Week 10:

SARIMA

Predictions

Financial time series

Last week

- ↪ different kinds of non-stationarity (deterministic vs. stochastic)
- ↪ ARIMA models
- ↪ unit root tests

ARIMA model

ARIMA(p, d, q):

$$\varphi(B)\underbrace{(1-B)^d}_{\Delta^d} Y_t = \alpha + \theta(B)\varepsilon_t$$

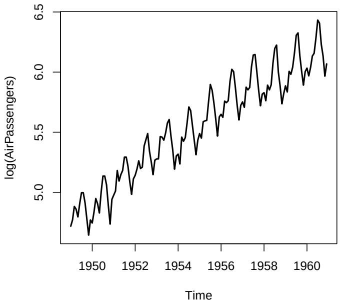
where

- $\{\varepsilon_t\}$ is WN
- $$\varphi(B) = 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p,$$
  $$\theta(B) = 1 + \theta_1 z + \cdots + \theta_q z^q,$$

  such that the roots of $\varphi(z)$ lie outside the unit circle
- Δ differencing operator

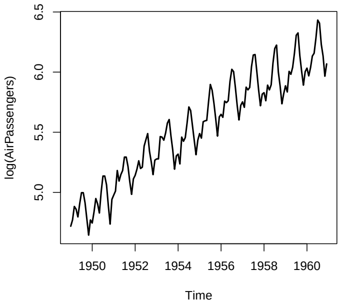
Stochastic seasonality

Example: log of monthly numbers of airline passengers in 1949 to 1960



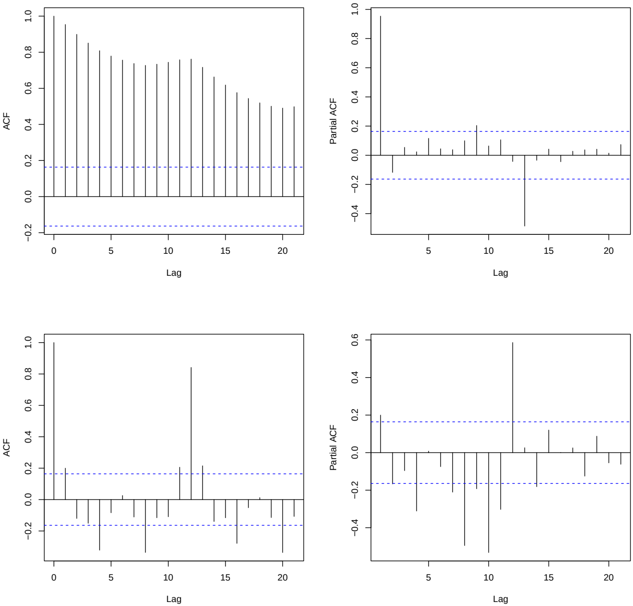
Stochastic seasonality

Example: log of monthly numbers of airline passengers in 1949 to 1960



$\rightsquigarrow$ need to model the trend as well as seasonality

Example (cont.)



SARIMA model for monthly observations

$\hookrightarrow$ ARIMA(P, D, Q) for each month (same models)

$$\Phi(B^{12})\Delta_{12}^{D} Y_t = \Theta(B^{12})\eta_t,$$

where $\Delta_{12} = 1 - B^{12}$ and $\{\eta_t\}$ random component (not WN)

$\hookrightarrow$ model $\{\eta_t\}$ via ARIMA(p, d, q) model

$$\varphi(B)\Delta^{d}\eta_t = \theta(B)\varepsilon_t,$$

where $\{\varepsilon_t\}$ is a white noise

SARIMA model for monthly observations

↪ ARIMA(P, D, Q) for each month (same models)

$$\Phi(B^{12})\Delta_{12}^{D} Y_t = \Theta(B^{12})\eta_t,$$

where $\Delta_{12} = 1 - B^{12}$ and $\{\eta_t\}$ random component (not WN)

