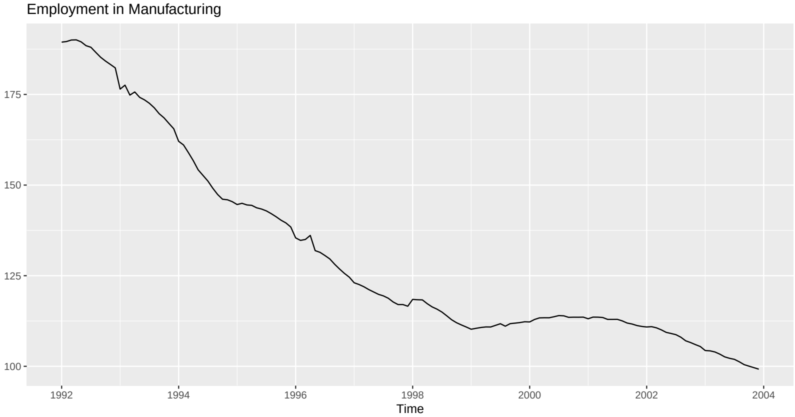
Applied Time Series Econometrics

Oxana Malakhovskaya, NRU HSE

October 8, 2019

Trends in economic time series

- A frequently observed feature of economic and financial time series is that they show trending behaviour.



Trends in economic time series (2)

- The earliest approach to work with trended series involved augmenting the model with a determinisic time trend.
- Nelson and Plosser(1982) showed that this method could represent a misspecification of the model's dynamics. They suggested that a researcher should use a stochastic trend instead.
- AR(1) equation that we analysed last time was:

$$y_t = \phi y_{t-1} + \nu_t, \qquad \nu_t \sim iid(0, \sigma^2) \qquad |\phi| < 1$$

- Now we move from $|\phi| < 1$ to $|\phi| = 1$ and we have a `random walk` process:

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

What is the problem?

- Sample moments do not have fixed limits, as they do for stationary processes but converge to random variables.
- The least squares estimator of ϕ is superconsistent with convergence rate greater than the usual rate of $\sqrt{T}$ for stationary processes.
- The asymptotic distribution of the least squares estimator is non-standard on contrast to asymptotic normality result for stationary processes.

Models of trends

Two possible specifications of model trends in the data are:

$$
\begin{aligned}
&1) \quad y_t = \beta_0 + \beta_1 t + \varepsilon_t \quad \varepsilon_t \sim iid(0, \sigma_\varepsilon^2) \\
&2) \quad y_t = \mu + y_{t-1} + \nu_t \quad \nu_t \sim iid(0, \sigma_\nu^2)
\end{aligned}
$$

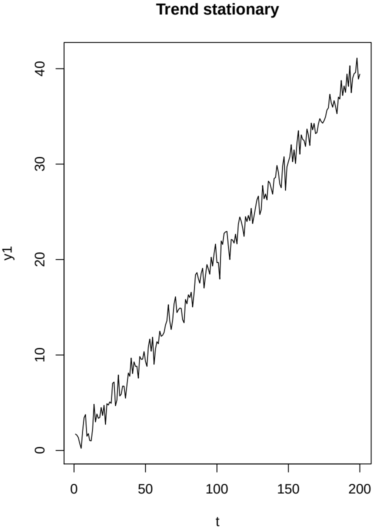
1) A deterministic trend assumes stationary (transitory) deviations around a deterministic trend.

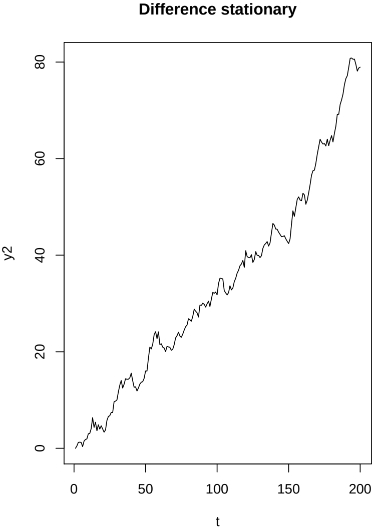
2) A stochasic trend assumes that only changes in the variable $(y_t - y_{t-1})$ are stationary. Process 2) is titled a `random walk with drift` process.

