Time Series Analysis

Non-stationarity

Binzhi Chen

Department of Economics
University of Birmingham

March 4, 2020

Introduction

In previous lectures, we go through time series analysis based on one important assumption, which is stationarity.

However, sometimes that is not the case in the reality world. Most of financial data are non-stationarity.

Now we will introduce two types of non-stationarity

Trend stationary

Consider a time series includes a deterministic trend:

$$y_t = \alpha + \delta t + \psi(L)\epsilon_t \tag{1}$$

Where $\sum_{j=1}^{\infty} |\psi_j| < \infty$, roots of $\psi(z) = 0$ are outside the unit circle, and ϵ_t is a white noise.

Why this process is non-stationarity?

This process is called *trend stationary* since it is stationary if it subtracts the linear function of time trend $\alpha + \delta t$ in this process.

The procedure that removing the $\mathcal{G}(t, \delta)$ to make the process stationary is called *detrending*

Unit Root Process

The second specification is a *unit root* process, which is also called *difference-stationarity*

A stochastic process is said to contain a unit root if,

$$(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) y_t = c + (1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q)\epsilon_t$$

can be factored as,

$$(1 - L)(1 - \widetilde{\phi}_1 L - \widetilde{\phi}_2 L^2 - \cdots - \widetilde{\phi}_{p-1} L^{p-1}) y_t = c + (1 + \theta_1 L + \theta_2 L^2 + \cdots +)\epsilon_t$$

Unit Root Process

Or a more simple form,

$$(1-L)y_t = \mu + \psi(L)\epsilon_t$$

Where $\mu = c/(1 - \widetilde{\phi}_1 L - \widetilde{\phi}_2 L^2 - \cdots - \widetilde{\phi}_{p-1} L^{p-1})$

and $\psi(L) = \frac{(1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q)}{(1-\widetilde{\phi}_1 L-\widetilde{\phi}_2 L^2-\cdots-\widetilde{\phi}_{p-1} L^{p-1})}$

The most common unit root process is *random walk*:

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

Or random walk with a drift

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

Why this process is non-stationarity?

Unit Root Process

We may have two unit roots in a process:

$$y_t = 2y_{t-1} - y_{t-2} + \epsilon_t$$

Or,

$$(1-L)(1-L)y_t = \epsilon_t$$

If a non-stationary series, y_t must be difference d times before it becomes

stationary, then it is said to be integrated of order d.

We write $y_t \sim$I(d). If $y_t \sim$I(d), then $\Delta^d y_t \sim$I(0)

One unit roots in a process is $y_t \sim$I(1).
Two unit roots in a process is $y_t \sim$I(2).

Unit Root Process

I(1) and I(2) series can wander a long way from their mean value and cross this mean value rarely.

I(0) series should cross the mean frequently.

The majority of economic and financial series contain a single unit root, although some are stationary and consumer prices have been argued to have 2 unit roots.

Why Linear Time Trends

In reality, most of financial data follows an exponential trend than a linear trend,

$$y_t = e^{\delta t}$$

If we take logs we will get the linear trends,

$$log(y_t) = \delta t$$

Thus, it is common to take logs before attempting to describe a time series includes a deterministic trend.

Why Differencing Unit Root Process

It is also common to take logs before attempting to describe unit root process.

Another step we need to take is differencing.

Definition:

Let $\Delta y_t = y_t - y_{t-1}$:

$$\begin{aligned} y_t &= y_{t-1} + \mu + \epsilon_t \\ \Delta y_t &= \mu + \epsilon_t \end{aligned}$$

Why Differencing Unit Root Process

For small changes, the first difference of the log of a variable is approximately the same as the percentage change in the variable:

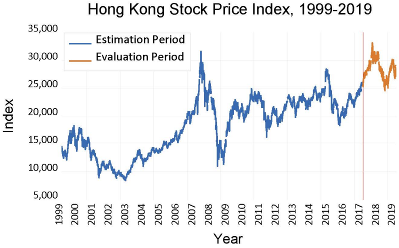
$$
\begin{aligned}
(1-L)log(y_t) &= log(y_t/y_{t-1}) \\
&= log[1+(y_t-y_{t-1})/y_{t-1}] \\
&= (y_t-y_{t-1})/y_{t-1}
\end{aligned}
$$

Where $log(1+x)=x$.

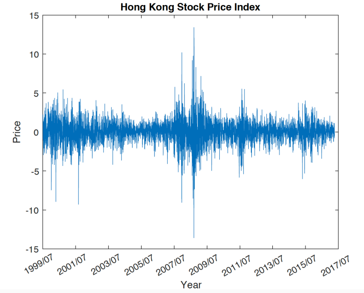
Thus, if the logs of a variable are specified to follow a unit root process, the assumption is that the rate of growth of the series is a stationary stochastic process.