↪ model $\{\eta_t\}$ via ARIMA(p, d, q) model

$$\varphi(B)\Delta^{d}\eta_t = \theta(B)\varepsilon_t,$$

where $\{\varepsilon_t\}$ is a white noise

Then

$$\varphi(B)\Delta^{d}\Phi(B^{12})\Delta_{12}^{D} Y_t = \Theta(B^{12})\underbrace{\varphi(B)\Delta^{d}\eta_t}_{\theta(B)\varepsilon_t} = \Theta(B^{12})\theta(B)\varepsilon_t,$$

i.e the resulting model

$$\varphi(B)\Phi(B^{12})\Delta^{d}\Delta_{12}^{D} Y_t = \Theta(B^{12})\theta(B)\varepsilon_t$$

multiplicative seasonal process SARIMA $(p, d, q) \times (P, D, Q)_{12}$

Example

SARIMA $(0,1,1) \times (0,1,1)_{12}$ (so called airline model)

$$(1-B)(1-B^{12})Y_t = (1+\Theta_1 B^{12})(1+\theta_1 B)\varepsilon_t$$

or equivalently

$$Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \Theta_1\varepsilon_{t-12} + \theta_1\Theta_1\varepsilon_{t-13}.$$

Example

SARIMA $(0,1,1) \times (0,1,1)_{12}$ (so called airline model)

$$(1-B)(1-B^{12})Y_t = (1+\Theta_1 B^{12})(1+\theta_1 B)\varepsilon_t$$

or equivalently

$$Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \Theta_1\varepsilon_{t-12} + \theta_1\Theta_1\varepsilon_{t-13}.$$

- How to choose $(p,d,q) \times (P,D,Q)$ for a given data?
  d, D: exploratory graphs
  p, q, P, Q typically based on information criteria

Example

SARIMA $(0, 1, 1) \times (0, 1, 1)_{12}$ (so called airline model)

$$(1 - B)(1 - B^{12})Y_t = (1 + \Theta_1 B^{12})(1 + \theta_1 B)\varepsilon_t$$

or equivalently

$$Y_t - Y_{t-1} - Y_{t-12} + Y_{t-13} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \Theta_1\varepsilon_{t-12} + \theta_1\Theta_1\varepsilon_{t-13}.$$

- How to choose $(p, d, q) \times (P, D, Q)$ for a given data?
  d, D: exploratory graphs
  p, q, P, Q typically based on information criteria
- seasonality for different s number of seasons analogously

Example (cont.)

```
> a = auto.arima(y, d = 1,D=1)
> a
Series: y
ARIMA(0,1,1)(0,1,1)[12]

Coefficients:
                  ma1              sma1
      -0.401828016756  -0.556944838448
s.e.   0.089643846165   0.073099677314

sigma^2 = 0.00137126000392:  log likelihood = 244.7
AIC=-483.4   AICc=-483.21   BIC=-474.77
```

+ model verification

Predictions

Predictions for ARMA(p, q)

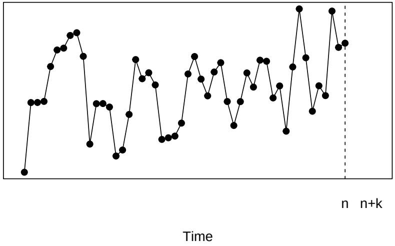
Consider a stationary and invertible model

$$Y_t = \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

and data $Y_1, \ldots, Y_n$ from this process

Aim: Construct a prediction $\widehat{Y}_{n+k} = \widehat{Y}_{n+k}(n)$ of Y_{n+k} based on information known upon time n for $k \geq 1$

$\hookrightarrow$ linear prediction with minimal error



Theoretical prediction

$\widehat{Y}_{n+k}$ linear in the whole past $\{Y_s, s \leq n\} = \{Y_s\}_{s=-\infty}^{n}$

- stationarity $\rightsquigarrow$

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

with $\psi_0 = 1$ and

$$Y_{n+k} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n+k-j} = \varepsilon_{n+k} + \psi_1 \varepsilon_{n+k-1} + \cdots + \psi_k \varepsilon_n + \psi_{k+1} \varepsilon_{n-1} + \ldots$$

Theoretical prediction

$\widehat{Y}_{n+k}$ linear in the whole past $\{Y_s, s \leq n\} = \{Y_s\}_{s=-\infty}^{n}$