Simulations: code

```r
bet0 <- 0.1
bet1 <- 0.2
mu <- 0.3
y1 <- rep(0,times = 200)
y2 <- rep(0,times = 200)
eps <- rnorm(200)
nu <- rnorm(200)
for (i in 2:200) {
  y2[i] <- mu +  y2[i - 1] + nu[i]
}
t <- 1:200
y1 <- bet0 + bet1*t + eps
par(mfrow = c(1,2))
plot(t,y1, type = "l", main = "Trend stationary")
plot(t,y2, type = "l", main = "Difference stationary")
```

Simulations: plot





TS and DS processses(1)

The formal differernce between two types of processes becomes obvious if we solve the stochstic trend model backward recursively:

$$
\begin{aligned}
1)\quad & y_t = \beta_0 + \beta_1 t + \varepsilon_t \\
2)\quad & y_t = \mu + y_{t-1} + \nu_t = \\
& = \mu + (\mu + y_{t-2} + \nu_{t-1}) + \nu_t = \\
& = y_0 + \mu t + \nu_t + \nu_{t-1} + \cdots + \nu_1
\end{aligned}
$$

The disturbance term is the cumulative sum of the shocks to the system:

$$\varepsilon_t = \sum_{i=1}^{t} \nu_i$$

The effect of a shock does not decay over time.

TS and DS processses(2)

To make series that follow TS and DS processes stationary, we use different techniques.

TS process:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$
$$\tilde{y}_t = y_t - \beta_1 t = \beta_0 + \varepsilon_t$$

DS process:

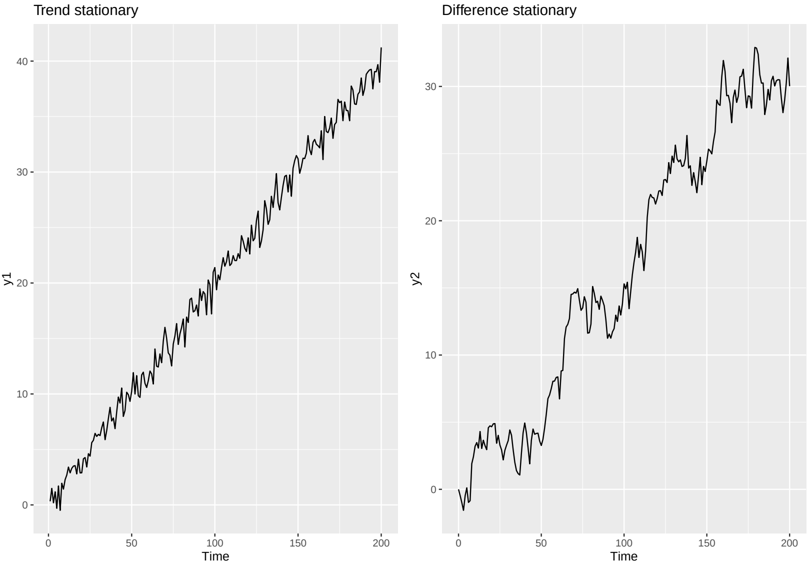
$$y_t = \mu + y_{t-1} + \nu_t$$
$$\Delta y_t = y_t - y_{t-1} = \mu + \nu_t$$

$\tilde{y}$ and Δy_t can be represented by ARMA models. We may use the `arima.sim` function to simulate them. In our example, it is sufficient to differentiate DS process once to make it stationary. So this process is referred to as integrated of order one (I(1)) process.

Alternative simulation: code

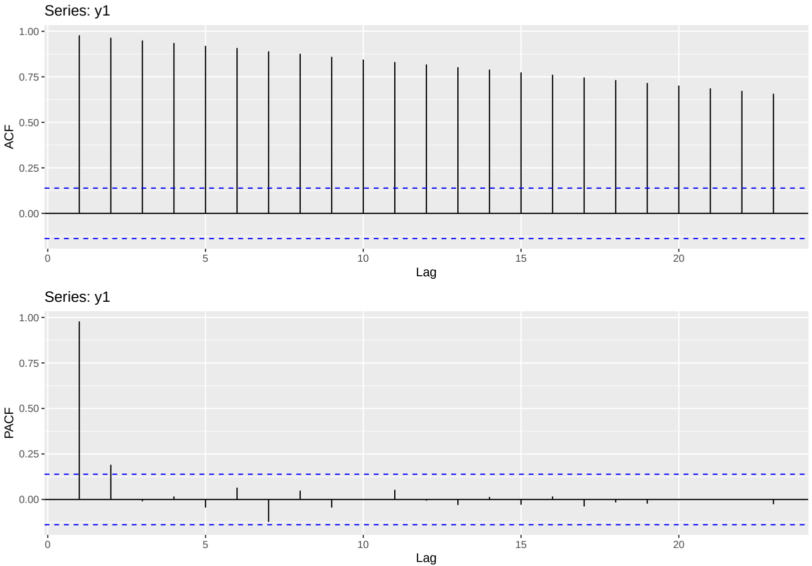
```r
library(gridExtra)
bet0 <- 0.1
bet1 <- 0.2
mu <- 0.3
y_tilde <- arima.sim(model = list(order = c(0,0,0)),
                     n = 200, mean = bet0)
y1 <- y_tilde + bet1*(1:200)
y2 <- arima.sim(model= list(order = c(0,1,0)),
                n = 200, mean = mu)
plot1 <- autoplot(y1) + ggtitle("Trend stationary")
plot2 <- autoplot(y2) + ggtitle("Difference stationary")
grid.arrange(plot1, plot2, ncol=2)
```

