ARIMA processes

Stationary processes

Stationary processes: Plan

- Definition

Stationary processes: Plan

- Definition
- Random walk

Stationary processes: Plan

- Definition
- Random walk
- Autocovariance function

Stationary processes: Plan

- Definition
- Random walk
- Autocovariance function
- ACF, PACF

Stationary processes

A stochastic process whose characteristics do not change over time

Stationary processes

A stochastic process whose characteristics do not change over time

Weak or wide-sense stationarity

A process (y_t) is said to be weakly stationary, if for each t and k:

$$\begin{cases} \mathbb{E}(y_t) = \mu \\ \mathrm{Cov}(y_t, y_{t+k}) = \gamma_k \end{cases}$$

Stationary processes

A stochastic process whose characteristics do not change over time

Weak or wide-sense stationarity

A process (y_t) is said to be weakly stationary, if for each t and k:

$$\begin{cases} \mathbb{E}(y_t) = \mu \\ \mathrm{Cov}(y_t, y_{t+k}) = \gamma_k \end{cases}$$

Strong or strict-sense stationarity

A process (y_t) is said to be strictly stationary, if for each k joint distribution of a r.v. $(y_t, y_{t+1}, y_{t+2}, \ldots, y_{t+k})$ does not depend on t

Stationary process: example

Independent observations

The quantities (y_t) are independent and equally distributed with finite expectation μ_y and finite variance σ_y^2

Stationary process: example

Independent observations

The quantities (y_t) are independent and equally distributed with finite expectation μ_y and finite variance σ_y^2

$$\mu_y = \mathbb{E}(y_t)$$

Stationary process: example

Independent observations

The quantities (y_t) are independent and equally distributed with finite expectation μ_y and finite variance σ_y^2

$$\mu_y = \mathbb{E}(y_t)$$

$$\gamma_0 = \mathrm{Cov}(y_t, y_t) = \mathrm{Var}(y_t) = \sigma_y^2$$

Stationary process: example

Independent observations

The quantities (y_t) are independent and equally distributed with finite expectation μ_y and finite variance σ_y^2

$$\mu_y = \mathbb{E}(y_t)$$

$$\gamma_0 = \mathrm{Cov}(y_t, y_t) = \mathrm{Var}(y_t) = \sigma_y^2$$

$$\gamma_k = \mathrm{Cov}(y_t, y_{t+k}) = 0, \text{ for } k \geq 1$$

Non-Stationary Process Example

Random Walk

$$\begin{cases} y_0 = \mu \\ y_t = y_{t-1} + u_t, \text{ for } t \geq 1 \end{cases},$$

where u_t is white noise

Non-Stationary Process Example

Random Walk

$$\begin{cases} y_0 = \mu \\ y_t = y_{t-1} + u_t, \text{ for } t \geq 1 \end{cases},$$

where u_t is white noise

Explicitly: $y_t = \mu + u_1 + u_2 + \ldots + u_t$.

Non-Stationary Process Example

Random Walk

$$\begin{cases} y_0 = \mu \\ y_t = y_{t-1} + u_t, \text{ for } t \geq 1 \end{cases},$$

where u_t is white noise

Explicitly: $y_t = \mu + u_1 + u_2 + \ldots + u_t$.

$$\mu_y = \mathbb{E}(y_t)$$

Non-Stationary Process Example

Random Walk

$$\begin{cases} y_0 = \mu \\ y_t = y_{t-1} + u_t, \text{ for } t \geq 1 \end{cases},$$

where u_t is white noise

Explicitly: $y_t = \mu + u_1 + u_2 + \ldots + u_t$.

$$\mu_y = \mathbb{E}(y_t)$$

$$\gamma_0 = \mathrm{Cov}(y_t, y_t) = \mathrm{Var}(y_t) = \mathrm{Var}(\mu + u_1 + \ldots + u_t) = t\sigma_u^2$$

Non-Stationary Process Example

Random Walk

$$\begin{cases} y_0 = \mu \\ y_t = y_{t-1} + u_t, \text{ for } t \geq 1 \end{cases},$$

where u_t is white noise

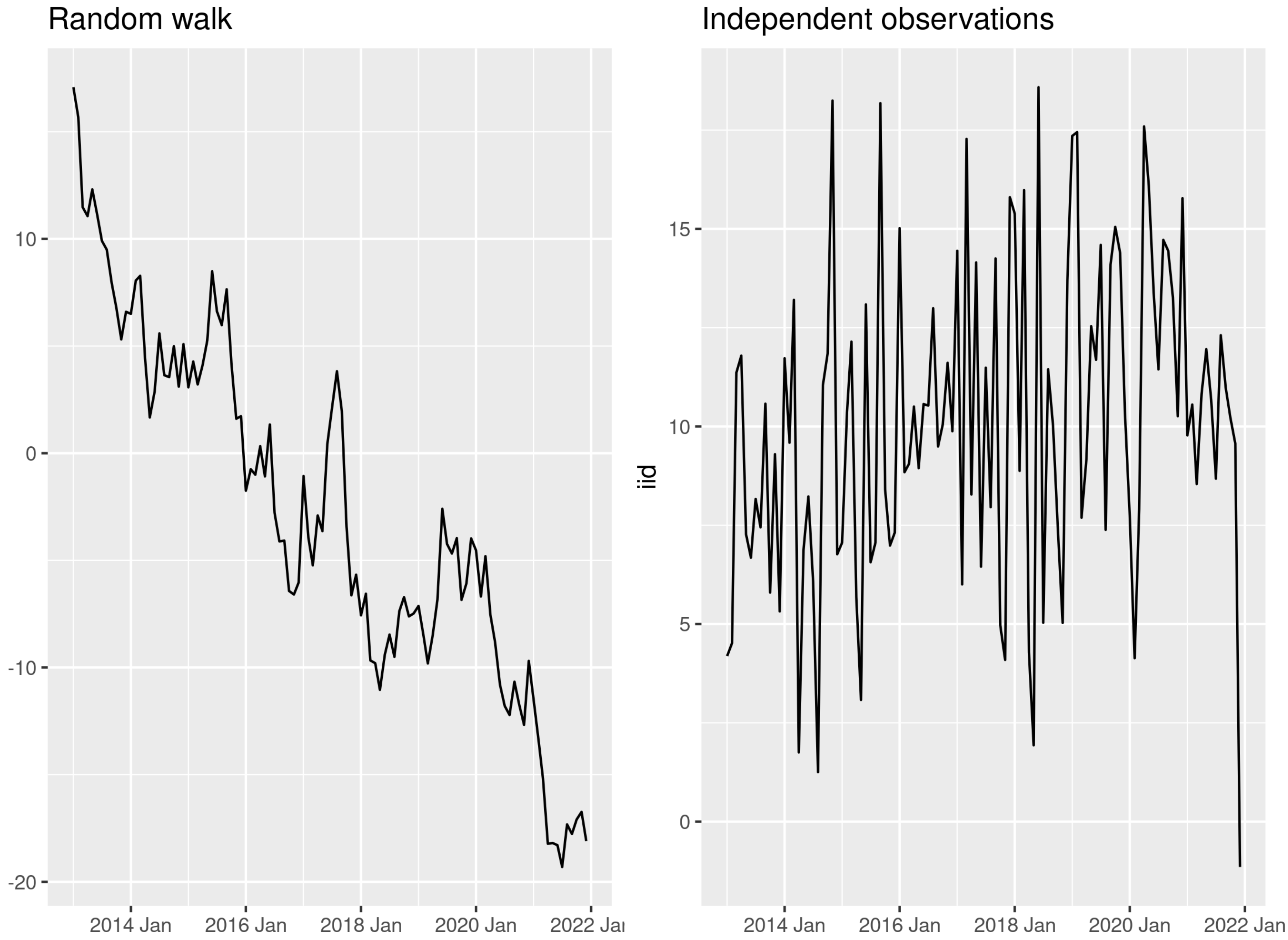
Explicitly: $y_t = \mu + u_1 + u_2 + \ldots + u_t$.

$$\mu_y = \mathbb{E}(y_t)$$

$$\gamma_0 = \mathrm{Cov}(y_t, y_t) = \mathrm{Var}(y_t) = \mathrm{Var}(\mu + u_1 + \ldots + u_t) = t\sigma_u^2$$

$$\gamma_k = \mathrm{Cov}(y_t, y_{t+k}) = \mathrm{Cov}(y_t, y_t + u_{t+1} + \ldots + u_{t+k}) = \mathrm{Var}(y_t)$$

Random Walk vs Random Sample



Autocovariance function

Definition

For a stationary process (y_t), the function $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$ is called autocovariance

Autocovariance function

Definition

For a stationary process (y_t), the function $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$ is called autocovariance

Definition

For a stationary process (y_t), the function $\rho_k = \mathrm{Corr}(y_t, y_{t+k})$ is called autocorrelation

ACF: autocorrelation function

$$\rho_k = \mathrm{Corr}(y_t, y_{y+j}) = \frac{\mathrm{Cov}(y_t, y_{y+j})}{\sqrt{\mathrm{Var}(y_t)\,\mathrm{Var}(y_{t+k})}} =$$

ACF: autocorrelation function

$$\rho_k = \operatorname{Corr}(y_t, y_{y+j}) = \frac{\operatorname{Cov}(y_t, y_{y+j})}{\sqrt{\operatorname{Var}(y_t)\operatorname{Var}(y_{t+k})}} = \frac{\gamma_k}{\sqrt{\gamma_0 \gamma_0}} = \frac{\gamma_k}{\gamma_0}$$

Partial Correlation

Definition

$$\mathrm{pCorr}(U, D; R_1, R_2, \ldots, R_n) = \mathrm{Corr}(U^*, D^*), \text{ where}$$
$$U^* = U - Best(U; R_1, R_2, \ldots, R_n),$$
$$D^* = D - Best(D; R_1, R_2, \ldots, R_n)$$

Partial Correlation

Definition

$$\mathrm{pCorr}(U, D; R_1, R_2, \ldots, R_n) = \mathrm{Corr}(U^*, D^*), \text{ where}$$
$$U^* = U - Best(U; R_1, R_2, \ldots, R_n),$$
$$D^* = D - Best(D; R_1, R_2, \ldots, R_n)$$

The values U^* and D^* are the versions of U and D uninfluenced by the covariates $R_1, \ldots, R_n$

$$\mathrm{Cov}(U^*, R_i) = 0, \quad \mathrm{Cov}(D^*, R_i) = 0.$$

PACF

Definition

For a stationary process (y_t) the function

$$\varphi_{kk} = \mathrm{pCorr}(y_t, y_{t+k}; y_{t+1}, \ldots, y_{t+k-1})$$

is called partial autocorrelation

ACF and PACF: Intuition

For stationary process

- ACF:

$$\rho_k = \mathrm{Corr}(y_t, y_{t+k})$$

  Joint strength of relationship between y_t and y_{t+k}

- PACF:

$$\varphi_{kk} = \mathrm{pCorr}(y_t, y_{t+k}; y_{t+1}, \ldots, y_{t+k-1})$$

  Strength of relationship between y_t and y_{t+k} with the links through intermediate observations being broken

Stationary processes: Summary

- Constants $\mathbb{E}(y_t)$, $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$

Stationary processes: Summary

- Constants $\mathbb{E}(y_t)$, $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$
- The random sample is stationary

