Applied Regression

Time Series - Part-2

Module 10   Lecture - 10-2

ANOTHER WAY TO ESTIMATE

Another way to estimate the parameters in regression with autocorrelated errors

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

is the maximum likelihood technique. Without going into mathematical details, we can use the software to calculate all the estimates.

Revisit - Example-2 We will fit the regression model with time series errors to the toothpaste market share data originally analyzed in Example-2. The SAS procedure for fitting regression models with time series errors is SAS PROC AUTOREG. It produces two outputs (1) MLR model results with few more things (2) First Order (requested) model results.

(1) Ordinary Least Square

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	$< .0001$

Durbin-Watson Statistics

Order	DW	Pr $<$ DW	Pr $>$ DW
1	1.1358	0.0098	0.9902
2	1.6450	0.2182	0.7818
3	1.8265	0.4704	0.5296
4	1.6144	0.3831	0.6169

Estimates of Autoregressive Parameters

Lag	Coefficient	SE	t-Value
1	-0.409437	0.221275	-1.85

(2) Maximum Likelihood Method of the Regression with autocorrelated model of lag-1

Durbin-Watson Statistics

Order	DW	Pr < DW	Pr > DW
1	1.8924	0.3472	0.6528
2	1.9913	0.5265	0.4735
3	2.0576	0.6995	0.3005
4	1.3841	0.2058	0.7942

Parameter Estimates

Variable	DF	Estimate	SE	t-stat	p-value
Intercept	1	26.3322	1.4777	17.82	$< .0001$
Price	1	-23.5903	1.7222	-13.70	$< .0001$
AR1	1	-0.4323	0.2203	-1.96	0.0663

(1) Notice that the autocorrelation parameter (or the lag one autocorrelation) is estimated to be 0.4094, which is very similar to the value obtained before.

(2) Durbin-Watson statistic does not exhibit any positive autocorrelation structure.

Point and Intervals Prediction - One Step

We now consider how to obtain predictions of new observations. These are actually forecasts of future values at some lead time. These are popularly known as forecasts as oppose to prediction because we are going out the range of the time. As we learn in MLR model it is not always a good idea to predict outside the range of x values.

First, it is very tempting to ignore the autocorrelation in the data when making predictions of future values (forecasting), and simply substitute the conditional maximum likelihood estimates into the regression equation:

$$\hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 \, x_t$$

Now. suppose that we are at the end of the current time period, T, and we wish to obtain a prediction or forecast for period (T + 1). Using the above equation, this results in

$$\hat{y}_{T+1} = \hat{\beta}_0 + \hat{\beta}_1 \, x_{T+1}$$

assuming that the value of the predictor variable in the next time period

x_{T+1} is known.

Unfortunately, this naive approach isn't correct. From Cochran-Orcutt model, we know that the observation at time period t is

$$y_t = \phi\, y_{t-1} + \beta_0(1-\phi) + \beta_1(x_t - \phi\, x_{t-1}) + (\epsilon_t - \phi\, \epsilon_{t-1})$$

So at the end of the current time period T the next observation is

$$y_{T+1} = \phi\, y_T + \beta_0(1-\phi) + \beta_1(x_{T+1} - \phi\, x_T) + a_{T+1}$$

Assume that the future value of the regressor variable x_{T+1} is known. Obviously, at the end of the current time period, both y_T and x_T are known. The random error at time (T + 1) a_{T+1} hasn't been observed yet, and because we have assumed that the expected value of the errors is zero, the best estimate we can make of a_{T+1} is $a_{T+1} = 0$ (as in MLR model). This suggests that a reasonable forecast of the observation in time period (T + 1) that we can make the end of the current time period T is

$$\hat{y}_{T+1} = \hat{\phi}\, y_T + \hat{\beta}_0(1-\hat{\phi}) + \hat{\beta}_1(x_{T+1} - \hat{\phi}\, x_T)$$

Notice that this forecast is likely to be very different than the naive forecast obtained by ignoring the autocorrelation.

