Trend Stationary and Unit Root processes

(Hamilton: Chapters 16-17)

Isabel Casas

Syddansk Universitet

- Estimation of trend-stationary processes
- Estimation of unit root processes
- Dickey–Fuller test
- Phillips–Perron unit root test
- Augmented Dickey–Fuller test

SYDDANSK UNIVERSITET

Estimation of trend-stationary processes

If the process looks like:

$$y_t = \alpha + \delta t + \epsilon_t \qquad \epsilon_t \sim IID(0, \sigma^2)$$

- We obtain $\hat{\alpha}, \hat{\delta}$ using OLS.
- If the ϵ_t is Gaussian, then we can use the usual t-test and F-test for inference.
- If ϵ_t is non-Gaussian, we have to use different t-test and F-test because although the estimators and consistent and asymptotically normal, they do not follow the usual theory.
- In fact, in this case the convergence rate of $\hat{\alpha}$ and $\hat{\delta}$ are different.

Estimation of trend-stationary processes

If the process looks like:

$$y_t = \alpha + \delta t + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} + \epsilon_t \qquad \epsilon_t \sim IID(0, \sigma^2) \quad (1)$$

with roots of $1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z = 0$ outside the unit circle.

- We can estimate $\alpha, \delta, \phi_1, \ldots, \phi_p$ using OLS
- The estimators are consistent and asymptotically normal with convergence rate $\sqrt{T}$

Detrending US GNP

We use the log of the US GNP. $y_t = log(GNP)$

Detrending US GNP

Assume that the process is of the type $y_t = \alpha + \delta t + \epsilon_t$

```
> t = 1:length(y)
> m <- lm(y ~ 1 + t)
> c <- round(coef(m), 3)
> c
```

```
(Intercept)           t
      5.405       0.018
```

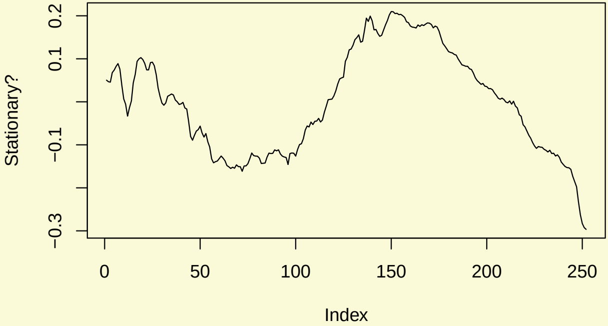
```
> y.detrend <- resid(m)
```

$$y_t = 5.405 + 0.018t + \epsilon_t$$

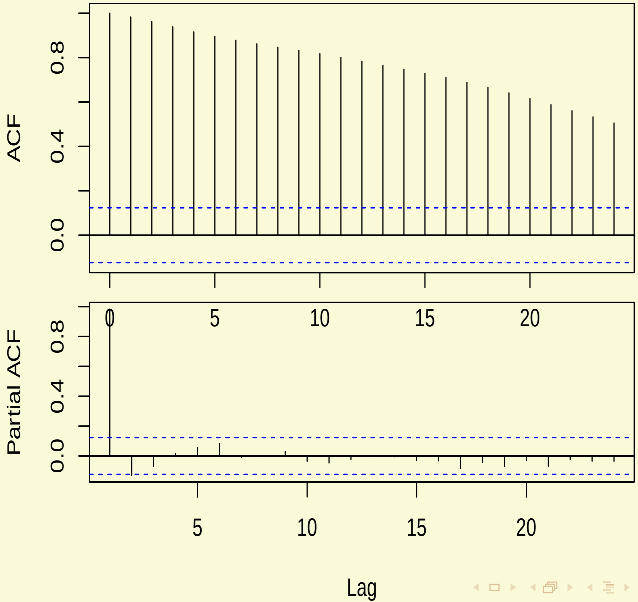
Detrending US GNP

Plot the detrended series

```
> plot(y.detrend, type = "l", ylab = "Stationary?")
```



Detrending US GNP



Estimation of unit roots processes

- If $y_t = \phi y_{t-1} + \epsilon_t \quad \epsilon_t \sim IID(0, \sigma^2)$, then the OLS estimate

$$\hat{\phi} = \frac{\sum_{t=2}^{T} y_{t-1} y_t}{\sum_{i=2}^{T} y_{t-1}^2}$$

  is asymptotically normal:

$$\sqrt{T}(\hat{\phi} - \phi) \to^d N(0, 1 - \phi^2)$$

- if y_t is an unit root process ($\phi = 1$), Q: Do you see any problem?

