Unit root nonstationarity and long-memory

by Hedibert Lopes

Based on Tsay’s Analysis of Financial Time Series (3rd edition)
Outline

1. Random walk
2. Mean reverting
3. Predictability
4. Random walk with drift
5. Trend-stationarity
6. ARIMA($p, 1, q$)
7. Unit-root tests
8. Examples
   - US GDP
   - S&P500 returns
9. Long memory models
UNIT ROOT NONSTATIONARITY

In some studies, interest rates, foreign exchange rates, or the price series of an asset are of interest.

These series tend to be nonstationary.

For a price series, the nonstationarity is mainly due to the fact that there is no fixed level for the price.

In the time series literature, such a nonstationary series is called unit-root nonstationary time series.

The best known example of unit-root nonstationary time series is the random-walk model.
Random walk

A time series \( \{p_t\} \) is a random walk if it satisfies

\[
p_t = p_{t-1} + a_t
\]

where \( p_0 \) is a real number denoting the starting value of the process and \( \{a_t\} \) is a white noise series.

If \( p_t \) is the log price of a particular stock at date \( t \), then \( p_0 \) could be the log price of the stock at its initial public offering (IPO) (i.e., the logged IPO price).

If \( a_t \) has a symmetric distribution around zero, then conditional on \( p_{t-1} \), \( p_t \) has a 50-50 chance to go up or down, implying that \( p_t \) would go up or down at random.
Random walk

Mean reverting

Predictability

Random walk with drift

Trend-stationarity

ARIMA($p$, 1, $q$)

Unit-root tests

Examples

US GDP

S&P500 returns

Long memory models

**Example**

```r
set.seed(1234)
n = 1000
p = rep(0,n)
a = rep(0,n)
p[1] = 0
a[1] = 1
for (t in 2:n){
a[t] = sample(c(-1,1), size=1)
p[t] = p[t-1] + a[t]
}

par(mfrow=c(2,2))
ts.plot(p[1:50], ylab="")
abline(h=0, lty=2)
ts.plot(p[1:100], ylab="")
abline(h=0, lty=2)
ts.plot(p[1:200], ylab="")
abline(h=0, lty=2)
ts.plot(p[1:1000], ylab="")
abline(h=0, lty=2)
```
**Random walk**

**Mean reverting**

**Predictability**

**Random walk with drift**

**Trend-stationarity**

**ARIMA**($p$, 1, $q$)

**Unit-root tests**

**Examples**

US GDP

S&P500 returns

**Long memory models**
AR(1) and Random Walk

If we treat the random-walk model as a special AR(1) model, then the coefficient of $p_{t-1}$ is unity, which does not satisfy the weak stationarity condition of an AR(1) model.

A random-walk series is, therefore, not weakly stationary, and we call it a **unit-root nonstationary time series**.

The random-walk model has widely been considered as a statistical model for the movement of *logged stock prices*. Under such a model, the stock price is not predictable or mean reverting.
**Mean Reverting**

The 1-step-ahead forecast of model $p_t = p_{t-1} + a_t$, at the forecast origin $h$ is

$$\hat{p}_h(1) = E(p_{h+1}|p_h, p_{h-1}, \ldots) = p_h.$$  

The 2-step-ahead forecast is

$$\hat{p}_h(2) = E(p_{h+2}|p_h, p_{h-1}, \ldots) = E(p_{h+1}|p_h, p_{h-1}, \ldots) = p_h,$$

which again is the log price at the forecast origin.

In fact, for any forecast horizon $l > 0$, we have

$$\hat{p}_h(l) = p_h$$

Therefore, the process is not mean reverting.
The MA representation of the random-walk model is

\[ pt = a_t + a_{t-1} + a_{t-2} + \cdots \]

First, the \( l \)-step ahead forecast error is

\[ e_h(l) = p_{h+l} - \hat{p}_h(l) \]
\[ = p_{h+l} - p_h \]
\[ = a_{h+l} + a_{h+l-1} + \cdots + a_{h+1} \]

so that

\[ V[e_h(l)] = l\sigma_a^2 \to \infty \quad \text{as} \quad l \to \infty. \]

This result says that the usefulness of point forecast \( \hat{p}_h(l) \) diminishes as \( l \) increases, which again implies that the model is not predictable.
**Memory**

The impact of any past shock $a_{t-i}$ on $p_t$ does not decay over time.

