Università di Pavia

Unit roots

Eduardo Rossi

The question of whether to detrend or to difference a time series prior to further analysis depends on whether the time series is TSP (trend-stationary process) or DSP (difference-stationary process).

- A time series is said to have a unit root if we can write it as

$$Y_t = DT_t + u_t$$

$$u_t = \phi u_{t-1} + \varepsilon_t$$

with $\phi = 1$, ε_t stationary and DT_t a deterministic component. Δu_t is stationary and ΔY_t is stationary around the change in the deterministic part. In this case $Y_t \sim I(1)$.

- $Y_t \sim I(0)$, and DT_t is a linear trend, then

$$Y_t = DT_t + \varepsilon_t$$

$$\Delta Y_t = \alpha + \varepsilon_t - \varepsilon_{t-1}$$

$$\alpha = DT_t - DT_{t-1}$$

Thus ΔY_t has a moving-average unit root. MA unit roots arise if a stationary series differenced. This is the overdifferencing.

Two classes of tests, called *unit root tests*, have been developed to answer this question:

- Tests using $I(1)$ as the null hypothesis are based on the unit autoregressive root as the null.
- Tests using $I(0)$ as the null hypothesis are based on the unit moving-average root as the null.

Conceptually the unit root tests are straightforward. In practice, however, there are a number of difficulties:

- Unit root tests generally have nonstandard and non-normal asymptotic distributions.
- These distributions are functions of standard Brownian motions, and do not have convenient closed form expressions. Consequently, critical values must be calculated using simulation methods.
- The distributions are affected by the inclusion of deterministic terms, e.g. constant, time trend, dummy variables, and so different sets of critical values must be used for test regressions with different deterministic terms.

Dickey-Fuller Tests set up:

$$\begin{aligned} Y_t &= \phi Y_{t-1} + \varepsilon_t \\ Y_t &= \alpha + \phi Y_{t-1} + \varepsilon_t \\ Y_t &= \alpha + \beta t + \phi Y_{t-1} + \varepsilon_t \end{aligned}$$

For instance,

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

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

$$\begin{aligned} y_t &= \mu_t + u_t \\ \mu_t &= \mu_{t-1} + \epsilon_t \quad \epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2) \\ \phi(L)u_t &= \theta(L)\eta_t \quad \eta_t \sim i.i.d.(0, \sigma_\eta^2) \end{aligned}$$

with $Cov(\epsilon_t, \eta_t) = 0$. The hypotheses are

$$\begin{aligned} H_0 : \sigma_\epsilon^2 &= 0 \quad (\mu_t = \mu_0) \\ H_1 : \sigma_\epsilon^2 &> 0 \quad (\mu_t = \mu_0 + \sum_{i=1}^{t} \epsilon_i) \end{aligned}$$

Consider OLS estimation of a Gaussian AR(1) process

$$Y_t = \phi Y_{t-1} + \epsilon_t$$

$$\epsilon_t \sim i.i.d. N\left(0, \sigma^2\right)$$

$$Y_0 = 0$$

the OLS estimate of ϕ is given by

$$\widehat{\phi}_T = \frac{\sum_1^T Y_{t-1} Y_t}{\sum_1^T Y_{t-1}^2} = \phi + \frac{\sum_1^T Y_{t-1}\epsilon_t}{\sum_1^T Y_{t-1}^2}$$

$$\sqrt{T}\left(\widehat{\phi}_T - \phi\right) \xrightarrow{d} N\left(0, \left(1-\phi^2\right)\right)$$

when $\rho = 1$, $Asyvar\left[\sqrt{T}\left(\widehat{\phi}_T - \phi\right)\right] = 0$, $\sqrt{T}\left(\widehat{\phi}_T - \phi\right) \xrightarrow{p} 0$.