Stationary processes: Summary

- Constants $\mathbb{E}(y_t)$, $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$
- The random sample is stationary
- Random walk is non-stationary

Stationary processes: Summary

- Constants $\mathbb{E}(y_t)$, $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$
- The random sample is stationary
- Random walk is non-stationary
- Autocovariance function

Stationary processes: Summary

- Constants $\mathbb{E}(y_t)$, $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$
- The random sample is stationary
- Random walk is non-stationary
- Autocovariance function
- Partial correlation — correlation with the effect of a set of controlling random variables removed

Stationary processes: Summary

- Constants $\mathbb{E}(y_t)$, $\gamma_k = \mathrm{Cov}(y_t, y_{t+k})$
- The random sample is stationary
- Random walk is non-stationary
- Autocovariance function
- Partial correlation — correlation with the effect of a set of controlling random variables removed
- In the time series, we removed the effect of intermediate observations

MA Process

MA Process: Plan

- Definition and notations with lags

MA Process: Plan

- Definition and notations with lags
- Stationarity

MA Process: Plan

- Definition and notations with lags
- Stationarity
- Predictability

MA Process: Plan

- Definition and notations with lags
- Stationarity
- Predictability
- Invertibility

Lag operator

Definition

For the process (y_t) defined at $t \in \mathbb{Z}$, lagged process Ly_t is the same sequence of values with a shifted index,

$$Ly_t = y_{t-1}$$

Lag operator

Definition

For the process (y_t) defined at $t \in \mathbb{Z}$, lagged process Ly_t is the same sequence of values with a shifted index,

$$Ly_t = y_{t-1}$$

$$L^2 y_t = L \cdot L \cdot y_t = L \cdot y_{t-1} = y_{t-2}$$

Lag operator

Definition

For the process (y_t) defined at $t \in \mathbb{Z}$, lagged process Ly_t is the same sequence of values with a shifted index,

$$Ly_t = y_{t-1}$$

$$L^2 y_t = L \cdot L \cdot y_t = L \cdot y_{t-1} = y_{t-2}$$

$$\Delta y_t = y_t - y_{t-1} = (1 - L) y_t$$

Lag operator

Definition

For the process (y_t) defined at $t \in \mathbb{Z}$, lagged process Ly_t is the same sequence of values with a shifted index,

$$Ly_t = y_{t-1}$$

$$L^2 y_t = L \cdot L \cdot y_t = L \cdot y_{t-1} = y_{t-2}$$

$$\Delta y_t = y_t - y_{t-1} = (1 - L) y_t$$

$$\Delta_{12} y_t = y_t - y_{t-12} = (1 - L^{12}) y_t$$

MA process

Definition

Process (y_t), which can be represented as

$$y_t = \mu + u_t + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q},$$

where $\alpha_q \neq 0$ and (u_t) is white noise, is called the $MA(q)$ process

MA — Moving Average

MA process

Definition

Process (y_t), which can be represented as

$$y_t = \mu + u_t + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q},$$

where $\alpha_q \neq 0$ and (u_t) is white noise, is called the $MA(q)$ process

MA — Moving Average

Example $MA(1)$ process:

$$y_t = 5 + u_t + 0.3 u_{t-1},$$

where (u_t) is some white noise

Notation with lags

MA with lag polynomial

Process (y_t), which can be represented as

$$y_t = \mu + P(L)u_t,$$

where $P(L)$ is a polynomial of degree q in lag L with $P(0) = 1$, and (u_t) is white noise, is called $MA(q)$ a process

Notation with lags

MA with lag polynomial

Process (y_t), which can be represented as

$$y_t = \mu + P(L)u_t,$$

where $P(L)$ is a polynomial of degree q in lag L with $P(0) = 1$, and (u_t) is white noise, is called $MA(q)$ a process

An example $MA(2)$ process:

$$y_t = 5 + (1 - 0.2L + 0.3L^2)u_t,$$

where (u_t) is white noise

ACF and Forecasts

Traditionally $MA(q)$ process is evaluated assuming joint normality of (y_t).

ACF and Forecasts

Traditionally $MA(q)$ process is evaluated assuming joint normality of (y_t).

Zero $\rho_k = 0$ for $k > q$ implies independence of y_t and y_{t+k}.

ACF and Forecasts

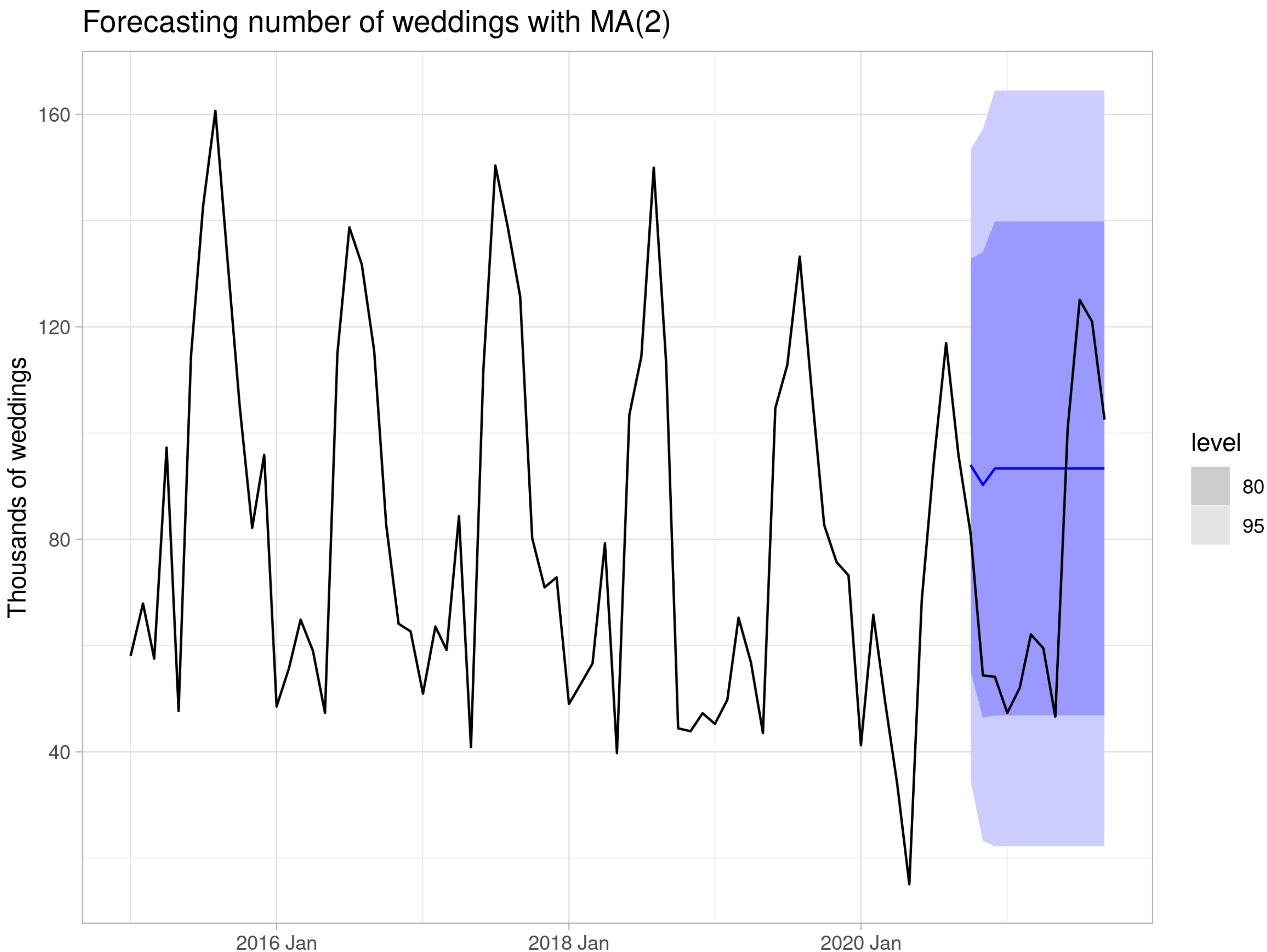
Traditionally $MA(q)$ process is evaluated assuming joint normality of (y_t).

Zero $\rho_k = 0$ for $k > q$ implies independence of y_t and y_{t+k}.

Forecasts more than q steps ahead are exactly the same.

$$(y_{T+q+1} \mid \mathcal{F}_T) \sim (y_{T+q+2} \mid \mathcal{F}_T) \sim (y_{T+q+3} \mid \mathcal{F}_T) \sim \ldots$$

Predictions for $MA(2)$



$MA(\infty)$

Definition

Process (y_t), which can be represented as

$$y_t = \mu + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \dots,$$

where (u_t) is white noise, an infinite number of $\alpha_i \neq 0$ and $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$, is called the $MA(\infty)$ process

$MA(\infty)$

Definition

Process (y_t), which can be represented as

$$y_t = \mu + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \dots,$$

where (u_t) is white noise, an infinite number of $\alpha_i \neq 0$ and $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$, is called the $MA(\infty)$ process

$MA(\infty)$:

$$y_t = 5 + u_t + 0.5u_{t-1} + 0.5^2 u_{t-2} + 0.5^3 u_{t-3} + \dots$$

$MA(\infty)$

Definition

Process (y_t), which can be represented as

$$y_t = \mu + u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \dots,$$

where (u_t) is white noise, an infinite number of $\alpha_i \neq 0$ and $\sum_{i=1}^{\infty} \alpha_i^2 < \infty$, is called the $MA(\infty)$ process

$MA(\infty)$:

$$y_t = 5 + u_t + 0.5 u_{t-1} + 0.5^2 u_{t-2} + 0.5^3 u_{t-3} + \dots$$

And this is not allowed:

$$y_t = 5 + u_t + \frac{1}{\sqrt{2}} u_{t-1} + \frac{1}{\sqrt{3}} u_{t-2} + \frac{1}{\sqrt{4}} u_{t-3} + \dots$$

Convergences

Theorem

If a $\sum_{i=0}^{\infty} \alpha_i^2 < \infty$ and (u_t) is a zero-mean stationary process, then the sequence of partial sums y_t^q of the form

$$y_t^q = \mu + \sum_{i=0}^{q} \alpha_i u_{t-i}$$

converges for $q \to \infty$ in mean, in probability, and in distribution

Nuance: the convergence of the weighted sum is guaranteed for the stationary (u_t)

Convergences

Theorem

If a $\sum_{i=0}^{\infty} \alpha_i^2 < \infty$ and (u_t) is a zero-mean stationary process, then the sequence of partial sums y_t^q of the form

$$y_t^q = \mu + \sum_{i=0}^{q} \alpha_i u_{t-i}$$

converges for $q \to \infty$ in mean, in probability, and in distribution

Nuance: the convergence of the weighted sum is guaranteed for the stationary (u_t)

Bonus

…and the resulting process (y_t) is stationary

Wald's Theorem

Theorem

If (y_t) is a stationary process, then it can be represented as:

$$y_t = \sum_{i=0}^{\infty} \alpha_i u_{t-i} + r_t,$$

where

- (u_t) — white noise,
- $\sum \alpha_i^2 < \infty$,
- r_t is a linear predictable random process,
- $\operatorname{Cov}(u_t, r_t) = 0$

Predictable Process

Correct definition

A process (r_t) is called linearly predictable if

- (r_t) is stationary,
- $r_t = \beta_0 + \beta_1 r_{t-1} + \beta_2 r_{t-2} + \ldots + \beta_p r_{t-p}$