The series has a strong memory as it remembers all of the past shocks.

In economics, the shocks are said to have a permanent effect on the series.

The strong memory of a unit-root time series can be seen from the sample ACF of the observed series.

The sample ACFs are all approaching 1 as the sample size increases.
Random walk
Mean reverting
Predictability
Random walk with drift
Trend-stationarity
ARIMA($p$, 1, $q$)
Unit-root tests
Examples
US GDP
S&P500 returns
Long memory models

Based on 100,000 observations!
**Random walk with drift**

Log return series of a market index tends to have a small and positive mean, so the model for the log price is

\[ p_t = \mu + p_{t-1} + a_t \]

The constant term \( \mu \) represents the time trend of the log price \( p_t \) and is often referred to as the *drift of the model*.

Assume that the initial log price is \( p_0 \):

\[
\begin{align*}
    p_1 &= \mu + p_0 + a_1, \\
    p_2 &= \mu + p_1 + a_2 = 2\mu + p_0 + a_2 + a_1, \\
    p_3 &= \mu + p_2 + a_3 = 3\mu + p_0 + a_3 + a_2 + a_1, \\
    &\vdots \\
    p_t &= \mu t + p_0 + \sum_{i=1}^{t} a_i.
\end{align*}
\]
The log price consists of a time trend and a pure random-walk process:

\[ p_t = p_0 + \mu t + \sum_{i=1}^{t} a_i. \]

The conditional standard deviation of \( p_t \), \( \sigma_a \sqrt{t} \), grows at a slower rate than the conditional expectation of \( p_t \), \( p_0 + \mu t \).

Let \( n = 10,000 \) and \( a_t \) iid \( N(0, 0.0637^2) \) so

\[ p_t = \mu + p_{t-1} + a_t \]

where \( p_0 = 0 \) and \( \mu = 0 \) or \( \mu = 0.0103 \).
Random walk

Mean reverting

Predictability

Random walk with drift

Trend-stationarity

ARIMA($p$, 1, $q$)

Unit-root tests

Examples

US GDP

S&P500 returns

Long memory models

RW with drift
**INTERPRETATION OF THE CONSTANT TERM**

**MA(q) model:** the constant term is the mean of the series.

**Stationary AR(p) or ARMA(p, q) models:** the constant term is related to the mean via

$$\mu = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p}.$$

**Random walk with drift:** the constant term becomes the time slope of the series.
**Trend-stationarity**

The simplest trend-stationary time series model,

\[ p_t = \beta_0 + \beta_1 t + r_t, \]

where \( r_t \) is a stationary time series, say a stationary AR\((p)\).

Major difference between the two models:

- **Random walk with drift model**
  \[ E(p_t) = p_0 + \mu t \quad \text{and} \quad V(p_t) = \sigma_a^2 t. \]

- **Trend-stationary model**
  \[ E(p_t) = \beta_0 + \beta_1 t \quad \text{and} \quad V(p_t) = V(r_t) \]

with \( V(r_t) \) finite and time invariant.
Consider an ARMA model. If one extends the model by allowing the AR polynomial to have 1 as a characteristic root, then the model becomes the well-known autoregressive integrated moving-average (ARIMA) model.

An ARIMA model is said to be unit-root nonstationary because its AR polynomial has a unit root.

Like a random-walk model, an ARIMA model has strong memory because the $\psi_i$ coefficients in its MA representation do not decay over time to zero, implying that the past shock $a_{t-i}$ of the model has a *permanent effect* on the series.
A conventional approach for handling unit-root nonstationarity is to use *differencing*.

A time series $y_t$ is said to be an ARIMA($p$, 1, $q$) process if the change series

$$c_t = y_t - y_{t-1} = (1 - B)y_t$$

follows a stationary and invertible ARMA($p$, $q$) model.

Price series are believed to be nonstationary, but the log return series,

$$r_t = \log(P_t) - \log(P_{t-1}),$$

is stationary.
To test whether the log price $p_t$ of an asset follows a random walk or a random walk with drift, we employ the models

$$p_t = \phi_1 p_{t-1} + e_t$$  \hspace{1cm} (1)$$

$$p_t = \phi_0 + \phi_1 p_{t-1} + e_t$$  \hspace{1cm} (2)$$

where $e_t$ denotes the error term, and consider the null hypothesis

$$H_0 : \phi_1 = 1$$

versus the alternative hypothesis

$$H_a : \phi_1 < 1.$$  

This is the well-known unit-root testing problem (Dickey and Fuller, 1979).  

