Lecture 4. Unit Root Tests for Stationarity

Narek Ohanyan

American University of Armenia

February 20, 2024

Outline

Trend-stationary processes

The time series y_t is called **trend stationary** if it is stationary around a deterministic trend.

Consider the following time series model

$$\begin{aligned} y_t &= \alpha + \beta t + x_t \\ x_t &= \phi x_{t-1} + e_t \qquad e_t \sim \mathrm{WN}\left(0, \sigma^2\right) \end{aligned}$$

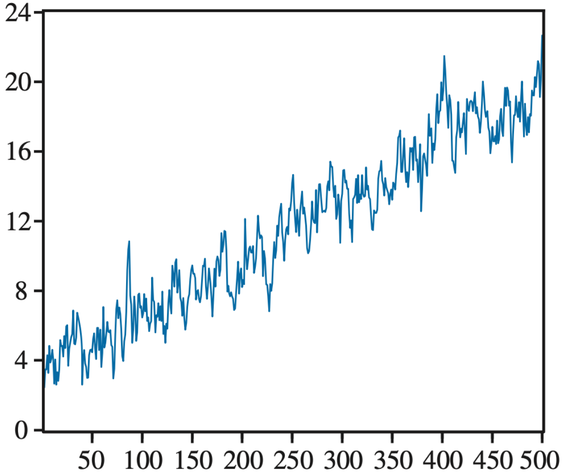
In other words, y_t is trend stationary if after removing the trend $\alpha + \beta t$, the resulting series x_t is stationary.

We can also write the model as

$$y_t = \gamma_0 + \phi y_{t-1} + \gamma_1 t + e_t$$

where $\gamma_0 = \alpha\left(1 - \phi\right) + \phi\beta$ and $\gamma_1 = \beta\left(1 - \phi\right)$.

An example of a trend-stationary process



(c) $y_t = 1 + 0.01t + 0.7y_{t-1} + v_t$

Random-walk processes

The time series y_t is called a **random walk** if it has the following form

$$y_t = y_{t-1} + e_t \qquad e_t \sim \text{WN}\left(0, \sigma^2\right)$$

The random walk process may be written as

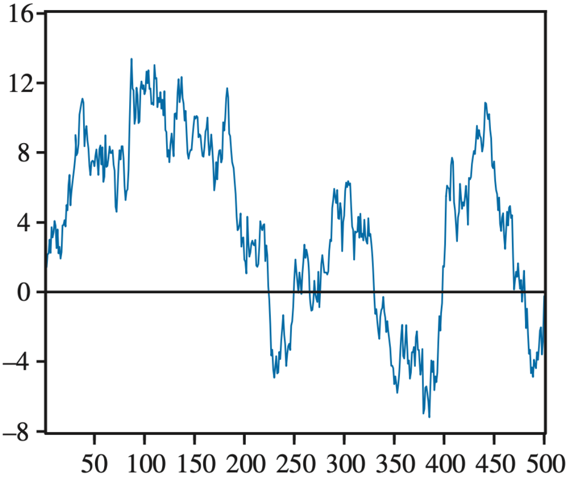
$$y_t = y_0 + \sum_{i=1}^{t} e_i$$

The term $\sum_{i=1}^{t} e_i$ is called the **stochastic trend**.

Random walk process has the following properties

$$\mathrm{E}\left(y_t\right) = y_0 \qquad \mathrm{Var}\left(y_t\right) = t\sigma^2 \qquad \mathrm{Cov}\left(y_t, y_{t-k}\right) = \left(t-k\right)\sigma^2$$

An example of a random walk process



(d) $y_t = y_{t-1} + v_t$

Random-walk with drift

The time series y_t is called a **random walk with drift** if it has the following form

$$y_t = \alpha + y_{t-1} + e_t \qquad e_t \sim \mathrm{WN}\left(0, \sigma^2\right)$$

The random walk process may be written as

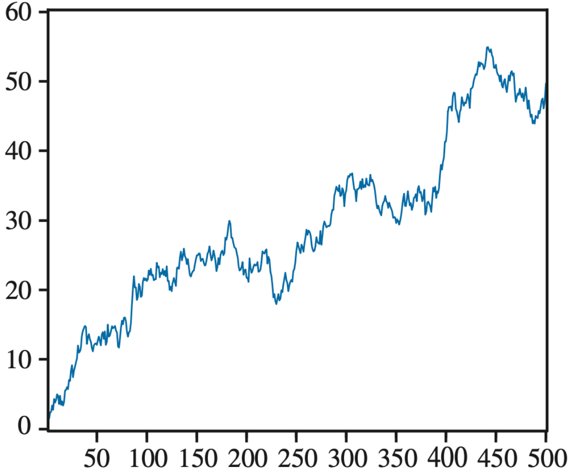
$$y_t = y_0 + \alpha t + \sum_{i=1}^{t} e_i$$

The term αt is called the **deterministic trend**.