Invertibility condition

Characteristic representation

The equation $MA(q)$ of the process satisfies the invertibility condition if the characteristic polynomial $\phi(\lambda)$ has all roots $|\lambda_i| < 1$

Invertibility condition

Characteristic representation

The equation $MA(q)$ of the process satisfies the invertibility condition if the characteristic polynomial $\phi(\lambda)$ has all roots $|\lambda_i| < 1$

Lag representation

The equation $MA(q)$ of the process satisfies the invertibility condition if all roots of the lag polynomial $P(L)$ are $|\ell_i| > 1$

Example of invertible $MA(1)$ notation

$$y_t = 5 + u_t + 0.5u_{t-1}, \quad \sigma_u^2 = 4$$

Example of invertible $MA(1)$ notation

$$y_t = 5 + u_t + 0.5u_{t-1}, \quad \sigma_u^2 = 4$$

$$\lambda^1 + 0.5 \cdot \lambda^0$$

$$\phi(\lambda) = \lambda + 0.5$$

Example of invertible $MA(1)$ notation

$$y_t = 5 + u_t + 0.5u_{t-1}, \quad \sigma_u^2 = 4$$

$$\lambda^1 + 0.5 \cdot \lambda^0$$

$$\phi(\lambda) = \lambda + 0.5$$

$$\lambda_1 = -0.5$$

Example of noninvertible $MA(1)$ notation

$$y_t = 5 + u_t + 2u_{t-1}, \quad \sigma_u^2 = 1$$

Example of noninvertible $MA(1)$ notation

$$y_t = 5 + u_t + 2u_{t-1}, \quad \sigma_u^2 = 1$$

$$\lambda^1 + 2 \cdot \lambda^0$$

$$\phi(\lambda) = \lambda + 2$$

Example of noninvertible $MA(1)$ notation

$$y_t = 5 + u_t + 2u_{t-1}, \quad \sigma_u^2 = 1$$

$$\lambda^1 + 2 \cdot \lambda^0$$

$$\phi(\lambda) = \lambda + 2$$

$$\lambda_1 = -2$$

Nuance

Difference

Stationarity is a property of the (y_t) process itself.

Invertibility is a property of the equation (process notation) for (y_t).

Nuance

Difference

Stationarity is a property of the (y_t) process itself.

Invertibility is a property of the equation (process notation) for (y_t).

$MA(q)$ has a single notation when invertible

MA process: Summary

- $MA(q)$ — weighting of several white noises

MA process: Summary

- $MA(q)$ — weighting of several white noises
- $MA(q)$ is a stationary process

MA process: Summary

- $MA(q)$ — weighting of several white noises
- $MA(q)$ is a stationary process
- ACF vanishes sharply, $PACF$ tends to zero

MA process: Summary

- $MA(q)$ — weighting of several white noises
- $MA(q)$ is a stationary process
- ACF vanishes sharply, $PACF$ tends to zero
- Invertibility condition: roots of the characteristic polynomial $|\lambda_i| < 1$ or roots of the lag polynomial $|\ell_i| > 1$.

ARMA equation

ARMA equation: Plan

- Definition

ARMA equation: Plan

- Definition
- Non-uniqueness of solutions

About the purpose of old problems

Goal: A simple equation for a wide variety of processes

About the purpose of old problems

Goal: A simple equation for a wide variety of processes

Problems:

- Non-uniqueness of equation for one process

About the purpose of old problems

Goal: A simple equation for a wide variety of processes

Problems:

- Non-uniqueness of equation for one process

Requirement: invertibility of the equation

About the purpose of old problems

Goal: A simple equation for a wide variety of processes

Problems:

- Non-uniqueness of equation for one process

Requirement: invertibility of the equation

- $MA(\infty)$ has infinite number of parameters

About the purpose of old problems

Goal: A simple equation for a wide variety of processes

Problems:

- Non-uniqueness of equation for one process

Requirement: invertibility of the equation

- $MA(\infty)$ has infinite number of parameters

Let's try adding lags y_t to the equation!

New problem

$y_t - y_{t-1} = u_t - u_{t-1}$, where (u_t) is white noise

New problem

$$y_t - y_{t-1} = u_t - u_{t-1}, \text{ where } (u_t) \text{ is white noise}$$

Solutions:

New problem

$$y_t - y_{t-1} = u_t - u_{t-1},\text{ where } (u_t) \text{ is white noise}$$

Solutions:

- $y_t = u_t$;

New problem

$$y_t - y_{t-1} = u_t - u_{t-1},$$ where (u_t) is white noise

Solutions:

- $y_t = u_t$;
- $y_t = u_t - 0.7$;

New problem

$$y_t - y_{t-1} = u_t - u_{t-1}, \text{ where } (u_t) \text{ is white noise}$$

Solutions:

- $y_t = u_t$;
- $y_t = u_t - 0.7$;
- $y_t = u_t - 0.8$

New problem

$$y_t - y_{t-1} = u_t - u_{t-1}, \text{ where } (u_t) \text{ is white noise}$$

Solutions:

- $y_t = u_t$;
- $y_t = u_t - 0.7$;
- $y_t = u_t - 0.8$

New problem

$$y_t - y_{t-1} = u_t - u_{t-1}, \text{ where } (u_t) \text{ is white noise}$$

Solutions:

- $y_t = u_t$;
- $y_t = u_t - 0.7$;
- $y_t = u_t - 0.8$

Infinite number of solutions

ARMA equation

Definition

Equation

$$y_t = c + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + u_t + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q},$$

where (u_t) is white noise, we'll call an $ARMA$ equation

ARMA — Autoregression and Moving Average

ARMA equation

Definition

Equation

$$y_t = c + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + u_t + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q},$$

where (u_t) is white noise, we'll call an $ARMA$ equation

ARMA — Autoregression and Moving Average

Definition

An equation of the form $P(L)y_t = c + Q(L)u_t$, where (u_t) is white noise, $P(L)$ and $Q(L)$ are lag polynomials with $P(0) = Q(0) = 1$, we'll call an $ARMA$ equation

ARMA equation: Summary

An equation is not a process!

ARMA equation: Summary

An equation is not a process!

Why?

- One equation has many solutions

ARMA equation: Summary

An equation is not a process!

Why?

- One equation has many solutions
- One process can be described by several equations

ARMA process

ARMA process

- Equation irreducibility

ARMA process

- Equation irreducibility
- Solution structure

ARMA process

- Equation irreducibility
- Solution structure
- ARMA process

Irreducibility of Equation

Definition

$ARMA$ an equation of the form $P(L)y_t = c + Q(L)u_t$ is called irreducible, if the polynomials $P(L)$ and $Q(L)$ do not have common roots

Irreducibility of Equation

Definition

$ARMA$ an equation of the form $P(L)y_t = c + Q(L)u_t$ is called irreducible, if the polynomials $P(L)$ and $Q(L)$ do not have common roots

Reducible Equation:

$$y_t - y_{t-1} = u_t - u_{t-1} \text{ or } (1-L)y_t = (1-L)u_t$$

Irreducibility of Equation

Definition

$ARMA$ an equation of the form $P(L)y_t = c + Q(L)u_t$ is called irreducible, if the polynomials $P(L)$ and $Q(L)$ do not have common roots

Reducible Equation:

$$y_t - y_{t-1} = u_t - u_{t-1} \text{ or } (1-L)y_t = (1-L)u_t$$

Irreducible equation:

$$y_t - y_{t-1} = u_t - 0.5u_{t-1} \text{ or } (1-L)y_t = (1-0.5L)u_t$$

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Let's try different initial conditions:

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Let's try different initial conditions:

- $y_0 = 0$

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Let's try different initial conditions:

- $y_0 = 0$

$y_1 = u_1, y_2 = u_2 + 0.5u_1, y_3 = u_3 + 0.5u_2 + 0.25u_1, \ldots$

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Let's try different initial conditions:

- $y_0 = 0$

$y_1 = u_1, y_2 = u_2 + 0.5u_1, y_3 = u_3 + 0.5u_2 + 0.25u_1, \ldots$

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Let's try different initial conditions:

- $y_0 = 0$

$y_1 = u_1, y_2 = u_2 + 0.5u_1, y_3 = u_3 + 0.5u_2 + 0.25u_1, \ldots$

- $y_0 = 2u_1$

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Let's try different initial conditions:

- $y_0 = 0$

$y_1 = u_1, y_2 = u_2 + 0.5u_1, y_3 = u_3 + 0.5u_2 + 0.25u_1, \ldots$

- $y_0 = 2u_1$

$y_1 = 2u_1, y_2 = u_2 + u_1, y_3 = u_3 + 0.5u_2 + 0.5u_1, \ldots$

Initial Conditions

Irreducible $ARMA$ equation:

$$y_t = 0.5 y_{t-1} + u_t, \text{ where } (u_t) \text{ is white noise}$$

Let's try different initial conditions:

- $y_0 = 0$

$y_1 = u_1, y_2 = u_2 + 0.5u_1, y_3 = u_3 + 0.5u_2 + 0.25u_1, \dots$

- $y_0 = 2u_1$

$y_1 = 2u_1, y_2 = u_2 + u_1, y_3 = u_3 + 0.5u_2 + 0.5u_1, \dots$

The initial conditions also determine the past y_t!

ARMA Solutions

Theorem I

Any $ARMA$ equation with at least one y_t lag has an infinite number of solutions

ARMA Solutions

Theorem I

Any $ARMA$ equation with at least one y_t lag has an infinite number of solutions

Theorem II

In order to obtain a unique solution of an $ARMA$ equation of the form $P(L)y_t = c + Q(L)u_t$, it suffices to specify the initial conditions in an amount equal to the power of $P(L)$

ARMA Solutions

Theorem I

Any $ARMA$ equation with at least one y_t lag has an infinite number of solutions

Theorem II

In order to obtain a unique solution of an $ARMA$ equation of the form $P(L)y_t = c + Q(L)u_t$, it suffices to specify the initial conditions in an amount equal to the power of $P(L)$

$$y_t = 0.6y_{t-1} + 0.08y_{t-2} + u_t$$

ARMA Solutions

Theorem I

Any $ARMA$ equation with at least one y_t lag has an infinite number of solutions

Theorem II

In order to obtain a unique solution of an $ARMA$ equation of the form $P(L)y_t = c + Q(L)u_t$, it suffices to specify the initial conditions in an amount equal to the power of $P(L)$

$$y_t = 0.6y_{t-1} + 0.08y_{t-2} + u_t \text{ and } y_0 = u_0, y_1 = u_0 + 4$$

And how many stationary solutions?

Correct theorem

If an $ARMA$ equation $P(L)y_t = c + Q(L)u_t$ is irreducible, then it

- has exactly one stationary solution if the lag polynomial $P(\ell)$ has all roots $|\ell_i| \neq 1$;
- has no stationary solutions if the lag polynomial $P(\ell)$ has a root with $|\ell_i| = 1$

And how many stationary solutions?

Correct theorem

If an $ARMA$ equation $P(L)y_t = c + Q(L)u_t$ is irreducible, then it

- has exactly one stationary solution if the lag polynomial $P(\ell)$ has all roots $|\ell_i| \neq 1$;
- has no stationary solutions if the lag polynomial $P(\ell)$ has a root with $|\ell_i| = 1$

- $y_t = 0.5y_{t-1} + u_t, P(L) = 1 - 0.5L, \ell_1 = 2$:

And how many stationary solutions?