Original Series and Dlog Seires

Graphs of *Exponential Series* and *Differencing-Log Series*.



(a) Exponential Series



(b) Differencing-Log Series

Use the Right Approach to Each Process

Although trend-stationary and difference-stationary series are both "trending" over time, the correct approach to induce stationarity needs to be used in each case.

Differencing a trend-stationary series would "remove" the non stationarity, but will introduce a non-invertible MA(1) structure into the errors. Consider a trend stationary series,

$$y_t = \alpha + \delta t + \epsilon_t$$

Subtract $y_{t-1} = \alpha + \delta(t-1) + \epsilon_{t-1}$ we will get,

$$\Delta y_t = \delta + \epsilon_t + \epsilon_{t-1}$$

Which is indeed a stationary process but the process will have some problems as it includes a non-invertible MA(1) structure into the errors.

Use the Right Approach to Each Process

Conversely, detrending a series which has stochastic trend, will not remove the stochastic non-stationarity. Consider a difference stationary series,

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

If we subtract $\alpha + \delta t$ on both side, it would no help for the series as it still has a unit root in the AR part.

Why do we need to test for Non-Stationarity?

The stationarity or otherwise of a series can strongly influence its behaviour and properties - e.g. persistence of shocks will be in
finite for nonstationary series

Spurious regressions. If two variables are trending over time, a regression of one on the other could have a high R^2 even if the two are totally unrelated (More examples in workshop).

If the variables in the regression model are not stationary, then it can be proved that the standard assumptions for asymptotic analysis will not be valid. In other words, the usual "t-ratios" will not follow a t-distribution, so we cannot validly undertake hypothesis tests about the regression parameters.

DF Test

The early and pioneering work on testing for a unit root in time series was done by Dickey and Fuller (Dickey and Fuller 1979, Fuller 1976).
The basic objective of the test is to test the null hypothesis that $\phi = 1$ in:

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

$$H_0 : \phi = 1$$
$$H_1 : \phi < 1$$

We also use the regression:

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

$$H_0 : \psi = 0$$
$$H_1 : \psi < 0$$

A test of $\phi = 1$ is equivalent to a test of $\psi = 0$ since $\psi = \phi - 1$

DF Test

We can write:

$$H_0 : y_t = \phi y_{t-1} + \epsilon_t$$

and the alternatives may be expressed as:

$$H_1 : \Delta y_t = \psi y_{t-1} + \mu + \lambda t + \epsilon_t$$

1. $\mu = \lambda = 0$, This is a test for a random walk against a stationary AR(1).

2. $\lambda = 0$, This is a test for a random walk against a stationary AR(1) with drift.

3. $\mu \neq 0$ and $\lambda \neq 0$, This is a test for a random walk against a stationary AR(1) with drift and time trend.

DF Test

In each case, the tests are based on the t-ratio on the y_{t-1} term in the estimated regression of Δy_t on y_{t-1}, plus a constant in case 2 or a constant and trend in case 3.

The test statistics are defined as: test statistic $= \frac{\hat{\psi}}{SE(\hat{\psi})}$

The test statistic does not follow the usual t-distribution under the null, since the null is one of non-stationarity, but rather follows a non-standard distribution. Critical values are derived from Monte Carlo experiments in, for example, Fuller (1976). Relevant examples of the distribution are shown in table below.

DF Test

Significance level	CV for constant but no trend	CV for constant and trend
10%	-2.57	-3.12
5%	-2.86	-3.41
1%	-3.43	-3.96

The null hypothesis of a unit root is rejected in favour of the stationary alternative in each case if the test statistic is more negative than the critical value.

ADF Test

The tests above are only valid if ϵ_t is white noise. In particular, ϵ_t will be autocorrelated if there was autocorrelation in the dependent variable of the regression (Δy_t) which we have not modelled. The solution is to "augment" the test using p lags of the dependent variable.

The alternative model is now written:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + \epsilon_t$$

The same critical values from the DF tables are used as before. A problem now arises in determining the optimal number of lags of the dependent variable. There are two ways to deal with it:

1. IC
2. Schwert $P_{max} = [12 \cdot (T/100)^{1/4}]$. More details in workshop.

PP Test

Phillips and Perron have developed a more comprehensive theory of unit root nonstationarity. The tests are similar to ADF tests, but they incorporate an automatic correction to the DF procedure to allow for autocorrelated residuals.

The tests usually give the same conclusions as the ADF tests, and the calculation of the test statistics is complex.

