Trends and DF tests

Trend stationary (TS) series (Figs. 1 & 2)

A trend stationary series fluctuates around a deterministic trend (the mean of the series) with no tendency for the amplitude of the fluctuations to increase or decrease.

Linear deterministic trend: $y_t = \alpha + \beta t + \varepsilon_t$ $\quad \varepsilon_t \sim iid(0,\sigma^2)$, $t=1,2,...,T$

Quadratic deterministic trend: $y_t = \alpha + \beta t + \gamma t^2 + \varepsilon_t$ $\quad \varepsilon_t \sim iid(0,\sigma^2)$, $t=1,2,...,T$

In general: $y_t=f(t)+ \varepsilon_t$ $\quad \varepsilon_t \sim iid(0,\sigma^2)$, $t=1,2,...,T$

The mean is time dependent, but the variance is constant. ε_t can be any stationary series. I.e. it does not have to be a white-noise.

The trend can be removed by fitting a deterministic polynomial time trend: the residuals from the fitted trend will give us the de-trended series.

An AR(1) process with a linear deterministic trend can be represented as

$$y_t - \alpha - \beta t = \phi_1(y_{t-1} - \alpha - \beta(t-1)) + \varepsilon_t \qquad \varepsilon_t \sim iid(0,\sigma^2),\ t=1,2,...,T$$

where the deterministic trend is subtracted from y_t. In practice, however, α and β are unknown and have to be estimated. The model can be rewritten as

$$y_t = (1-\phi_1)\alpha + \phi_1 \beta + (1-\phi_1)\beta t + \phi_1 y_{t-1} + \varepsilon_t$$

which includes an intercept and a trend, that is

$$y_t = \alpha^* + \beta^* t + \phi_1 y_{t-1} + \varepsilon_t$$

where $\alpha^*=(1-\phi_1)\alpha + \phi_1\beta$, and $\beta^*= (1-\phi_1)\beta$

If $|\phi_1|<1$, the AR process is stationary around a deterministic trend.

Difference Stationary (DS) series (or Integrated series) & Stochastic Trends

If a non-stationary series can be made stationary by differencing d times we say that the series is integrated to order d and write I(d). The process is also known as a Difference-stationary-process (**DSP**).

A stationary series is therefore, integrated of order zero, I(0). A white noise series is I(0).

So y_t is I(d) if the series $w_t=\Delta^d y_t$ is stationary. Δ is the difference operator, i.e. $\Delta y_t = y_t - y_{t-1};$
$\Delta^2 y_t = \Delta\Delta y_t = \Delta(y_t - y_{t-1}) = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$ etc.

y_t is I(1) if the series $w_t=\Delta y_t = y_t - y_{t-1}$ is stationary.
y_t is I(2) if the series $w_t=\Delta^2 y_t = y_t - 2y_{t-1} + y_{t-2}$ is stationary.

Random Walk (Fig. 3)

y_t is a random walk if $y_t = y_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim \text{iid}(0, \sigma_\varepsilon^2)$.

This is an AR(1) with $\phi = 1$ in $y_t = \phi y_{t-1} + \varepsilon_t$. This series is said to have a unit root, or to be integrated of order 1, I(1).

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

Let the process start at t = 0 with a value y_o which is assumed to be fixed. Then

$$y_1 = y_0 + \varepsilon_1$$
$$y_2 = y_1 + \varepsilon_2 = y_o + \varepsilon_1 + \varepsilon_2$$
.
.
$$y_t = y_0 + \varepsilon_1 + \varepsilon_2 + .. + \varepsilon_t = y_0 + \sum_{\tau=1}^{t} \varepsilon_\tau \qquad (1)$$

y_t is expressed as a function of the initial value y_0 and a partial sum series $\sum_{\tau=1}^{t} \varepsilon_\tau$ called the stochastic trend. All stochastic shocks ε_τ can change the level of y_t permanently. While in a stationary series the effect of a shock tends to zero as time passes.

Summary of properties of simple random walk

$E(y_t) = y_0 + t \text{ (times) } 0 = y_0$ [fixed] .

$Var(y_t) = Var(\sum_{\tau=1}^{t} \varepsilon_\tau) = t\sigma^2$ time dependent, ie. $Var(y_t)$ has a trend. So y_t is non-stationary.

But $\Delta y_t = \varepsilon_t$ which is stationary.

