04 Time series with unit root

Andrius Buteikis, andrius.buteikis@mif.vu.lt
http://web.vu.lt/mif/a.buteikis/

Difference Stationary (DS) Time Series

For a long time each economic time series used to be decomposed into a deterministic trend and a stationary process, i.e. it was assumed that all time series are TS series.

However, later it appeared that most economic series belonged to another class of *difference stationary* series.

- If Y_t is stationary, or its residuals ϵ_t in the decomposition $Y_t = T_t + S_t + \epsilon_t$ are stationary, then Y_t is called a **Trend Stationary (or TS) series**;
- If Y_t is not a TS series buts its differences $\Delta Y_t = Y_t - Y_{t-1}$ **are stationary**, then it is called a **Difference Stationary (or DS) series**;

An Example of a DS Proces

Recall that AR(1):

$$Y_t = \phi Y_{t-1} + \epsilon_t \quad \epsilon_t \sim WN(0, \sigma_\epsilon^2)$$

is stationary if $|\phi| < 1$. Therefore it is a TS series.

However, if $\phi = \pm 1$, the process is no longer stationary. Indeed, since:

$$Y_t = Y_{t-1} + \epsilon_t = Y_{t-2} + \epsilon_{t-1} + \epsilon_t = ... = Y_0 + \epsilon_1 + ...\epsilon_t$$

then $\mathbb{E}Y_t = \mathbb{E}Y_0 = const$ but $Var(Y_t) = \sigma_\epsilon^2 \cdot t \neq const$.

Note that whatever *deterministic* trend T_t we try to remove, $Var(Y_t - T_t) = Var(Y_t) \neq const$, thus this time series is **not a TS series**.

On the other hand, $\Delta Y_t = \epsilon_t$ *is* stationary, thus a $Y_t = Y_0 + \sum_{i=1}^{t} \epsilon_i$ **is a DS process**.

It will be useful to have three expressions for the TS series
$Y_t = \alpha_0 + \alpha_1 t + \phi Y_t + \epsilon_t$, $|\phi| < 1$:

▸ $Y_t - \beta_0 - \beta_1 t = \phi(Y_{t-1} - \beta_0 - \beta_1(t-1)) + \epsilon_t$

where $\beta_0 = (\alpha_0(1-\phi) - \alpha_1\phi)/(1-\phi)^2$ and $\beta_1 = \alpha_1/(1-\phi)$. Thus, the *deviations* of Y_t from the straight line $\beta_0 + \beta_1 t$ are stationary, i.e., Y_t is a TS process.

▸ $\begin{cases} Y_t = \beta_0 + \beta_1 t + u_t \\ u_t = \phi u_{t-1} + \epsilon_t \end{cases}$

In other words, Y_t is a straight line plus stationary $AR(1)$ shocks, i.e. it is again a TS process (note: solve the first equation for u_t and substitute it to the second line).

▸ The above given TS series can be expressed as $Y_t = T_t + u_t$, where T_t is a *deterministic* function and u_t is stationary.

Another way to define a DS series is to say that it contains a **stochastic** or random **trend**: the process $Y_t = (Y_0 + \alpha t + \epsilon_1 + ... + \epsilon_{t-1}) + \epsilon_t$ has a 'random' trend (the expression in parenthesis).

Unit Roots

AR(1) time series Y_t is described by $Y_t = \phi Y_{t-1} + w_t$, where w_t is a WN.

If the root of an inverse characteristic equation $1 - \phi z = 0$, namely, $z = 1/a$ equals one (i.e. if $\phi = 1$), then we say that Y_t has a unit root.

Definition.

We say that an *AR(1) process* $Y_t = \alpha + \phi Y_{t-1} + \epsilon_t$, $\epsilon_t \sim WN(0, \sigma^2)$, **has a unit root** if $\phi = 1$.

This unit root process is not stationary and is also referred to as the *random walk* model. The coefficient α is called a *drift*. If Y_t has a unit root, then ΔY_t will be stationary. For this reason, series with unit roots are referred to as *difference stationary* (DS) series.

If Y_t is an AR(p) process: $Y_t + \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + w_t$, then it is called a process with a unit root if **at least one** root of the inverse characteristic equation $1 - \phi_1 z - ... - \phi_p z^p = 0$ equals 1.
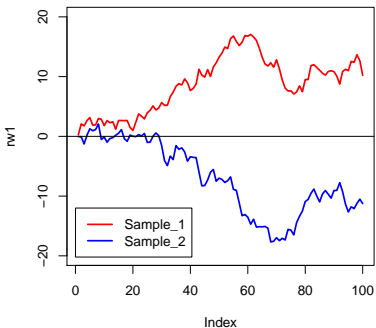
NOTE: this is **not** the same as $\phi_1 = 1$!

The trajectories of the process with unit root make long 'excursions' up and down. Below we present a couple of examples:
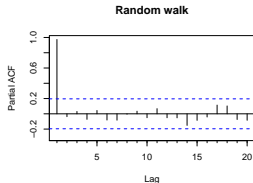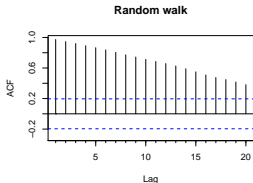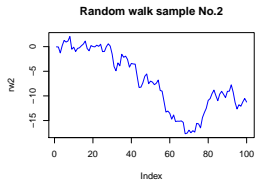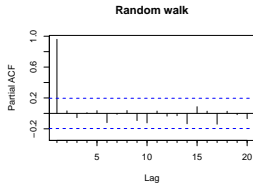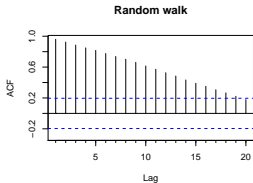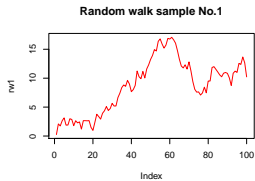

Stationary AR(1) with $\phi = 0.6$
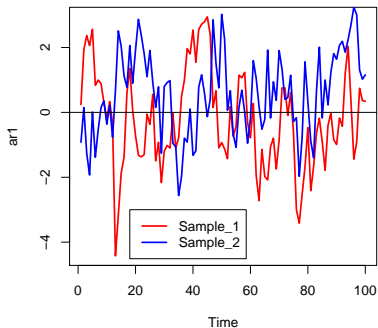
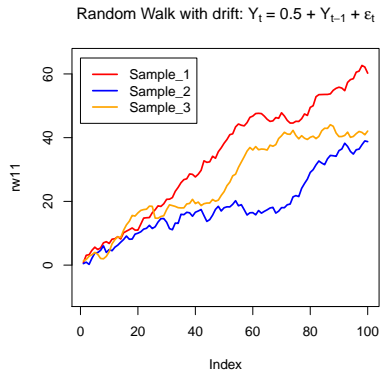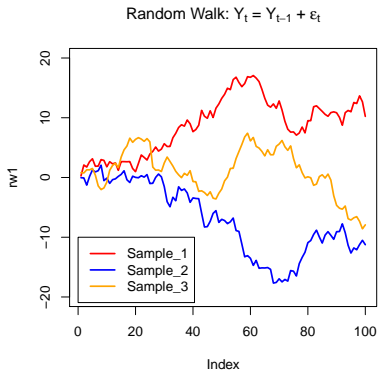Non–stationary AR(1) with $\phi = 1$, $\alpha = 0$

The sample ACF of the process decays very slowly and the first bar of the sample PACF is almost equal to 1 (other bars are almost zero).

