ECON/FIN 250: Forecasting in Finance and Economics: Section 7: Unit Roots & Dickey-Fuller Tests

Patrick Herb
Brandeis University
Spring 2016
Course Overview

1. Key Objectives
2. ARIMA Processes
3. Testing for Unit Roots
1. Key Objectives

2. ARIMA Processes

3. Testing for Unit Roots
Key Objectives

- ARIMA Processes
- Dickey-Fuller Tests
Stochastic Trends

Random Walk

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \] (1)
Random Walk

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hspace{1cm} (1)

Random Walk with Drift

\[ y_t = \delta + y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hspace{1cm} (2)
Stochastic Trends

Geometric Random Walk

\[ \log(y_t) = \log(y_{t-1}) + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \] (3)

Better for series where percentage changes are constant magnitude over time: stock prices, GDP...

Patrick Herb (Brandeis University)
Stochastic Trends

Geometric Random Walk

\[ \log(y_t) = \log(y_{t-1}) + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  (3)

Geometric Random Walk with Drift

\[ \log(y_t) = \delta + \log(y_{t-1}) + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  (4)

Better for series where percentage changes are constant magnitude over time: stock prices, GDP...
Basic Random Walk Properties

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hspace{1cm} (5)

Starting at \( y_0 \)

\[ y_t = y_0 + \sum_{i=1}^{t} \epsilon_i \]  \hspace{1cm} (6)
Basic Random Walk Properties

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \quad (5) \]

Starting at \( y_0 \)

\[ y_t = y_0 + \sum_{i=1}^{t} \epsilon_i \quad (6) \]

\[ E[y_t] = y_0 \quad (7) \]
Basic Random Walk Properties

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \] (5)

Starting at \( y_0 \)

\[ y_t = y_0 + \sum_{i=1}^{t} \epsilon_i \] (6)

\[ E[y_t] = y_0 \] (7)

\[ \text{Var}[y_t] = \sum_{i=1}^{T} E[\epsilon_i^2] = t\sigma^2 \] (8)

\[ \lim_{t \to \infty} \text{Var}[y_t] = \infty \] (9)
Basic Random Walk Properties

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \tag{5} \]

Starting at \( y_0 \)

\[ y_t = y_0 + \sum_{i=1}^{t} \epsilon_i \tag{6} \]

\[ E[y_t] = y_0 \tag{7} \]

\[ \text{Var}[y_t] = \sum_{i=1}^{T} E[\epsilon_i^2] = t\sigma^2 \tag{8} \]

\[ \lim_{t \to \infty} \text{Var}[y_t] = \infty \tag{9} \]
$y_t = \delta + y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$ \hspace{1cm} (10)

Starting at $y_0$

$y_t = y_0 + t\delta + \sum_{i=1}^{T} \epsilon_i$ \hspace{1cm} (11)
Basic Random Walk with Drift Properties

\[ y_t = \delta + y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \] \hspace{1cm} (10)

Starting at \( y_0 \)

\[ y_t = y_0 + t\delta + \sum_{i=1}^{T} \epsilon_i \] \hspace{1cm} (11)

\[ E[y_t] = y_0 + t\delta \] \hspace{1cm} (12)
Basic Random Walk with Drift Properties

\[ y_t = \delta + y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hspace{1cm} (10)

Starting at \( y_0 \)

\[ y_t = y_0 + t\delta + \sum_{i=1}^{T} \epsilon_i \]  \hspace{1cm} (11)

\[ E[y_t] = y_0 + t\delta \]  \hspace{1cm} (12)

\[ \text{Var}[y_t] = \sum_{i=1}^{T} E[\epsilon_i^2] = t\sigma^2 \]  \hspace{1cm} (13)
Basic Random Walk with Drift Properties

\[ y_t = \delta + y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hspace{1cm} (10)

Starting at \( y_0 \)

\[ y_t = y_0 + t\delta + \sum_{i=1}^{T} \epsilon_i \]  \hspace{1cm} (11)

\[ E[y_t] = y_0 + t\delta \]  \hspace{1cm} (12)

\[ Var[y_t] = \sum_{i=1}^{T} E[\epsilon_i^2] = t\sigma^2 \]  \hspace{1cm} (13)

\[ \lim_{t \to \infty} Var[y_t] = \infty \]  \hspace{1cm} (14)
Random Walk Forecasts

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hfill (15)

Forecast \( h \) steps ahead starting at time \( T \)

\[ y_{T+h} = y_T + \sum_{i=1}^{h} \epsilon_{T+i} \]  \hfill (16)
Random Walk Forecasts

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hspace{1cm} (15)