- stationarity $\rightsquigarrow$

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

with $\psi_0 = 1$ and

$$Y_{n+k} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n+k-j} = \varepsilon_{n+k} + \psi_1 \varepsilon_{n+k-1} + \cdots + \psi_k \varepsilon_n + \psi_{k+1} \varepsilon_{n-1} + \ldots$$

- If we know $\{Y_s\}_{s=-\infty}^{n}$ + invertibility $\rightsquigarrow$ we know $\{\varepsilon_s\}_{s=-\infty}^{n}$

Theoretical prediction

$\widehat{Y}_{n+k}$ linear in the whole past $\{Y_s, s \le n\} = \{Y_s\}_{s=-\infty}^{n}$

- stationarity $\rightsquigarrow$

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

with $\psi_0 = 1$ and

$$Y_{n+k} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n+k-j} = \varepsilon_{n+k} + \psi_1 \varepsilon_{n+k-1} + \cdots + \psi_k \varepsilon_n + \psi_{k+1} \varepsilon_{n-1} + \ldots$$

- If we know $\{Y_s\}_{s=-\infty}^{n}$ + invertibility $\rightsquigarrow$ we know $\{\varepsilon_s\}_{s=-\infty}^{n}$
- $\rightsquigarrow$ prediction $\widehat{Y}_{n+k}$ linear in $\{\varepsilon_s\}_{s=-\infty}^{n}$ such that

$$\mathrm{E}(Y_{n+k} - \widehat{Y}_{n+k})^2 \to \min$$

Theoretical prediction

$\widehat{Y}_{n+k}$ linear in the whole past $\{Y_s, s \leq n\} = \{Y_s\}_{s=-\infty}^{n}$

- stationarity $\rightsquigarrow$

$$Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$$

with $\psi_0 = 1$ and

$$Y_{n+k} = \sum_{j=0}^{\infty} \psi_j \varepsilon_{n+k-j} = \varepsilon_{n+k} + \psi_1 \varepsilon_{n+k-1} + \cdots + \psi_k \varepsilon_n + \psi_{k+1} \varepsilon_{n-1} + \ldots$$

- If we know $\{Y_s\}_{s=-\infty}^{n}$ + invertibility $\rightsquigarrow$ we know $\{\varepsilon_s\}_{s=-\infty}^{n}$
- $\rightsquigarrow$ prediction $\widehat{Y}_{n+k}$ linear in $\{\varepsilon_s\}_{s=-\infty}^{n}$ such that

$$\mathrm{E}(Y_{n+k} - \widehat{Y}_{n+k})^2 \to \min$$

- solution

$$\widehat{Y}_{n+k} = \psi_k \varepsilon_n + \psi_{k+1} \varepsilon_{n-1} + \cdots = \sum_{j=k}^{\infty} \psi_j \varepsilon_{n+k-j}$$

Theoretical prediction (cont.)

Error of the prediction

$$e_{n+k} = Y_{n+k} - \widehat{Y}_{n+k} = \sum_{j=0}^{k-1} \psi_j \varepsilon_{n+k-j}$$

and $e_{n+1} = \varepsilon_t$ (innovation).

Variance of the error

$$\mathrm{Var}\, e_{n+k} = \sigma^2 \sum_{j=0}^{k-1} \psi_j^2$$

(formula useful for prediction intervals)

Practical construction of predictions

Model

$$Y_t = \varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} + \varepsilon_t + \cdots + \theta_q \varepsilon_{t-q}$$

1. Estimate parameters $\widehat{\varphi}_1, \ldots, \widehat{\theta}_q$
2. Compute recursively

$$\widehat{Y}_{n+k} = \widehat{\varphi}_1 \widehat{Y}_{n+k-1} + \cdots + \widehat{\varphi}_p \widehat{Y}_{n+k-p} + \widehat{\varepsilon}_{n+k} + \cdots + \widehat{\theta}_q \widehat{\varepsilon}_{n+k-q},$$

where

$$\widehat{Y}_{n+j} = \begin{cases} Y_{n+j} & j \leq 0, \\ \widehat{Y}_{n+j} & j > 0, \end{cases}$$

and

$$\widehat{\varepsilon}_{n+j} = \begin{cases} 0 & j > 0, \\ Y_{n+j} - \widehat{Y}_{n+j}(n+j-1) & j \leq 0, \end{cases}$$