An example of two different random walk (or unit root) processes:

- Random walk: $Y_t = Y_{t-1} + \epsilon_t$
- Random walk with drift: $Y_t = \alpha + Y_{t-1} + \epsilon_t$.

The random walk with drift exhibits trend behavior.

Integrated Time Series

A series, which become stationary after first differencing is said to be *integrated of order one* and denoted *I(1)*.

If Y_t is an AR(1) process with unit root, then the process of differences $\Delta Y_t = Y_t - Y_{t-1}$ is a *WN* process.

If Y_t is an AR(p) process with *one* unit root, then the process of differences is a stationary **AR(p-1)** process. Indeed:

$$
\begin{aligned}
w_t &= (1 - a_1 L - ... - a_p L^p) Y_t \\
&= (1 - b_1 L - ... - b_{p-1} L^{p-1})(1 - L) Y_t \\
&= (1 - b_1 L - ... - b_{p-1} L^{p-1}) \Delta Y_t
\end{aligned}
$$

The previously generated random walk (i.e. an AR(1) process with $\phi = 1$) and its first differences alongside their ACF and PACF:
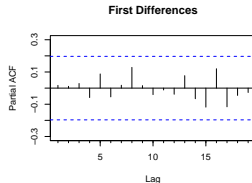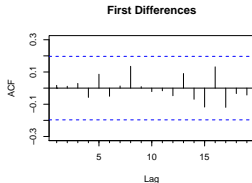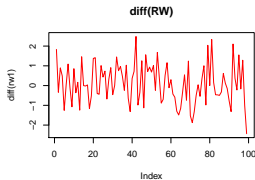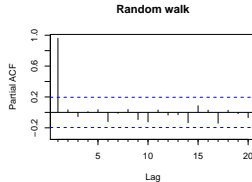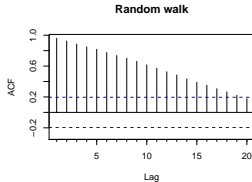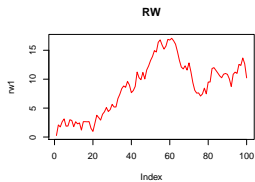
The ARIMA(p, d, q) Process

If ΔY_t is described by a stationary ARMA(p,q) model, we say that Y_t is described by an autoregressive *integrated* moving average model of order *p, 1, q*, i.e. an **ARIMA(p,1,q)** model:

$$\Phi(L)(1-L)Y_t = \alpha + \Theta(L)\epsilon_t$$

If only the series of *d*th differences is a stationary ARMA(p,q) process, then the series Y_t is the *d*th order integrated series and denoted by *I(d)*.

Then, we say that Y_t is described by an autoregressive *integrated* moving average model of order *p, d, q*, i.e. an **ARIMA(p,d,q)** model:

$$\Phi(L)(1-L)^d Y_t = \alpha + \Theta(L)\epsilon_t$$

The symbol *I(0)* is used to denote a *stationary* series.

Rewriting an AR(1) Model for Differences

If we subtract Y_{t-1} from both sides of the equation in AR(1):
$Y_t = \alpha + \phi Y_{t-1} + \epsilon_t$, we get:

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

where $\rho = \phi - 1$.

Note that if $\phi = 1 \Rightarrow \rho = 0$, then ΔY_t fluctuates around α randomly.

For future reference, note that we can test for $\rho = 0$ to see if a series has a unit root.

Furthermore, a time series will be stationary if
$-1 < \phi < 1 \Rightarrow -2 < \rho < 0$.

We will refer to this as the *stationary condition*.

Rewriting an AR(p) Model for Differences

If we subtract Y_{t-1} from the AR(p):

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$$

we get:

$$\Delta Y_t = \alpha + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + ... + \gamma_{p-1} \Delta Y_{t-p+1} + \epsilon_t$$

where the coefficients in this regression $\rho, \gamma_1, ..., \gamma_{p-1}$ are functions of $\phi_1, ..., \phi_p$. for instance, $\rho = \phi_1 + ... + \phi_p - 1$.

Note that both equations are AR models.

As in the AR(1) case $\rho = 0$ implies that the AR(p) time series **contains** a unit root. If $-2 < \rho < 0$, then the series is stationary.

In other words, if $\rho = 0$, then our equation involves only ΔY_t:

$$\Delta Y_t = \alpha + \gamma_1 \Delta Y_{t-1} + ... + \gamma_{p-1} \Delta Y_{t-p+1} + \epsilon_t$$

Additional Remarks

> The AR(p) process has a first order unit root if:
> (1) it is not stationary, but
> (2) its first differences are stationary;
>
> In AR(1) case this is equivalent to $\phi = 1$ or $\rho = 0$;
> In AR(p) case this is equivalent to $\phi_1 + ... + \phi_p = 1$ or $\rho = 0$.

- We say that the $ARMA(p+1, q)$ process Y_t has a unit root if at least one of the $p + 1$ autoregressive roots of $\Phi(z) = 0$ equals 1. This is equivalent to saying that the differenced process ΔY_t is $ARMA(p, q)$. The original $ARMA(p+1, q)$ process Y_t is now called the $ARIMA(p, 1, q)$ process.

> If the dth differences $\Delta^d Y_t$ make a stationary ARMA(p, q) process, then the process Y_t is called an ARIMA(p,d,q) process.

If Y_t is a ARIMA(1, 1, 1): $(1 - 0.4L)(1 - L)Y_t = (1 + 0.3L)\epsilon_t$.

Then, we can rewrite it as:

$$\frac{1 - 0.4L}{1 + 0.3L} \cdot (1 - L)Y_t = \epsilon_t \Rightarrow (1 - 0.4L)(1 - 0.3L + (0.3L)^2 - ...)(1 - L)Y_t = \epsilon_t$$

The infinite order polynomial on the left-hand-side can be approximated by a finite polynomial $AR(k)$ process. The popular `Dickey-Fuller` unit root test always interprets the process under consideration as an AR process.

With the exception of a case called cointegration (a topic for later), we do **not** want to include unit root variables in regression models.

If a unit root in Y_t is present, we will want to difference it and use ΔY_t. In order to do so, we must know first *if* Y_t has a unit root.

As noted previously, a drifting unit root series exhibits trend behavior.

Question: Could we simply examine time series plots for such trending to determine if it indeed has a unit root?

Answer: No.

We will explain this with an example.