**Unit-root test**

To test whether the log price $p_t$ of an asset follows a random walk or a random walk with drift, we employ the models

$$p_t = \phi_1 p_{t-1} + e_t$$  \hspace{1cm} (1)$$

$$p_t = \phi_0 + \phi_1 p_{t-1} + e_t$$  \hspace{1cm} (2)$$

where $e_t$ denotes the error term, and consider the null hypothesis

$$H_0 : \phi_1 = 1$$

versus the alternative hypothesis

$$H_a : \phi_1 < 1.$$  

This is the well-known unit-root testing problem (Dickey and Fuller, 1979).  

**Unit-root test**

To test whether the log price $p_t$ of an asset follows a random walk or a random walk with drift, we employ the models

$$p_t = \phi_1 p_{t-1} + e_t$$  \hspace{1cm} (1)$$

$$p_t = \phi_0 + \phi_1 p_{t-1} + e_t$$  \hspace{1cm} (2)$$

where $e_t$ denotes the error term, and consider the null hypothesis

$$H_0 : \phi_1 = 1$$

versus the alternative hypothesis

$$H_a : \phi_1 < 1.$$  

This is the well-known unit-root testing problem (Dickey and Fuller, 1979).
**Dickey-Fuller Test**

A convenient test statistic is the $t$ ratio of the least-squares (LS) estimate of $\phi_1$ under the null hypothesis.

For equation (1), the LS method gives

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{T} p_{t-1}p_t}{\sum_{t=1}^{T} p_{t-1}^2} \quad \text{and} \quad \hat{\sigma}_e^2 = \frac{\sum_{t=1}^{T} (p_t - \hat{\phi}_1 p_{t-1})^2}{T - 1}$$

where $p_0 = 0$. The $t$ ratio is

$$DF \equiv t \text{ ratio} = \frac{\hat{\phi}_1 - 1}{\text{std}(\hat{\phi}_1)} = \frac{\sum_{t=1}^{T} p_{t-1}e_t}{\hat{\sigma}_e \sqrt{\sum_{t=1}^{T} p_{t-1}^2}}$$

which is commonly referred to as the Dickey-Fuller (DF) test.
If \( \{e_t\} \) is a white noise series with finite moments of order slightly greater than 2, then the DF statistic converges to a function of the standard Brownian motion as \( T \to \infty \); see Chan and Wei (1988) and Phillips (1987) for more information.

If \( \phi_0 = 0 \) but equation (2) is employed anyway, then the resulting \( t \) ratio for testing \( \phi_1 = 1 \) will converge to another nonstandard asymptotic distribution.

If \( \phi_0 \neq 0 \) and equation (2) is used, then the \( t \) ratio for testing \( \phi_1 = 1 \) is asymptotically normal.
An AR($p$) model

\[ y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t, \]

or

\[(1 - \phi_1 B - \phi_2 B^2 + \cdots \phi_p B^p) y_t = \Phi(B) y_t = \phi_0 + \epsilon_t,\]

can be rewritten as

\[ \Delta y_t = \pi x_{t-1} + \sum_{i=1}^{p-1} \tilde{\phi}_i \Delta y_{t-i} + \epsilon_t, \]

where

- \( \Delta y_t = (1 - B) y_t = y_t - y_{t-1}, \)
- \( \pi = -\Phi(1) = \phi_1 + \phi_2 + \cdots + \phi_p - 1, \)
- \( \tilde{\phi}_s \) are functions of the \( \phi_s \).

A (single) unit root implies that \( \Phi(1) = \pi = 0 \) or that

\[ \phi_1 + \phi_2 + \cdots + \phi_p = 1. \]
The model

\[ y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \varepsilon_t, \]

can be rewritten as

\[ \Delta y_t = \pi x_{t-1} + \tilde{\phi}_2 \Delta y_{t-2} + \tilde{\phi}_2 \Delta y_{t-2} + \varepsilon_t, \]

where

\[ \pi = \phi_1 + \phi_2 + \phi_3 - 1 \]

\[ \tilde{\phi}_1 = -(\phi_2 + \phi_3) \]

\[ \tilde{\phi}_2 = -\phi_3 \]
Augmented DF Test

To verify the existence of a unit root in an AR\((p)\) process, one may perform the test