where $\widehat{Y}_k(k-1)$ is prediction of Y_k based on data up to time $k-1$

Example: AR(2) model

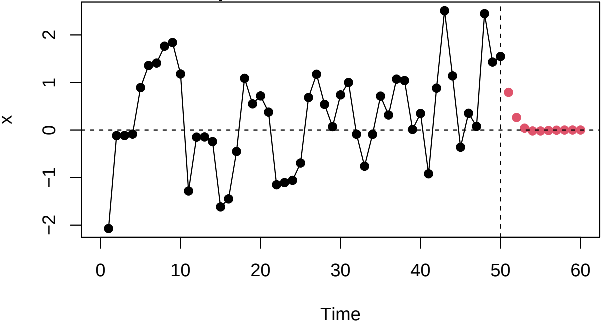
Model

$$Y_t = \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \varepsilon_t$$

and data $Y_1, \ldots, Y_n$.

Then

$$\widehat{Y}_{n+1} = \widehat{\varphi}_1 Y_n + \widehat{\varphi}_2 Y_{n-1},$$
$$\widehat{Y}_{n+2} = \widehat{\varphi}_1 \widehat{Y}_{n+1} + \widehat{\varphi}_2 Y_n,$$
$$\widehat{Y}_{n+3} = \widehat{\varphi}_1 \widehat{Y}_{n+2} + \widehat{\varphi}_2 \widehat{Y}_{n+1},$$
$$\vdots$$



Example: ARMA(1,1) model

Model

$$Y_t = \varphi_1 Y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

and data $Y_1, \ldots, Y_n$:

$$\begin{aligned} \widehat{Y}_{n+1} &= \widehat{\varphi}_1 Y_n + \widehat{\theta}_1 \widehat{\varepsilon}_n, \\ \widehat{Y}_{n+2} &= \widehat{\varphi}_1 \widehat{Y}_{n+1}, \\ \widehat{Y}_{n+3} &= \widehat{\varphi}_1 \widehat{Y}_{n+2}, \\ &\vdots \end{aligned}$$

where $\widehat{\varepsilon}_n$ is computed recursively from

$$\varepsilon_t = Y_t - \varphi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1}$$

Set $\varepsilon_0 = 0$ and Y_0 arbitrary (reasonably), then

$$\begin{aligned} \widehat{\varepsilon}_1 &= Y_1 - \widehat{\varphi}_1 Y_0 - 0, \\ \widehat{\varepsilon}_2 &= Y_2 - \widehat{\varphi}_1 Y_1 - \widehat{\theta}_1 \widehat{\varepsilon}_1, \\ &\vdots \\ \widehat{\varepsilon}_n &= Y_n - \widehat{\varphi}_1 Y_{n-1} - \widehat{\theta}_1 \widehat{\varepsilon}_{n-1} \end{aligned}$$

Example: prediction

Further examples $\rightsquigarrow$ read in the book: page 167–169

Interval predictions

Theoretical prediction:

$$\widehat{Y}_{n+k} = \sum_{j=k}^{\infty} \psi_j \varepsilon_{t+k-j}$$

with error

$$e_{n+k} = Y_{n+k} - \widehat{Y}_{n+k} = \sum_{j=0}^{k-1} \psi_j \varepsilon_{n+k-j}$$

Assumption: ε_i are iid from $\mathrm{N}(0, \sigma^2)$

Let $\mathcal{F}_n = \sigma\{Y_s, s \leq n\} = \sigma\{\varepsilon_s, s \leq n\}$. Then

- e_{n+k} is independent of $\mathcal{F}_n$
- $$e_{n+k} \sim \mathrm{N}\left(0, \sigma^2 \sum_{j=0}^{k-1} \psi_j^2\right)$$

Interval predictions (cont.)

Hence

- $$Y_{n+k} - \widehat{Y}_{n+k} \sim \mathrm{N}\left(0, \sigma^2 \sum_{j=0}^{k-1} \psi_j^2\right)$$
  $$\frac{Y_{n+k} - \widehat{Y}_{n+k}}{\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}} \sim \mathrm{N}(0, 1)$$
- If σ^2 and ψ_j were known $\rightsquigarrow$ prediction interval with confidence $1 - \alpha$
  $$\left(\widehat{Y}_{n+k} - u_{1-\alpha/2}\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}, \widehat{Y}_{n+k} + u_{1-\alpha/2}\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}\right)$$

Interval predictions (cont.)