Comparing TS and DS Process Graphically

Let us examine three models:

1. Random walk with a drift (DS process with a stochastic trend):

$$Y_t = \delta + Y_{t-1} + \epsilon_t = Y_0 + \delta \cdot t + \sum_{j=1}^{t} \epsilon_j, \quad \epsilon_t \sim WN(0, \sigma^2)$$

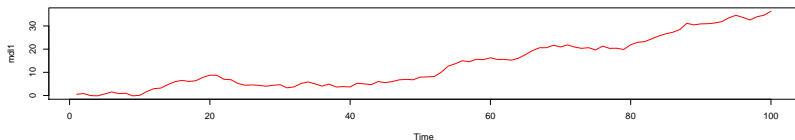2. The process with a linear trend and WN disturbances:

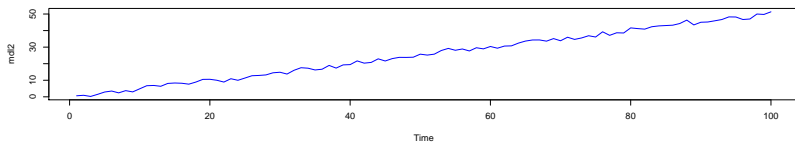$$Y_t = \alpha + \delta \cdot t + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$$

3. The process with a linear trend and AR(1) disturbances (a TS process with a deterministic trend):

$$\begin{cases} Y_t = \alpha + \delta \cdot t + u_t \\ u_t = \phi u_{t-1} + \epsilon_t, \quad |\phi| < 1, \quad \epsilon_t \sim WN(0, \sigma^2) \end{cases}$$

DS with a stochastic trend – random walk with a drift

TS with a deterministic trend – process with linear trend and WN disturbances

TS with a deterministic trend – process with linear trend and AR(1) disturbances , $\phi = 0.6$
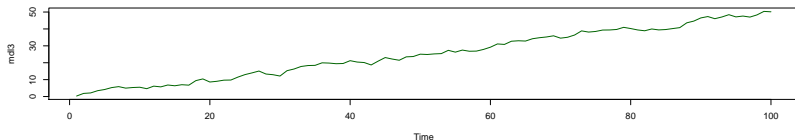
All three models have very similar paths but their properties are essentially different. Thus, looking at time series plots alone is not enough to tell whether a series has a unit root - we need some statistical procedures to decide upon it.

The forecasts of TS models tend to the trend line while a DS model does not possess this **mean reverting** property (at any point the process begins anew).

The OLS estimate of δ in the DS process:
$$Y_t = \delta + Y_{t-1} + \epsilon_t = Y_0 + \delta \cdot t + \epsilon_1 + ... + \epsilon_t, \quad \epsilon_t \sim WN$$
is inconsistent because the error process $\sum_{i=1}^{t} \epsilon_i$ is not a WN.

To estimate δ consistently, difference Y_t.

The OLS estimate of δ in the stationary TS process:
$$\begin{cases} Y_t = \alpha + \delta \cdot t + u_t \\ u_t = \phi u_{t-1} + \epsilon_t, \quad |\phi| < 1, \epsilon_t \sim WN \end{cases}$$
is consistent.

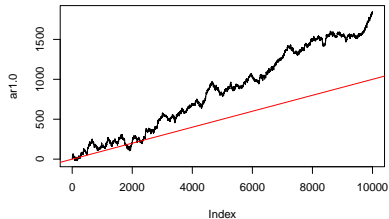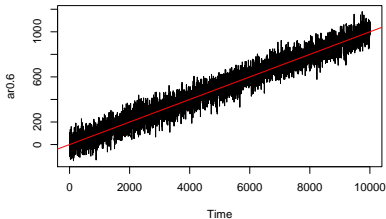Below is an illustration of the above. In the DS case, when the sample size increases, the δ estimate does not converge to its true value (the red line).
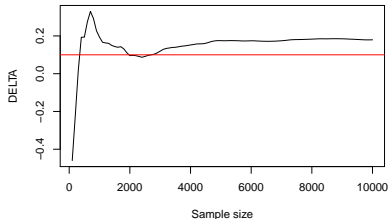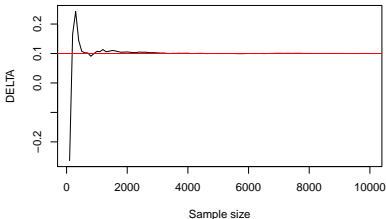
Summary:

1. The **nonstationary** time series on which we focus are those containing a **unit root**, i.e. these series contain a **stochastic trend**. But if we difference these time series, the resulting time series will be **stationary**. For this reason they are also called **difference stationary**.

2. The **stationary** time series on which we focus have $-2 < \rho < 0$ in the AR(p) model. However, these series can exhibit trend behavior through incorporation of a **deterministic trend**. In this case, they are referred to as **trend stationary**.

If we add a deterministic trend to the AR(p) model, we obtain a very special *general model* that is commonly used in univariate time series analysis:

$$\Delta Y_t = \alpha + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + ... + \gamma_{p-1} \Delta Y_{t-p+1} + \delta t + \epsilon_t$$

We refer to the above as the **AR(p) with deterministic trend model** and use it later. We use namely this form of an AR process because it is more convenient for testing the hypothesis for unit root, i.e. $H_0 : \rho = 0$.

Testing in the AR(p) with deterministic trend model

While it is possible to use OLS to estimate the coefficients of the AR(p) model for δY_t, in the case where we face the possibility of a unit root there is one important complication that occurs in the AR(p) model that was not present earlier.

To understand it, let us divide the coefficients in the model into two groups:

1. $\alpha, \gamma_1, ..., \gamma_{p-1}$ and δ
2. ρ

In other words, we consider hypothesis tests involving ρ *separately* from those involving the other coefficients.

Testing significance of $\alpha, \gamma_1, ..., \gamma_{p-1}$

Many sophisticated statistical criteria and testing methods exist to determine the appropriate lag length in an AR(p) model. Nonetheless, simply looking at the t-ratio or p-values in regression outputs can be quite informative.

Alternatively, a more common route is to proceed *sequentially*, that is, to choose a maximum lag length p_{max} and then sequentially drop lag lengths if the relevant coefficients are insignificant.

The Sequential Procedure:

1. Begin with an AR(p_{max})
2. If the p_{max}-th lag is insignificant, we reduce the model to an AR($p_{max} - 1$)
3. If the $p_{max} - 1$-th lag is insignificant, then we drop it and examine an AR($p_{max} - 2$), etc.

▶ Generally, start with a fairly large choice for p_{max}.

Testing significance of δ

In the AR(p) with deterministic trend model, we also have to worry about testing whether $\delta = 0$.

This can be accomplished in the standard way by checking whether its p-value is less than the level of significance (e.g. 0.05).

This test can be done at any stage, but it is common to carry it out *after* following the sequential procedure for choosing *p*.

Note that in hypothesis testing, ρ is different from other coefficients - thus we must treat it differently.

Testing involving ρ

A correct way of testing for a unit root, i.e. the null hypothesis

$$H_0 : \rho = 0$$

is known as the **Dickey-Fuller (or ADF) test**. They recommend using the same t-ratio for testing $\rho = 0$, but *correcting* the p-value.

We can motivate the Dickey-Fuller test in terms of the following: testing could be done by comparing a test statistic (here , the t-ratio) to a critical value to determine whether the former was either 'small' (in which case we accept the null hypothesis) or 'large', i.e. 'far from zero' (in which case the null hypothesis was rejected).

In the standard stationary case, the critical values are taken from statistical tables of the Student distribution. Dickey and Fuller demonstrated that in the *unit root* case, this is **incorrect**.