Correct theorem

If an $ARMA$ equation $P(L)y_t = c + Q(L)u_t$ is irreducible, then it

- has exactly one stationary solution if the lag polynomial $P(\ell)$ has all roots $|\ell_i| \neq 1$;
- has no stationary solutions if the lag polynomial $P(\ell)$ has a root with $|\ell_i| = 1$

- $y_t = 0.5y_{t-1} + u_t, P(L) = 1 - 0.5L, \ell_1 = 2$: one stationary solution;

And how many stationary solutions?

Correct theorem

If an $ARMA$ equation $P(L)y_t = c + Q(L)u_t$ is irreducible, then it

- has exactly one stationary solution if the lag polynomial $P(\ell)$ has all roots $|\ell_i| \neq 1$;
- has no stationary solutions if the lag polynomial $P(\ell)$ has a root with $|\ell_i| = 1$

- $y_t = 0.5y_{t-1} + u_t, P(L) = 1 - 0.5L, \ell_1 = 2$: one stationary solution;
- $y_t = y_{t-1} + u_t, P(L) = 1 - L, \ell_1 = 1$: no stationary solutions

AR process

Definition

$AR(p)$ process with equation

$$y_t = c + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + u_t,$$

where (u_t) is white noise and $\beta_p \neq 0$, is the solution of this equation in the form of $MA(\infty)$ with respect to (u_t)

AR process

Definition

$AR(p)$ process with equation

$$y_t = c + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + u_t,$$

where (u_t) is white noise and $\beta_p \neq 0$, is the solution of this equation in the form of $MA(\infty)$ with respect to (u_t)

Definition with lags

$AR(p)$ process with equation

$$P(L)y_t = c + u_t,$$

where (u_t) is white noise, $P(L)$ has power p and $P(0) = 1$, is the solution of this equation in the form of $MA(\infty)$ with respect to (u_t)

Definitions by different authors

In our definition of $AR(p)$, the process is necessarily stationary.

Definitions by different authors

In our definition of $AR(p)$, the process is necessarily stationary.

Some authors do not include the requirement of stationarity in the definition of $AR(p)$.

ARMA process

Definition

$ARMA(p,q)$ process with equation

$$y_t = c + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + u_t + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q},$$

where (u_t) is white noise, $\beta_p \neq 0$, $\alpha_q \neq 0$ and the equation is irreducible, is the solution of this equation in the form of $MA(\infty)$ with respect to (u_t)

ARMA process

Definition

$ARMA(p, q)$ process with equation

$$y_t = c + \beta_1 y_{t-1} + \ldots + \beta_p y_{t-p} + u_t + \alpha_1 u_{t-1} + \ldots + \alpha_q u_{t-q},$$

where (u_t) is white noise, $\beta_p \neq 0$, $\alpha_q \neq 0$ and the equation is irreducible, is the solution of this equation in the form of $MA(\infty)$ with respect to (u_t)

Definition with lags

$ARMA(p, q)$ process with equation

$$P(L)y_t = c + Q(L)u_t,$$

where (u_t) is white noise, $P(L)$ and $Q(L)$ have powers p and q, respectively, are irreducible and $P(0) = Q(0) = 1$, is the solution of this equation in the form $MA(\infty)$ with respect to (u_t)

What about non-uniqueness?

The same $ARMA(p, q)$ process (y_t) can be described by different equations!

What about non-uniqueness?

The same $ARMA(p,q)$ process (y_t) can be described by different equations!

Invertibility

If the series (y_t) is an $ARMA(p,q)$ process with the equation $P(L)y_t = c + Q(L)u_t$, then this equation will be unique if the MA part satisfies the invertibility condition.

What about non-uniqueness?

The same $ARMA(p, q)$ process (y_t) can be described by different equations!

Invertibility

If the series (y_t) is an $ARMA(p, q)$ process with the equation $P(L)y_t = c + Q(L)u_t$, then this equation will be unique if the MA part satisfies the invertibility condition.

Condition for invertibility of the $ARMA$ equation:

- the characteristic polynomial $\phi_{MA}(\lambda)$ has all roots $|\lambda_i| < 1$;
- the lag polynomial $Q(L)$ has all roots $|\ell_i| > 1$

ARMA process: Summary

- The process is stationary by definition

ARMA process: Summary

- The process is stationary by definition
- AR and MA processes are special cases of the $ARMA$ process

ARMA process: Summary

- The process is stationary by definition
- AR and MA processes are special cases of the $ARMA$ process
- Theoretical ACF and $PACF$ decrease exponentially

ARMA process: Summary

- The process is stationary by definition
- AR and MA processes are special cases of the $ARMA$ process
- Theoretical ACF and $PACF$ decrease exponentially
- The invertibility condition guarantees uniqueness

ARMA process: Summary

- The process is stationary by definition
- AR and MA processes are special cases of the $ARMA$ process
- Theoretical ACF and $PACF$ decrease exponentially
- The invertibility condition guarantees uniqueness
- An irreducible equation either has a unique stationary solution or it does not exist

ARIMA process

ARIMA process: Plan

- Stationarity of $ARMA$

ARIMA process: Plan

- Stationarity of $ARMA$
- Definition of $ARIMA$

ARIMA process: Plan

- Stationarity of $ARMA$
- Definition of $ARIMA$
- Differencing

Nuances

- Process $y_t \sim ARMA(p, q)$ is stationary by definition:

Nuances

- Process $y_t \sim ARMA(p,q)$ is stationary by definition: $\mathbb{E}(y_t) = \mu_y$, $\mathrm{Var}(y_t) = \gamma_0$, $\mathrm{Cov}(y_t, y_{t-k}) = \gamma_k$

Nuances

- Process $y_t \sim ARMA(p,q)$ is stationary by definition: $\mathbb{E}(y_t) = \mu_y$, $\mathrm{Var}(y_t) = \gamma_0$, $\mathrm{Cov}(y_t, y_{t-k}) = \gamma_k$
- In the canonical notation $ARMA(p,q)$ of the process $P(L)y_t = c + Q(L)u_t$ for the polynomial $P(L)$ all roots $|\ell| > 1$

Nuances

- Process $y_t \sim ARMA(p, q)$ is stationary by definition: $\mathbb{E}(y_t) = \mu_y$, $\mathrm{Var}(y_t) = \gamma_0$, $\mathrm{Cov}(y_t, y_{t-k}) = \gamma_k$
- In the canonical notation $ARMA(p, q)$ of the process $P(L)y_t = c + Q(L)u_t$ for the polynomial $P(L)$ all roots $|\ell| > 1$
- When estimating the $ARMA(p, q)$ process by the maximum likelihood method, these restrictions are imposed a priori

What to do with non-stationary processes?

Definition

The random process (y_t) is called the $ARIMA(p, 1, q)$ w.r.t. the white noise process (u_t), if (y_t) is non-stationary, but Δy_t is a stationary $ARMA(p, q)$ process w.r.t. the white noise (u_t)

What to do with non-stationary processes?

Definition

The random process (y_t) is called the $ARIMA(p, 1, q)$ w.r.t. the white noise process (u_t), if (y_t) is non-stationary, but Δy_t is a stationary $ARMA(p, q)$ process w.r.t. the white noise (u_t)

Definition

The random process (y_t) is called the $ARIMA(p, 2, q)$ w.r.t. the white noise process (u_t), if (y_t) and (Δy_t) are non-stationary, but $\Delta^2 y_t$ is a stationary $ARMA(p, q)$ process w.r.t. the white noise (u_t)

What to do with non-stationary processes?

Definition

The random process (y_t) is called the $ARIMA(p,1,q)$ w.r.t. the white noise process (u_t), if (y_t) is non-stationary, but Δy_t is a stationary $ARMA(p,q)$ process w.r.t. the white noise (u_t)

Definition

The random process (y_t) is called the $ARIMA(p,2,q)$ w.r.t. the white noise process (u_t), if (y_t) and (Δy_t) are non-stationary, but $\Delta^2 y_t$ is a stationary $ARMA(p,q)$ process w.r.t. the white noise (u_t)

$\Delta y_t = y_t - y_{t-1}$ and $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$

What to do with non-stationary processes?

Definition

The random process (y_t) is called the $ARIMA(p,1,q)$ w.r.t. the white noise process (u_t), if (y_t) is non-stationary, but Δy_t is a stationary $ARMA(p,q)$ process w.r.t. the white noise (u_t)

Definition

The random process (y_t) is called the $ARIMA(p,2,q)$ w.r.t. the white noise process (u_t), if (y_t) and (Δy_t) are non-stationary, but $\Delta^2 y_t$ is a stationary $ARMA(p,q)$ process w.r.t. the white noise (u_t)

$\Delta y_t = y_t - y_{t-1}$ and $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$

ARIMA — AutoRegressive Integrated Moving Average

How to choose?

$ARIMA(p, 0, q)$ or $ARIMA(p, 1, q)$ or $ARIMA(p, 2, q)$

How to choose?

$ARIMA(p, 0, q)$ or $ARIMA(p, 1, q)$ or $ARIMA(p, 2, q)$

- Analyse the graph:

How to choose?

$ARIMA(p, 0, q)$ or $ARIMA(p, 1, q)$ or $ARIMA(p, 2, q)$

- Analyse the graph: stationary process graph oscillates around its mean with a constant deviation

How to choose?

$ARIMA(p, 0, q)$ or $ARIMA(p, 1, q)$ or $ARIMA(p, 2, q)$

- Analyse the graph: stationary process graph oscillates around its mean with a constant deviation
- Evaluate all the models and choose the best one by cross-validation

How to choose?

$ARIMA(p, 0, q)$ or $ARIMA(p, 1, q)$ or $ARIMA(p, 2, q)$

- Analyse the graph: stationary process graph oscillates around its mean with a constant deviation
- Evaluate all the models and choose the best one by cross-validation Time consuming!

How to choose?

$ARIMA(p, 0, q)$ or $ARIMA(p, 1, q)$ or $ARIMA(p, 2, q)$

- Analyse the graph: stationary process graph oscillates around its mean with a constant deviation
- Evaluate all the models and choose the best one by cross-validation Time consuming!
- You cannot use AIC!

How to choose?

$ARIMA(p,0,q)$ or $ARIMA(p,1,q)$ or $ARIMA(p,2,q)$

- Analyse the graph: stationary process graph oscillates around its mean with a constant deviation
- Evaluate all the models and choose the best one by cross-validation Time consuming!
- You cannot use AIC!
  $\ln L(y_1,\ldots,y_n \mid \theta)$ and $\ln L(y_2,\ldots,y_n \mid \theta, y_1)$ and $\ln L(y_3,\ldots,y_n \mid \theta, y_1, y_2)$ incomparable!

How to choose?