Random walk process has the following properties

$$\mathrm{E}\left(y_t\right) = y_0 + \alpha t \qquad \mathrm{Var}\left(y_t\right) = t\sigma^2 \qquad \mathrm{Cov}\left(y_t, y_{t-k}\right) = \left(t-k\right)\sigma^2$$

An example of a random walk with drift process



(e) $y_t = 0.1 + y_{t-1} + v_t$

Outline

Spurious regressions

Regressions of non-stationary time series may produce **spurious** results.

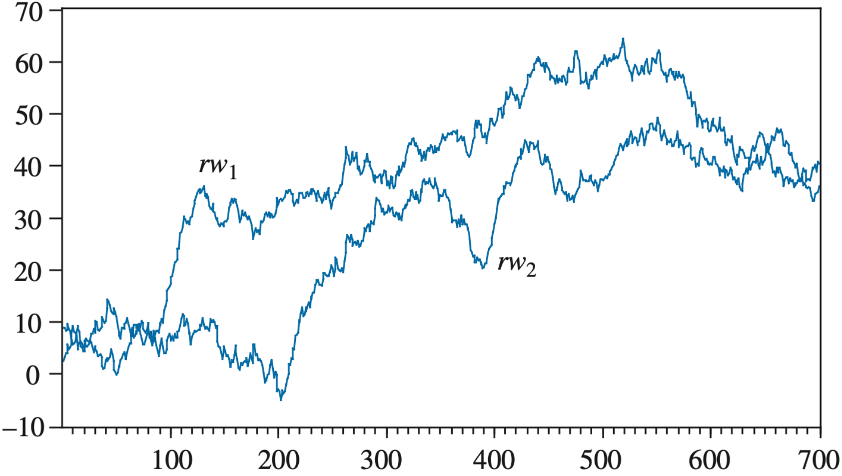
Consider two independent random walks

$$
\begin{aligned}
y_t &= y_{t-1} + e_t \qquad & e_t \sim \mathrm{WN}\left(0, \sigma^2\right) \\
x_t &= x_{t-1} + u_t \qquad & u_t \sim \mathrm{WN}\left(0, \sigma^2\right)
\end{aligned}
$$

The regression of y_t on x_t will likely produce a high R^2 and significant t-statistics.

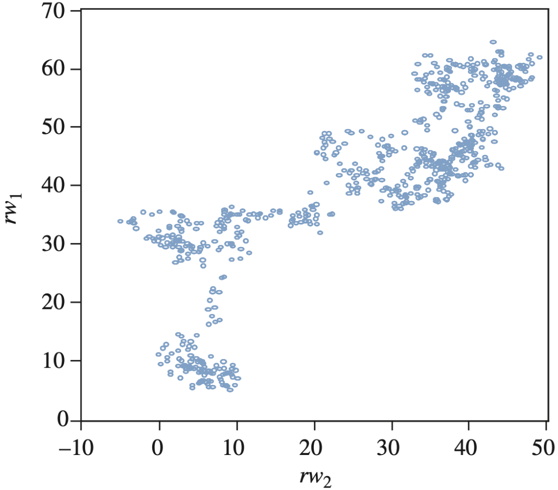
The problem is that both y_t and x_t are non-stationary, and the regression is spurious.

Example of two independent random walks



(a) Time series

Scatter plot of two independent random walks



(*b*) Scatter plot

Outline

Unit root tests

Recall that an AR(1) process is given by

$$y_t = \phi y_{t-1} + e_t \qquad e_t \sim \mathrm{WN}\left(0, \sigma^2\right)$$

When $\phi = 1$, the process is called a **unit root** process.

The most popular test for unit root is the **Dickey-Fuller test** for the hypothesis

$$\mathrm{H}_0 : \phi = 1 \qquad \text{vs.} \qquad \mathrm{H}_1 : \phi < 1$$

Other tests include the **Phillips-Perron test**, the **Kwiatkowski-Phillips-Schmidt-Shin test**, and the **Elliott-Rothenberg-Stock test**.