\[ H_0 : \beta_1 = 1 \quad \text{versus} \quad H_a : \beta < 1 \]

using the regression

\[
y_t = c_t + \beta y_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta y_{t-i} + e_t, \]

\[ \Delta y_t = y_t - y_{t-1} \quad \text{and} \]

\[
c_t = \begin{cases} 
0 & \text{no constant, no trend} \\
\omega_0 \neq 0 & \text{constant, no trend} \\
\omega_0 + \omega_1 t & \text{constant, trend}
\end{cases}
\]
The $t$ ratio of $\hat{\beta} - 1$,

$$\text{ADF-test} = \frac{\hat{\beta} - 1}{\text{std}(\hat{\beta})}$$

where $\hat{\beta}$ denotes the least-squares estimate of $\beta$, is the well-known augmented Dickey-Fuller (ADF) unit-root test.
Consider the log series of U.S. quarterly GDP from 1947.I to 2008.IV.

The series exhibits an upward trend, showing the growth of the U.S. economy, and has high sample serial correlations.

The first differenced series, representing the growth rate of U.S. GDP, seems to vary around a fixed mean level, even though the variability appears to be smaller in recent years.

With $p = 10$, the ADF test statistic is $-1.61$ with a p-value 0.45, indicating that the unit-root hypothesis cannot be rejected.
Random walk
Mean reverting
Predictability
Random walk with drift
Trend-stationarity
ARIMA$(p, 1, q)$
Unit-root tests
Examples
US GDP
S&P500 returns
Long memory models
**R CODE**

```r
install.packages("tseries")
library("tseries")
da=read.table("http://faculty.chicagobooth.edu/ruey.tsay/teaching/fts3/q-gdp4708.txt",header=T)
n  = nrow(da)
da[c(1,n),]
# year mon day gdp
#1 1947 1 1 237.2
#248 2008 10 1 14200.3

gdp=log(da[,4])
date = da[,1]+da[,2]/12

par(mfrow=c(2,2))
plot(date,gdp,xlab="Year",ylab="log-gdp",type="l")
plot(date[2:n],diff(gdp),xlab="Year",ylab="diff log-gdp",type="l")
acf(gdp,main="log-gdp")
pacf(diff(gdp),main="diff log-gdp")

m1=ar(diff(gdp),method="mle")
m1
#Coefficients:
# 1 2 3 4 5 6 7 8 9 10
# 0.4045 0.1883 -0.1158 0.0215 -0.1602 0.0739 0.0853 -0.0993 0.1538 0.0959
#
#Order selected 10 sigma^2 estimated as 8.68e-05

adf.test(gdp,k=10)
#
# Augmented Dickey-Fuller Test
#data:  gdp
#Dickey-Fuller = -0.37047, Lag order = 10, p-value = 0.9872
#alternative hypothesis: stationary
```
S&P500 RETURNS

Consider the log series of the S&P 500 index from January 3, 1950, to April 16, 2008, for 14,462 observations.

Testing for a unit root in the index is relevant if one wishes to verify empirically that the index follows a random walk with drift. To this end, we use $c_t = \omega_0 + \omega_1 t$ in applying the ADF test.

Furthermore, we choose $p = 15$ based on the sample PACF of the first differenced series.

The resulting test statistic is $-1.998$ with a p-value of 0.602. Thus, the unit-root hypothesis cannot be rejected at any reasonable significance level.
Random walk
Mean reverting
Predictability
Random walk with drift
Trend-stationarity
ARIMA($p, 1, q$)
Unit-root tests
Examples
US GDP
S&P500 returns
Long memory models
Random walk
Mean reverting
Predictability
Random walk with drift
Trend-stationarity
ARIMA($p, 1, q$)
Unit-root tests
Examples
US GDP
S&P500 returns
Long memory models

install.packages("tseries")
library("tseries")
sp5=log(da[,7])
date = da[,1]+(da[,2]*30+da[,3]/31)/366

pdf(file="adftest1.pdf",width=10,height=8)
par(mfrow=c(2,2))
plot(date,da[,7],xlab="Days",ylab="SP500",type="l",main="Log price")
plot(date,sp5,xlab="Days",ylab="SP500",type="l",main="Log returns")
acf(sp5,main="Log returns")
pacf(sp5,main="Log returns")
dev.off()