To explain, assume that:

$$Y_t = \phi Y_{t-1} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$$

and test the null $H_0 : \phi = 1$ (unit root) against an alternative stationary hypothesis: $H_1 : (-1 <)\phi < 1$.

In doing so, we used to apply the t-ratio statistics:

$$\tau = \frac{\widehat{\phi} - 1}{\widehat{s.e.(\hat{\phi})}}$$

Recall that if H_0 is true, τ has the Student distribution (thus, if $\tau < -1.645$, we reject H_0). However, in the unit root case, the probability density function of τ is *shifted* to the left and the 0.05 quantile is *more negative than* -1.645.

A rough rule of thumb can be used that will not lead you too far wrong if your number of observations is moderately large (e.g. T > 50).

Rule of Thumb when Testing for Unit Root

The approximate rule is given in the following strategy for testing a unit root:

1. Use the strategy outlined above to estimate the AR(p) with deterministic trend model. Record the t-statistics corresponding to ρ (i.e. the t-ratio coefficient on Y_{t-1});

2. If the final version of your model **includes a deterministic trend**, the Dickey-Fuller critical value is approximately -**3.45**. If the t-ratio on ρ is **more negative than** -3.45, **reject the unit root hypothesis** and conclude that the series is stationary. Otherwise, conclude that the series has a unit root.

3. If the final version of your model **does not include a deterministic trend**, the Dickey-Fuller critical value is approx. -2.89. If the t-statistic on ρ is **more negative than** -2.89, **reject the unit root hypothesis** and conclude that the series is stationary.

Unit Root with `auto.arima`

Testing for an (autoregressive) unit root is a time consuming procedure. However, it appears that the `auto.arima` function allows us to automate this procedure well (we shall restrict ourselves to the non-seasonal data). A non-seasonal ARIMA(p,d,q) process is given by:

$$\Phi(L)(1 - L)^d Y_t = \alpha + \Theta(L)\epsilon_t$$

The main task in automatic ARIMA forecasting is selecting an appropriate model order, that is the values of p, q and d.

If d is known, we can select the orders p and q via a information criterion (e.g. AIC).

To choose d, `auto.arima` uses unit root tests. However, Dickey-Fuller test is based on a null hypothesis that **a unit root exists**, i.e. $H_0 : \rho = 0$, which biases results towards more differences rather than fewer differences, i.e. we get a large value d.

Therefore, the default behavior of `auto.arima` is based on the **Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test** with null hypothesis of **no unit root**, i.e. Y_t is stationary: we test the data for a unit root.

If KPSS test statistic is *significant*, i.e. there exists a unit root in Y_t, $d = 1$, we test the *differenced* data for a unit root and so on.

Example: Stock prices on the NYSE

The data file nyse.txt contains monthly data from 1952:1 through 1996:1 on a major stock price index sp (= StockPrice) provided by the New York Stock Exchange:

```
suppressPackageStartupMessages({require("forecast")})
txt1 <- "http://uosis.mif.vu.lt/~rlapinskas/(data%20R&GRETL/"
txt2 <- "nyse.txt"
sp <- read.csv(url(paste0(txt1, txt2)),
                    header = TRUE, as.is = TRUE)
sp <- ts(sp, start = c(1952, 1), freq = 12)
```
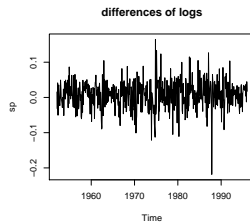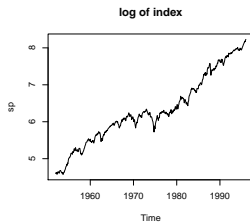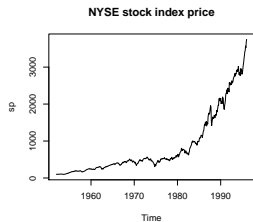
We shall apply the discussed properties of auto.arim to the logarithm of the data:

```
sp.kpss <- auto.arima(log(sp), test = "kpss")
sp.kpss
```

```
## Series: log(sp)
## ARIMA(0,1,0) with drift
##
## Coefficients:
##          drift
##         0.0069
## s.e.    0.0018
##
## sigma^2 estimated as 0.00165:  log likelihood=942.79
## AIC=-1881.59   AICc=-1881.56   BIC=-1873.05
```

Similar results can be obtained using the ADF test:

```
auto.arima(log(sp), test = "adf")
```

```
## Series: log(sp)
## ARIMA(0,1,0) with drift
##
## Coefficients:
##          drift
##         0.0069
## s.e.    0.0018
##
## sigma^2 estimated as 0.00165:  log likelihood=942.79
## AIC=-1881.59   AICc=-1881.56   BIC=-1873.05
```

Unit Root - Summary

Testing for an (autoregressive) unit root is a time consuming procedure.

To begin, we rewrite our basic model

$$Y_t = \alpha + \delta t + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$$

as

$$\Delta Y_t = \alpha + \delta t + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + ... + \gamma_{p-1} \Delta Y_{t-p+1} + \epsilon_t$$

The latter form is more convenient since the unit root hypothesis can be simply formulated as $H_0 : \rho = 0$; the former model can also have a multicollinearity problem.

When testing for unit root, it is important to understand the meaning of the hypothesis - for example, the function `ur.df` in R package `urca` suggests three options: `type = "none"`, `"drift"` and `"trend"` when carrying out the Augmented Dickey-Fuller (ADF) unit root test.

The ADF test can handle more complex models than the Dickey-Fuller test, and it is also more powerful.

1. "none" - our time series Y_t resembles a stationary AR(1) process with zero mean (however, we suspect that it could be a random walk):

$$H_1 : Y_t = \phi Y_{t-1} + w_t \quad \sim \quad Y_t - Y_{t-1} = (\phi - 1)Y_{t-1} + w_t, |\phi| < 1$$
$$H_0 : Y_t = Y_{t-1} + w_t \quad \sim \quad Y_t - Y_{t-1} = 0 \cdot Y_{t-1} + w_t$$

In this case, the null $\phi = 1$ means that our process is a random walk without drift. If the hypothesis is rejected, we conclude that the process is stationary AR(1) with zero mean.
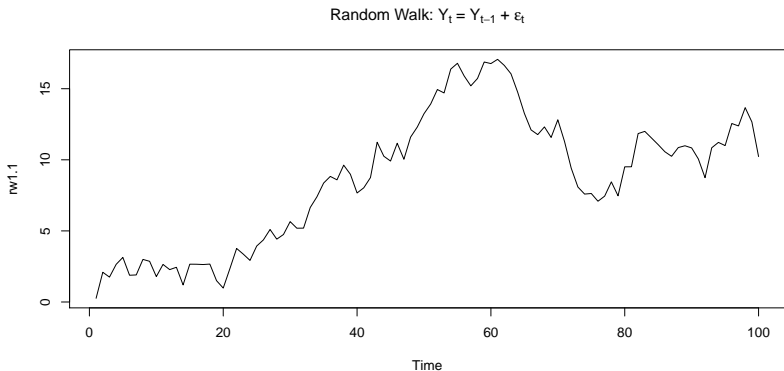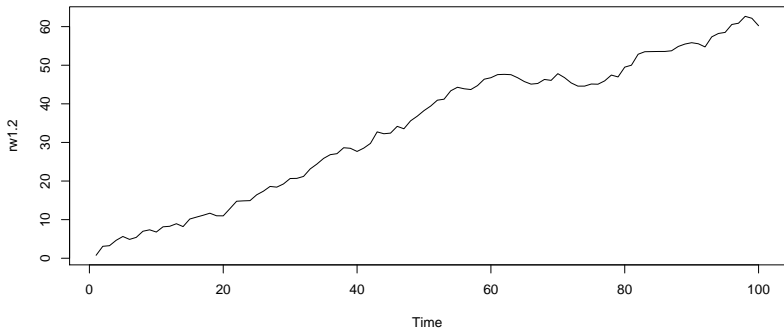