SYDDANSK UNIVERSITET

Estimation of unit roots processes

- If we have a unit root process, we have to be more careful
- We cannot use the the "usual" asymptotics
- This generalises to all I(1) processes.
- Therefore, it is important that we test for unit roots.
- It can be proven that $\hat{\phi}$ is super-consistent: $\hat{\phi} \to \phi$ at rate T rather than $\sqrt{T}$
- $\hat{\phi}$ is not asymptotically normal. Therefore, the t-test is different.
- The test statistics of a t-test follows a Dickey–Fuller (DF) distribution and it does not have a close form. So quantiles of the distribution must be computed numerically

Example: GNP

Assume that the process is of the type $y_t = y_{t-1} + \phi y_{t-1} + \epsilon_t$ for $y_t = log(GNP)$

```
> uroot <- arima(y, order = c(1, 1, 0))
> coef(uroot)

      ar1
0.8309014
```

```
> z = diff(y)
> uroot2 = arima(z, order = c(1, 0, 0), include.mean = F)
> coef(uroot2)

      ar1
0.8309006
```

Unit root test

- Conceptually the unit root test is a t-test on $H_0 : \phi = 1$
- Basically it tests the null hypothesis that a time series is I(1) against the time series is I(0)
- In practice, there are a few issues we have to take into account
  - Unit root tests generally have nonstandard and non-normal asymptotic distributions
  - The distributions are functionals of the Brownian motion and do not have consistent closed forms. So the critical value has to be computed numerically
  - These distributions are affected by deterministic trends (constant, time trend, dummy variables)

DF test - Case 1

Case 1: No constant or time trend.

True process	Estimated process	Innovations
$y_t = y_{t-1} + \epsilon_t$	$y_t = \phi y_{t-1} + \epsilon_t$	$\epsilon_t \sim IIDN(0, \sigma^2)$

- The hypothesis

$$
\begin{aligned}
H_0 :& \phi = 1 \quad (\phi(z) = 0 \text{ has a unit root}) \\
H_1 :& \phi < 1 \quad (\phi(z) = 0 \text{ has roots larger than unity})
\end{aligned}
$$

- In practice, the DF test is perfromed by reparameterising:

$$
\begin{aligned}
y_t - y_{t-1} =& (\phi - 1) y_{t-1} + \epsilon_t \\
\Delta y_t =& \pi y_{t-1} + \epsilon_t
\end{aligned}
$$

- Q: What are the hypothesis then?

DF test - Case 1

- The DF test statistics (it doesn't need the standard error):

$$DF = T\hat{\pi} \sim \text{ Table B.5 - distribution}$$

- The OLS test statistics:

$$t_{\pi=0} = \frac{\hat{\pi}}{se(\hat{\pi})} \sim \text{ Table B.6 - distribution}$$

Both statistics can be used.

DF test - Case 2

Case 2: Constant but not time trend.

True process	Estimated process	Innovations
$\Delta y_t = \epsilon_t$	$\Delta y_t = \alpha + \pi y_{t-1} + \epsilon_t$	$\epsilon_t \sim IIDN(0, \sigma^2)$

- Tests Statistics:

$$DF = T\hat{\pi} \sim \text{ Table B.5 - distribution}$$

$$t_{\pi=0} = \frac{\hat{\pi}}{se(\hat{\pi})} \sim \text{ Table B.6 - distribution}$$

DF test - Case 3

Case 3: Constant but no time trend.

True process	Estimated process	Innovations
$\Delta y_t = \alpha + \epsilon_t$	$\Delta y_t = \alpha + \pi y_{t-1} + \epsilon_t$	$\epsilon_t \sim IID(0, \sigma^2)$

- The OLS estimator is consistent and asymptotically normal
- $t_{\pi=0} = \frac{\hat{\pi}}{\hat{\sigma}_{\hat{\pi}}} \sim N(0, 1)$

DF test - Case 4

Case 4: Constant and time trend.