The unit root coefficient converges at a faster rate (T) than a coefficient for a stationary regression.
To obtain a nondegenerate asymptotic distribution for $\widehat{\phi}_T$ in the unit root case, we have to multiply by T rather by $\sqrt{T}$.

$$T\left(\widehat{\phi}_T - 1\right) = T\frac{\sum_1^T Y_{t-1}\epsilon_t}{\sum_1^T Y_{t-1}^2}$$

Under the unit root null, $\{Y_t\}$ is not stationary and the usual sample moments do not converge to fixed constants.
Phillips (1987) showed that the sample moments of $\{Y_t\}$ converge to random functions of Brownian motion:

$$T^{-3/2}\sum_{t=}^{T} Y_{t-1} \xrightarrow{d} \int_0^1 W(r)\mathrm{d}r$$

$$T^{-2}\sum_{t=}^{T} Y_{t-1} \xrightarrow{d} \int_0^1 W(r)^2\mathrm{d}r$$

$$T^{-1}\sum_{t=1}^{T} Y_{t-1}\epsilon_t \xrightarrow{d} \int_0^1 W(r)\mathrm{d}W(r)$$

where $W(r)$ denotes a standard Brownian motion (*Wiener process*) defined on the unit interval.

A Wiener process $W(\cdot)$ is a continuous-time stochastic process, associating each date $r \in [0,1]$ a scalar random variable $W(r)$ that satisfies:

1. $W(0) = 0$
2. For any date $0 \leq t_1 \leq \ldots \leq t_k \leq 1$ the changes

   $$(W_{t_2} - W_{t_1}), (W_{t_3} - W_{t_2}), \ldots, (W_{t_k} - W_{t_{k-1}})$$

   are independent normal with

   $$W(t) - W(s) \sim N(0, (t - s))$$

3. $W(r)$ is continuous in r.

Under the null of unit root:

$$T\left(\widehat{\phi}_T - 1\right) = \frac{T^{-1}\sum_{t=1}^{T} Y_{t-1}\varepsilon_t}{T^{-2}\sum_{t=1}^{T} Y_{t-1}^2} \xrightarrow{d} \frac{\int_0^1 W(r)\, dW(r)}{\int_0^1 W(r)^2\, dr} = \frac{\frac{1}{2}\left[W(1)^2 - 1\right]}{\int_0^1 W(r)^2\, dr}$$

This is known as *Dickey-Fuller distribution*.

This result is based on:

- the numerator

$$T^{-1}\sum_{t=1}^{T} Y_{t-1}\varepsilon_t \xrightarrow{d} \int_0^1 W(r)\, dW(r) = \frac{1}{2}\left[W(1)^2 - 1\right] \sim \frac{1}{2}\left[\chi_1^2 - 1\right]$$

- the denominator has a nonzero asymptotic variance and

$$T^{-2}\sum_{t=1}^{T} Y_{t-1}^2 \xrightarrow{d} \int_0^1 W(r)^2\, dr$$

Since the normalized bias $T\left(\widehat{\phi}_T - 1\right)$ has a well defined limiting distribution that does not depend on nuisance parameters it can also be used as a test statistic for the null hypothesis $H_0 : \phi = 1$.

We can use the distribution for testing the unit root null hypothesis $H_0 : \phi = 1$, or we can normalize it with the standard error of the OLS estimator and construct the t-statistic testing $\phi = 1$.
The test statistic is

$$t_{\widehat{\phi}} = \frac{\left(\widehat{\phi} - 1\right)}{SE\left(\widehat{\phi}\right)} \xrightarrow{d} \frac{\int_0^1 W(r)\, dW(r)}{\left[\int_0^1 W(r)^2\, dr\right]^{1/2}} = \frac{\frac{1}{2}\left[W(1)^2 - 1\right]}{\left[\int_0^1 W(r)^2\, dr\right]^{1/2}}$$

The asymptotic distribution of the $t-$statistics under $\phi = 1$ is not the standard $t-$distribution, and so conventional critical values are not valid. Quantiles of the distribution must be computed by numerical approximation or by simulation.

- The numerator of the distribution is skewed to the right, being a χ_1^2 minus its expectation;
- the ratio is skewed to the left.