$ARIMA(p,0,q)$ or $ARIMA(p,1,q)$ or $ARIMA(p,2,q)$

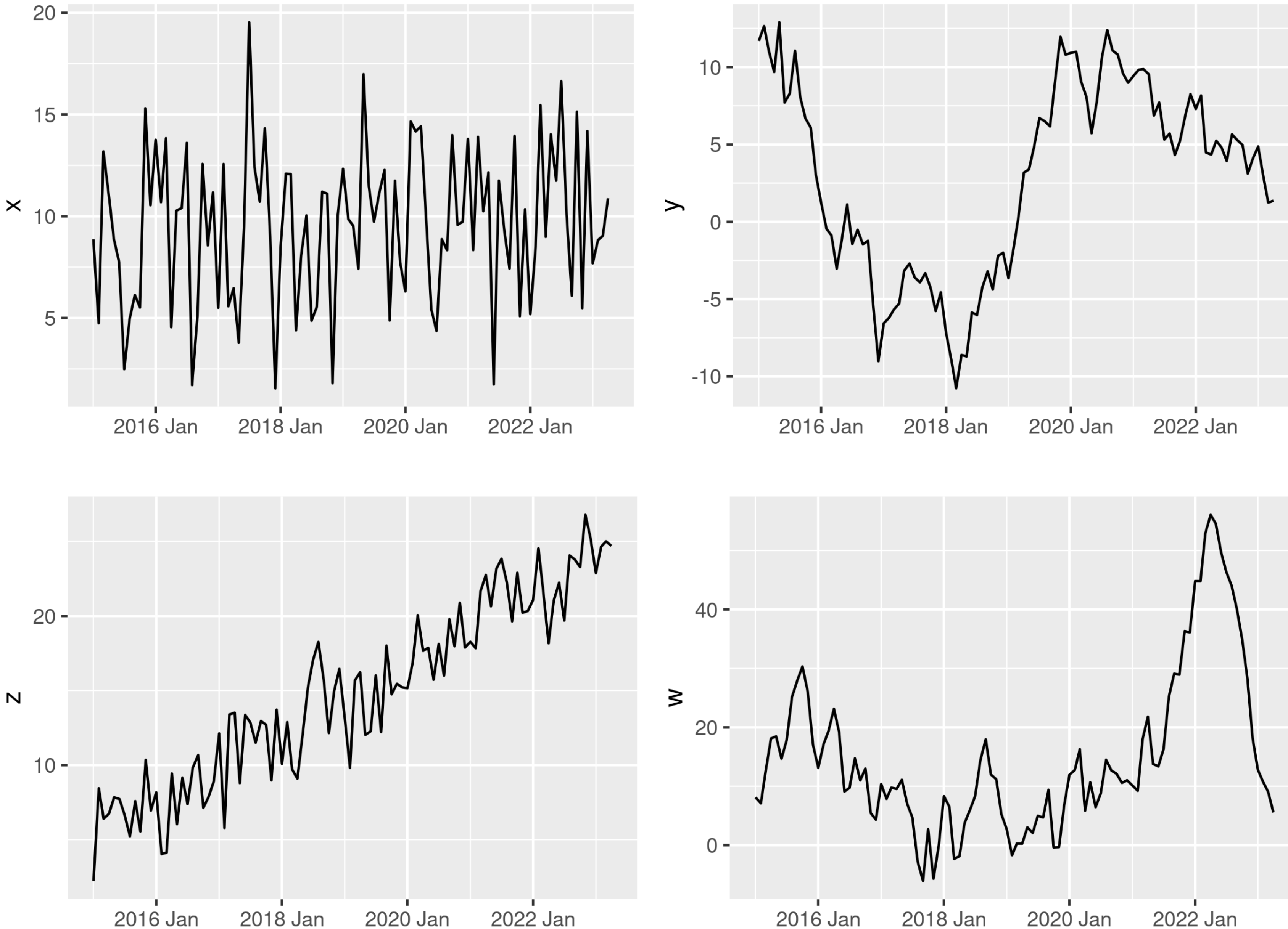
- Analyse the graph: stationary process graph oscillates around its mean with a constant deviation
- Evaluate all the models and choose the best one by cross-validation Time consuming!
- You cannot use AIC!
  $\ln L(y_1,\ldots,y_n \mid \theta)$ and $\ln L(y_2,\ldots,y_n \mid \theta, y_1)$ and $\ln L(y_3,\ldots,y_n \mid \theta, y_1, y_2)$ incomparable!
- There are unit root tests!

How to choose?

$ARIMA(p, 0, q)$ or $ARIMA(p, 1, q)$ or $ARIMA(p, 2, q)$

- Analyse the graph: stationary process graph oscillates around its mean with a constant deviation
- Evaluate all the models and choose the best one by cross-validation Time consuming!
- You cannot use AIC!
  $\ln L(y_1, \ldots, y_n \mid \theta)$ and $\ln L(y_2, \ldots, y_n \mid \theta, y_1)$ and $\ln L(y_3, \ldots, y_n \mid \theta, y_1, y_2)$ incomparable!
- There are unit root tests!
  ADF, KPSS, PP, …

Analysing the graphs



ARIMA: Summary

- $ARMA$ is for stationary series

ARIMA: Summary

- $ARMA$ is for stationary series
- Sometimes Δy_t or $\Delta^2 y_t$ is stationary

ARIMA: Summary

- $ARMA$ is for stationary series
- Sometimes Δy_t or $\Delta^2 y_t$ is stationary
- Choose between $ARMA$ and $ARIMA$

SARIMA process

SARIMA process: Plan

- Seasonal $ARMA$

SARIMA process: Plan

- Seasonal $ARMA$
- Seasonal $ARIMA$

SARIMA process: Plan

- Seasonal $ARMA$
- Seasonal $ARIMA$
- Choosing between models

Seasonality and $ARIMA$

Using $ARMA$ and $ARIMA$ models, we can model seasonality!

Seasonality and $ARIMA$

Using $ARMA$ and $ARIMA$ models, we can model seasonality!

$$MA(12) : y_t = c + u_t + a_1 u_{t-1} + a_2 u_{t-2} + \ldots + a_{12} u_{t-12}.$$

$$ARIMA(12,1,0) : \Delta y_t = c + u_t + b_1 \Delta y_{t-1} + \ldots + b_{12} \Delta y_{t-12}.$$

ARMA should be economical!

Let's focus on non-zero coefficients!

ARMA should be economical!

Let's focus on non-zero coefficients!

Definition

If the stationary $ARMA$ model for y_t can be written with fewer parameters as

$$P_{non}(L)P_{seas}(L^{12})y_t = c + Q_{non}(L)Q_{seas}(L^{12})u_t,$$

where the degrees of the lag polynomials are $\deg P_{non} = p$, $\deg P_{seas} = P$, $\deg Q_{non} = q$, $\deg Q_{seas} = Q$, then it is also called $SARMA(p, q)(P, Q)[12]$

Examples

- $SARMA(1, 0)(0, 2)[12]$

$$(1 - b_1 L)y_t = c + (1 + d_1 L^{12} + d_2 L^{24})u_t;$$

Examples

- $SARMA(1,0)(0,2)[12]$

$$(1 - b_1 L)y_t = c + (1 + d_1 L^{12} + d_2 L^{24})u_t;$$

- $SARMA(0,2)(1,0)[12]$

$$(1 - f_1 L^{12})y_t = c + (1 + a_1 L + a_2 L^2)u_t;$$

Examples

- $SARMA(1,0)(0,2)[12]$

$$(1 - b_1 L)y_t = c + (1 + d_1 L^{12} + d_2 L^{24})u_t;$$

- $SARMA(0,2)(1,0)[12]$

$$(1 - f_1 L^{12})y_t = c + (1 + a_1 L + a_2 L^2)u_t;$$

- $SARMA(1,2)(2,1)[12]$

$$(1 - f_1 L^{12} - f_2 L^{24})(1 - b_1 L^1)y_t = c + (1 + a_1 L + a_2 L^2)(1 + d_1 L^{12})u_t$$

SARIMA

By analogy with the difference $\Delta y_t = y_t - y_{t-1}$, we can consider the seasonal difference $\Delta_{12} y_t = y_t - y_{t-12}$

SARIMA

By analogy with the difference $\Delta y_t = y_t - y_{t-1}$, we can consider the seasonal difference $\Delta_{12} y_t = y_t - y_{t-12}$

Definition

If the series $z_t = \Delta^{d} \Delta_{12}^{D} y_t$ is described by the stationary model $SARMA(p, q)(P, Q)[12]$, then y_t is said to be described by the $SARIMA(p, d, q)(P, D, Q)[12]$ model

SARIMA

By analogy with the difference $\Delta y_t = y_t - y_{t-1}$, we can consider the seasonal difference $\Delta_{12} y_t = y_t - y_{t-12}$

Definition

If the series $z_t = \Delta^{d} \Delta_{12}^{D} y_t$ is described by the stationary model $SARMA(p, q)(P, Q)[12]$, then y_t is said to be described by the $SARIMA(p, d, q)(P, D, Q)[12]$ model

d is the number of times the first difference should be taken $\Delta = 1 - L$;

D is the number of times the seasonal difference should be taken $\Delta_{12} = 1 - L^{12}$;

SARIMA

By analogy with the difference $\Delta y_t = y_t - y_{t-1}$, we can consider the seasonal difference $\Delta_{12} y_t = y_t - y_{t-12}$

Definition

If the series $z_t = \Delta^d \Delta_{12}^D y_t$ is described by the stationary model $SARMA(p, q)(P, Q)[12]$, then y_t is said to be described by the $SARIMA(p, d, q)(P, D, Q)[12]$ model

d is the number of times the first difference should be taken $\Delta = 1 - L$;

D is the number of times the seasonal difference should be taken $\Delta_{12} = 1 - L^{12}$; $y_t \sim SARIMA(0, 0, 2)(1, 1, 2)[12]$ means that

$\Delta_{12} y_t \sim SARMA(0, 2)(1, 2)[12]$

How to choose?

$SARIMA(p,0,q)(P,0,Q)$ or $SARIMA(p,0,q)(P,1,Q)[12]$?

How to choose?

$SARIMA(p,0,q)(P,0,Q)$ or $SARIMA(p,0,q)(P,1,Q)[12]$?

- Analyse the graph!

How to choose?

$SARIMA(p,0,q)(P,0,Q)$ or $SARIMA(p,0,q)(P,1,Q)[12]$?

- Analyse the graph!
- Evaluate all these models and choose the best one by cross-validation

How to choose?

$SARIMA(p,0,q)(P,0,Q)$ or $SARIMA(p,0,q)(P,1,Q)[12]$?

- Analyse the graph!
- Evaluate all these models and choose the best one by cross-validation Time consuming!

How to choose?

$SARIMA(p,0,q)(P,0,Q)$ or $SARIMA(p,0,q)(P,1,Q)[12]$?

- Analyse the graph!
- Evaluate all these models and choose the best one by cross-validation Time consuming!
- You cannot use AIC!

How to choose?

$SARIMA(p, 0, q)(P, 0, Q)$ or $SARIMA(p, 0, q)(P, 1, Q)[12]$?

- Analyse the graph!
- Evaluate all these models and choose the best one by cross-validation Time consuming!
- You cannot use AIC! The conditional and unconditional likelihood functions contain different numbers of terms.

How to choose?

$SARIMA(p,0,q)(P,0,Q)$ or $SARIMA(p,0,q)(P,1,Q)[12]$?

- Analyse the graph!
- Evaluate all these models and choose the best one by cross-validation Time consuming!
- You cannot use AIC! The conditional and unconditional likelihood functions contain different numbers of terms.
- There are unit root tests!

How to choose?

$SARIMA(p,0,q)(P,0,Q)$ or $SARIMA(p,0,q)(P,1,Q)[12]$?