Also time dependent covariances and autocorrelations:

$$\gamma(s) = (t-s)\sigma^2\,, \quad \rho(s) = \sqrt{\frac{t-s}{t}}$$

Random walk covariances and autocorrelations

$$\text{Cov}(y_t, y_{t-1}) = \gamma(1) = E\left[\left(\sum_{\tau=1}^{t} \varepsilon_\tau\right)\left(\sum_{s=1}^{t-1} \varepsilon_s\right)\right] = E\left(\sum_{\tau=1}^{t-1} \varepsilon_\tau^2\right) = (t-1)\sigma^2$$

$$\rho(s) = \frac{\gamma(s)}{\sqrt{V(y_t)}\sqrt{V(y_{t-s})}} = \frac{(t-s)\sigma^2}{\sqrt{t\sigma^2}\sqrt{(t-s)\sigma^2}} = \frac{(t-s)}{\sqrt{t}\sqrt{(t-s)}} = \sqrt{\frac{t-s}{t}}$$

Random walk <u>with a drift</u> $y_t = \mu + y_{t-1} + \varepsilon_t$, $\Delta y_t = \mu + \varepsilon_t$

$$y_t = y_0 + t\,\mu + \sum_{\tau=1}^{t} \varepsilon_\tau$$

$E(y_t) = y_0 + t\,\mu$ a trend in the mean

$Var(y_t) = t\sigma^2$ a trend in the variance

Random walk with drift has a trend in both mean and variance.

It has a *deterministic trend* ($y_0 + t\,\mu$) plus a *stochastic trend* $\sum_{\tau=1}^{t} \varepsilon_\tau$.

Note: A random walk is a special case of an I(1) series. A differenced stationary process, say Δy_t, can be ARMA (p,q) in general: $\Phi(L)\Delta y_t = \Theta(L)\varepsilon_t$, but for a random walk Δy_t is iid(0, σ^2).

- In $y_t = y_{t-1} + \varepsilon_t$ the effects of a shock last for ever where in a stationary series $y_t = \phi_1 y_{t-1} + \varepsilon_t$, the effects of a shock $\rightarrow 0$ as time passes.
- An I(0) series will fluctuate around a mean and observations will cross this value frequently. An I(1) series will wander widely and will only rarely return to an earlier value.
- $\rho_k \rightarrow 0$ rapidly for an I(0) series. For an I(1) series this will stay around 1 for even very large k
- Standard distribution theory is <u>not valid</u> when we are dealing with non-stationary series. In particular, the Weak-Law-of-Large-Numbers (WLLN) does not hold.
  (WLLN: under certain conditions, the sample moments converge to population moments as the sample size $\rightarrow \infty$).

Differences between I(0) and I(1) series – Summary

I(0)	I(1)
Effects of a shock $\rightarrow 0$ as time passes	Effects of a shock lasts for ever
Observations will fluctuate around a mean and they will cross this value frequently	Observations will wander widely and will only rarely return to an earlier value
$\rho_k \rightarrow 0$ rapidly	ρ_k will stay around 1 for even very large k
Standard distribution theory can be used	Standard distribution theory cannot be used