The usual one-sided 5% critical value for standard normal is -1.645.
The one-sided 5% critical value for the DF distribution is -1.941.
Note: -1.645 is the 9.45% quantile of the DF distribution.

Thus using the conventional critical values can lead to considerable *overrejection* of the null hypothesis of a unit root.

When testing for unit roots, it is crucial to specify the null and alternative hypotheses appropriately to characterize the trend properties of the data at hand.

- If the observed data does not exhibit an increasing or decreasing trend, then the appropriate null and alternative hypotheses should reflect this.
- The trend properties of the data under the alternative hypothesis will determine the form of the test regression used.
- The type of deterministic terms in the test regression will influence the asymptotic distributions of the unit root test statistics.

The test regression is

$$Y_t = c + \phi Y_{t-1} + \epsilon_t$$

and includes a constant to capture the nonzero mean under the alternative. The hypotheses to be tested are

$$\begin{aligned} H_0 &: \quad c = 0, \phi = 1 \quad (Y_t \sim I(1) \text{ without drift}) \\ H_1 &: \quad |\phi| < 1 \quad (Y_t \sim I(0) \text{ with non zero mean}) \end{aligned}$$

This formulation is appropriate for non-trending economic and financial series like interest rates, exchange rates, and spreads. The test statistics $t_{\widehat{\phi}}$ and $T(\widehat{\phi} - 1)$ are computed from the above regression. Under $H_0 : \phi = 1, c = 0$ the asymptotic distributions of these test statistics are influenced by the presence, but not the coefficient value, of the constant in the test regression.

The regression is:

$$Y_t = c + \delta t + \phi Y_{t-1} + \epsilon_t$$

and includes a constant and deterministic time trend to capture the deterministic trend under the alternative. The hypotheses to be tested are

$$H_0 \quad : \quad \delta = 0, \phi = 1 \quad (Y_t \sim I(1) \text{ with drift})$$
$$H_1 \quad : \quad |\phi| < 1 \quad (Y_t \sim I(0) \text{ with deterministic trend}))$$

This formulation is appropriate for trending time series like asset prices or the levels of macroeconomic aggregates like real GDP. Under H_0 the asymptotic distributions of these test statistics are influenced by the presence but not the coefficient values of the constant and time trend in the test regression.

- The unit root tests described above are valid if the time series yt is well characterized by an AR(1) with white noise errors.
- Many economic and financial time series have a more complicated dynamic structure than is captured by a simple AR(1) model.
- Said and Dickey (1984) augment the basic autoregressive unit root test to accommodate general ARMA(p,q) models with unknown orders and their test is referred to as the augmented Dickey- Fuller (ADF) test.

The ADF test tests the null hypothesis that $\{Y_t\}$ is $I(1)$ against the alternative that it is $I(0)$, assuming that the dynamics in the data have an ARMA structure.

Suppose that the data were really generated from an AR(p) process

$$\left(1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p\right) Y_t = \varepsilon_t$$

where $\epsilon_t \sim i.i.d. \left(0, \sigma^2\right)$, and finite fourth moment.
It is helpful to write the autoregression in a slightly different form.
Define,

$$\rho \equiv \phi_1 + \phi_2 + \ldots + \phi_p$$

$$\zeta_i \equiv - \left[\phi_{j+1} + \phi_{j+2} + \ldots + \phi_p\right] \quad j = 1, 2, \ldots, p-1$$

$$
\begin{aligned}
& (1-\rho L)-\left(\zeta_1 L+\zeta_2 L^2+\cdots+\zeta_{p-1} L^{p-1}\right)(1-L) \\
= \quad & 1-\rho L-\zeta_1 L-\zeta_2 L^2-\cdots-\zeta_{p-1} L^{p-1}+\zeta_1 L^2+\zeta_2 L^3+\cdots+\zeta_{p-1} L^p \\
= \quad & 1-\left(\rho+\zeta_1\right) L-\left(\zeta_2-\zeta_1\right) L^2-\cdots-\left(\zeta_{p-1}-\zeta_{p-2}\right) L^{p-1}-\left(-\zeta_{p-1} L^p\right) \\
= \quad & 1-\phi_1 L-\phi_2 L^2-\cdots-\phi_{p-1} L^{p-1}-\phi_p L^p
\end{aligned}
$$