Random Walk: $Y_t = Y_{t-1} + \varepsilon_t$

2. "drift" - our time series resembles a stationary AR(1) process with nonzero mean (however, we suspect that it could be a random walk with drift):

$$H_1 : Y_t = \alpha + \phi Y_{t-1} + w_t \sim Y_t - Y_{t-1} = \alpha + (\phi - 1)Y_{t-1} + w_t, |\phi| < 1$$
$$H_0 : Y_t = \alpha + Y_{t-1} + w_t \sim Y_t - Y_{t-1} = \alpha + 0 \cdot Y_{t-1} + w_t$$

In this case, the null $\phi = 1$ means that our process is a random walk with drift. If the hypothesis is rejected, we conclude that the process is stationary AR(1) with nonzero mean.

Random Walk with drift: $Y_t = 0.5 + Y_{t-1} + \varepsilon_t$

3. `"trend"` - our time series resembles a stationary AR(1) process around a linear trend (however, we suspect that it could be a random walk with drift):

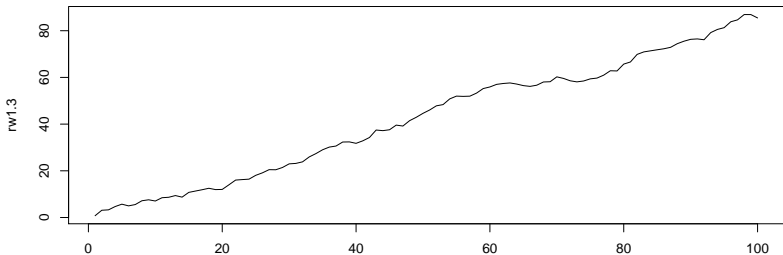$$H_1 : Y_t - a - b \cdot t = \phi(Y_{t-1} - a - b \cdot t) + w_t \sim$$
$$Y_t = [a(1-\phi) + b\phi] + b(1-\phi) \cdot t + \phi Y_{t-1} + w_t, |\phi| < 1$$
$$H_0 : Y_t = b + Y_{t-1} + w_t \sim$$
$$Y_t - Y_{t-1} = b + 0 \cdot Y_{t-1} + w_t$$

In this case, the null $\phi = 1$ means that our process is a random walk with drift. If the hypothesis is rejected, we conclude that our process is a stationary AR(1) around a line $a + bt$.

Random Walk with drift: $Y_t = 0.5 + 0.005\ t + Y_{t-1} + \varepsilon_t$

Example: Stock prices on the NYSE (continued)

We have to plot Y_t first and then decide upon the type of the model. In theory, our sequential procedure gives us the right model, though you should always be careful.

▶ We can manually specify the auto.arima model using dynlm:

```
suppressPackageStartupMessages({library(dynlm)})
sp.mdl <- dynlm(d(log(sp)) ~ L(log(sp)) + time(log(sp)))
round(summary(sp.mdl)$coefficients, 4)
```

```
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -2.1695     1.1246 -1.9292   0.0542
## L(log(sp))      -0.0155     0.0082 -1.8995   0.0580
## time(log(sp))    0.0012     0.0006  1.9362   0.0534
```

The t - statistics on ρ is -1.8995, which is not more negative than -3.42. Hence we can accept the hypothesis that NYSE stock price *log(price)* contains a unit root and makes a random walk with drift.

(i.e., $\phi = 1$ and thus $log(sp)$ is described (see the "trend" case above) by a model $\Delta log(sp)_t = b + w_t$)

▶ The same conclusion can be obtained in a simpler manner with:

```
suppressPackageStartupMessages({library(urca)})
sp.test <- ur.df(log(sp), type = "trend", lags = 0)
round(summary(sp.test)@testreg$coefficients, 4)
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0792     0.0388  2.0382   0.0420
## z.lag.1       -0.0155     0.0082 -1.8995   0.0580
## tt             0.0001     0.0000  1.9362   0.0534
```

```
summary(sp.test)@cval
```

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

where 5% critical value of tau3 is -3.41.

- ▶ Another approach in testing the unit root hypothesis is not to "look for the highest significant lag"" but search the model with minimum AIC (or BIC).

This often allows to automate the model selection procedure: below, the function ur.df begins with lags = 4 and goes down until the minimum AIC model (with lags = 1) is detected:

```
sp.test <- ur.df(log(sp), type = "trend",
                 lags = 4, selectlags = "AIC")
round(summary(sp.test)@testreg$coefficients, 4)
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.0856     0.0391  2.1864   0.0292
## z.lag.1      -0.0168     0.0082 -2.0463   0.0412
## tt            0.0001     0.0000  2.0525   0.0406
## z.diff.lag    0.0526     0.0438  1.2013   0.2302
```

Now the t - statistics with ρ is -2.046, but this doesn't changes our final conclusion. Though, the model suggested is an $ARIMA(1, 1, 0)$ with an insignificant z.diff.lag coefficient.

- ▶ The same conclusions can also be drawn from using `auto.arima`:

```r
library(forecast)
aa.mod = auto.arima(log(sp))
aa.mod
```

```
## Series: log(sp)
## ARIMA(0,1,0) with drift
##
## Coefficients:
##          drift
##         0.0069
## s.e.    0.0018
##
## sigma^2 estimated as 0.00165:  log likelihood=942.79
## AIC=-1881.59   AICc=-1881.56   BIC=-1873.05
```