Dickey-Fuller test

The Dickey-Fuller test is based on the following alternative models

- No constant, no trend

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

- Constant, no trend

$$\Delta y_t = \alpha + \gamma y_{t-1} + e_t$$

- Constant and trend

$$\Delta y_t = \alpha + \gamma y_{t-1} + \lambda t + e_t$$

The hypothesis of the tests is

$$\begin{aligned} \mathrm{H}_0 : \gamma = 0 \qquad &\Longleftrightarrow \qquad \mathrm{H}_0 : \phi = 1 \\ \mathrm{H}_1 : \gamma < 0 \qquad &\Longleftrightarrow \qquad \mathrm{H}_1 : \phi < 1 \end{aligned}$$

AR processes and the Dickey-Fuller tests

AR processes: $\lvert\rho\rvert < 1$	Setting $\rho = 1$	Dickey Fuller Tests
$y_t = \rho y_{t-1} + u_t$	$y_t = y_{t-1} + u_t$	Test with no constant and no trend
$y_t = \alpha + \rho y_{t-1} + v_t$ $\alpha = \mu(1-\rho)$	$y_t = y_{t-1} + v_t$ $\alpha = 0$	Test with constant and no trend
$y_t = \alpha + \rho y_{t-1} + \lambda t + v_t$ $\alpha = (\mu(1-\rho) + \rho\delta)$ $\lambda = \delta(1-\rho)$	$y_t = \delta + y_{t-1} + v_t$ $\alpha = \delta$ $\lambda = 0$	Test with constant and trend

Distribution of the test statistic

The Dickey-Fuller test equation is estimated by OLS, and the test statistic is

$$\tau = \frac{\hat{\gamma}}{\mathrm{se}(\hat{\gamma})}$$

where $\hat{\gamma}$ is the OLS estimate of γ, and $\mathrm{se}(\hat{\gamma})$ is the standard error of the estimate.

Under the null hypothesis, the test statistic has a non-standard distribution.

The null hypothesis is rejected if $\tau < \tau_c$, where τ_c depends on the type of the model used.

Critical values of the Dickey-Fuller test

Model	1%	5%	10%
$\Delta y_t = \gamma y_{t-1} + v_t$	-2.56	-1.94	-1.62
$\Delta y_t = \alpha + \gamma y_{t-1} + v_t$	-3.43	-2.86	-2.57
$\Delta y_t = \alpha + \lambda t + \gamma y_{t-1} + v_t$	-3.96	-3.41	-3.13
Standard critical values	-2.33	-1.65	-1.28

Note: These critical values are taken from R. Davidson and J. G. MacKinnon (1993), *Estimation and Inference in Econometrics*, New York: Oxford University Press, p. 708.

Augmented Dickey-Fuller test

The **Augmented Dickey-Fuller (ADF)** test is a generalization of the Dickey-Fuller test.

The ADF test includes additional lags of the dependent variable in the regression equation

$$\Delta y_t = \alpha + \gamma y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta y_{t-i} + e_t$$

Inclusion of additional lags aims to eliminate the *serial correlation* in the error term.

The number of lags p is chosen based on the AIC or BIC criteria.

Outline

Integrated processes

Recall that if y_t follows a random walk, then Δy_t is

$$\Delta y_t = y_t - y_{t-1} = e_t$$

which is stationary.

A time series y_t is called **integrated** of order one, denoted as $I(1)$, if its first difference is stationary.

In general, a time series y_t is called *integrated* of order d, denoted as $I(d)$, if its d-th difference is stationary

$$\Delta^d y_t = \Delta \left(\Delta^{d-1} y_t \right)$$

The order of integration d is the number of times the series must be differenced to obtain a stationary series.

ARIMA models

The **Autoregressive Integrated Moving Average (ARIMA)** model is a generalization of the ARMA model.

The ARIMA(p, d, q) model is given by

$$x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + e_t + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q} \qquad e_t \sim \text{WN}\left(0, \sigma^2\right)$$

where $x_t = \Delta^d y_t$.

ARIMA models are estimated by **Maximum Likelihood Estimation (MLE)**.

Prior to estimation, d must be determined by the ADF test, and the p and q are chosen based on the AIC or BIC criteria.