The autoregression can be written

$$\left\{(1-\rho L)-\left(\zeta_1 L+\zeta_2 L^2+\cdots+\zeta_{p-1} L^{p-1}\right)(1-L)\right\} Y_t=\varepsilon_t$$

or

$$Y_t=\rho Y_{t-1}+\zeta_1 \Delta Y_{t-1}+\cdots+\zeta_{p-1} \Delta Y_{t-p+1}+\varepsilon_t$$

Suppose that the process that generated Y_t contains a single unit root; that is, suppose one root of

$$1-\phi_1 z-\phi_2 z^2-\ldots-\phi_p z^p=0$$

is unity

$$1-\phi_1-\phi_2-\ldots-\phi_p=0$$

and all other roots are outside the unit circle. This implies

$$\rho=1$$

When $\rho = 1$

$$1-\phi_1 z-\phi_2 z^2-\ldots-\phi_p z^p = \left(1 - \zeta_1 z - \zeta_2 z^2 - \cdots - \zeta_{p-1} z^{p-1}\right)(1-z)$$

All the roots of

$$1 - \zeta_1 z - \zeta_2 z^2 - \cdots - \zeta_{p-1} z^{p-1} = 0$$

lie outside the unit circle. Under the null hypothesis that $\rho = 1$

$$\left(1 - \zeta_1 L - \zeta_2 L^2 - \cdots - \zeta_{p-1} L^{p-1}\right) \Delta Y_t = \varepsilon_t$$

ADF - No Constant

$$Y_t = \rho Y_{t-1} + \zeta_1 \Delta Y_{t-1} + \cdots + \zeta_p \Delta Y_{t-p+1} + \varepsilon_t$$

true process has $\rho = 1$.

Test statistic for $H_0 : \rho = 1$:

$$t_{\widehat{\rho}} = \frac{\widehat{\rho}_T - 1}{SE(\widehat{\rho})}$$

$$Z_{ADF} = \frac{T\,(\widehat{\rho}_T - 1)}{1 - \widehat{\zeta}_{1,T} - \cdots - \widehat{\zeta}_{p-1,T}}$$

has the same asymptotic distribution as in the case of regression

$$Y_t = \rho Y_{t-1} + \epsilon_t$$

when the true process is a RW.

The estimated regression contains a constant:

$$Y_t = \alpha + \rho Y_{t-1} + \zeta_1 \Delta Y_{t-1} + \cdots + \zeta_p \Delta Y_{t-p+1} + \varepsilon_t$$

but the true process is an AR(p) with a unit root with no drift:

$$\alpha = 0, \ \rho = 1$$

Test statistic for $H_0 : \rho = 1$:

$$t_{\widehat{\rho}} = \frac{\widehat{\rho}_T - 1}{SE(\widehat{\rho})}$$

$$Z_{ADF} = \frac{T\left(\widehat{\rho}_T - 1\right)}{1 - \widehat{\zeta}_{1,T} - \cdots - \widehat{\zeta}_{p-1,T}} \xrightarrow{L} \frac{\frac{1}{2}\left\{\left[W\left(1\right)\right]^2 - 1\right\} - W\left(1\right)\int W\left(r\right)dr}{\int\left[W\left(r\right)\right]^2 dr - \left[\int W\left(r\right)dr\right]^2}$$

this is the same limiting distribution for the statistic in the case of the estimate of ρ in a regression without lagged ΔY_t and with serially uncorrelated disturbances.