The function `auto.arima` runs through many models (including those for ΔY_t) and chooses the one with **minimum AIC**, thus the "best" model here is ARIMA(0,1,0) which means unit root.
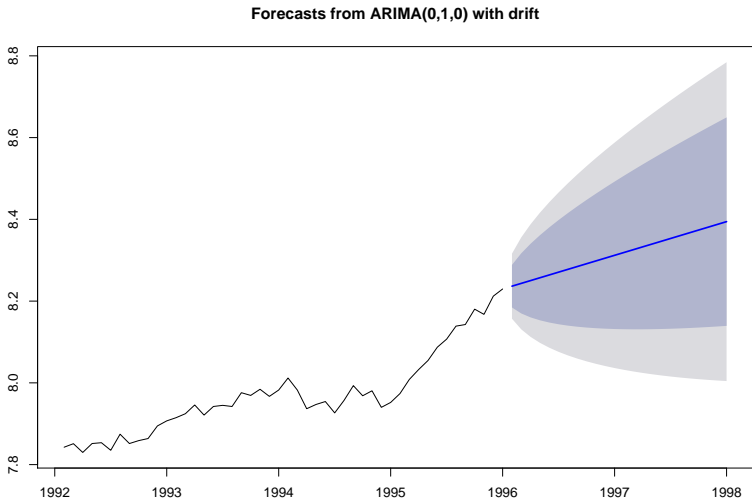
And lastly, we have to check whether the residuals of our final model make *WN*:

```
tsdiag(aa.mod)
```
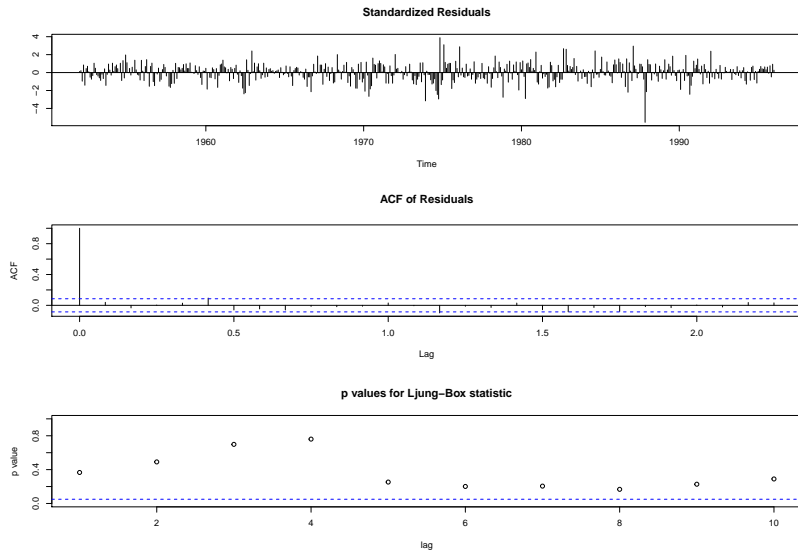
We can also forecast the data:

```
plot(forecast(aa.mod), include = 48)
```


Forecasts from ARIMA(0,1,0) with drift

Alternatively, if we check the differenced data:

```
sp.test.2 <- ur.df(diff(log(sp)), type = "none",
                   lags = 4, selectlags = "AIC")
summary(sp.test.2)@teststat
```

```
##                     tau1
## statistic -8.421185
```

```
summary(sp.test.2)@cval
```

```
##        1pct  5pct  10pct
## tau1 -2.58 -1.95 -1.62
```

Because the test statistic = -8.421185 is less than the critical value, we reject the unit root hypothesis and conclude that the **differenced series** is stationary.

Remark on deriving an AR(p) process in terms of its differences

In the examples $\epsilon_t \sim WN(0, \sigma^2)$.

AR(1)

Let's say our AR(1) model is of the following form:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \epsilon_t$$

We can rewrite the model:

$$Y_t = \alpha + \phi_1 Y_{t-1} \pm Y_{t-1} + \epsilon_t = \alpha + (\phi_1 - 1)Y_{t-1} + Y_{t-1} + \epsilon_t$$

Which leads to the model for the differences:

$$\Delta Y_t = \alpha + (\phi_1 - 1)Y_{t-1} + \epsilon_t$$

AR(2)

Let's say our AR(2) model is of the following form:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t$$

We rewrite the model:

$$Y_t = \alpha + \phi_1 Y_{t-1} \pm Y_{t-1} + \phi_2 Y_{t-2} \pm \phi_2 Y_{t-1} + \epsilon_t$$
$$= \alpha + (\phi_1 + \phi_2 - 1)Y_{t-1} + Y_{t-1} + (-\phi_2)\Delta Y_{t-1} + \epsilon_t$$

Which leads to the model for the differences:

$$\Delta Y_t = \alpha + (\phi_1 + \phi_2 - 1)Y_{t-1} - \phi_2 \Delta Y_{t-1} + \epsilon_t$$

AR(3)

our AR(3) model is of the following form:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \epsilon_t$$

We rewrite the model:

$$\begin{aligned}
Y_t &= \alpha + \phi_1 Y_{t-1} \pm Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} \pm \phi_3 Y_{t-2} + \epsilon_t \\
&= \alpha + (\phi_1 - 1)Y_{t-1} + Y_{t-1} + (\phi_2 + \phi_3)Y_{t-2} \pm (\phi_2 + \phi_3)Y_{t-1} + (-\phi_3)\Delta Y_{t-2} + \epsilon_t \\
&= \alpha + (\phi_1 + \phi_2 + \phi_3 - 1)Y_{t-1} + Y_{t-1} - (\phi_2 + \phi_3)\Delta Y_{t-1} - \phi_3 \Delta Y_{t-2} + \epsilon_t
\end{aligned}$$

Which leads to the model for the differences:

$$\Delta Y_t = \alpha + (\phi_1 + \phi_2 + \phi_3 - 1)Y_{t-1} - (\phi_2 + \phi_3)\Delta Y_{t-1} - \phi_3 \Delta Y_{t-2} + \epsilon_t$$

AR(p)

We see that we can generalize this for the AR(p) process:

$$Y_t = \alpha + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \epsilon_t, \quad \epsilon_t \sim WN(0, \sigma^2)$$

we get:

$$\Delta Y_t = \alpha + \rho Y_{t-1} + \gamma_1 \Delta Y_{t-1} + ... + \gamma_{p-1} \Delta Y_{t-p+1} + \epsilon_t$$

where:
$$\rho = \phi_1 + ... + \phi_p - 1$$
$$\gamma_j = - [\phi_{j+1} + ... + \phi_p], \quad j = 1, ..., p-1$$

Seasonal ARIMA models

So far in this section, we have restricted our attention to non-seasonal ARIMA models. A seasonal ARIMA model is formed by including additional seasonal terms in the ARIMA models we have seen so far: $SARIMA(p, d, q)(P, D, Q)_S$:

$$\Phi(L^S)\phi(L)(1 - L)^d(1 - L^S)^D(Y_t - \mu) = \Theta(L^S)\theta(L)\epsilon_t$$

where d is the integration order of the non-seasonal part of the model, and D is the integration order of the seasonal part of the model and:

▶ The non-seasonal components are:

$$AR: \phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$$
$$MA: \theta(L) = 1 + \theta_1 L + ... + \theta_q L^q$$

▶ The seasonal components are:

$$\text{Seasonal AR: } \Phi(L^S) = 1 - \Phi_1 L^S - ... - \Phi_p L^{S \cdot P}$$
$$\text{Seasonal MA: } \Theta(L^S) = 1 + \Theta_1 L^S + ... + \Theta_q L^{S \cdot Q}$$