Alternative simulation: plot



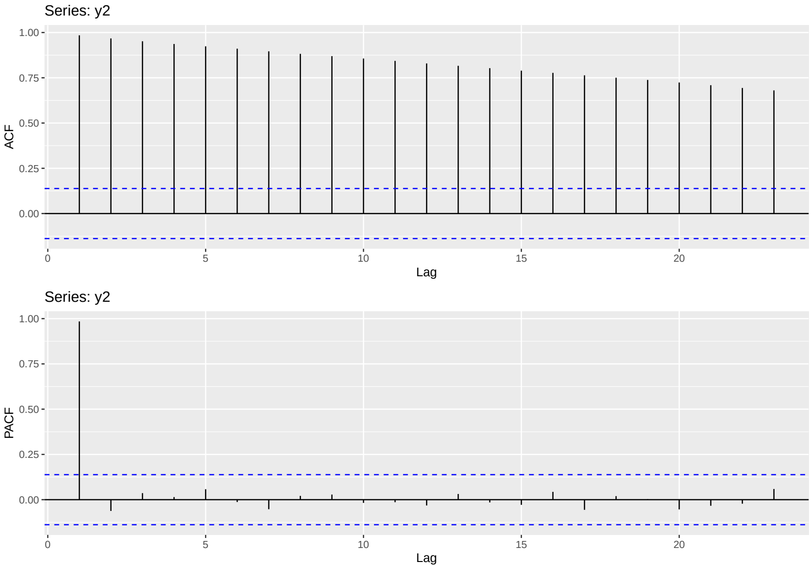
ACF and PACF for a TS process

```
grid.arrange(ggAcf(y1), ggPacf(y1), nrow=2)
```



ACF and PACF for a DS process

```
grid.arrange(ggAcf(y2), ggPacf(y2), nrow=2)
```