Hence

- $$Y_{n+k} - \widehat{Y}_{n+k} \sim \mathrm{N}\left(0, \sigma^2 \sum_{j=0}^{k-1} \psi_j^2\right)$$
$$\frac{Y_{n+k} - \widehat{Y}_{n+k}}{\sigma \sqrt{\sum_{j=0}^{k-1} \psi_j^2}} \sim \mathrm{N}(0, 1)$$
- If σ^2 and ψ_j were known $\rightsquigarrow$ prediction interval with confidence $1 - \alpha$
$$\left(\widehat{Y}_{n+k} - u_{1-\alpha/2}\sigma \sqrt{\sum_{j=0}^{k-1} \psi_j^2}, \widehat{Y}_{n+k} + u_{1-\alpha/2}\sigma \sqrt{\sum_{j=0}^{k-1} \psi_j^2}\right)$$
- In practice: replace σ^2 and ψ_j with estimates $\rightsquigarrow$ no exact prescribed confidence, but typically acceptable

Interval predictions (cont.)

Hence

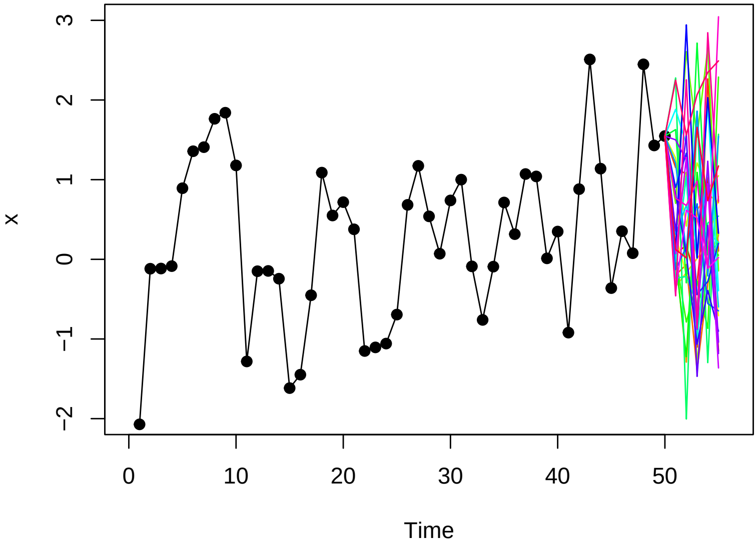
- $$Y_{n+k} - \widehat{Y}_{n+k} \sim \mathrm{N}\left(0, \sigma^2 \sum_{j=0}^{k-1} \psi_j^2\right)$$
  $$\frac{Y_{n+k} - \widehat{Y}_{n+k}}{\sigma \sqrt{\sum_{j=0}^{k-1} \psi_j^2}} \sim \mathrm{N}(0, 1)$$
- If σ^2 and ψ_j were known $\rightsquigarrow$ prediction interval with confidence $1 - \alpha$
  $$\left(\widehat{Y}_{n+k} - u_{1-\alpha/2}\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}, \widehat{Y}_{n+k} + u_{1-\alpha/2}\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}\right)$$
- In practice: replace σ^2 and ψ_j with estimates $\rightsquigarrow$ no exact prescribed confidence, but typically acceptable
- How to get estimates of the MA(∞) representation $Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$?

Interval predictions (cont.)