FIGURE 1 Trend Stationary Process

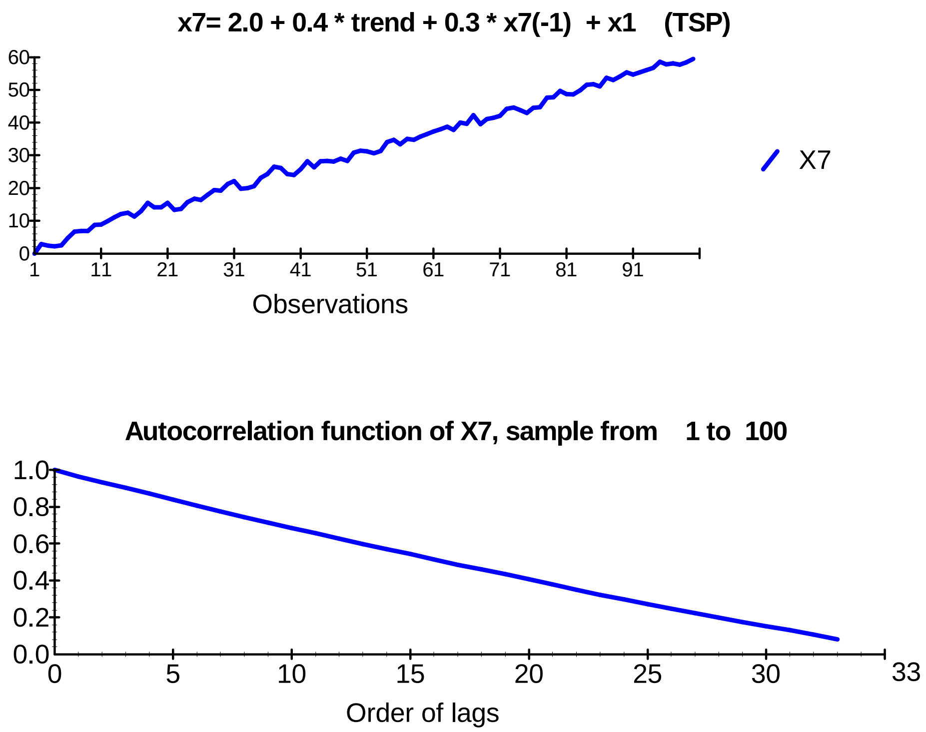
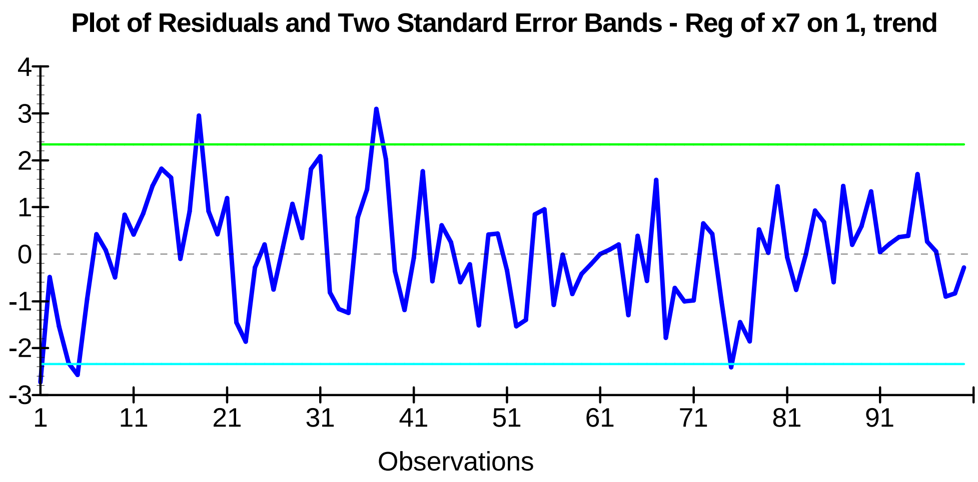
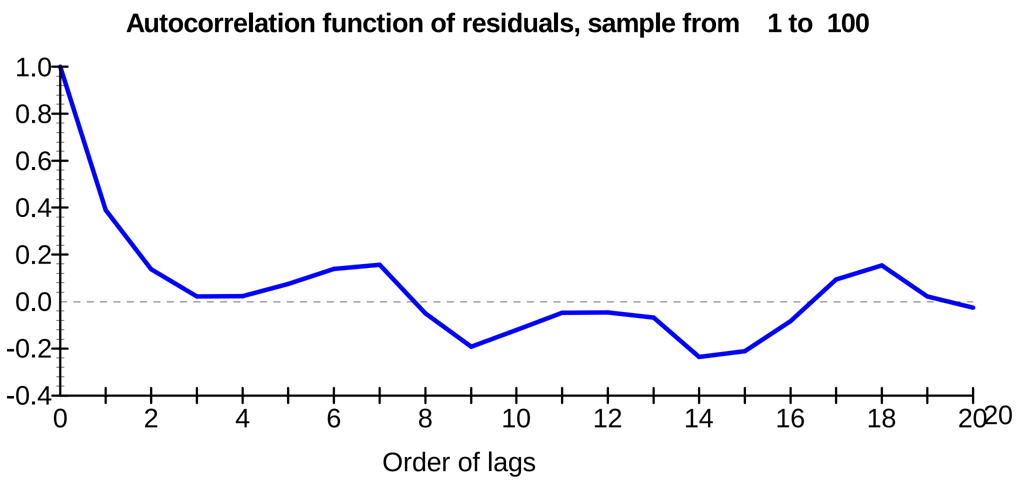
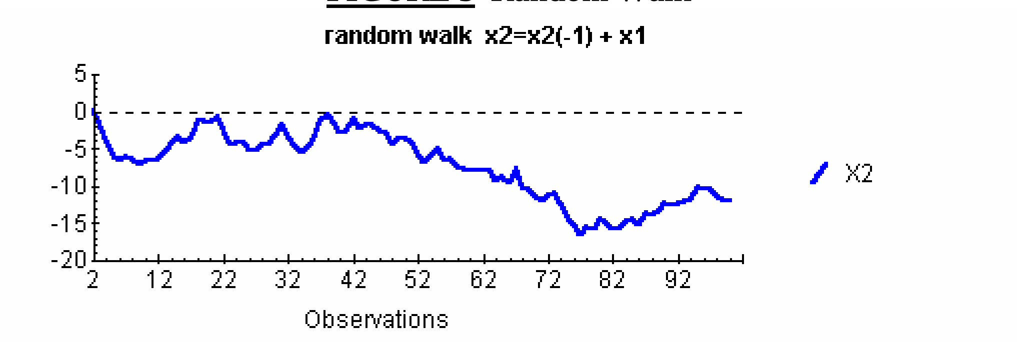