To find a prediction interval on the forecast, we need to find the variance of the prediction error. The one-step-ahead forecast error is

$$y_{T+1} - \hat{y}_{T+1} = a_{T+1}$$

assuming that all of the parameters in the forecasting model are known. The variance of the one-step ahead forecast error is

$$Var(a_{T+1}) = \sigma_a^2$$

Using the variance of the one-step-ahead forecast error, we can construct a $100(1-\alpha)\%$ prediction interval for the lead-one forecast and hence the

$$100(1-\alpha)\% \text{ PI is } \quad \hat{y}_{T+1} \ \pm \ z_{\alpha/2}\,\sigma_a$$

To actually compute an interval, σ_a we must replace it by its estimate, resulting in

$$100(1-\alpha)\% \text{ PI is } \quad \hat{y}_{T+1} \ \pm \ z_{\alpha/2}\,\hat{\sigma}_a$$

as the PI. Because σ_a and the model parameters in the forecasting equation been replaced by estimates, the probability level of the above PI is only approximate.

<u>Point and Intervals Prediction - Two Step</u>

$$\begin{aligned} y_{T+2} &= \phi\, y_{T+1} + \beta_0(1-\phi) + \beta_1(x_{T+2} - \phi\, x_{T+1}) + a_{T+2} \\ &= \phi\big[\phi\, y_T + \beta_0(1-\phi) + \beta_1(x_{T+1} - \phi\, x_T) + a_{T+1}\big] \\ &\qquad + \beta_0(1-\phi) + \beta_1(x_{T+2} - \phi\, x_{T+1}) + a_{T+2} \end{aligned}$$

Assume that the future value of the regressor variables x_{T+1} and x_{T+2} are known. At the end of the current time period, both y_T and x_T are known. The random errors at time (T + 1) and (T + 2) haven't been observed yet, and because we have assumed that the expected value of the errors is zero, the best estimate we can make of both a_{T+1} and a_{T+2} are zero. This suggests that the forecast of the observation in time period (T + 2) made at the end of the current time period T is

$$\hat{y}_{T+2} = \hat{\phi}\, \hat{y}_{T+1} + \hat{\beta}_0(1 - \hat{\phi}) + \hat{\beta}_1(x_{T+2} - \hat{\phi}\, x_{T+1})$$

Now to compute the prediction interval, we need to account for the variance. The two step forecast error is

$$y_{T+2} - \hat{y}_{T+2} = a_{T+2} + \phi\, a_{T+1}$$

assuming that all estimated parameters are actually known. The variance of the two-step ahead forecast error is

$$Var(a_{T+2} + \phi\, a_{T+1}) = \sigma_a^2 + \phi^2\, \sigma_a^2 = (1 + \phi^2)\, \sigma_a^2$$

Using the variance of the two-step-ahead forecast error, we can construct a $100(1-\alpha)\%$ PI for the lead-two forecast as

$$100(1-\alpha)\% \text{ PI is } \quad \hat{y}_{T+2} \ \pm \ z_{\alpha/2}\, (1 + \phi^2)\, \sigma_a$$

To actually compute the PI, both σ_a and ϕ must be replaced by estimates, resulting in

$$100(1-\alpha)\% \text{ PI is } \quad \hat{y}_{T+2} \ \pm \ z_{\alpha/2}\,(1+\hat{\phi}^2)\,\hat{\sigma}_a$$

as the PI. Because σ_a and ϕ have been replaced by estimates, the probability level on the above PI is only approximate.

In general, if we want to forecast τ periods ahead, the forecasting equation is

$$\hat{y}_{T+\tau} = \hat{\phi}\,\hat{y}_{T+\tau-1} + \hat{\beta}_0(1-\hat{\phi}) + \hat{\beta}_1(x_{T+\tau} - \hat{\phi}\,x_{T+\tau-1})$$