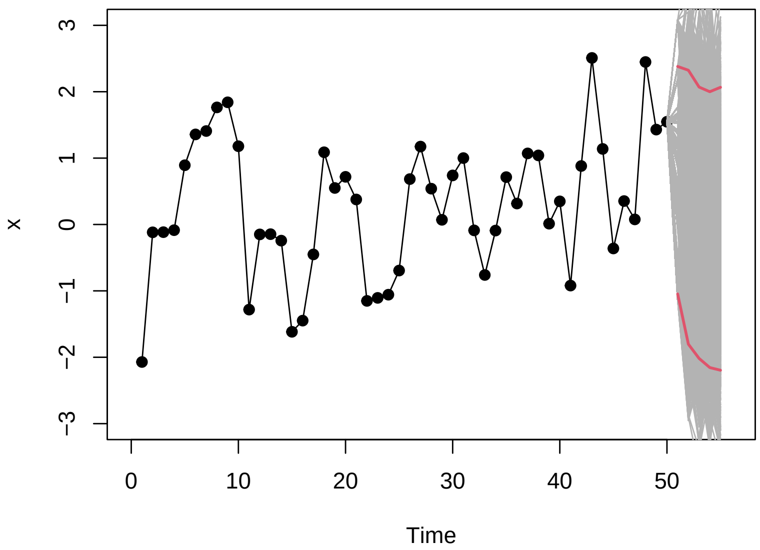
Hence

- $$Y_{n+k} - \widehat{Y}_{n+k} \sim \mathrm{N}\left(0, \sigma^2 \sum_{j=0}^{k-1} \psi_j^2\right)$$
  $$\frac{Y_{n+k} - \widehat{Y}_{n+k}}{\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}} \sim \mathrm{N}(0,1)$$
- If σ^2 and ψ_j were known $\rightsquigarrow$ prediction interval with confidence $1-\alpha$
  $$\left(\widehat{Y}_{n+k} - u_{1-\alpha/2}\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}, \widehat{Y}_{n+k} + u_{1-\alpha/2}\sigma\sqrt{\sum_{j=0}^{k-1} \psi_j^2}\right)$$
- In practice: replace σ^2 and ψ_j with estimates $\rightsquigarrow$ no exact prescribed confidence, but typically acceptable
- How to get estimates of the MA(∞) representation $Y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$?
  $\rightsquigarrow$ estimate AR and MA polynomials $\rightsquigarrow$ use Taylor expansion on their ratio

Prediction intervals based on bootstrap



Prediction intervals based on bootstrap



Prediction intervals based on bootstrap

Setting:

- $\hookrightarrow$ data $Y_1, \ldots, Y_n$,
- $\hookrightarrow$ estimated model parameters ϕ
- $\hookrightarrow$ residuals $\widehat{\varepsilon}_t \rightsquigarrow$ normalized residuals

$$\widehat{\varepsilon}_t^* = \widehat{\varepsilon}_t - \frac{1}{n}\sum_{t=1}^{n}\widehat{\varepsilon}_t$$

Prediction intervals based on bootstrap

Setting:

- $\hookrightarrow$ data $Y_1, \ldots, Y_n$,
- $\hookrightarrow$ estimated model parameters ϕ
- $\hookrightarrow$ residuals $\widehat{\varepsilon}_t \rightsquigarrow$ normalized residuals

$$\widehat{\varepsilon}_t^* = \widehat{\varepsilon}_t - \frac{1}{n} \sum_{t=1}^{n} \widehat{\varepsilon}_t$$

Idea: For $b = 1, \ldots, B$

1. Simulate possible future values for $Y_{n+1}^b, \ldots, Y_{n+k}^b$ from the original series $Y_1, \ldots, Y_n$ and $\widehat{\phi}$ (using the estimated model) and innovations sampled from $\widehat{\varepsilon}_1^*, \ldots, \widehat{\varepsilon}_n^*$.
2. For each $j = 1, \ldots, k$ compute 0.025 and 0.975 sample quantiles from Y_{n+j}^b to get the prediction intervals.

Prediction intervals based on bootstrap

Setting:

- $\hookrightarrow$ data $Y_1, \ldots, Y_n$,
- $\hookrightarrow$ estimated model parameters ϕ
- $\hookrightarrow$ residuals $\widehat{\varepsilon}_t \rightsquigarrow$ normalized residuals

$$\widehat{\varepsilon}_t^* = \widehat{\varepsilon}_t - \frac{1}{n} \sum_{t=1}^{n} \widehat{\varepsilon}_t$$

Prediction intervals based on bootstrap

Setting:

- $\hookrightarrow$ data $Y_1, \ldots, Y_n$,
- $\hookrightarrow$ estimated model parameters ϕ
- $\hookrightarrow$ residuals $\widehat{\varepsilon}_t \rightsquigarrow$ normalized residuals

$$\widehat{\varepsilon}_t^* = \widehat{\varepsilon}_t - \frac{1}{n}\sum_{t=1}^{n} \widehat{\varepsilon}_t$$