F test statistic for $H_0 : \alpha = 0, \rho = 1$ has the same asymptotic distribution as the F test in the DF test regression:

$$Y_t = \alpha + \rho Y_{t-1} + \epsilon_t$$

Alternative formulation:

$$\Delta Y_t = \alpha + \pi Y_{t-1} + \zeta_1 \Delta Y_{t-1} + \cdots + \zeta_{p-1} \Delta Y_{t-p+1} + \varepsilon_t$$

where $\pi \equiv \rho - 1$.

$$H_0 : \alpha = 0, \pi = 0$$

The estimated regression contains a constant:

$$Y_t = \alpha + \rho Y_{t-1} + \zeta_1 \Delta Y_{t-1} + \cdots + \zeta_p \Delta Y_{t-p+1} + \varepsilon_t$$

the true process is the same specification as estimated regression, $\alpha \neq 0, \rho = 1$.
$\widehat{\rho}_T$ converges at the rate $T^{3/2}$ to a Gaussian variable; the other estimated coefficients converge at the rate of $\sqrt{T}$.
Any t or F test involving any coefficients from the regressions can be compared with the usual t or F tables.

Estimated regression:

$$Y_t = \alpha + \delta t + \rho Y_{t-1} + \zeta_1 \Delta Y_{t-1} + \cdots + \zeta_p \Delta Y_{t-p+1} + \varepsilon_t$$

True process: α any value, $\rho = 1$ and $\delta = 0$.
$t_{\widehat{\rho}}$ and Z_{ADF} has the same asymptotic distribution as in the case of the DF regression:

$$Y_t = \alpha + \delta t + \rho Y_{t-1} + \epsilon_t$$

where the true process is $\alpha \neq 0, \delta = 0, \rho = 1$.

The following model was estimated by OLS for the interest rate data:

$$i_t = 0.195 + 0.96904 i_{t-1} + 0.355 \Delta i_{t-1} - 0.388 \Delta i_{t-2} + 0.276 \Delta i_{t-3} - 0.107 \Delta i_{t-4}$$

$T = 164$. For these estimates the augmented Dickey-Fuller ρ test

$$\frac{164}{1 - 0.355 + 0.388 - 0.276 + 0.107} (0.96904 - 1) = -5.74$$

This a test of $\rho = 1$ under the maintained hypothesis of $\alpha = 0$. In a sample of this size, the 95% of the time when there is really a unit root, the statistic $T\left(\widehat{\rho}_T - 1\right)$ will be above -13.8.
Since $-5.74 > -13.8$, the null hypothesis that the Treasury bill rate follows a fifth-order autoregression with no constant term, and a single unit root is accepted at the 5% level.

The OLS t test for this same hypothesis is

$$(0.96904 - 1) \,/\, (0.018604) = -1.66$$

Since $-1.66 > -2.89$, the null hypothesis of a unit root is accepted by the augmented Dickey-Fuller t test as well.

Phillips and Perron (1988) and Perron (1988) suggested non-parametric test statistics for the null hypothesis of a unit root that explicitly allows for weak dependence and heterogeneity of the error process.

A shortcoming of the PP tests is their low power if the true data-generating process is an AR(1)-process with a coefficient close to one.

- ADF and PP test are asymptotically equivalent.
- PP has better small sample properties than ADF.
- Both have low power against $I(0)$ alternatives that are close to being $I(1)$ processes.
- Power of the tests diminishes as deterministic terms are added to the test regression.

To improve the power of the unit root test, Elliott, Rothenberg and Stock (1996) proposed a local to unity detrending of the time series.

A drawback of the DF-type tests is that the nuisance parameters (i.e., the coefficients of the deterministic regressors) are either not defined or have a different interpretation under the alternative hypothesis of stationarity.

In the regression

$$\Delta y_t = \alpha + \delta t + \pi y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t$$

α represents a linear trend and δ represents a quadratic trend under the null hypothesis of integration, but these coefficients have the interpretation of a level and linear trend regressor under the alternative hypothesis of stationarity.

Solution: LM-type test, that has the same set of nuisance parameters under both the null and alternative hypothesis.

Higher polynomials than a linear trend are allowed.

The model

$$\begin{aligned} y_t &= \alpha + Z_t'\delta + x_t \\ x_t &= \pi x_{t-1} + \epsilon_t \end{aligned}$$

$\epsilon_t \sim i.i.d.N(0, \sigma^2), Z_t = (t, \ldots, t^p)$

1. Run the OLS regression

$$\Delta y_t = \delta' \Delta Z_t + u_t$$

and obtain $\widetilde{\delta}$.

