MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

IDSS

Applied Data Science Program

TIME SERIES

Munther A. Dahleh

Massachusetts Institute of Technology

IDSS MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

Learning Time-Series

- Time Series are everywhere
  - Finance, weather, control systems

- Questions:
  - Modeling
  - Forecasting
  - Decisions and policy

Learning Time-Series: Outline

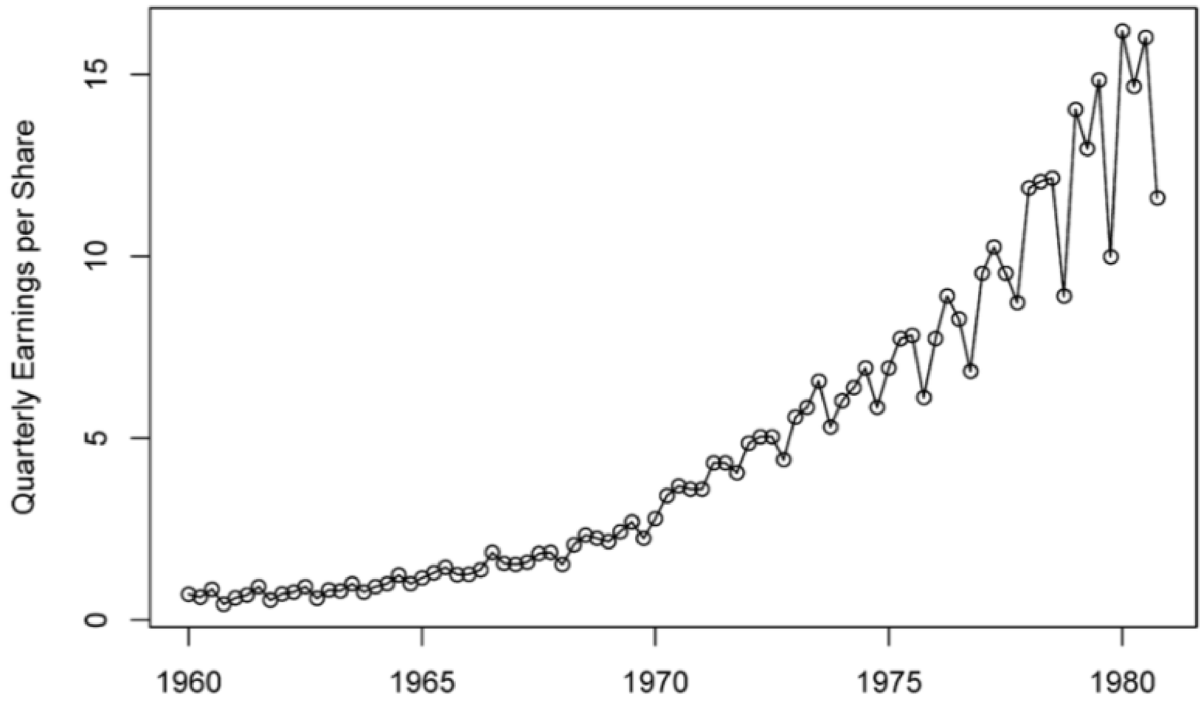
- Part I: Introduction to Time Series
  - Stationarity
  - Trends