Forecast \( h \) steps ahead starting at time \( T \)

\[ y_{T+h} = y_T + \sum_{i=1}^{h} \epsilon_{T+i} \]  \hspace{1cm} (16)

\[ E[y_{T+h} | \Omega_T] = y_T, \quad \forall h \]  \hspace{1cm} (17)
Random Walk Forecasts

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \tag{15}

Forecast \( h \) steps ahead starting at time \( T \)

\[ y_{T+h} = y_T + \sum_{i=1}^{h} \epsilon_{T+i} \]  \tag{16}

\[ E[y_{T+h}|\Omega_T] = y_T, \quad \forall h \]  \tag{17}

\[ Var[y_{T+h}|\Omega] = E[(y_{T+h} - E[y_{T+h}|\Omega_T])^2] \]  \tag{18}
Random Walk Forecasts

\[ y_t = y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2) \]  \hspace{1cm} (15)

Forecast \( h \) steps ahead starting at time \( T \)

\[ y_{T+h} = y_T + \sum_{i=1}^{h} \epsilon_{T+i} \]  \hspace{1cm} (16)

\[ E[y_{T+h}|\Omega_T] = y_T, \quad \forall h \]  \hspace{1cm} (17)

\[ \text{Var}[y_{T+h}|\Omega] = E[(y_{T+h} - E[y_{T+h}|\Omega_T])^2] \]  \hspace{1cm} (18)

\[ \text{Var}[y_{T+h}|\Omega_T] = E[(y_{T+h} - y_T)^2] = \sum_{i=1}^{h} \epsilon_{T+i}^2 = h\sigma^2_\epsilon \]  \hspace{1cm} (19)
Random Walk Forecast Intuitive Properties

- Best forecast is current value
- Longer forecasts don’t converge to the mean
- Forecast variance expand linearly in $h$ as $h\sigma_\epsilon^2$
- Forecast std’s expand as $\sqrt{h}$ or $\sqrt{h\sigma_\epsilon}$
ARIMA Processes

1. Key Objectives

2. ARIMA Processes

3. Testing for Unit Roots
Integrated Processes ARIMA($p,d,q$)

- An AR($p$) process is a unit root process if one of the roots of the lag operator polynomial is equal to one.
- Unit roots result in nonstationary behavior.
- Differencing a random walk process “integrates” or “undoes” the unit root.
Integrated Processes ARIMA(p,d,q)

- An AR(p) process is a unit root process if one of the roots of the lag operator polynomial is equal to one.
- Unit roots result in nonstationary behavior.
- Differencing a random walk process “integrates” or “undoes” the unit root.

\[
y_t - y_{t-1} = \epsilon_t \\
z_t = \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)
\]  

(20)  

(21)

The difference is a white noise process.
More Examples

Consider an ARIMA(1,1,0)

\[ y_t - y_{t-1} = \phi(y_{t-1} - y_{t-2}) + \epsilon_t \]  \hspace{1cm} (22)

\[ z_t = \phi z_{t-1} + \epsilon_t \]  \hspace{1cm} (23)

Difference is AR(1)
Consider an ARIMA(1,1,1)

\[ y_t - y_{t-1} = \phi(y_{t-1} - y_{t-2}) + \theta\epsilon_{t-1} + \epsilon_t \]  \hspace{1cm} (24)

\[ z_t = \phi z_{t-1} + \theta \epsilon_{t-1} + \epsilon_t \]  \hspace{1cm} (25)

Difference is ARMA(1,1)
Differencing vs. Detrending

Differencing a Stochastic Trend Model

\[ y_t = a_0 + y_{t-1} + \epsilon_t \]  \hspace{1cm} (26)

\[ y_t = y_0 + a_0 t + \sum_{i=1}^{t} \epsilon_i \]  \hspace{1cm} (27)

\[ y_{t-1} = y_0 + a_0(t-1) + \sum_{i=1}^{t-1} \epsilon_i \]  \hspace{1cm} (28)

\[ \Delta y_t = a_0 + \epsilon_t \]  \hspace{1cm} (29)

- The difference is stationary
- A series with a unit root can be transformed into a stationary series by differencing
Differencing vs. Detrending

Differencing a Deterministic Trend Model

\[ y_t = y_0 + a_1 t + \epsilon_t \]  \hspace{1cm} (30)
\[ y_{t-1} = y_0 + a_1 (t - 1) + \epsilon_{t-1} \]  \hspace{1cm} (31)
\[ \Delta y_t = a_1 + \epsilon_t + \epsilon_{t-1} \]  \hspace{1cm} (32)

- The difference is not stationary (not invertible)
- Need to detrend data with deterministic time trend, not (necessarily) difference
- A trend-stationary series can be transformed into a stationary series by removing the deterministic trend
ARIMA Notation and Methods

ARIMA\((p,d,q)\)
- \(p\) = AR
- \(q\) = MA
- \(d\) = differencing level

What is differencing level?
- \(d = 1\), \(y_t - y_{t-1}\) is ARMA\((p,q)\)
- \(d = 2\), \((y_t - y_{t-1}) - (y_{t-1} - y_{t-2})\) is ARMA\((p,q)\)
- \(d = 3\), keep going. Higher order differencing is rare in economics data
What do you do?

