Autocorrelation in Time Series Data

- The basic regression models considered so far have assumed that the random error terms ϵ_i are either uncorrelated random variables or independent normal random variables. What do we do if the error terms are correlated over time, such as time series data? Error terms correlated over time are said to be *autocorrelated* or *serially correlated.*

  When the error terms in the regression model are positively autocorrelated, the use of ordinary least squares procedures has a number of important consequences:

  1. The estimated regression coefficients are still unbiased, but they no longer have the minimum variance property and may be quite inefficient.

  2. MSE may seriously underestimate the variance of the error terms.

  3. $s\{b_k\}$ calculated according to ordinary least squares procedures may seriously underestimate the true standard deviation of the estimated regression coefficient.

- **First-Order Autoregressive Simple Linear Regression Model (AR(1)).**

$$Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$$

  where:

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

  2. ρ is the *autocorrelation* parameter such that $|\rho| < 1$.

  3. u_t (called *disturbances*) are independent $N(0, \sigma^2)$

- **First-Order Autoregressive Multiple Regression Model.**

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 X_{t2} + \cdots + \beta_{p-1} X_{t,p-1} + \epsilon_t$$

  where:

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

  2. ρ is a parameter such that $|\rho| < 1$.

  3. u_t are independent $N(0, \sigma^2)$

- **Properties of Error Terms.**

  1. $E(\epsilon_t) = 0$

  2. $V(\epsilon_t) = \dfrac{\sigma^2}{1-\rho^2}$

  3. $Cov(\epsilon_t, \epsilon_{t-1}) = \rho\left(\dfrac{\sigma^2}{1-\rho^2}\right)$

- **Durbin-Watson Test for Autocorrelation.**

$$H_0 : \rho = 0 \ versus \ H_1 : \rho > 0$$

  Test Statistic: $D = \dfrac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$, where $e_t = Y_t - \hat{Y}_t$ (residual).

- **Blaisdell Company Example (on page 488).** The data are stored in *BlaisdellCompany.csv.*

- **Remedial Measures for Autocorrelation.**

  1. *Addition of Predictor Variables.* One major cause of autocorrelated error terms is the omission from the model one or more key predictor variables that have time-ordered effects on the response variable.

2. *Use of Transformed Variables.* Only when use of additional predictor is not helpful in eliminating the problem of autocorrelated errors should a remedial action based on transformed variables be employed.

$$Y_t' = Y_t - \rho Y_{t-1}$$

a. Cochrane-Orcutt Procedure. (*see R commands for example.*)

i. Estimate ρ. Note: $\epsilon_t = \rho\epsilon_{t-1} + u_t$ (regression through the origin) and so $r = \frac{\sum_{t=2}^{n} e_{t-1}e_t}{\sum_{t=2}^{n} e_{t-1}^2}$.

ii. Fit the transformed model using the transformed variables Y' and X'.

iii. Test for need to iterate.

b. Hildreth-Lu Procedure. Chose ρ that minimizes the $SSE = \sum(Y_t' - \hat{Y}_t')^2$ for the transformed regression model. (*see R commands for example.*)

c. First Difference Procedure. Use $\rho = 1$, then fit a regression through the origin for the transformed variables (X', Y') to obtain b_1'. To obtain the estimates of the regression parameters for the original variables, use: $b_0 = \bar{Y} - b_1'\bar{X}$ and $b_1 = b_1'$.
(*see R commands for example.*)

- **Forecasting with Autocorrelated Error Terms.**
  Since $Y_t = \beta_0 + \beta_1 X_t + \epsilon_t$, where $\epsilon_t = \rho\epsilon_{t-1} + u_t$, then

$$Y_{n+1} = \beta_0 + \beta_1 X_{n+1} + \rho\epsilon_n + u_{n+1}$$

  1. Estiamte: $F_{n+1} = \hat{Y}_{n+1} + r * e_n$, where: $\hat{Y}_{n+1} = b_0 + b_1 X_{n+1}$ and $e_n = Y_n - \hat{Y}_n$.

  2. Prediction Interval: $F_{n+1} \pm t_{(1-\alpha/2;n-3)} s\{pred\}$

- For the Blaisdell Company example, the trade association has projected that deseasonalized industry sales in the first quarter of 2003 (i.e. quarter 21) will be $X_{21} = \$175.3$ million. Construct the 95% prediction interval using the

  1. Cochrane-Orcutt procedure.

  2. Hildreth-Lu procedure.

  3. First differences procedure.