Algorithm: For $b = 1, \ldots, B$

1. Simulate $\{Y_t^{(b)}\}$ from the same model using the estimated parameters and innovations resampled from $\widehat{\varepsilon}_1^*, \ldots, \widehat{\varepsilon}_n^*$.
2. Compute $\widehat{\phi}^{(b)}$.
3. Simulate possible future values for $Y_{n+1}^b, \ldots, Y_{n+k}^b$ from the original series $Y_1, \ldots, Y_n$ and $\widehat{\phi}^{(b)}$ and innovations sampled from $\widehat{\varepsilon}_1^*, \ldots, \widehat{\varepsilon}_n^*$.

For each $j = 1, \ldots, k$ compute 0.025 and 0.975 sample quantiles to get the prediction intervals.

Financial time series: Volatility modelling using ARCH and GARCH model

Motivation

Let Y_t follow a stationary and invertible ARMA(p, q) model and $\mathcal{F}_{t-1} = \sigma\{Y_s, \varepsilon_s, s \leq t-1\}$.

Then we can write

$$Y_t = \underbrace{\varphi_1 Y_{t-1} + \cdots + \varphi_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}}_{\mu(\mathcal{F}_{t-1})} + \varepsilon_t$$

where

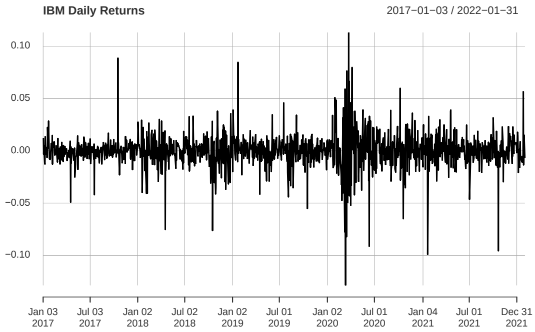
- ↪ μ is linear in $\mathcal{F}_{t-1}$
- ↪ $$\text{Var}\,(Y_t|\mathcal{F}_{t-1}) = \text{Var}\,(\varepsilon_t|\mathcal{F}_{t-1}) = \text{Var}\,\varepsilon_t = \sigma^2$$

  so the data are conditionally homoscedastic

However, many financial time series are conditionally heteroscedastic

Volatility

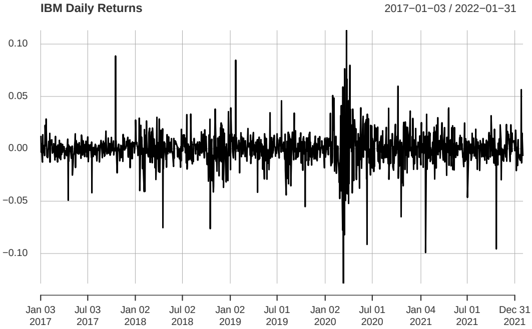
- conditional variance (of e.g. underlying asset return) ⇝ not directly observable
- important factor in options trading, risk management
- modelling volatility ⇝ improvement in prediction intervals



Characteristics of volatility for log-returns

- volatility clustering
- leverage effect (asymmetry of the impact of past positive and negative log returns)

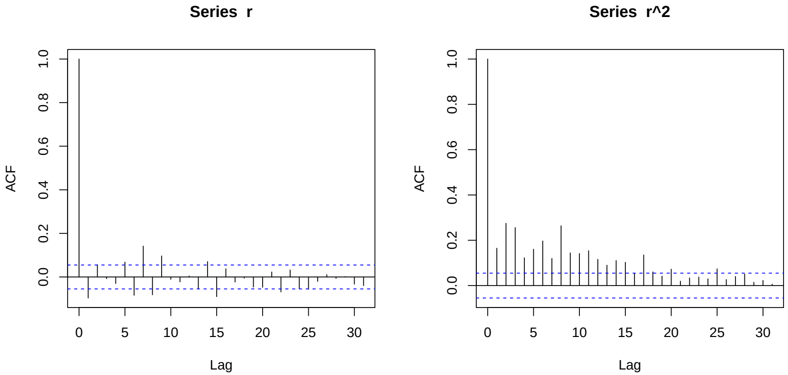
(read more in the book, part 8.1)