- If you know $d$
  - Difference $y_t$ $d$ times
  - Estimate ARMA components
  - Generate forecasts of $\hat{z}_t$
  - Add back together $\hat{y}_t = \hat{y}_{t-1} + \hat{z}_t$

- Problem: You often don’t know $d$

- In economics data, the question is often between $d = 1$ or $d = 0$
In Lag Notation

**ARMA(p,q)**

\[ \Phi(L)y_t = c + \Theta(L)\epsilon_t \]  

(33)

\[ \Phi(L) = 1 - \phi_1 L - \phi_2 L^2 + \ldots \phi^p L^p \]  

(34)

\[ \Theta(L) = 1 - \theta_1 L - \theta_2 L^2 + \ldots \theta^q L^q \]  

(35)

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

\[ \Phi(L)(1 - L)y_t = c + \Theta(L)\epsilon_t \]  

(36)

**ARIMA(p,d,q)**

\[ \Phi(L)(1 - L)^d y_t = c + \Theta(L)\epsilon_t \]  

(37)
Testing for Unit Roots

1. Key Objectives

2. ARIMA Processes

3. Testing for Unit Roots
Testing $d = 0$ or $d = 1$

- What about regressing $y_t = \phi y_{t-1} + \epsilon_t$?
- Test for $\phi = 1$
- Many people did this a while ago
- It turns out the distribution of $\phi$ is not t-distribution
- Proper tests have Dickey-Fuller-distribution
Dickey-Fuller Tests

\[ y_t = \phi y_{t-1} + \epsilon_t \quad (38) \]
\[ y_t - y_{t-1} = (\phi - 1)y_{t-1} + \epsilon_t \quad (39) \]
\[ \Delta y_t = \gamma y_{t-1} + \epsilon_t \quad (40) \]

- Regress \( y_t - y_{t-1} \) on \( y_{t-1} \)
- Testing is \( \gamma = 0 \) is equivalent to testing \( \phi = 1 \)
- The null hypothesis is \( \gamma = 0, \phi = 1, \) or \( \{y_t\} \) has a unit root
- The alternative is \( (\phi - 1) < 0 \)
- Failing to reject \( \rightarrow \) unit root
- Rejecting the null \( \rightarrow \) no unit root
- Alternative would be \( \phi < 1 \) \( \rightarrow \) stationary AR(1)
- Alternatives are important
Random Walk + Drift

$$\Delta y_t = a_0 + \gamma y_{t-1} + \epsilon_t$$  \hspace{1cm} (41)

Random Walk + Drift + Deterministic Time Trend

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \epsilon_t$$  \hspace{1cm} (42)

Test if $\gamma = 0$ to determine if process has unit root
Augmented Dickey-Fuller

Null: Random Walk + AR(p); Alternative: AR(p) + No Mean

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t$$ (43)

Null: Random Walk + Drift + AR(p); Alternative: AR(p) + Mean

$$\Delta y_t = a_0 + \gamma y_{t-1} + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t$$ (44)

Null: Random Walk + Drift + Deterministic Time Trend + AR(p); Alternative: AR(p) + Trend

$$\Delta y_t = a_0 + \gamma y_{t-1} + a_2 t + \sum_{i=2}^{p} \beta_i \Delta y_{t-i+1} + \epsilon_t$$ (45)
Dickey-Fuller Instructions

- Don’t worry about critical values and tests, they are performed by all good software.
- Do worry about interpretation and alternatives. No software does this.
- Basic Steps
  - Is your data trending over time? If yes, use form 3
  - If no obvious trend, then use form 1 or 2 depending on mean
Stata Code: Forms 1-3

1. `dfuller lgdp, noconstant reg lags(k)`
2. `dfuller lgdp, reg lags(k)`
3. `dfuller lgdp, trend reg lags(k)`

- The “reg” option prints out regression coefficients
- Skipping the “drift” option
- Try this with U.S. GDP data
Size and Power: Important

- Unit root testing can be difficult
- The Power of the Dickey-Fuller test can be low
  - Higher chance of Type II error
  - Failing to reject a false null hypothesis
  - Failing to reject a unit root when there is no unit root
- Often accept random walk null when time trend might be true model
- Think about your data