True process	Estimated process	Innovations
$\Delta y_t = \alpha + \epsilon_t$	$\Delta y_t = \alpha + \delta t + \pi y_{t-1} + \epsilon_t$	$\epsilon_t \sim IIDN(0, \sigma^2)$

- Test statistics:

$$DF = T\hat{\pi} \sim \text{ Table B.5 - distribution}$$

$$t_{\pi=0} = \frac{\hat{\pi}}{se(\hat{\pi})} \sim \text{ Table B.6 - distribution}$$

DF test – example

With US GNP data. Let us try Case 4.

$$H_0 : \Delta y_t = \alpha + \epsilon_t \quad \alpha > 0$$
$$H_1 : \Delta y_t = \alpha + \delta t + \pi y_{t-1} + \epsilon_t \quad \pi < 0$$

```
> gnp2 <- log(read.table("../data/USGNP_Hamilton.dat", h =
> T <- length(gnp2)
> z = gnp2 * 100
> delta_z = diff(z)
> t = (1:(T - 1))
> ols <- lm(delta_z ~ 1 + t + z[1:(T - 1)])
> coef <- round(coef(ols), 4)
```

$$\Delta y_t = 26.1479 + 0.0251t + -0.0353y_{t-1} + \epsilon_t$$

Note $\pi < 0$.

DF test – example

```
> summary(ols)
Call:
lm(formula = delta_z ~ 1 + t + z[1:(T - 1)])

Residuals:
     Min        1Q    Median        3Q       Max
-3.02558 -0.62807  0.08742  0.71295  2.48473

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  26.14793   17.33664   1.508    0.134
t             0.02508    0.01866   1.344    0.181
z[1:(T - 1)] -0.03526    0.02419  -1.458    0.147

Residual standard error: 1.067 on 132 degrees of freedom
Multiple R-squared: 0.02064,        Adjusted R-squared: 0.0
```

DF test – example

We are in case 4. So the 5% critical value from B.5 for a T=136 sample of this size is -20.9. The Dickey–Fuller test:

$$DF = T\hat{\pi} = -4.8008 > -20.9$$

so the null hypothesis is not rejected.

So the 5% critical value from B.6 is -3.44. The OLS t-statistics for $\pi = 0$:

$$t_{\pi=0} = \frac{-0.0353}{0.0242} = -1.4587 > -3.44$$

so the null hypothesis is not rejected.

DF test – summary

- The asymptotic properties of the OLS estimate $\hat{\phi}$ when $\phi = 1$ depends on whether the process contains a constant term or a time trend.
- Which of the test should we choose?
- If the true process has a trend but we exclude that from the estimated process, then $\hat{\phi}$ may be biased and the power of the test will be reduced.
- If the true process does not have a trend but this is included in the estimated process, then the size of the test is reduced.
- We would use Case 3 for a series with a trend and Case 2 for a series without a significant trend
- Table 17.1 of Hamilton displays a summary

DF test for serially correlated innovations

If we estimate by OLS

$$y_t = \alpha + \phi y_{t-1} + \epsilon \quad \epsilon_t = \phi(L)\eta_t$$

then the estimate $\hat{\phi}$ is inconsistent. However, if $phi = 1$, then the OLS estimate is consistent.

- If ϵ_t has an ARMA structure, this affects the power of the DF test negatively
- Phillips and Perron (1988) suggest to estimate this model with OLS and modify the statistics to take into account of serial correlation
- The Augmented Dickey–Fuller (ADF) instead includes terms with the lags of y_t to correct for this serial correlation
- The test is a negative number. The more negative, the stronger the rejection of H_0
- We add $\sum_{j=1}^{p} \rho_j \Delta y_{t-j}$ to the estimation model to capture the serial correlation of ϵ_t