- Analyse the graph!
- Evaluate all these models and choose the best one by cross-validation Time consuming!
- You cannot use AIC! The conditional and unconditional likelihood functions contain different numbers of terms.
- There are unit root tests!
  And rules of thumb…

STL decomposition and the power of seasonality

Step 1. Find the STL expansion of the series (y_t)

$$y_t = trend_t + seas_t + remainder_t$$

STL decomposition and the power of seasonality

Step 1. Find the STL expansion of the series (y_t)

$$y_t = trend_t + seas_t + remainder_t$$

Step 2. Calculate the strength of seasonality

$$F_{seas} = \max\left\{1 - \frac{\mathrm{sVar}(remainder)}{\mathrm{sVar}(seas + remainder)}, 0\right\}$$

STL decomposition and the power of seasonality

Step 1. Find the STL expansion of the series (y_t)

$$y_t = trend_t + seas_t + remainder_t$$

Step 2. Calculate the strength of seasonality

$$F_{seas} = \max\left\{1 - \frac{\mathrm{sVar}(remainder)}{\mathrm{sVar}(seas + remainder)}, 0\right\}$$

Step 3. If the strength of seasonality is above the threshold, then move to $\Delta_{12} y_t = y_t - y_{t-12}$

SARIMA: Summary

- Seasonal ARIMA is more compact

SARIMA: Summary

- Seasonal ARIMA is more compact
- The strength of seasonality from the STL expansion is used to decide if a seasonal difference $\Delta_{12} y_t$ is needed

Unit root tests: ADF test

ADF test: Plan

- Test assumptions

ADF test: Plan

- Test assumptions
- Test algorithm

ADF test: Plan

- Test assumptions
- Test algorithm
- Three variations of the test

Why do we need an stationarity tests?

We want to answer the questions:

Why do we need an stationarity tests?

We want to answer the questions:

- Should the $ARMA$ model be used for (y_t) or for (Δy_t)?

Why do we need an stationarity tests?

We want to answer the questions:

- Should the $ARMA$ model be used for (y_t) or for (Δy_t)?
- How to include a constant in a model?

Why do we need an stationarity tests?

We want to answer the questions:

- Should the $ARMA$ model be used for (y_t) or for (Δy_t)?
- How to include a constant in a model?

Why do we need an stationarity tests?

We want to answer the questions:

- Should the $ARMA$ model be used for (y_t) or for (Δy_t)?
- How to include a constant in a model?

Name "unit root test":

Why do we need an stationarity tests?

We want to answer the questions:

- Should the $ARMA$ model be used for (y_t) or for (Δy_t)?
- How to include a constant in a model?

Name "unit root test":

$$\Delta = 1 - L = P(L)$$

The equation $1 - \ell = 0$ has a root $\ell = 1$

ADF test

ADF — Augmented Dickey Fuller test

ADF test

ADF — Augmented Dickey Fuller test

Three variations of the test: without a constant, with a constant, with a trend

ADF with constant

$$\Delta y_t = c + \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

ADF with constant

$$\Delta y_t = c + \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

$\Delta y_t = m + x_t$;

(x_t) is a stationary $AR(p)$ process with $\mathbb{E}(x_t) = 0$;

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

ADF with constant

$$\Delta y_t = c + \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

$\Delta y_t = m + x_t$;

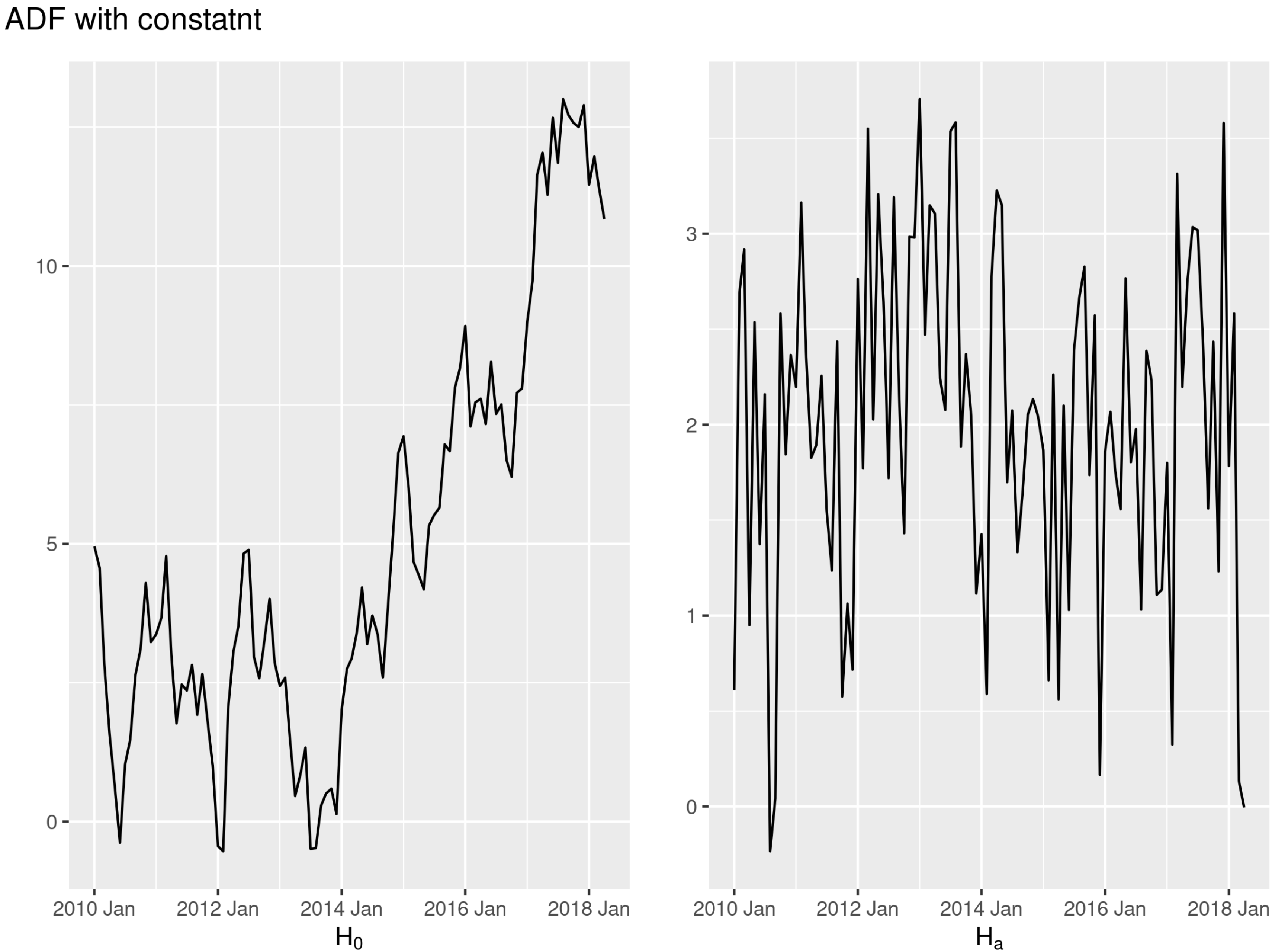
(x_t) is a stationary $AR(p)$ process with $\mathbb{E}(x_t) = 0$;

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

H_a: $\beta < 0$

(y_t) is a stationary $AR(p+1)$ process

ADF with constant: H_0 and H_a



ADF with constant: algorithm

Step 1. Evaluate regression

$$\widehat{\Delta y_t} = \hat{c} + \hat{\beta} y_{t-1} + \hat{d}_1 \Delta y_{t-1} + \ldots + \hat{d}_p \Delta y_{t-p}$$

ADF with constant: algorithm

Step 1. Evaluate regression

$$\widehat{\Delta y_t} = \hat{c} + \hat{\beta} y_{t-1} + \hat{d}_1 \Delta y_{t-1} + \ldots + \hat{d}_p \Delta y_{t-p}$$

Step 2. Calculate the t-statistics using the classic formula

$$ADF = \frac{\hat{\beta} - 0}{se(\hat{\beta})}$$

ADF with constant: algorithm

Step 1. Evaluate regression

$$\widehat{\Delta y_t} = \hat{c} + \hat{\beta} y_{t-1} + \hat{d}_1 \Delta y_{t-1} + \ldots + \hat{d}_p \Delta y_{t-p}$$

Step 2. Calculate the t-statistics using the classic formula

$$ADF = \frac{\hat{\beta} - 0}{se(\hat{\beta})}$$

Under true H_0, the distribution of the ADF-statistic converges to the special DF distribution with DF^c!

ADF with constant: algorithm

Step 1. Evaluate regression

$$\widehat{\Delta y_t} = \hat{c} + \hat{\beta} y_{t-1} + \hat{d}_1 \Delta y_{t-1} + \ldots + \hat{d}_p \Delta y_{t-p}$$

Step 2. Calculate the t-statistics using the classic formula

$$ADF = \frac{\hat{\beta} - 0}{se(\hat{\beta})}$$

Under true H_0, the distribution of the ADF-statistic converges to the special DF distribution with DF^c!

Step 3. We conclude:

If $ADF < DF^c$ then H_0 is rejected

ADF without constant

$$\Delta y_t = \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

ADF without constant

$$\Delta y_t = \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

(Δy_t) is a stationary $AR(p)$ process with $\mathbb{E}(\Delta y_t) = 0$;

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

ADF without constant

$$\Delta y_t = \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

(Δy_t) is a stationary $AR(p)$ process with $\mathbb{E}(\Delta y_t) = 0$;

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

H_a: $\beta < 0$

(y_t) is a stationary $AR(p+1)$ process with $\mathbb{E}(y_t) = 0$;

ADF without constant

$$\Delta y_t = \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

(Δy_t) is a stationary $AR(p)$ process with $\mathbb{E}(\Delta y_t) = 0$;