2. Compute $\widetilde{\psi}_x = y_1 - Z_1'\widetilde{\delta}$

3. Calculate the series $\widetilde{S}_t = y_t - \widetilde{\psi}_x - Z_t'\widetilde{\delta}$
4. the test regression is given by

$$\Delta y_t = \gamma' \Delta Z_t + \phi S_{t-1} + v_t$$

5. the test statistics for the null $\phi = 0$ are defined by

$$Z = \frac{T\widehat{\phi}}{\widehat{\omega}^2}$$

where

$$\widehat{\omega}^2 = \frac{T^{-1}\sum_t \widehat{\epsilon}_t^2}{T^{-1}\sum_t \widehat{\epsilon}_t^2 + 2T^{-1}\sum_{s=1}^{l}\sum_{t=s+1}^{T} \widehat{\epsilon}_t \widehat{\epsilon}_{t-s}}$$

where $\widehat{\epsilon}_t$ are the residuals from

$$x_t = \pi x_{t-1} + \epsilon_t$$

Kwiatkowski, Phillips, Schmidt and Shin [1992] proposed an LM test for testing trend and/or level stationarity (the KPSS test). That is, now the null hypothesis is a stationary process.

Taking the null hypothesis as a stationary process and the unit root as an alternative is in accordance with a conservative testing strategy.

Hence, if we then reject the null hypothesis, we can be pretty confident that the series indeed has a unit root. Therefore, if the results of the tests above indicate a unit root but the result of the KPSS test indicates a stationary process, one should be cautious and opt for the latter result.

KPSS consider the following model:

$$\begin{aligned} y_t &= \xi t + r_t + \epsilon_t \\ r_t &= r_{t-1} + u_t \end{aligned}$$

where r_t is a random walk and $u_t \sim (0, \sigma_u^2)$.

Null hypothesis:

$$H_0 : \sigma_u^2 = 0$$

or r_t is a constant.
Under the hypothesis of $u_t \sim i.i.d.(0, \sigma_u^2)$, the test statistic:

$$LM = \frac{\sum_{t=1}^{T} S_t^2}{\widehat{\sigma}_e^2}$$

$\widehat{\sigma}_e^2 = \sum_t e_t^2/T$ and S_t is the partial sum of e_t defined by

$$S_t = \sum_{i=1}^{t} e_i \quad t = 1, 2, \ldots, T$$

where e_t are the residuals from the regression of y_t on a constant and a time trend

$$e_t = y_t - \widehat{c} - \widehat{\beta} t$$

For testing the null of level stationarity, the test is constructed the same way except that e_t is obtained as the residual from a regression of y_t on an intercept only.

KPSS consider the case of a general error process. When the errors are i.i.d. $\widehat{\sigma}_e^2 \xrightarrow{p} \sigma^2$.
When the errors are not i.i.d. the appropriate denominator of the test statistic is an estimate of the *long-run variance* σ^2 not σ_e^2:

$$\sigma^2 = \lim_{T\to\infty} T^{-1} E(S_T^2)$$

A consistent estimator of σ^2 is given by:

$$\widehat{\sigma}_{Tl}^2 = \frac{1}{T}\sum_t e_t^2 + \frac{2}{T}\sum_s^l w_{sl} \sum_{t=s+1}^T e_t e_{t-s}$$

w_{sl} is an optimal weighting function that corresponds to the choice of a spectral window.

KPSS use the Bartlett window (Newey and West (1987)):

$$w_{sl} = 1 - \frac{\tau}{l+1}$$

for consistency of $\widehat{\sigma}^2_{Tl}$ it is necessary that $l \to \infty$ as $T \to \infty$.

Test statistic:

$$LM = \frac{\sum_{t=1}^{T} S_t^2}{\widehat{\sigma}^2_{Tl}}$$

KPSS derive the asymptotic distribution of the modified statistic and tabulate the critical values by simulation.