Moreover, log-returns are often

- uncorrelated, but not independent
- heavy-tailed

Conditionally heteroscedastic white noise



A sequence $\{e_t\}$ can be a dependent white noise

- $\{e_t\}$ are uncorrelated and Var $e_t = \sigma^2$
- the conditional variance

$$\text{Var}\,[e_t | e_{t-1}, e_{t-2}, \ldots] \text{ is a function of the past observations}$$

General time series model

If $\mathcal{F}_{t-1} = \sigma\{Y_s, s \leq t-1\}$

$$Y_t = \mu(\mathcal{F}_{t-1}) + \sigma(\mathcal{F}_{t-1})\varepsilon_t$$

where ε_t are iid (0,1) and ε_t is independent of $\mathcal{F}_s$ for $s \leq t$

Then

$$\begin{aligned} \mathrm{E}[Y_t|\mathcal{F}_{t-1}] &= \mu(\mathcal{F}_{t-1}) + \sigma(\mathcal{F}_{t-1})\mathrm{E}[\varepsilon_t|\mathcal{F}_{t-1}] = \mu(\mathcal{F}_{t-1}), \\ \mathrm{Var}\,[Y_t|\mathcal{F}_{t-1}] &= \sigma(\mathcal{F}_{t-1})^2. \end{aligned}$$

- ARMA: model for μ linear in $\mathcal{F}_{t-1}$
- we now focus on nonlinear modelling of $\sigma(\mathcal{F}_{t-1})$

$$e_t = \sigma_t\varepsilon_t, \quad \sigma_t = \sigma(\mathcal{F}_{t-1})$$

application: either to residuals from a fitted ARMA (or regression model), or series with $\mu \equiv 0$ (white noise)

ARCH model

ARCH (autoregressive conditional heteroscedasticity) by Engle (1982). He observed that

- ↪ Financial time series are heteroscedastic, i.e., their volatility changes in time.
- ↪ The volatility is a simple quadratic function of past prediction errors.

ARCH(r) model:

$$e_t = \sigma_t \varepsilon_t,$$
$$\sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \cdots + \alpha_r e_{t-r}^2$$

where ε_t are iid

$$\mathrm{E}\varepsilon_t = 0, \quad \mathrm{Var}\,\varepsilon_t = 1$$

and

$$\alpha_0 > 0, \quad \alpha_1, \ldots, \alpha_r \in [0, 1), \quad \sum_{i=1}^{r} \alpha_i < 1 \tag{A}$$

Conditional expectation

Let $\boldsymbol{X}$, $\boldsymbol{Y}$ be random vectors with finite means. Then

$$\begin{aligned} \mathrm{E}\boldsymbol{X} &= \mathrm{E}[\mathrm{E}(\boldsymbol{X}|\boldsymbol{Y})], \\ \mathrm{Var}\,\boldsymbol{X} &= \mathrm{E}\mathrm{Var}\,[\boldsymbol{X}|\boldsymbol{Y}] + \mathrm{Var}\,\mathrm{E}[\boldsymbol{X}|\boldsymbol{Y}], \end{aligned}$$

and

$$\mathrm{E}[\boldsymbol{X}g(\boldsymbol{Y})|\boldsymbol{Y}] = g(\boldsymbol{Y})\mathrm{E}[\boldsymbol{X}|\boldsymbol{Y}]$$

for any measurable function g. Furthermore,

$$\mathrm{Cov}\,(\boldsymbol{X}, \boldsymbol{Y}) = \mathrm{Cov}\,(\boldsymbol{X}, \mathrm{E}[\boldsymbol{Y}|\boldsymbol{X}]).$$

If $\mathcal{F} \subset \mathcal{G}$

$$\mathrm{E}[\boldsymbol{X}|\mathcal{F}] = \mathrm{E}[\mathrm{E}[\boldsymbol{X}||\mathcal{G}]|\mathcal{F}] = \mathrm{E}[\mathrm{E}[\boldsymbol{X}||\mathcal{F}]|\mathcal{G}].$$