- Part II: Models of Time Series

- Part 3: Learning Time Series

Introduction

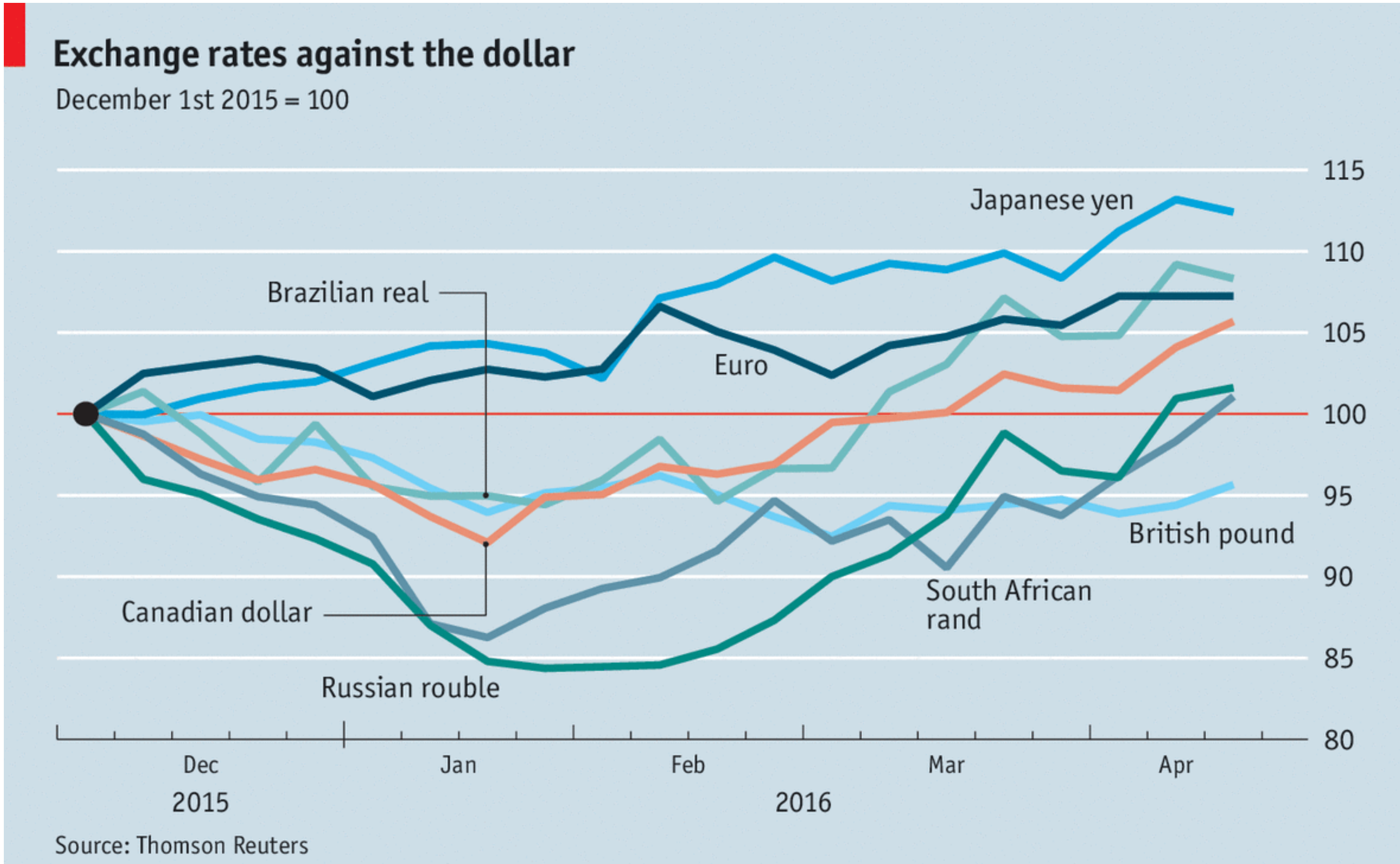
Johnson & Johnson quarterly earnings per share