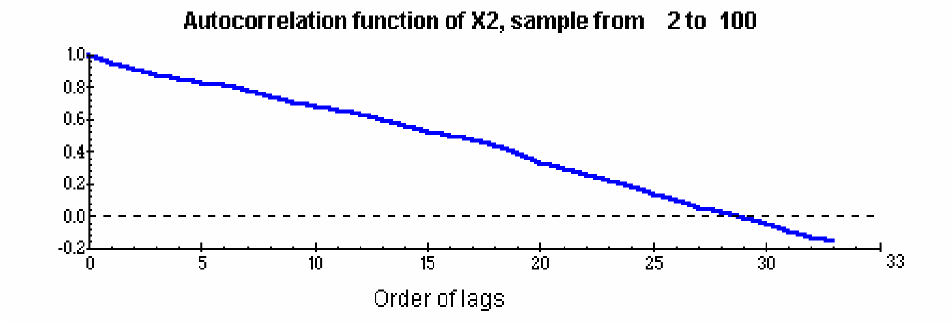
FIGURE 2 Detrended series from Figure 1.





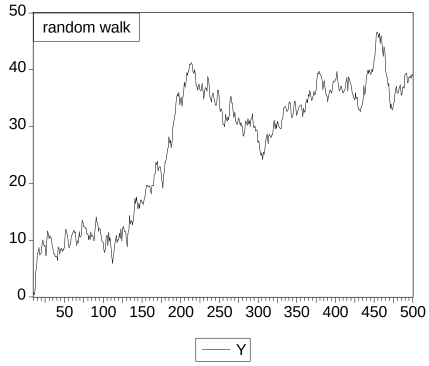
<u>FIGURE 3</u> Random Walk

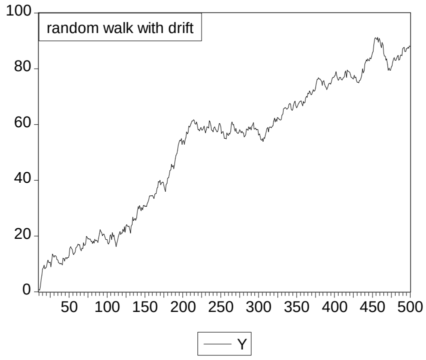




Detecting non stationarity by visual inspection of plots and correlograms

A simple random walk and a random walk with drift



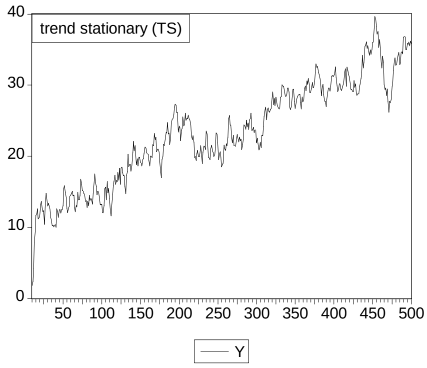


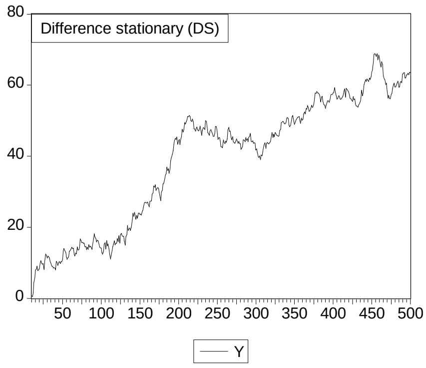
The <u>simple random walk</u> process shows no particular tendency to increase or decrease over time, nor it shows any tendency to revert to a given mean value (e.g. exchange rates)
The time path of the <u>random walk with drift</u> is dominated by the deterministic trend (e.g. money supply, GNP). The series can deviate from the deterministic trend for long periods.