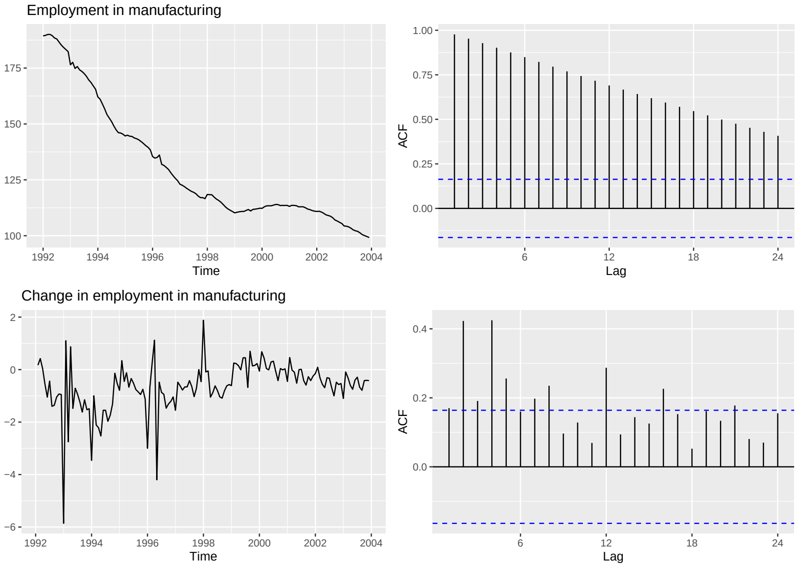


Identifying non-stationary series

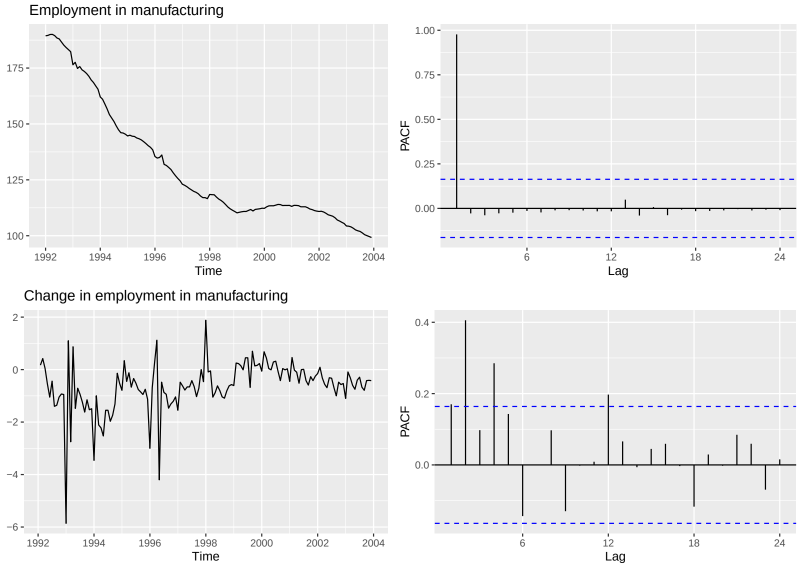
We can use several graphical hints to identify a non-stationary series.

- time plot
- The ACF of non-stationary data decreases slowly
- The partial autocorrelation for the first lag is often large and positive.

ACF for a series in levels and differences



PACF for a series in levels and differences



Second-order differencing

- Occasionally the differenced data will not appear stationary and we have to difference the data a second time:

$$
\begin{aligned}
\Delta^2 y_t &= \Delta y_t - \Delta y_{t-1} = \\
&= (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = \\
&= y_t - 2y_{t-1} + y_{t-2}
\end{aligned}
$$

- In practice, it is almost never necessary to go beyond second-order differences, and second-order differencing is rather rare.

Seasonal differencing

- A seasonal difference between an observation and the corresponding observation from the previous year.

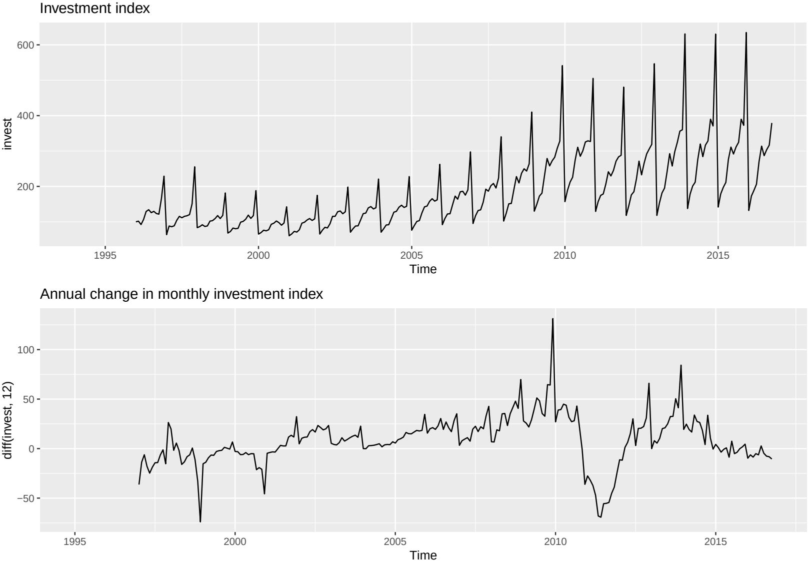
$$\Delta^s y_t = y_t - y_{t-m},$$

where m is a number of seasons ($m = 12$ for monthly data, $m = 4$ for quarterly data).

Example with investment index series (Russian data)

```
plot1 <-autoplot(invest) + ggtitle("Investment index")
plot2 <-autoplot(diff(invest,12)) +
  ggtitle("Annual change in monthly investment index")
grid.arrange(plot1, plot2, nrow = 2)
```

Seasonal differencing:example



Seasonally differenced series is closer to be stationary.