adf.test(sp5,k=10)
# Augmented Dickey-Fuller Test
data:  sp5
#Dickey-Fuller = -1.9482, Lag order = 10, p-value = 0.6004
#alternative hypothesis: stationary

m2=ar(diff(sp5),method="mle")
m2$x.mean
# [1]  0.0002988915
m2
#Call:
ar(x = diff(sp5), method = "mle")
#
#Coefficients:
# [1] 1 2
# 0.0721 -0.0387
#
#Order selected 2 sigma^2 estimated as 8.068e-05

adf.test(diff(sp5),k=10)
# Augmented Dickey-Fuller Test
data:  diff(sp5)
#Dickey-Fuller = -37.537, Lag order = 10, p-value = 0.01
#alternative hypothesis: stationary
Long memory models

There exist some time series whose ACF decays slowly to zero at a polynomial rate as the lag increases. These processes are referred to as long-memory time series.

One such example is the fractionally differenced process defined by

$$(1 - B)^d x_t = a_t - 0.5 < d < 0.5,$$

where $\{a_t\}$ is a white noise series, and

$$(1 - B)^d = \sum_{k=0}^{\infty} (-1)^k \frac{d(d - 1) \cdots (d - k + 1)}{k!} B^k$$
The study of long memory originated in the 1950s in the field of hydrology, where studies of the levels of the river Nile (Hurst, 1951) demonstrated anomalously fast growth of the rescaled range of the time series.

After protracted debates about whether this was a transient (finite time) effect, the mathematical pioneer Benoît B. Mandelbrot showed that if one retained the assumption of stationarity, novel mathematics would then be essential to sufficiently explain the Hurst effect.

In doing so he rigorously defined (Mandelbrot and Van Ness, 1968; Mandelbrot and Wallis, 1968) the concept of long memory.

Some properties

1. If $d < 0.5$, then $x_t$ is a weakly stationary process and has the infinite MA representation

$$x_t = a_t + \sum_{i=1}^{\infty} \psi_i a_{t-i}$$

where

$$\psi_k = \frac{(k + d - 1)!}{k!(d - 1)!}$$

2. If $d > -0.5$, then $x_t$ is invertible and has the infinite AR representation

$$x_t = \sum_{i=1}^{\infty} \pi_i x_{t-i} + a_t$$

where

$$\pi_k = \frac{(k - d - 1)!}{k!(-d - 1)!}$$
3. For $-0.5 < d < 0.5$, the ACF of $x_t$ is

$$\rho_k = \frac{d(1 + d) \cdots (k - 1 + d)}{(1 - d)(2 - d) \cdots (k = d)}, \quad k = 1, 2, \ldots.$$ 

In particular, $\rho_1 = d/(1 - d)$ and

$$\rho_k \approx \frac{(-d)!}{(d - 1)!} k^{2d-1} \quad \text{as} \quad k \to \infty.$$

4. For $-0.5 < d < 0.5$, the PACF of $x_t$ is $\phi_{k,k} = d/(k - d)$ for $k = 1, 2, \ldots$.

5. For $-0.5 < d < 0.5$, the spectral density function $f(\omega)$ of $x_t$, which is the Fourier transform of the ACF of $x_t$, satisfies

$$f(\omega) \sim \omega^{-2d}, \quad \text{as} \quad \omega \to 0,$$

where $\omega \in [0, 2\pi]$ denotes the frequency.
Of particular interest here is the behavior of ACF of $x_t$ when $d < 0.5$.

The property says that $\rho_k \sim k^{2d-1}$, which decays at a polynomial, instead of exponential, rate.

For this reason, such an $x_t$ process is called a long-memory time series.

A special characteristic of the spectral density function is that the spectrum diverges to infinity as $\omega \to 0$.

However, the spectral density function of a stationary ARMA process is bounded for all $\omega \in [0, 2\pi]$. 
If the fractionally differenced series \((1 - B)^d x_t\) follows an ARMA\((p, q)\) model, then \(x_t\) is called an ARFIMA\((p, d, q)\) process, which is a generalized ARIMA model by allowing for noninteger \(d\).

In practice, if the sample ACF of a time series is not large in magnitude, but decays slowly, then the series may have long memory.

For the pure fractionally differenced model, one can estimate \(d\) using either a maximum-likelihood method or a regression method with logged periodogram at the lower frequencies.
Bayesian ARFIMA