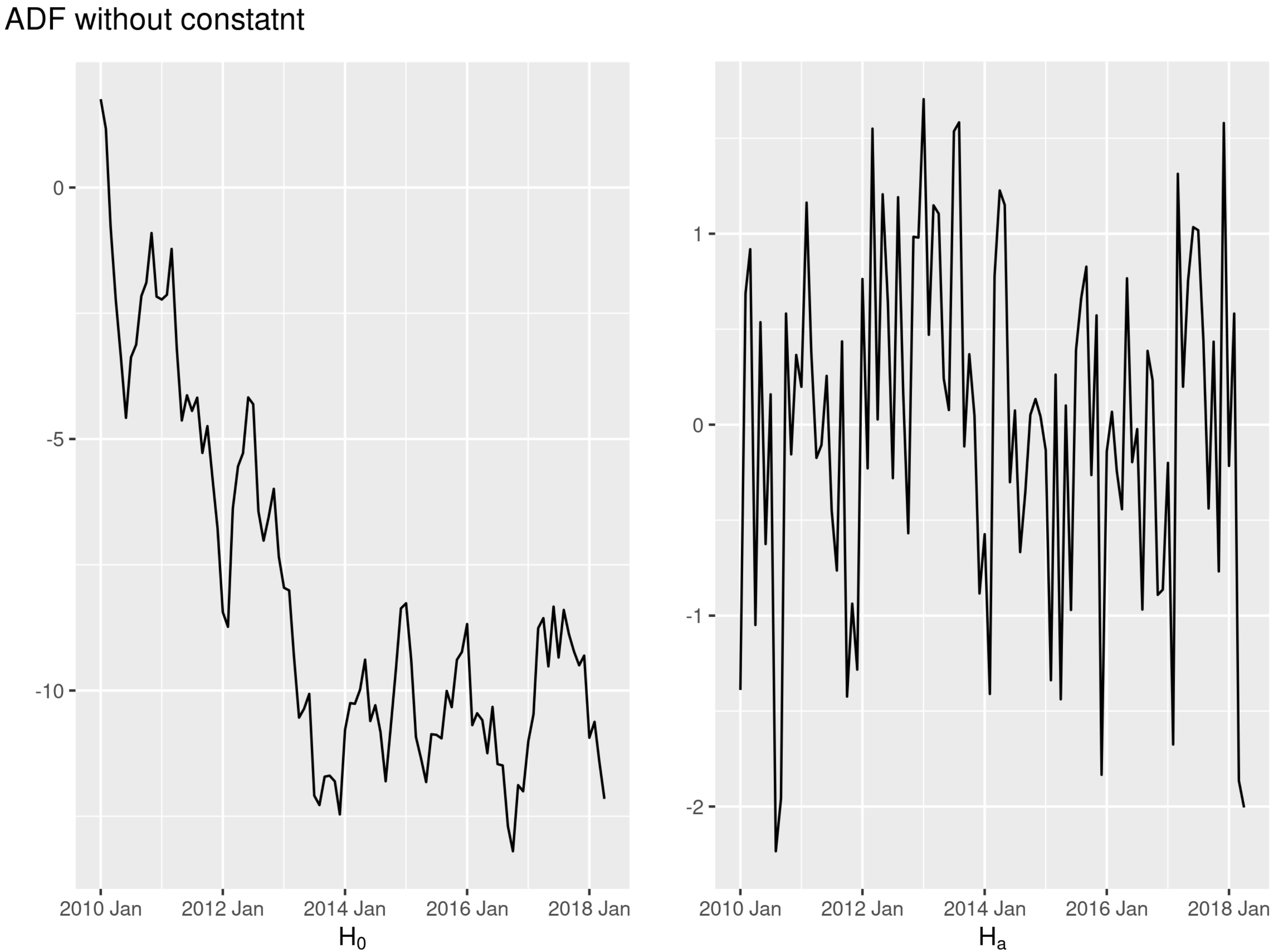
$y_t = y_0 + \sum_{i=1}^{t} \Delta y_i$

H_a: $\beta < 0$

(y_t) is a stationary $AR(p+1)$ process with $\mathbb{E}(y_t) = 0$;

The algorithm will have regression without a constant and another distribution DF^0

ADF without constant: H_0 and H_a



ADF with trend

$$\Delta y_t = c + gt + \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

ADF with trend

$$\Delta y_t = c + gt + \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

$\Delta y_t = k_1 + k_2 t + x_t$;

(x_t) is a stationary $AR(p)$ process with $\mathbb{E}(x_t) = 0$;

$y_t = y_0 + m_1 t + m_2 t^2 + \sum_{i=1}^{t} x_i$

ADF with trend

$$\Delta y_t = c + gt + \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

$\Delta y_t = k_1 + k_2 t + x_t$;

(x_t) is a stationary $AR(p)$ process with $\mathbb{E}(x_t) = 0$;

$y_t = y_0 + m_1 t + m_2 t^2 + \sum_{i=1}^{t} x_i$

H_a: $\beta < 0$

$y_t = m_1 + m_2 t + x_t$;

(x_t) is a stationary $AR(p+1)$ process with $\mathbb{E}(x_t) = 0$;

ADF with trend

$$\Delta y_t = c + gt + \beta y_{t-1} + d_1 \Delta y_{t-1} + \ldots + d_p \Delta y_{t-p} + u_t,$$

H_0: $\beta = 0$

$\Delta y_t = k_1 + k_2 t + x_t$;

(x_t) is a stationary $AR(p)$ process with $\mathbb{E}(x_t) = 0$;

$y_t = y_0 + m_1 t + m_2 t^2 + \sum_{i=1}^{t} x_i$

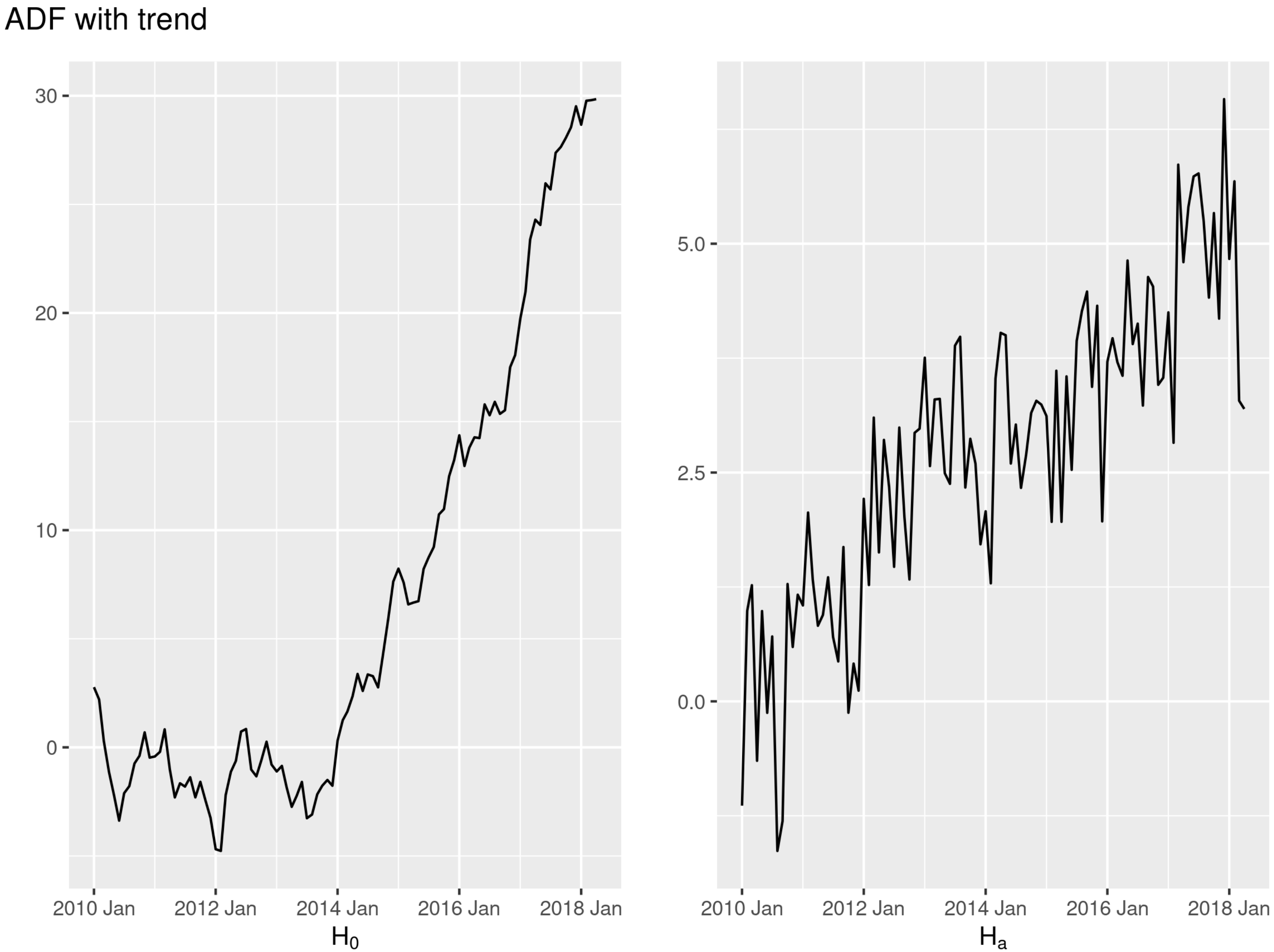
H_a: $\beta < 0$

$y_t = m_1 + m_2 t + x_t$;

(x_t) is a stationary $AR(p+1)$ process with $\mathbb{E}(x_t) = 0$;

The algorithm will have a regression with a constant and a trend and another distribution DF^{ct}

ADF with trend: H_0 and H_a



ADF test: Summary

- Applicable for making a decision about the transition to Δy_t

ADF test: Summary

- Applicable for making a decision about the transition to Δy_t
- Three variants of the ADF test with different assumptions

Unit root tests: KPSS test

KPSS test: Plan

- Long-term variance

KPSS test: Plan

- Long-term variance
- Prerequisites for the test

KPSS test: Plan

- Long-term variance
- Prerequisites for the test
- Two variations of the test

KPSS test

KPSS — Kwiatkowski–Phillips–Schmidt–Shin test

KPSS test

KPSS — Kwiatkowski–Phillips–Schmidt–Shin test

Two variations of the test: with a constant, with a trend

Long-term variance

Definition

For a stationary process (y_t), the quantity λ^2 is called long-term variance if

$$\operatorname{Var}(\bar{y}) = \frac{\lambda^2}{T} + o(1/T)$$

or

$$\lim_{T \to \infty} T \operatorname{Var}(\bar{y}) = \lambda^2,$$

where $\bar{y} = (y_1 + \ldots + y_T)/T$.

Long-term variance

Definition

For a stationary process (y_t), the quantity λ^2 is called long-term variance if

$$\mathrm{Var}(\bar{y}) = \frac{\lambda^2}{T} + o(1/T)$$

or

$$\lim_{T \to \infty} T\,\mathrm{Var}(\bar{y}) = \lambda^2,$$

where $\bar{y} = (y_1 + \ldots + y_T)/T$.

Motivation

For independent observations with the constant variance

$$\mathrm{Var}(\bar{y}) = \frac{\sigma^2}{T}, \text{ where } \sigma^2 = \mathrm{Var}(y_i)$$