For example, an $SARIMA(1, 1, 1)(1, 1, 1)_4$ model:

$$(1 - \Phi_1 L^4)(1 - \phi_1 L)(1 - L)(1 - L^4)Y_t = (1 + \Theta_1 L^4)(1 + \theta_1 L)\epsilon_t$$
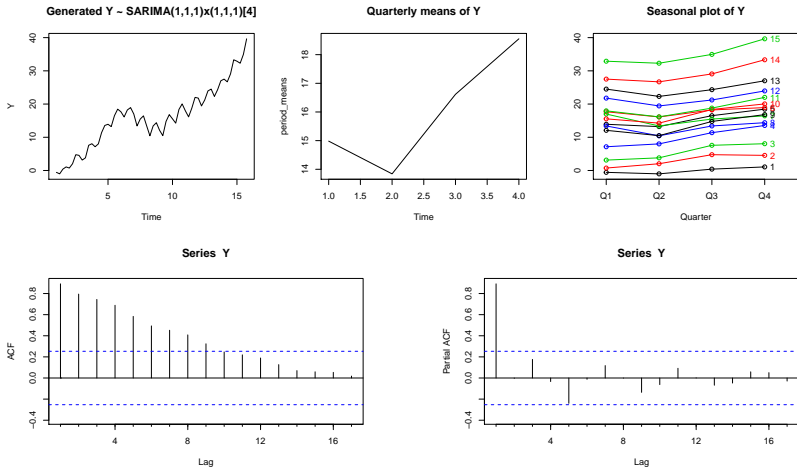
can be rewritten as:

$$\begin{aligned}
Y_t = {} & (1 + \phi_1)Y_{t-1} - \phi_1 Y_{t-2} + (1 + \Phi_1)Y_{t-4} \\
& - (1 + \phi_1 + \Phi_1 + \phi_1\Phi_1)Y_{t-5} + (\phi_1 + \phi_1\Phi_1)Y_{t-6} \\
& - \Phi_1 Y_{t-8} + (\Phi_1 + \phi_1\Phi_1)Y_{t-9} - \phi_1\Phi_1 Y_{t-10} \\
& + \epsilon_t + \theta_1\epsilon_{t-1} + \Theta_1\epsilon_{t-4} + \theta_1\Theta_1\epsilon_{t-5}
\end{aligned}$$

or, rearranging some elements, in terms of differences:

$$\begin{aligned}
\Delta Y_t = {} & \phi_1\Delta Y_{t-1} + (1 + \Phi_1)\Delta Y_{t-4} - (\phi_1 + \phi_1\Phi_1)\Delta Y_{t-5} \\
& - \Phi_1\Delta Y_{t-8} + \phi_1\Phi_1\Delta Y_{t-9} + \epsilon_t + \theta_1\epsilon_{t-1} + \Theta_1\epsilon_{t-4} + \theta_1\Theta_1\epsilon_{t-5}
\end{aligned}$$

with $\theta_1 = 0.2, \phi_1 = 0.2, \Theta_1 = 0.4, \Phi_1 = -0.3$ and $Y_0 = 0$:

Note: `forecast::seasonplot()` lets us plot the seasonal plot of each year of a `ts()` object (e.g. the quarterly data of each year as a separate series in a plot).

If we test for a unit root then, no matter which type of model we select (drift, trend, none), we still do not reject the null hypothesis $H_0 : \rho = 0$, i.e. the series has a unit root.

```
tst1 <- ur.df(Y, type = "none", lags = 10, selectlags = "AIC")
tst2 <- ur.df(Y, type = "drift", lags = 10, selectlags = "AIC")
tst3 <- ur.df(Y, type = "trend", lags = 10, selectlags = "AIC")
summary(tst1)@teststat
```

```
##                tau1
## statistic 1.577294
```

```
summary(tst2)@teststat
```

```
##                tau2      phi1
## statistic 0.4622241 1.270438
```

```
summary(tst3)@teststat
```

```
##                tau3      phi2      phi3
## statistic -2.771195 3.545955 4.849967
```

Now for the differentiated series:
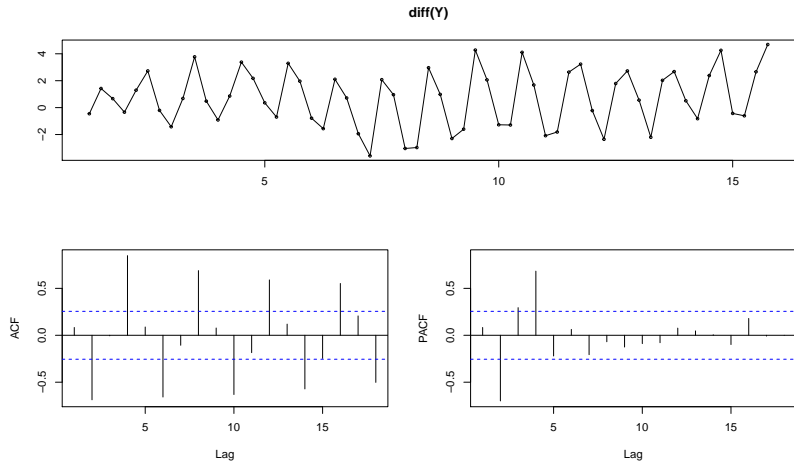
We can, very clearly, see the seasonal component every 4th lag in the decaying ACF plot and the first four significant lags in the PACF.

If we test the differences - we still have a unit root:

```
tst4 <- summary(ur.df(diff(Y), type = "none",
                      lags = 10, selectlags = "AIC"))
tst4@teststat
```

```
##                   tau1
## statistic -0.7193586
```

```
tst4@cval
```

```
##      1pct 5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Because the test statistic is greater than the critical value, we do not
reject the null hypothesis of a unit root.

If we test the seasonal differences - we still have a unit root:

```
tst4 <- summary(ur.df(diff(Y, 4), type = "none",
                      lags = 10, selectlags = "AIC"))
tst4@teststat
```

```
##                   tau1
## statistic -0.8321717
```

```
tst4@cval
```

```
##      1pct  5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Because the test statistic is greater than the critical value, we do not reject the null hypothesis of a unit root.

However, if we take the **seasonal** differences of the **differenced data** as well:

```
tst5 <- summary(ur.df(diff(diff(Y), 4), type = "none",
                      lags = 10, selectlags = "AIC"))
tst5@teststat


## tau1
## statistic -3.891775

tst5@cval


## 1pct 5pct 10pct
## tau1 -2.6 -1.95 -1.61
```

Since test statistic = -3.891 is *less* than the critical value, we reject the
null hypothesis and conclude that the seasonal differences of the
differenced series is a stationary process.

As mentioned before, Dickey-Fuller test is likely to be biased towards more differences (i.e. larger d value).

Alternatively, we can test via KPSS test:

```
tseries::kpss.test(Y)$p.value
```

```
## Warning in tseries::kpss.test(Y): p-value smaller than printe
```

```
## [1] 0.01
```

Since p-value < 0.05, we reject the null hypothesis **of no unit root**, i.e. our data has a unit root. Now let's look at the differences.

```
tseries::kpss.test(diff(Y))$p.value
```

```
## Warning in tseries::kpss.test(diff(Y)): p-value greater than
## value
```

```
## [1] 0.1
```

```
tseries::kpss.test(diff(Y, 4))$p.value
```

```
## [1] 0.06119338
```

Since p-value > 0.05 for either seasonal or first-order differences, we do not reject the null hypothesis **of no unit root**, i.e. the differenced data (either seasonally or first-order differenced) is stationary.