Phillips–Perron Unit Root Test

- The aforementioned test statistics have to be corrected to accommodate serially correlated errors
- It corrects the $t_{\pi=0}$ and DF statistics.
- The modified statistics are denoted by Z_t and Z_π

$$Z_\pi = T\hat{\pi} - \frac{1}{2}\frac{T^2\hat{\sigma}^2_{\hat{\pi}}}{s^2_T}(\hat{\lambda}^2 - \hat{\gamma}_0) \sim \text{ Table B.5 – distribution}$$

$$Z_t = \left(\frac{\hat{\gamma}_0}{\hat{\lambda}^2}\right)^{1/2} t_{\pi=0} - \frac{1}{2}\left(\frac{\lambda^2 - \hat{\gamma}_0}{\hat{\lambda}}\right)\left(\frac{T\hat{\sigma}_{\hat{\pi}}}{\hat{s}_T}\right) \sim \text{ Table B.6 – distribu}$$

Phillips–Perron Unit Root Test

$$
\begin{aligned}
\hat{\gamma}_0 &= \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_t^2 \\
\hat{\lambda}^2 &= \hat{\gamma}_0 + 2\sum_{t=1}^{q}\{1 - j/(q+1)\}\hat{\gamma}_j \\
\hat{\gamma}_j &= \frac{1}{T}\sum_{t=j+1}^{T}\hat{\epsilon}_t\hat{\epsilon}_{t-j} \\
s_T &= \frac{1}{T-2}\sum\hat{\epsilon}_T^2 \\
t_{\pi=0} &= \frac{\hat{\pi}}{se(\hat{\pi})}
\end{aligned}
$$

ADF - Case 1

Case 1: No constant or time trend with serially correlated innovations.

True process	Estimated process
$Deltay_t = \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + \epsilon_t$	$\Delta y_t = \pi y_{t-1} + \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + \eta_t$

- The classical OLS t and F tests of ρ_j are valid
- Tests for π:

$$Z_{DF} = \frac{T}{1 - \hat{\rho}_1 - \hat{\rho}_2 - \ldots - \hat{\rho}_p} \hat{\pi} \sim \text{ Table B.5 – distribution}$$

$$t_{\pi=0} = \frac{\hat{\pi}}{se(\hat{\pi})} \sim \text{ Table B.6 – distribution}$$

ADF test - Case 2

Case 2: Constant but not time trend with correlated innovations.

True process	Estimated process
$\Delta y_t = \epsilon_t$	$\Delta y_t = \alpha + \pi y_{t-1} + \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + \epsilon_t$

- The classical OLS t and F tests of ρ_j are valid
- Tests for π:

$$Z_{DF} = \frac{T}{1 - \hat{\rho}_1 - \hat{\rho}_2 - \ldots - \hat{\rho}_p} \hat{\pi} \sim \text{ Table B.5 – distribution}$$

$$t_{\pi=0} = \frac{\hat{\pi}}{se(\hat{\pi})} \sim \text{ Table B.6 – distribution}$$

ADF test - Case 3

Case 3: Constant without time trend with serially correlated innovations.

True process	Estimated process
$\Delta y_t = \alpha + \epsilon_t$	$\Delta y_t = \alpha + \pi y_{t-1} + \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + \epsilon_t$

- The classical OLS t and F tests of ρ_j are valid
- Same for the π

ADF test - Case 4

Case 4: Constant and time trend with serially correlated innovations.

True process	Estimated process
$\Delta y_t = \alpha + \epsilon_t$	$\Delta y_t = \alpha + \delta t + \pi y_{t-1} + \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + \epsilon_t$

- The classical OLS t and F tests of ρ_j are valid
- Tests for π:

$$Z_{DF} = \frac{T}{1 - \hat{\rho}_1 - \hat{\rho}_2 - \ldots - \hat{\rho}_p}\hat{\pi} \sim \text{ Table B.5 – distribution}$$

$$t_{\pi=0} = \frac{\hat{\pi}}{se(\hat{\pi})} \sim \text{ Table B.6 – distribution}$$