Combined differencing

- If after seasonal differencing non-stationarity remains, it can be removed with an ordinary first differencing.

$$
\begin{aligned}
\Delta(\Delta^s y_t) &= \Delta^s y_t - \Delta^s y_{t-1} = \\
&= (y_t - y_{t-12}) - (y_{t-1} - y_{t-13}) = \\
&= (y_t - y_{t-1}) - (y_{t-12} - y_{t-13}) = \\
&= \Delta^s(\Delta y_t)
\end{aligned}
$$

- It does not make difference which differencing is done first (seasonal or ordinary one).
- If seasonality is strong, it is recommended to start with seasonal differerncing. There is a chance that the resulting series will be stationary so first differerncing will be useless.

Non-stationarity and unit roots: random walk

Why referring to a series with a stochastic trend we use a "unit root" term?

$$
\begin{aligned}
y_t &= \mu + y_{t-1} + \nu_t \\
y_t - y_{t-1} &= \mu + \nu_t \\
y_t - Ly_t &= \mu + \nu_t \\
(1 - L)y_t &= \mu + \nu_t
\end{aligned}
$$

The relevant polynomial for $\phi_p(L) = 1 - L$ is

$$1 - z = 0$$

The only root is given by $z = 1$.

Non-stationarity and unit roots: ARIMA model

For more complicated process the general idea is the same. A stochastic trend is equivalent to a unit root presence in the AR polynomial.

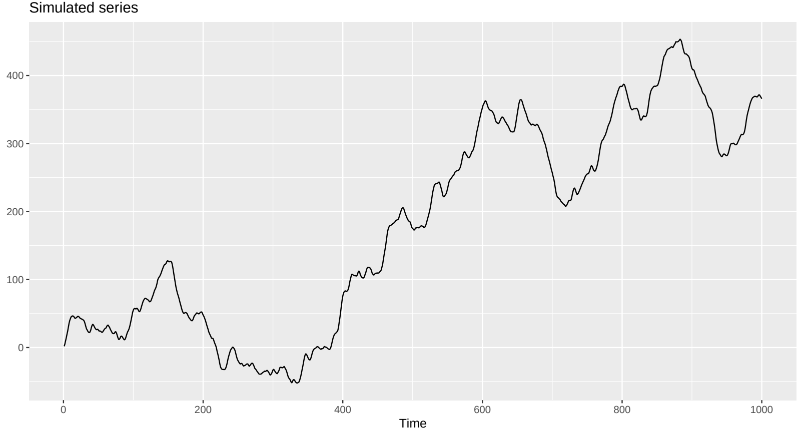
$$(1 - 0.5L)(1 - 0.8L)(1 - L)y_t = \nu_t$$
$$(1 - 2.3L + 1.7L^2 - 0.4L^3)y_t = \nu_t$$
$$y_t = 2.3y_{t-1} - 1.7y_{t-2} + 0.4y_{t-3} + \nu_t$$

The AR polynomial is as follows: $(1 - z)(1 - 0.5z)(1 - 0.8z) = 0$, and obviously it has a unit root: $z_1 = 1, z_2 = 2, z_3 = 1.25$.

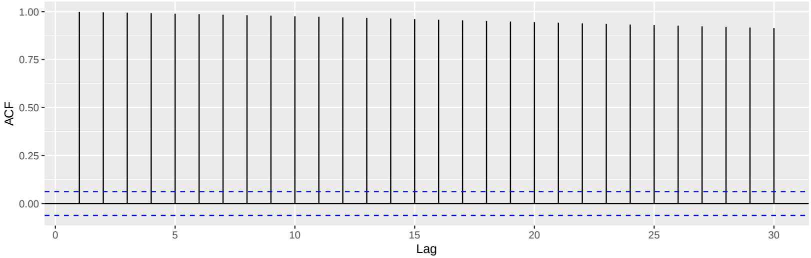
The process is non-stationary.

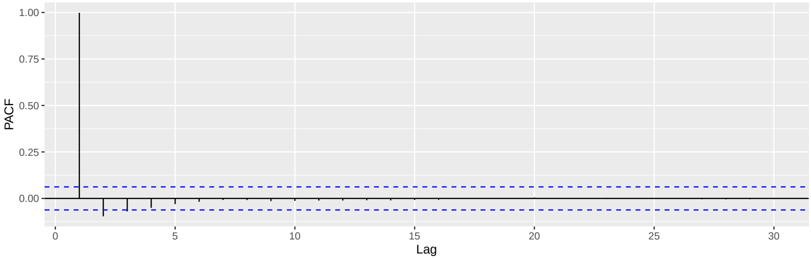
Simulated series

$$y_t = 2.3y_{t-1} - 1.7y_{t-2} + 0.4y_{t-3} + \nu_t$$



ACF and PACF for the simulated series





ARIMA representation

The same model can be represented as:

$$
\begin{gathered}
(1-0.5L)(1-0.8L)(1-L)y_t = \nu_t \\
(1-1.3L+0.4L^2)(1-L)y_t = \nu_t \\
\Delta y_t - 1.3\Delta y_{t-1} + 0.4\Delta y_{t-2} = \nu_t \\
\Delta y_t = 1.3\Delta y_{t-1} - 0.4\Delta y_{t-2} + \nu_t
\end{gathered}
$$