In small samples it is not always easy to distinguish between a random walk and a random walk with drift. A small absolute value of the drift μ, or a large variance of the shock ε could mask the long run properties of the random walk with drift.

Also it is not easy to distinguish between a stationary AR process (with deterministic trend) and a random walk (with drift).

Trend Stationary (TS) and Difference Stationary (DS) series

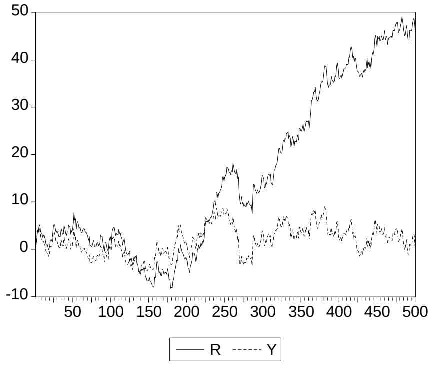


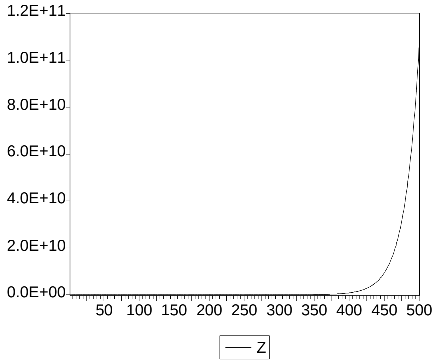


Consider now the following three series:

1) stationary AR(1) process $y_t = 0.05 + 0.95y_{t-1} + \varepsilon_t$
2) random walk with drift $r_t = 0.05 + r_{t-1} + \varepsilon_t$
3) explosive series $z_t = 0.05 + 1.05z_{t-1} + \varepsilon_t$

sample 1 500





An informal test for stationarity is based on inspection of the autocorrelation function.
The correlogram of a stationary AR series should decline exponentially, while for a nonstationary series it declines very slowly. Below are reported selected autocorrelation coefficients for the y and r series.

AUTOCORRELATIONS

	y: AR(1)	**r: RW**
1	0.951	0.995
3	0.856	0.982
5	0.783	0.972
11	0.596	0.939
15	0.496	0.917
36	0.127	0.770

The patterns of the two sets of autocorrelations are clearly different, confirming the stationarity of the y series and the nonstationarity of the random walk.

Testing for nonstationarity (or for the presence of unit roots)

The presence of unit roots means that standard distribution theory is not valid.
So it is important to test for stationarity of the series prior to any estimation, in order to use the appropriate procedure for de-trending.

Dickey and Fuller (DF) tests and Augmented DF tests

Single hypothesis tests

Consider an AR(1) process

$$y_t = \rho y_{t-1} + \varepsilon_t \quad \varepsilon_t \sim iid(0,\sigma^2)$$

If $\rho=1$ the equation defines a simple random walk and y is nonstationary. The null hypothesis for testing nonstationarity is H_0: $\rho = 1$. The test of this hypothesis is a unit root test.

A simple way to test the null hypothesis is to respecify the AR(1) equation as

$$y_t - y_{t-1} = (\rho-1)\, y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t \qquad (1)$$

thus the hypothesis H_0: $\rho = 1$ is now equivalent to a test of H_0: $\gamma = 0$, and we test against $\rho < 1$ $(\gamma < 0)$ (rejection region is on the left). We do not consider the case where $|\rho| > 1$ since this is an explosive case and we would not expect economic series to be explosive.