ADF – Example

Following the US GNP example, now assuming that the innovation are serially correlated. Case 4 with $p = 3$

```
> delta_z = diff(z)
> delta_z_1 = diff(z[1:(T - 1)])
> delta_z_2 = diff(z[1:(T - 2)])
> delta_z_3 = diff(z[1:(T - 3)])
> t = (4:(T - 1))
> ols2 <- lm(delta_z[4:(T - 1)] ~ 1 + t + z[3:(T - 2)] + de
+     3)] + delta_z_3)
> coef2 <- round(coef(ols2), 4)
```

$$\Delta y_t = 35.0167 - 0.0009t - 0.0481 y_{t-1} + 0.282 \Delta y_{t-1} + 0.1441 \Delta y_{t-2} - 0.1068 \Delta y_{t-3} + \epsilon_t$$

ADF – Example

```
> summary(ols2)

Call:
lm(formula = delta_z[4:(T - 1)] ~ 1 + t + z[3:(T - 2)] + delta_z_1[3:(T -
    2)] + delta_z_2[2:(T - 3)] + delta_z_3)

Residuals:
     Min       1Q   Median       3Q      Max
-2.92634 -0.59988  0.02939  0.67163  2.70052

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          35.01667   17.07917   2.050  0.04241 *
t                     0.03607    0.01847   1.953  0.05308 .
z[3:(T - 2)]         -0.04814    0.02387  -2.017  0.04585 *
delta_z_1[3:(T - 2)]  0.28197    0.08850   3.186  0.00182 **
delta_z_2[2:(T - 3)]  0.14415    0.09084   1.587  0.11505
delta_z_3            -0.10677    0.08794  -1.214  0.22698
---
Signif. codes:  0 âĂŸ***âĂŹ 0.001 âĂŸ**âĂŹ 0.01 âĂŸ*âĂŹ 0.05 âĂŸ.âĂŹ 0.1 âĂŸ âĂŹ 1

Residual standard error: 0.9909 on 126 degrees of freedom
Multiple R-squared: 0.1635,        Adjusted R-squared: 0.1303
F-statistic: 4.926 on 5 and 126 DF,  p-value: 0.0003756
```

ADF – Example

The 5% critical value of B.5 for Case 4 and T= 133 is -20.9

$$Z_{DF} = \frac{136}{1 - 0.282 - 0.1441 - 0.1068} = 199.794329366828 > -20.9$$

so we do not reject H_0.

The 5% critical value of B.6 for Case 4 and T= 133 is -3.44

$$t_{\pi=0} = -2.0042 > -3.44$$

so we do not reject $\pi = 0$

Choosing the Lag Length for the ADF Test

Practical issue: the choice of p!!!

- If p is too small then the remaining serial correlation in the errors will bias the test.
- If p is too large then the power of the test will suffer.
- Monte Carlo experiments suggest it is better to error on the side of including too many lags.

Choosing the Lag Length for the ADF Test

Ng and Perron "Unit Root Tests in ARMA Models with Data-Dependent Methods for the Selection of the Trun- cation Lag", JASA, 1995.

1. Set an upper bound $p_{\max}$ for p.
2. Estimate the ADF test regression with $p = p_{\max}$.
3. If the absolute value of the t-statistic of the coefficient of Δy_{t-p} is in absolute value greater than 1.6 then set $p = p_{\max}$ and perform the unit root test. Otherwise, $p_{\max} = p_{\max} - 1$ and start 1.

Q: What $p_{\max}$ do we choose as the start?

Schwert (1989) proposes

$$p_{\max} = \left[12 \left(\frac{T}{100} \right)^{1/4} \right]$$

Choosing the Lag Length for the ADF Test

Ng and Perron "Lag Length Selection and the Construc- tion of Unit Root Tests with Good Size and Power", ECTA, 2001.

- Choose p as the value that minimises the $MAIC(p)$.

$$MAIC(p) = \log(\hat{\sigma}_p^2) + \frac{2(\tau_T(p) + p)}{T - p_{\max}}$$

$$\tau_T(p) = \frac{\hat{\pi}^2 \sum_{t=p_{\max}+1}^{T} y_{t-1}}{\hat{\sigma}_p^2}$$

$$\hat{\sigma}_p^2 = \frac{1}{T - p_{\max}} \sum_{t=p_{\max}+1}^{T} \hat{\epsilon}_t^2$$