x_{t-1}) + (\epsilon_t - \phi\, \epsilon_{t-1}) \\
&= \beta'_0 + \beta_1 x'_t + a_t
\end{aligned}
$$

where $\beta'_0 = \beta_0(1-\phi)$ and $x'_t = (x_t - \phi\, x_{t-1})$. Notice that the error terms at in the transformed model are independent random variables. Unfortunately, this new re-parameterized model contains an unknown parameter ϕ and it is also no longer linear in the unknown parameters because it involves products of parameters. However, the first-order autoregressive process $\epsilon_t = \phi\, \epsilon_{t-1} + a_t$ can be viewed as a simple linear regression through the origin and the parameter ϕ can be estimated by obtaining the residuals of an OLS regression of y_t on x_t and then regressing e_t on e_{t-1}. The OLS regression of e_t on e_{t-1} results in

$$
\hat{\phi} = \frac{\sum_{i=2}^{T} e_t . e_{t-1}}{\sum_{i=2}^{T} e_t^2}
$$

Using $\hat{\phi}$ as an estimate of ϕ, we can calculate the transformed response and predictor variables as

$$y'_t = y_t - \hat{\phi}\, y_{t-1} \quad \text{and } x'_t = x_t - \hat{\phi}\, x_{t-1}$$

Now apply ordinary least squares to the transformed data. This will result in estimates of the transformed slope the intercept and a new set of residuals. The DurbinWatson test can be applied to these new residuals from the transformed model. If this test indicates that the new residuals are uncorrelated, then no additional analysis is required. However, if positive autocorrelation is still indicated, then another iteration is necessary. In the second iteration ϕ is estimated with new residuals that are obtained by using the regression coefficients from the new transformed model with the original regressor and response variables. This iterative procedure may be continued as necessary until the residuals indicate that the error terms in the transformed model are uncorrelated. Usually only one or two iterations are sufficient to produce uncorrelated errors.

Summary of the Process:

Step - 1. Use a MLR (or SLR) model to find the estimated regression plane.

Step - 2. Use the above predicted equation to find the errors and then plot and test for autocorrelation.

Step - 3. If autocorrelation is present use the errors to estimate ϕ.

Step - 4. Find the new (transformed) variables (y'_t & x'_t).

Step - 5. Refit the model with new transformed variables and find the new residuals.

Step - 6. Plot and test the new residuals for positive auto correlation.

Step - 7. If DW test is not rejected then stop, else repeat from step - 4.

Example-2. Table below presents data on the market share of a particular brand of toothpaste for 20 time periods and the corresponding selling price per pound.

t	Share	Price
1	3.63	0.97
2	4.2	0.95
3	3.33	0.99
4	4.54	0.91
5	2.89	0.98
...	...	....
...	...	....
17	7.18	0.83
18	7.9	0.79
19	8.45	0.76
20	8.23	0.78

A simple linear regression model is fit to these data. The DurbinWatson statistic for the residuals from this model is d = 1.13582, (p-value=0.0098) and the 5% critical values are $d_L = 1.20$ and $d_U = 1.41$, so there is evidence to support the conclusion that the residuals are positively autocorrelated.

Parameter Estimates

Variable	DF	Estimate	SE	t-stat	p-value
Intercept	1	26.90989	1.10988	24.25	$< .0001$
Price	1	-24.28977	1.29781	-18.72	$< .0001$

We use the CochraneOrcutt method to estimate the model parameters. The autocorrelation coefficient can be estimated using the residuals. Autocorrelation was estimated as:

$$\hat{\phi} = \frac{\sum_{i=2}^{T} e_t.e_{t-1}}{\sum_{i=2}^{T} e_t^2} = \frac{1.3547}{3.3083} = 0.409$$

Using $\hat{\phi} = 0.409$, we can calculate the transformed response and predictor variables as

$$y_t' = y_t - 0.409 \times y_{t-1} \quad \text{and } x_t' = x_t - 0.409 \times x_{t-1} \quad \text{for } t = 2, 3, \ldots 20.$$

These transformed variables are now used for fitting a regression model and following results are obtained.

Parameter Estimates

Variable	DF	Estimate	SE	t-stat	p-value
Intercept	1	15.80427	0.95022	16.63	$< .0001$
xprime	1	-24.13099	1.90934	-12.64	$< .0001$

The residuals from the transformed model are also calculated and Durbin-Watson statistic for the transformed model is d = 1.93 (p-value=0.3896) and the 5% critical values from Table A.6 are $d_L = 1.18$ and $d_U = 1.40$, so we conclude that there is no problem with autocorrelated errors in the transformed model. The CochraneOrcutt method has been effective in removing the autocorrelation.

Remark: The true slope in the transformed model β_1 is equal to the true slope in the original model β_1 . A comparison of the slopes in the two models reveals that they are close but not same as there is one less observation in the second model. However, if the standard errors are compared, the Cochrane-Orcutt method produces an estimate of the slope that has a larger standard error (= 1.9093) compare to the standard error of the ordinary least squares estimate (=1.2978). This reflects the fact that if the errors are autocorrelated and OLS is used, the standard errors of the model coefficients are likely to be underestimated.