Time Series Analysis

Linear Trend: Time series pattern graphed in downward or upward straight line.

$y_t = b_0 + b_1 t + \in$ Where b_0 = Intercept, b_1 = Slope coefficient and $\in$ is error term.

Time series data with exponential growth, time series can be expressed as

$y_t = e^{b_0 + b_1 t}$ or $\ln(y_t) = b_0 + b_1 t$

When Positive exponential growth period random variables grow at constant rate and have convex curve shape on time series plot.

If the data distributes equally above and below the regression line, then ***linear regression*** is appropriate, when data plots in curvilinear path or residuals persistently positive or negative over the time period ***log linear model*** is appropriate.

For time series models without serial correlation DW (Durbin-Watson) value equals to 2. A DW significantly different from 2 represents, existence of serial correlation in time series models.

Limitations of Linear models:
Serial correlation: where residuals are persistently positive or negative.
When serial correlation persistent auto regression models (AR) are recommended.

Auto regression (AR model): When the dependent variable is regressed against one or more lagged values of itself then the resultant model is known as ***Auto Regressive*** (AR) model.

$x_t = b_o + b_1 x_{t-1} + \in_t$

Where x_{t-1} = Value of time series at t-1 period.

In AR models past values of the variable are used to predict current value of the variable.

AR models are valid when and only time series is covariance stationary.

Covariance Stationary:

- *Constant and finite expected value*: expected value is constant over time period.
- *Constant and finite variance*: Volatility around the mean.
- *Constant and finite variance* between values at any *given lag*.

Autocorrelation and model fit:
When AR model is specified properly residuals will not exhibit serial correlation. To verify autocorrelation among residuals

First define the AR model, and then calculate autocorrelations among residuals.
Test the autocorrelations are significantly different from zero.
Standard error is $\mathbf{1}/\sqrt{T}$
Where T = number of observations.
Test statistic = Correlation term / standard error with n-2 degrees of freedom

$$= \rho \in_{,}, \varepsilon_{t-1} \;/\; (\mathbf{1}/\sqrt{T}\,)$$

Mean Reversion:
Time series tendency towards mean, when time series value is over the mean at one period will have decline in next value towards mean.

In AR (1) model mean- reverting level is = $\frac{b_0}{(1-b_1)}$

Root Mean Square Error: Square root of average squared errors is used to compare accuracies of forecasted out of sample data.

If sample data is 24 months of samples, first 12 months of data (in sample) is used to predict next 12 of the data (out Sample). Researchers may come up with AR (1) model or AR (2) model from in-sample to predict out-Sample. The model with lowest RMSE has lowest forecast errors.

Non-stationary coefficients: When financial and economic data are dynamic then the coefficient of time series may change from one period to another. Changes in economic environment also make time series unreliable for longer time spans.

Random walk: When time series exhibits predicted value is value from one period plus a random error then it is known and random walk.

$x_t = x_{t-1} + \varepsilon$ Here b_o = 0 and b_1 = 1

Where expected error term is zero, variance of error term is constant, and no serial correlation between errors.

Random walk with drift: Random walk with constant drift that is intercept term b_o not equals to 0. $x_t = b_0 + b_1 x_{t-1} + \varepsilon$ Where b_1 = 1

Covariance Stationary: Both Random walk and Random walk with drift are not covariant stationary. In both cases $b_1 = 1$ and $b_0 = 0$ for random walk and not equal to 0 for random walk with drift.

To become covariant stationary mean reversion must be finite, in the random walk mean reversion infinite since in $\frac{b_0}{(1-b_1)}$; $b_1 = 1$.

Unit Root and Non-Stationary: If the lag coefficient is 1, the series has unit root and not covariant stationary.
To determine covariant stationary or not two methods are used. 1. Run AR model to determine auto correlation, 2. Perform Dickey Fuller Test
Dickey Fuller test is performed by first differentiating the lags.

First differentiating is subtracting value of the time series from immediate preceding value to define new independent value y.

$y_t = x_t - x_{t-1} = \varepsilon$ Then regress AR (1) model for y.

Seasonality and correction: A pattern that repeat itself from one period to other. Monthly sales data or quarterly peak sales are examples of seasonality.

When seasonality persists in data the time series data would be misspecified until AR model take seasonality into consideration. To adjust seasonality factor ***additional lag*** of dependent variable must be added for the seasoned period.

ARCH Model (Autoregressive conditional heteroskadasticity) : If variance of residual in one period dependent on variance of residuals from preceding period then ARCH exists. In this condition, the standard errors for regression coefficients will be invalid and hypothesis tests will be give wrong conclusions.

To test ARCH (1), estimate time series with squared error term on first lag of squared residuals. To test ARCH generalized ***lease square*** method is used.

Cointegration:

When working with two time series in regression,

1. If both time series do not have unit root then regression is valid.
2. If one has unit root and other does not then the regression must be rejected.
3. If both time series have unit root and cointegrated, then the regression is valid.
4. If both time series have unit root and not cointegrated, then the regression must be rejected.