So, the **DF** test and the **KPSS** test give different results for the differenced series, which means that we cannot draw any concrete conclusions whether the differenced data is stationary or if it has a unit root. We **do know** from the test results that the initial time series has a unit root.

We can specify our model via `Arima`:

```
seas_arima <- forecast::Arima(Y,
                order = c(1, 1, 1),
                seasonal = list(order = c(1, 1, 1), period = 4),
                include.mean = FALSE)
seas_arima
```

```
## Series: Y
## ARIMA(1,1,1)(1,1,1)[4]
##
## Coefficients:
##           ar1     ma1     sar1    sma1
##       -0.4736  0.8007  -0.2299  0.3738
## s.e.   0.2905  0.2225   0.5315  0.4795
##
## sigma^2 estimated as 0.8797:  log likelihood=-72.61
## AIC=155.21   AICc=156.44   BIC=165.25
```

We can also do it via auto.arima (note that only seasonal differences are selected because we are using KPSS for unit root testing):

```
seas_arima_auto <- forecast::auto.arima(Y)
capture.output(summary(seas_arima_auto))[2]
```

```
## [1] "ARIMA(2,0,0)(0,1,0)[4] with drift "
```

```
plot.ts(Y)
lines(fitted(seas_arima), col = "red", lty = 2)
lines(fitted(seas_arima_auto), col = "blue", lty = 2)
legend(x = 1, y = 30, c("actual", "fitted", "fitted_auto"),
col = c("black", "red", "blue"), lty = c(1, 2, 2), cex = 0.7)
```
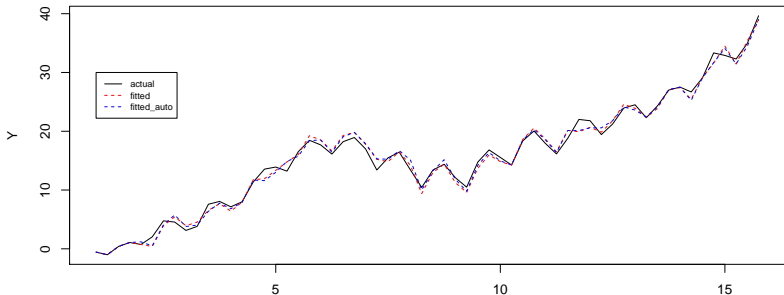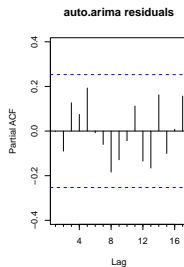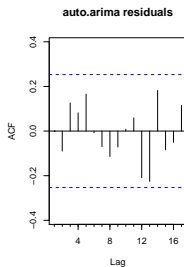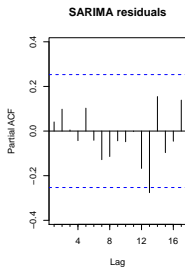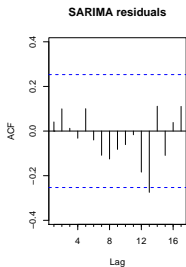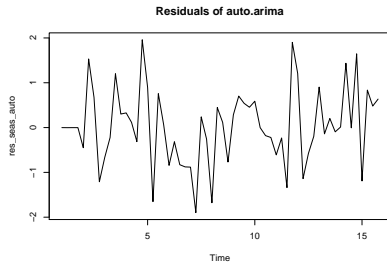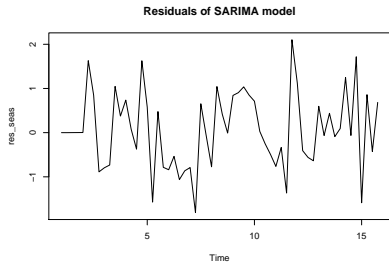
Note: it is important that Y_t is a *ts*() object with the correct frequency, i.e. `ts(y, freq = 4)`. Otherwise `auto.arima` cannot determine the frequency of a vector variable and will try to fit a non-seasonal *ARIMA* model which will still have the seasonal effect in the model residuals.

The smallest AIC value is of our manually specified model, although it is similar to the AIC from `auto.arima`.
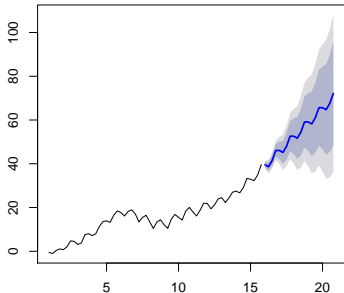
```
seas_arima$aic
```
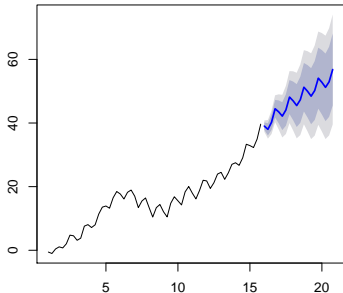
```
## [1] 155.2138
```

```
seas_arima_auto$aic
```

```
## [1] 155.67
```


Forecasts from ARIMA(1,1,1)(1,1,1)[4]   Forecasts from ARIMA(2,0,0)(0,1,0)[4] with drift

If we decide to use the Augmented Dickey-Fuller instead of the KPSS test for unit root testing in `auto.arima`:
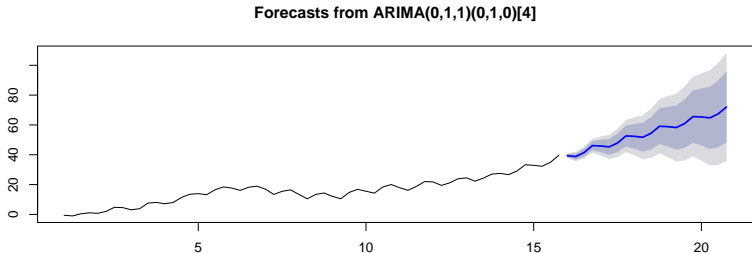
```r
seas_arima_auto_2 <- auto.arima(Y, test = "adf", max.d = 1)
capture.output(summary(seas_arima_auto_2))[2]
```

```
## [1] "ARIMA(0,1,1)(0,1,0)[4] "
```
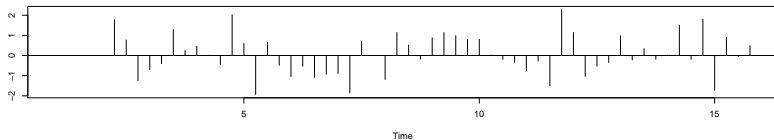
```r
seas_arima_auto_2$aic
```

```
## [1] 149.9099
```
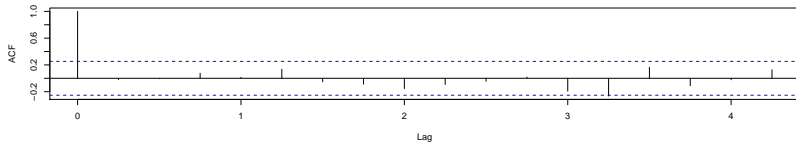
```r
plot(forecast(seas_arima_auto_2, 20))
```


Forecasts from ARIMA(0,1,1)(0,1,0)[4]

```
tsdiag(seas_arima_auto_2)
```


Standardized Residuals

ACF of Residuals

p values for Ljung–Box statistic