KPSS with constant

$$y_t = c + rw_t + x_t,$$

KPSS with constant

$$y_t = c + rw_t + x_t,$$

H_0: $rw_t = 0$

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

KPSS with constant

$$y_t = c + rw_t + x_t,$$

H_0: $rw_t = 0$

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

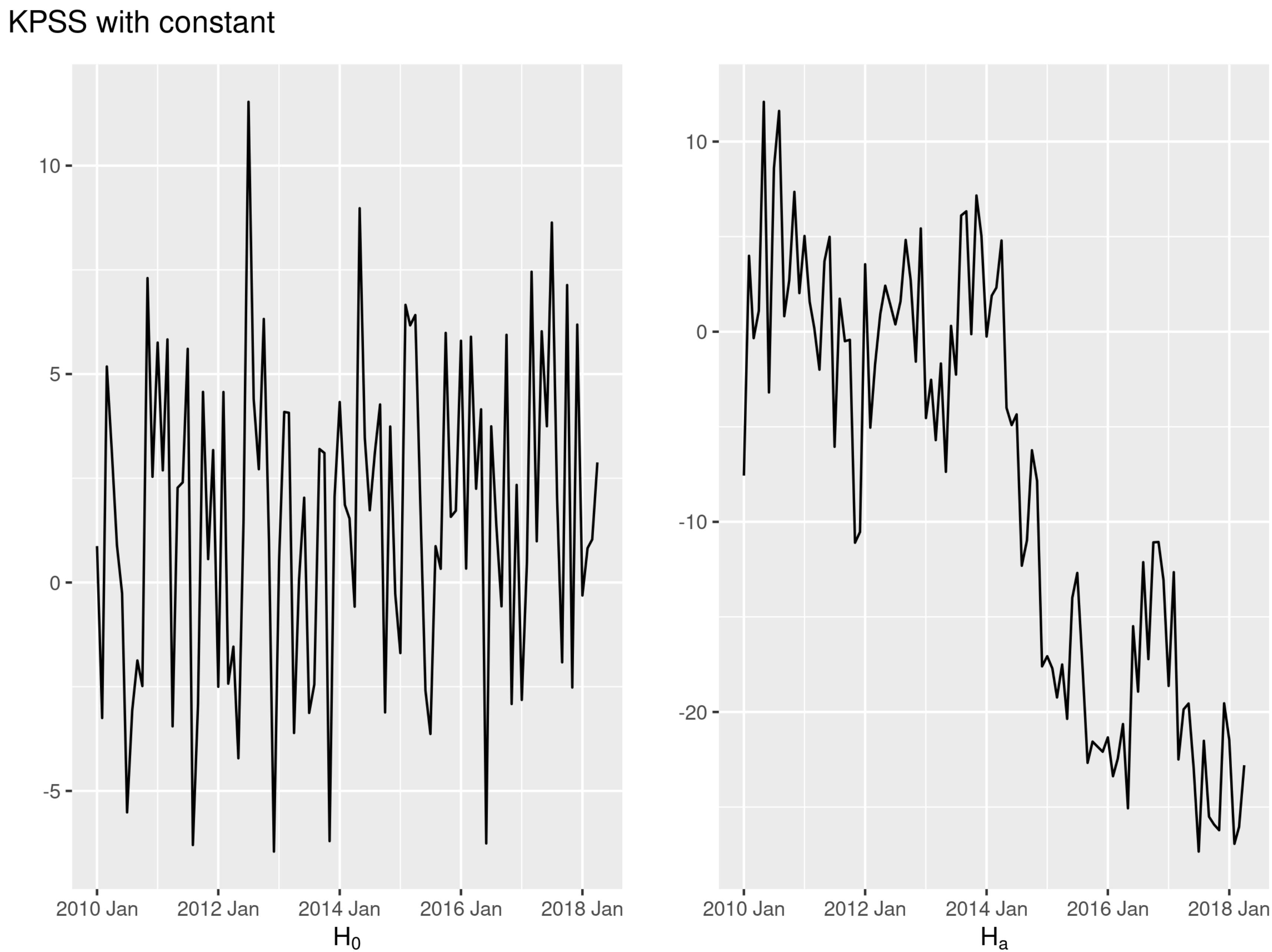
H_a: $rw_t = rw_{t-1} + u_t$

$rw_0 = 0$;

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

(u_t) is white noise independent of (x_t)

KPSS with constant: H_0 and H_a



KPSS with constant: algorithm

Step 1. Evaluate regression on a constant

$$\widehat{y_t} = \hat{c}$$

KPSS with constant: algorithm

Step 1. Evaluate regression on a constant

$$\widehat{y_t} = \hat{c}$$

Step 2. Calculate $KPSS$ statistics

$$KPSS = \frac{\sum_{t=1}^{T} S_t^2}{T^2 \hat{\lambda}^2},$$

where S_t is the accumulated sum of residuals, $S_t = \hat{u}_1 + \ldots + \hat{u}_t$,

and $\hat{\lambda}^2$ is a consistent estimator of the long-term variance.

KPSS with constant: algorithm

Step 1. Evaluate regression on a constant

$$\widehat{y_t} = \hat{c}$$

Step 2. Calculate $KPSS$ statistics

$$KPSS = \frac{\sum_{t=1}^{T} S_t^2}{T^2 \hat{\lambda}^2},$$

where S_t is the accumulated sum of residuals, $S_t = \hat{u}_1 + \ldots + \hat{u}_t$, and $\hat{\lambda}^2$ is a consistent estimator of the long-term variance.

Under true H_0, the distribution of the $KPSS$-statistic converges to a special distribution with $KPSS^c$!

KPSS with constant: algorithm

Step 1. Evaluate regression on a constant

$$\widehat{y_t} = \hat{c}$$

Step 2. Calculate $KPSS$ statistics

$$KPSS = \frac{\sum_{t=1}^{T} S_t^2}{T^2 \hat{\lambda}^2},$$

where S_t is the accumulated sum of residuals, $S_t = \hat{u}_1 + \ldots + \hat{u}_t$,

and $\hat{\lambda}^2$ is a consistent estimator of the long-term variance.

Under true H_0, the distribution of the $KPSS$-statistic converges to a special distribution with $KPSS^c$!

Step 3. We conclude:

If $KPSS > KPSS^c$ then H_0 is rejected

KPSS with trend

$$y_t = c + bt + rw_t + x_t,$$

KPSS with trend

$$y_t = c + bt + rw_t + x_t,$$

H_0: $rw_t = 0$

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

KPSS with trend

$$y_t = c + bt + rw_t + x_t,$$

H_0: $rw_t = 0$

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

H_a: $rw_t = rw_{t-1} + u_t$

$rw_0 = 0$;

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

(u_t) is white noise independent of (x_t)

KPSS with trend

$$y_t = c + bt + rw_t + x_t,$$

H_0: $rw_t = 0$

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

H_a: $rw_t = rw_{t-1} + u_t$

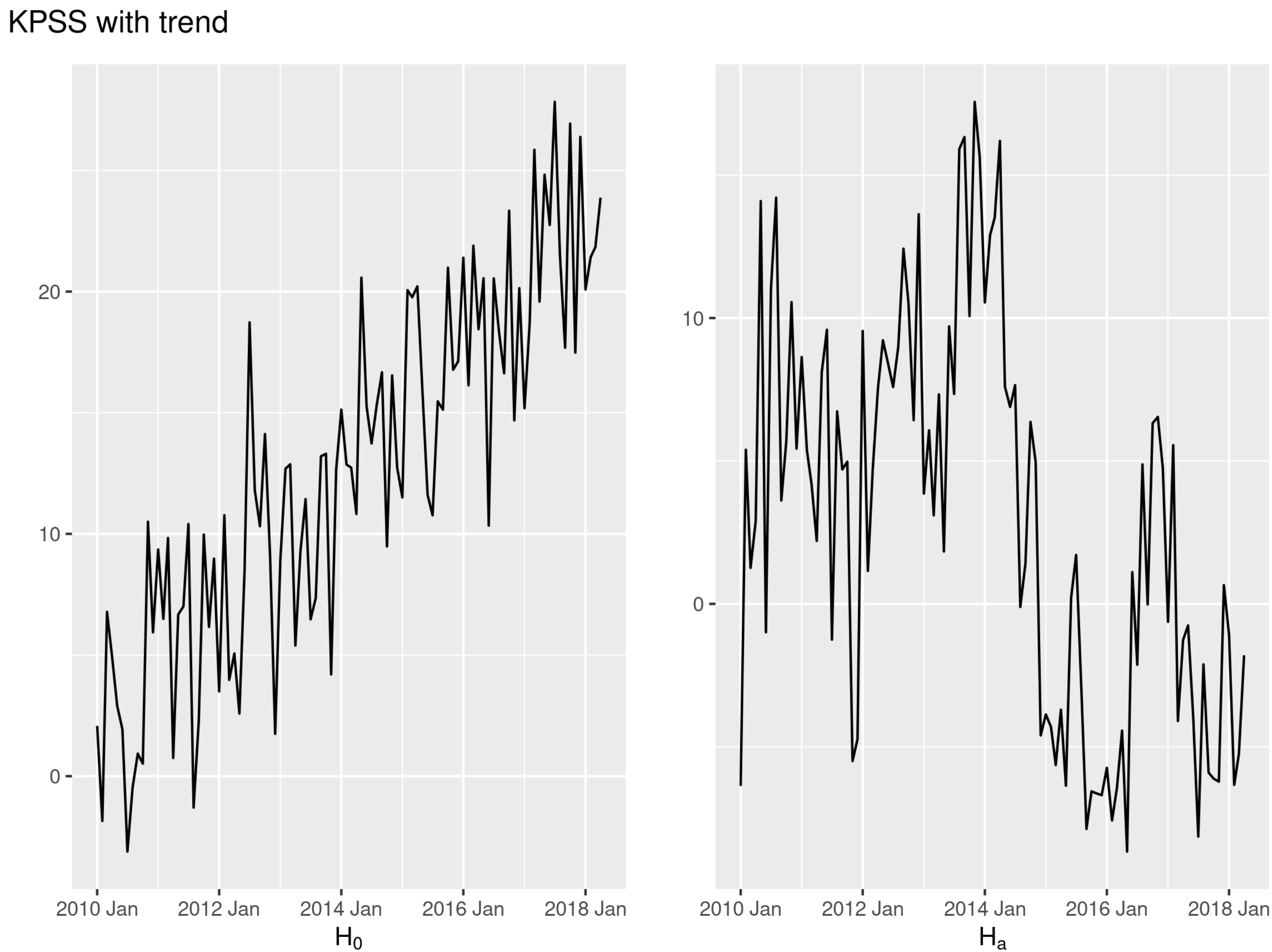
$rw_0 = 0$;

(x_t) is a stationary process with $\mathbb{E}(x_t) = 0$;

(u_t) is white noise independent of (x_t)

The first step of the algorithm will have a regression on a constant and a trend and the statistic under null will have another special distribution $KPSS^{ct}$

KPSS with trend: H_0 and H_a



Terminology

$$A. \quad y_t = a + bt + x_t$$

(y_t) — trend stationary (stationary around the trend)

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

Terminology

$$A. \quad y_t = a + bt + x_t$$

(y_t) — trend stationary (stationary around the trend)

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

Recipe: Estimate regression $a + bt$ with $ARMA$ errors for (y_t).

Terminology

$$A. \quad y_t = a + bt + x_t$$

(y_t) — trend stationary (stationary around the trend)

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

Recipe: Estimate regression $a + bt$ with $ARMA$ errors for (y_t).

$$B. \quad y_t = a + \sum_{i=1}^{t} x_i \text{ or } y_t = a + bt + \sum_{i=1}^{t} x_i$$

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

(y_t) — difference stationary (stationary in differences)

Terminology

$$A. \quad y_t = a + bt + x_t$$

(y_t) — trend stationary (stationary around the trend)

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

Recipe: Estimate regression $a + bt$ with $ARMA$ errors for (y_t).

$$B. \quad y_t = a + \sum_{i=1}^{t} x_i \text{ or } y_t = a + bt + \sum_{i=1}^{t} x_i$$

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

(y_t) — difference stationary (stationary in differences)

Recipe: evaluate $ARMA$ for (Δy_t).

Terminology

$$A. \quad y_t = a + bt + x_t$$

(y_t) — trend stationary (stationary around the trend)

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

Recipe: Estimate regression $a + bt$ with $ARMA$ errors for (y_t).

$$B. \quad y_t = a + \sum_{i=1}^{t} x_i \text{ or } y_t = a + bt + \sum_{i=1}^{t} x_i$$

(x_t) — a stationary process with $\mathbb{E}(x_t) = 0$

(y_t) — difference stationary (stationary in differences)

Recipe: evaluate $ARMA$ for (Δy_t).

Both (y_t) are non-stationary!

KPSS test: Summary

- Applicable for making a decision about the transition to Δy_t

KPSS test: Summary

- Applicable for making a decision about the transition to Δy_t
- Two versions of the KPSS test with different assumptions