This is an $ARMA(2,0)$ representation for Δy_t so y_t follows the $ARIMA(2,1,0)$ model and can be simulated accordingly:

```
y3 <- arima.sim(model = list(ar = c(1.3,  -0.4),
                        order = c(2,1,0)), n = 1000)
```

Unit root tests

To formally determine the order of integration which is of great importance for the analysis, several statistical tests have been developed for investigating it.

- Augmented Dickey-Fuller (ADF) test
- Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test
- Seasonal unit root test

The idea of the Augmented Dickey-Fuller test

Consider an AR(p) model:

$$
\begin{gathered}
y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \nu_t \\
y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p} = \mu + \nu_t \\
(1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) y_t = \mu + \nu_t
\end{gathered}
$$

The relevant polynomial is:

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

It has a unit root if and only if:

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

The idea of the ADF test is to check up if $\sum_{i=1}^{p} \phi_i = 1$.

ADF test

$$
\begin{aligned}
y_t &= \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \nu_t \\
y_t - y_{t-1} &= \mu + (\phi_1 + \phi_2 + \cdots + \phi_p - 1) y_{t-1} - \\
&- (\phi_1 + \cdots + \phi_p) y_{t-1} + \phi_1 y_{t-1} + (\phi_2 + \cdots + \phi_p) y_{t-2} - \\
&- (\phi_2 + \cdots + \phi_p) y_{t-2} + \phi_2 y_{t-2} + (\phi_3 + \cdots + \phi_p) y_{t-3} - \\
&- (\phi_3 + \cdots + \phi_p) y_{t-3} + \phi_3 y_{t-3} + (\phi_4 + \cdots + \phi_p) y_{t-4} - \\
&\dots \\
&- (\phi_{p-1} + \phi_p) y_{t-(p-1)} + \phi_{p-1} y_{t-(p-1)} + \phi_p y_{t-p} + \nu_t
\end{aligned}
$$

The equation can be written as:

$$
\Delta y_t = \mu + \beta y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-(p-1)},
$$

where $\beta = \sum_{i=1}^{p} \phi_i - 1$, and $\zeta_j = -\sum_{i=j+1}^{p} \phi_i$.

We test if β is equal to zero.

ADF-test

$$\Delta y_t = \mu + \beta y_{t-1} + \zeta_1 \Delta y_{t-1} + \zeta_2 \Delta y_{t-2} + \cdots + \zeta_{p-1} \Delta y_{t-(p-1)},$$
$$H_0 : \beta = 0$$
$$H_A : \beta < 0$$

The test statistics does not have an asymptotic standard normal distribution, but it has a non-standard limiting distribution.

To implement the ADF-test we use `adf.test()` function from the `tseries` package or `ur.df()` from `urca` package.

```
library(tseries)
adf.test(y3, k = 10)
```

tseries::adf.test function: output

Augmented Dickey-Fuller Test

data: y3 Dickey-Fuller = -2.3431, Lag order = 10, p-value = 0.4333
alternative hypothesis: stationary

- The adf.test function modifies the data before performing the ADF test. First it automatically detrends the data; then it recenters the data, giving it a mean of zero.
- If either detrending or recentering is undesireable for the application in hand, we use `ur.df()` instead.

urca::ur.df function: output

```r
library(urca)
urtest1 <- ur.df(y3, lags = 10, type = 'trend')
summary(urtest1)
```

Value of test-statistic is: -2.3431 2.1577 2.7919

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-3.96	-3.41	-3.12
phi2	6.09	4.68	4.03
phi3	8.27	6.25	5.34

Literature

1. Hyndman, R. J. and Athanasopoulos, G. (2013) Forecasting: principles and practice, Otexts
2. Lütkepohl, H., and Krätzig, M. (2004) Applied Time Series Econometrics, New York, Cambridge University Press
3. Martin, V., Hurn, S., and HarriS,D. (2013) Econometric Modelling with Time Series: Specification, Estimating and Testing, New York, Cambridge University Press