The equation can also include a constant term

$$\Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t \qquad (2)$$

and a constant and a trend variable

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \varepsilon_t \qquad (3)$$

The test is performed by estimating with OLS one of the three regressions above, and compare the t-statistic on the coefficient γ with the appropriate critical values.

t-ratio for $\hat{\gamma}$ in (1) [i.e. no constant equation] is called $\hat{\tau}_{nc}$

t-ratio for $\hat{\gamma}$ in (2) [no trend equation] is called $\hat{\tau}_c$

t-ratio for $\hat{\gamma}$ in (3) [with trend equation] is called $\hat{\tau}_{ct}$

In each of the three cases H_0 is $\gamma = 0$ (a unit root), against H_1 is $\gamma < 0$ (stationary series). If H_0 is rejected then:

In (1) y_t is stationary with zero mean: $y_t = \rho y_{t-1} + \varepsilon_t$

In (2) y_t is stationary with a non-zero mean in:

$$y_t - \mu = \rho(y_{t-1} - \mu) + \varepsilon_t \text{ or } y_t = \alpha + \rho y_{t-1} + \varepsilon_t \text{ where } \alpha = \mu(1-\rho)$$

In (3) y_t is stationary around a deterministic trend in :

$$y_t - a - bt = \rho(y_{t-1} - a - b(t-1)) + \varepsilon_t$$

$$\text{or } y_t = \alpha + \beta t + \rho y_{t-1} + \varepsilon_t \text{ where } \alpha = a(1-\rho) + b\rho, \text{ and } \beta = b(1-\rho) .$$

Critical values are reported in Fuller, W.A., 1976, Introduction to Statistical Time Series, and Dickey and Fuller (1981), Econometrica (report critical values for some sample sizes).

Note: τ does not have the Student t distribution under H_0: $\rho = 1$. The distribution of τ is shifted to the left, relative to that of a Student t or N(0,1). Larger negative values are needed to reject the null hypothesis. See table below: to reject H_0 at the 5% level we need $\tau<-1.95$. The 5% one-sided c.v. for a N(0,1), or Student t with large T, is –1.645. Thus inappropriate use of Std. Normal c.v's would lead to over-rejection of the null.

Asymptotic critical values for unit root tests

Test statistic	1%	5%	10%
N(0,1)	-2.33	-1.645	-1.128
τ_{nc}	-2.56	-1.94	-1.62
τ_c	-3.43	-2.86	-2.57
τ_{ct}	-3.96	-3.41	-3.13

J. Mackinnon provides a method for calculating c.v.'s for all sample sizes, T, using response surface regressions) (see J. Mackinnon, "Critical values for CI tests" in Engle and Granger (ed.), Long run economic relationships, OUP, 1991, Chapter 13) .

Example: $\mathrm{CV_T} = \phi_\infty + \frac{\phi_1}{\mathrm{T}} + \frac{\phi_2}{\mathrm{T}^2}$, for given values of ϕ_∞, ϕ_1 and ϕ_2, for T=106, the 5% critical value for the constant plus trend case as in equation (3) (N = 1 row in Mackinnon's table) is C.V. = $-3.4126 - \frac{4.039}{106} - \frac{17.83}{106^2} = -3.4523$. The null hypothesis of $H_0 : \gamma = 0$ is rejected in favour of the alternative $H_1 : \gamma < 0$ if $t_\gamma = \frac{\hat{\gamma} - 0}{se(\hat{\gamma})} \leq \mathrm{CV_T}$.

Response surfaces for critical values of cointegration tests

n	model	%point	ϕ_∞	ϕ_1	ϕ_2
1	Costant + trend	1			
		5	-3.4126	-4.039	-17.83
		10			

The **Augmented Dickey-Fuller (ADF)** test is used if the errors ε_t in (1)-(2)-(3) are serially correlated (y_t is an AR(p)). The test is carried out by expanding the DF regressions with lagged difference terms. So, for example, regression (3) becomes:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \varepsilon_t \qquad (4)$$

where sufficient lagged first differences should be included to secure an approximate white noise error term in the ADF regression. A test for the existence of a unit root is calculated as a t-statistic on γ, $t_\gamma = \frac{\hat{\gamma} - 0}{se(\hat{\gamma})}$, using the same critical values as those for the DF tests. In general, an AR(p) process can be reparameterised as

$$\Delta y_t = \gamma y_{t-1} + \sum_{j=1}^{p-1} \delta_j \Delta y_{t-j} + \varepsilon_t$$

For example, we can rewrite the AR(2) process

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

in differenced form as

$$\Delta y_t = (\phi_1 + \phi_2 - 1) y_{t-1} - \phi_2 \Delta y_{t-1} + \varepsilon_t$$

$$\Delta y_t = \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \varepsilon_t$$

where $\gamma=\phi_1+\phi_2-1$ and $\delta_1 = -\phi_2$.