The τ-step-ahead forecast error is (assuming that the estimated model parameters are known)

$$y_{T+\tau} - \hat{y}_{T+\tau} = a_{T+\tau} + \phi\,a_{T+\tau-1} + \ldots + \phi^{\tau-1}\,a_{T+1}$$

and the variance of the τ-step-ahead forecast error is

$$
\begin{aligned}
Var(a_{T+\tau} + \phi\, a_{T+\tau-1} + \ldots + \phi^{\tau-1}\, a_{T+1}) &= (1 + \phi^2 + \ldots + \phi^{2(\tau-1)})\, \sigma_a^2 \\
&= \frac{1-\phi^{2\tau}}{1-\phi^2}
\end{aligned}
$$

A $100(1-\alpha)\%$ PI for the lead-τ forecast is

$$\hat{y}_{T+\tau} \;\pm\; z_{\alpha/2} \left(\frac{1-\phi^{2\tau}}{1-\phi^2}\right) \sigma_a$$

Replacing σ_a and ϕ by estimates, the approximate $100(1-\alpha)\%$ PI is actually computed as

$$\hat{y}_{T+\tau} \;\pm\; z_{\alpha/2} \left(\frac{1-\hat{\phi}^{2\tau}}{1-\hat{\phi}^2}\right) \hat{\sigma}_a$$

Example-3 The data in the table below give the percentage share of market of a particular brand of canned peaches y_t for the past 15 months and the relative selling price x_t.

a. Fit a simple linear regression model to these data. Plot the residuals versus time. Is there any indication of autocorrelation?

Parameter Estimates

Variable	DF	Estimate	SE	t-stat	p-value
Intercept	1	24.59405	1.20560	20.40	$< .0001$
x	1	-0.08918	0.01368	-6.52	$< .0001$

b. Use the DurbinWatson test to determine if there is positive autocorrelation in the errors. What are your conclusions?

D-W Statistics

Durbin-Watson d	0.818
Pr $<$ DW	0.0015
Pr $>$ DW	0.9985
Number of Observations	15
1st Order Autocorrelation	0.541

c. Use one iteration of the CochraneOrcutt procedure to estimate the regression coefficients. Find the standard errors of these regression coefficients.

As the estimated auto correlation is 0.541,

we create $y_t' = y_t - (0.541) \times y_{t-1} \quad \& \quad x_t' = x_t - (0.541) \times x_{t-1}$

This transformed regression run results as

Parameter Estimates

Variable	DF	Estimate	SE	t-stat	p-value
Intercept	1	12.08788	0.55435	21.81	$< .0001$
xprime	1	-0.11050	0.01403	-7.87	$< .0001$

d. Is there positive autocorrelation remaining after the first iteration? Would you conclude that the iterative parameter estimation technique has been successful?

D-W Statistics

Durbin-Watson d	0.903
Pr < DW	0.0095
Pr > DW	0.9905
Number of Observations	14
1st Order Autocorrelation	0.298

As d = 0.903 and p-value is 0.0095, $H_0 : \phi = 0$ is rejected and hence we conclude that auto correlation is still present.

Revisit Example-3 - Using the previous results, estimate the Market Share for future time period 21 and also find a 95% confidence interval for the same time period. Assume that Price at time 21 is 0.77 and $\hat{\sigma_a} = 0.39843$.

Solution: Note that the estimate for a future market share can be found as

$$\hat{y}_{T+1} = \hat{\phi}\, y_T + \hat{\beta}_0(1 - \hat{\phi}) + \hat{\beta}_1(x_{T+1} - \hat{\phi}\, x_T)$$

From the output before we see that

$\hat{\phi} = 0.4094$(note that we changed the sign) $y_T = 8.23$, $\hat{\beta}_0 = 26.9099$

$\hat{\beta}_1 = -24.2898$, $x_{T+1} = 0.77$, $x_T = 0.78$

Hence

$$\begin{aligned}\hat{y}_{T+1} &= 0.4094 \times 8.23 \,+\, 26.9099 \times (1 - 0.4094) \\ &\quad + \,(-24.2898)(0.77 \,-\, 0.4094 \times 0.78) = 8.31\end{aligned}$$

The 95% confidence interval is

$$\hat{y}_{T+1} \;\pm\; z_{\alpha/2} \,.\, \hat{\sigma}_a = 8.31 \pm 1.96 \times 0.39843 = (7.5291, 9.0909)$$