A random walk process

A simple random walk model

A random walk is defined as a process where the current value of a variable is composed of the past value plus an error term defined as a white noise (a normal variable with zero mean and variance one).

Algebraically a random walk is represented as follows:

$$y_t = y_{t-1} + \epsilon_t$$

The implication of a process of this type is that the best prediction of y for next period is the current value, or in other words the process does not allow to predict the change $(y_t - y_{t-1})$. That is, the change of y is absolutely random.

It can be shown that the mean of a random walk process is constant but its variance is not. Therefore a random walk process is nonstationary, and its variance increases with t.

In practice, the presence of a random walk process makes the forecast process very simple since all the future values of y_{t+s} for $s > 0$, is simply y_t.

A random walk model with drift

A drift acts like a trend, and the process has the following form:

$$y_t = y_{t-1} + a + \epsilon_t$$

For $a > 0$ the process will show an upward trend. The third graph presented in the previous page was generated assuming $a = 1$. This process shows both a deterministic trend and a stochastic trend. Using the general solution for the previous process,

$$y_t = y_0 + at + \sum_{t=1}^{n} \epsilon_t$$

where $y_0 + at$ is the deterministic trend and $\sum_{t=1}^{n} \epsilon_t$ is the stochastic trend.

The relevance of the random walk model is that many economic time series follow a pattern that resembles a trend model. Furthermore, if two time series are independent random walk processes then the relationship between the two does not have an economic meaning. If one still estimates a regression model between the two the following results are expected:

(a) High R^2
(b) Low Durbin-Watson statistic (or high ρ)
(c) High t ratio for the slope coefficient

This indicates that the results from the regression are spurious. A regression in terms of the changes can provide evidence against previous spurious results. If the coefficient of a regression

$$\Delta y_t = a_0 + \beta_1 \Delta x_t + \nu_t$$

is not significant, then this is an indication that the relationship between y and x is spurious, and one should proceed by selecting other explanatory variable.

If the Durbin-Watson test is passed then the two series are *conintegrated*, and a regression between them is appropriate.