Under a unit root $\phi_1+\phi_2=1$. Remember that $\phi_1+\phi_2<1$ is a necessary condition for stationarity of an AR(2). This is why when we test H_0: $\gamma=0$ against H_1:$\gamma<0$, we test the existence of a unit root.

Limit of the unit root tests: Problems with power and size.
[Power= probability of rejecting null when false; size = level of significance, probability of type I error, i.e. reject null when true]

Low power (a) in small samples, (b) for ρ close to one, (c) in the presence of structural break. So, failure to reject H_0 provides only weak evidence in favour of the random walk hypothesis.

The size of the test may be affected by under specifying the lag length or using the incorrect equation for the test (wrong deterministic components).

Lag Length in ADF regressions can be chosen in various ways:

(**a**) Minimise a selection criteria, like the Akaike Information Criteria (AIC) or the Schwarz Criteria (SC); (**b**) select a large value for the lag order, use the standard normal distribution to test the significance of the last autoregressive coefficient, and if the null hypothesis is accepted reduce the lag order by one; (**c**) use different values for p and see if the results of the ADF test are robust; (**d**) Select the smallest value for p such that the errors are approximately white noise.

Choice of the deterministic components.

Suggestion: plot the series, if it looks like a simple r.w., use regression with constant term. If it looks like a r.w. with drift use regression with constant and time trend.
So, for inflation use model (2), for output use model (3).

Phillips and Perron (PP) nonparametric tests

Like the ADF test, the Phillips-Perron test is a test of the hypothesis $\gamma = 0$ in one of the three equations above. Unlike the ADF test, there are no lagged difference terms. Instead, the equation is estimated by ordinary least squares (with the optional inclusion of constant and time trend) and then the t-statistic of the coefficient is corrected for serial correlation in ε_t using the Newey-West procedure for adjusting the standard errors. The critical values for these tests are identical to those for the ADF equivalents. In some cases this test is more powerful than the ADF test and is less sensitive to over-specifying the lag length.

Test for a second unit root

This is done by using the model: $\Delta^2 y_t = \alpha + \gamma \Delta y_{t-1} + \sum_{j=1}^{p} \delta_j \Delta^2 y_{t-j} + \varepsilon_t$

and test for $H_0 : \gamma = 0$ against $H_1 : \gamma < 0$ using the same MacKinnon's critical values used to test for one unit root. Note that in this case only equation (1) and (2) are sensible.

Unit root tests and structural breaks

A trend stationary series which exhibits a structural break either in the intercept (level) or the slope of the deterministic trend (growth) cannot easily be distinguished from a nonstationary process. Thus, a unit root test which does not account for breaks in the series has very low power. Perron (1989, ‘The Great Crash, the oil price shock and the unit root hypothesis’, Econometrica, 57, 1361-1401) derives critical values for unit root tests which take account of shifts in the trend or intercept.

Nelson and Plosser (1982, ‘Trends and random walks in macroeconomic time series: some evidence and implications’, J. of Monetary Economics, 10, 139-162), found many US macroeconomic t.s. I(1). Perron (1989) showed that many of these series are I(0) but with structural breaks.

Joint tests

It is possible to test for a unit root and the deterministic components jointly.
Take equation (2) and (3) on its own or modified according to the ADF specifications. The various joint tests are:

F ratio (restricted vs unrestricted RSS) for testing $\alpha = \gamma = 0$ in (2) is called Φ_1 .
F ratio for testing $\alpha = \beta = \gamma = 0$ in (3) is called Φ_2 .
F ratio for testing $\beta = \gamma = 0$ in (3) is called Φ_3 .

$\Delta z_t = \alpha + \gamma z_{t-1} + \varepsilon_t$	H_0: $\alpha = \gamma = 0$	Φ_1
$\Delta z_t = \alpha + \beta t + \gamma z_{t-1} + \varepsilon_t$	H_0: $\alpha = \beta = \gamma = 0$	Φ_2
$\Delta z_t = \alpha + \beta t + \gamma z_{t-1} + \varepsilon_t$	H_0: $\beta = \gamma = 0$	Φ_3

The distribution of these F-type tests is again non-standard, critical values are reported in Dickey and Fuller (1981). For a sample of these critical values see Gujarati Table D.7, appendix D.