Criticism of DF and PP Test

Main criticism is that the power of the tests is low if the process is stationary but with a root close to the non-stationary boundary.
The tests are poor at deciding if $\phi = 1$ or $\phi = 0.95$, especially with small sample sizes.

One way to get around this is to use a stationarity test as well as the unit root tests we have looked at.

Consider a stationarity test:

$$H_0 : y_t \text{ is stationary}$$
$$H_1 : y_t \text{ is non-stationary}$$

One such stationarity test is the KPSS test (Kwaitowski, Phillips, Schmidt and Shin, 1992).

KPSS Test

The KPSS test assumes stationarity under the null. It can be used to test whether the series has a deterministic trend vs a stochastic trend:

$$y_t = \delta t + r_t + \epsilon_t$$

$$r_t = r_{t-1} + u_t$$

Where $\epsilon_t \sim (0, \sigma_\epsilon^2)$ is uncorrelated with $u_t \sim (0, \sigma_u^2)$

In this case the null hypotheses and the alternative can be expressed as:

$H_0 : \sigma_u^2 = 0$, which is trend-stationary
$H_0 : \sigma_u^2 = 0, \delta = 0$, which is level-stationary
$H_1 : \sigma_u^2 \neq 0$, which is difference-stationary

KPSS Test

Under the assumption of normality of ϵ_t and u_t , a one sided LM test of the null of stationarity can be constructed with:

$$KPSS = T^{-2} \sum_{t=1}^{T} \frac{S_t^2}{S(l)^2}$$

Where $S(l)^2$ is the long run variance of ϵ_t estimated as:

$$S(l)^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{\epsilon_t}^2 + \frac{2}{T} \sum W(S,l) \sum_{t=s+1}^{T} \hat{\epsilon}_t \hat{\epsilon_{t-s}}$$

where $W(S,l)$ is a kernel function (Bartlett, Spectral Quadratic). We also need to specify the number of lags l which should grow with T.

Under H_0, $\hat{\epsilon}_t$ can be estimated by OLS.

KPSS Test Steps

1. Regress y_t on a constant and a trend. get the OLS residuals $\hat{\epsilon}_t$
2. Calculate the Partial Sum of the residuals $S_t \sum_{i=1}^{T} \hat{\epsilon}_i$
3. Compute the KPSS statistic as:

$$KPSS = T^{-2} \sum_{t=1}^{T} \frac{S_t^2}{S(l)^2}$$

4. Reject H_0 when the KPSS statistic is larger than the relevant critical value.

5. Note that the distribution of the KPSS statistic converges to a different distribution depending on whether the model is trend, level or zero-mean stationary. The distribution is non-standard and can be derived by using Brownian Motions appealing to the FCLT and the CMT.

Unit Root Tests with Structural Breaks

The standard Dickey-Fuller-type unit root tests as well as the KPSS test presented above do not perform well if there are structural breaks in the series

The tests have low power in such circumstances and they fail to reject the unit root null hypothesis when it is incorrect as the slope parameter in the regression of y_t on y_{t-1} is biased towards unity

The larger the break and the smaller the sample, the lower the power of the test

Unit root tests are also oversized in the presence of structural breaks

Unit Root Tests with Structural Breaks

Perron (1989) demonstrates that after allowing for structural breaks in the tests, a whole raft of macroeconomic series may be stationary

He argues that most economic time series are best characterised by broken trend stationary processes, i.e. a deterministic trend but with a structural break.

Unit Root Tests with Structural Breaks

Perron (1989) proposes three test equations differing dependent on the type of break that is thought to be present:

1. A "crash" model that allows a break in the level (i.e. the intercept).
2. A "changing growth" model that allows for a break in the growth rate (i.e. the slope).
3. A model that allows for both types of break to occur at the same time, changing both the intercept and the slope of the trend.

Define the break point in the data as T_b and D_t is a dummy variable defined as:

$$D_t = \begin{cases} 0 & \text{if} \quad t < T_b \\ 1 & \text{if} \quad t > T_b \end{cases}$$

Unit Root Tests with Structural Breaks

The equation for the third (most general) version of the test is

$$\Delta y_t = \psi y_{t-1} + \mu + \alpha_1 D_t + \alpha_2 (t - T_b) D_t + \lambda t + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + \epsilon_t$$

For the crash only model, set $\alpha_2 = 0$

For the changing growth only model, set $\alpha_1 = 0$

In all three cases, there is a unit root with a structural break at T_b under the null hypothesis and a series that is a stationary process with a break under the alternative

A limitation of this approach is that it assumes that the break date is known in advance

It is possible, however, that the date will not be known and must be determined from the data.