Massachusetts Institute of Technology

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

Exchange Rates Against the Dollar

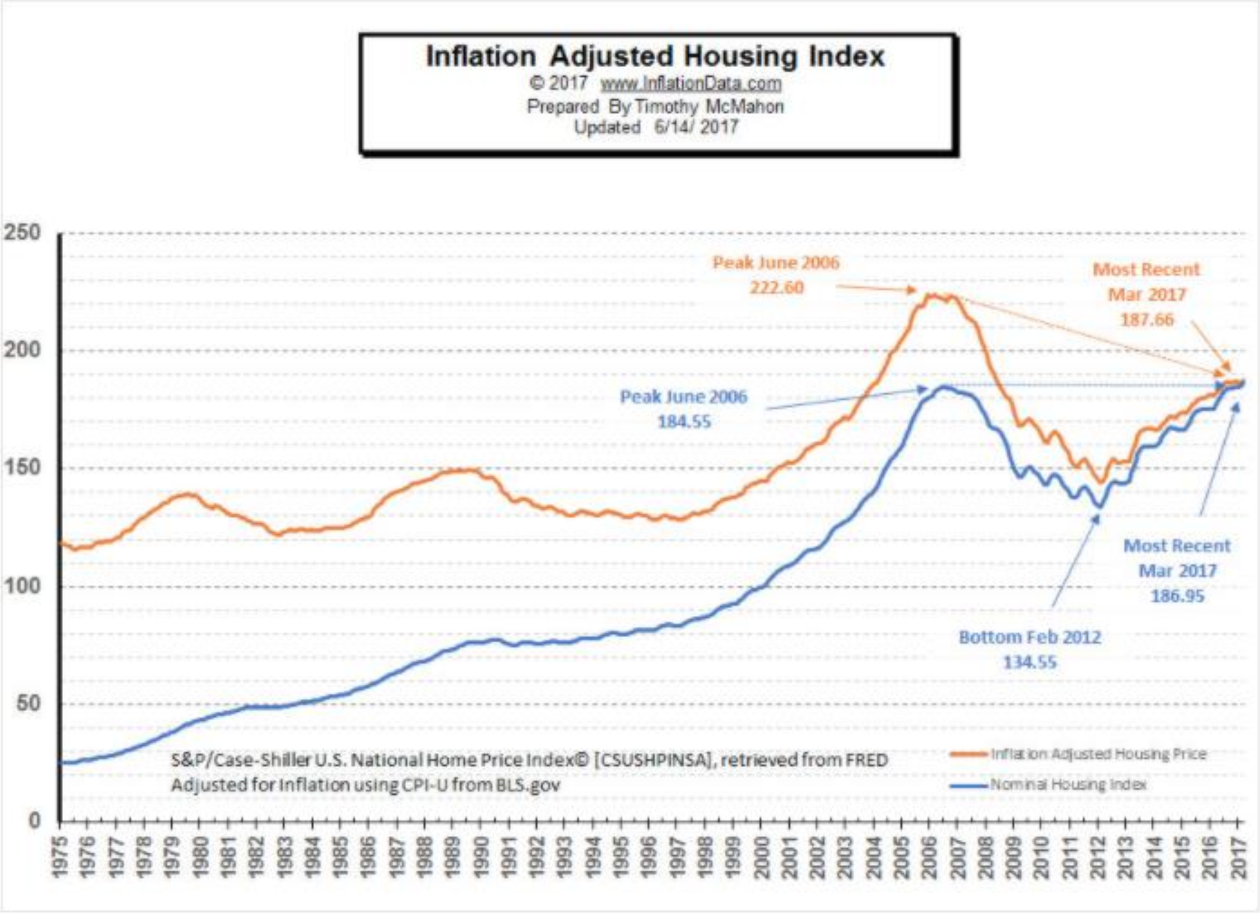


Economist.com

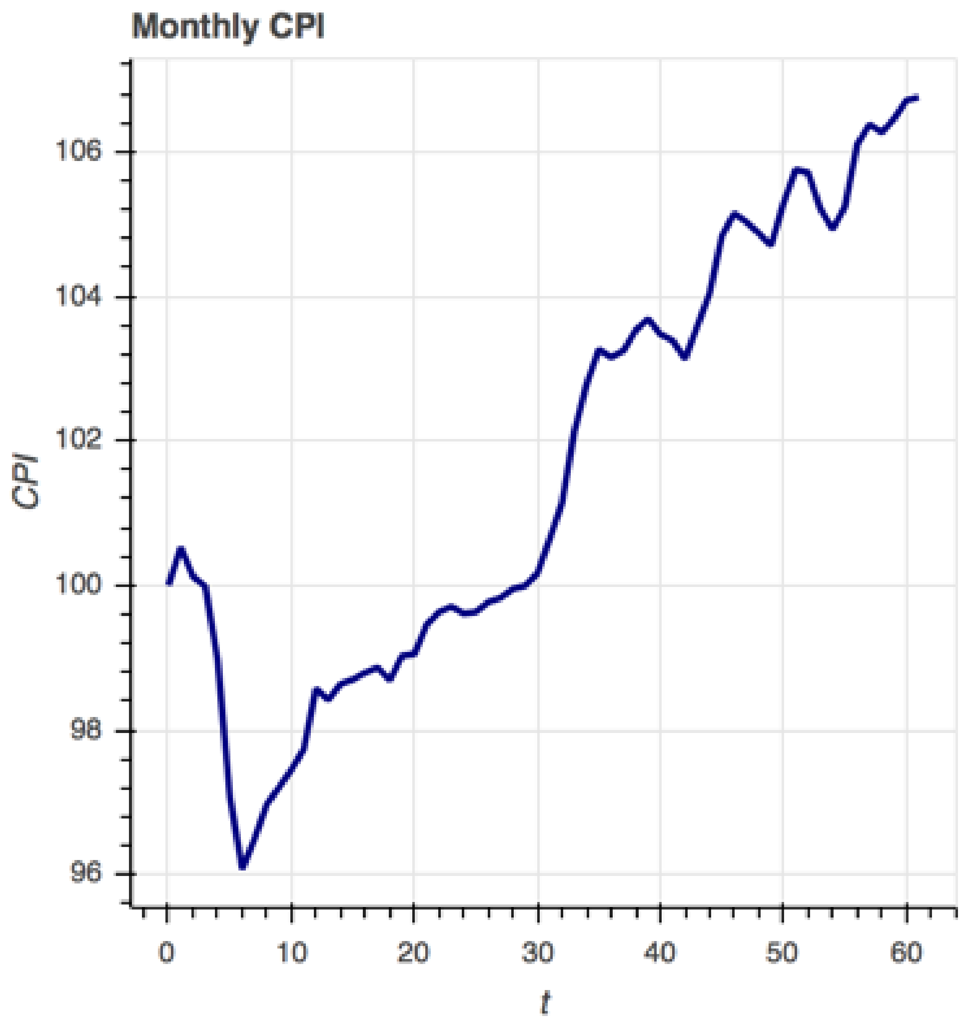
Massachusetts Institute of Technology

MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

House Prices



Consumer Price Index (billion Prices)



Massachusetts Institute of Technology

IDSS MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

…What is new here?

- What you measure today depends on yesterday
  - Could be simple or complicated dependence
  - Memory can be high
  - Memory is typically unknow
- The variation in the data can be due to a time-varying average: trends
  - Linear, quadratic (deterministic)
  - Periodic (seasonal)
- Transformation of the data may help

Stationarity

- Time series need to have some structure
- Stationarity: Some variables are ‘constant’ over time
  - Mean
  - Covariance
- What happens if such variables are changing
  - Transform to a stationary series
  - Assume slow variation

Mean and Autocovariance: Stationarity

- Mean of process is constant over time

$$\mathbf{E}(X_t) = \mu$$

- Sample Mean (constant over. λ)

$$\hat{\mu} = \frac{1}{N-\lambda}\sum_{i=\lambda}^{N-1} X_i$$

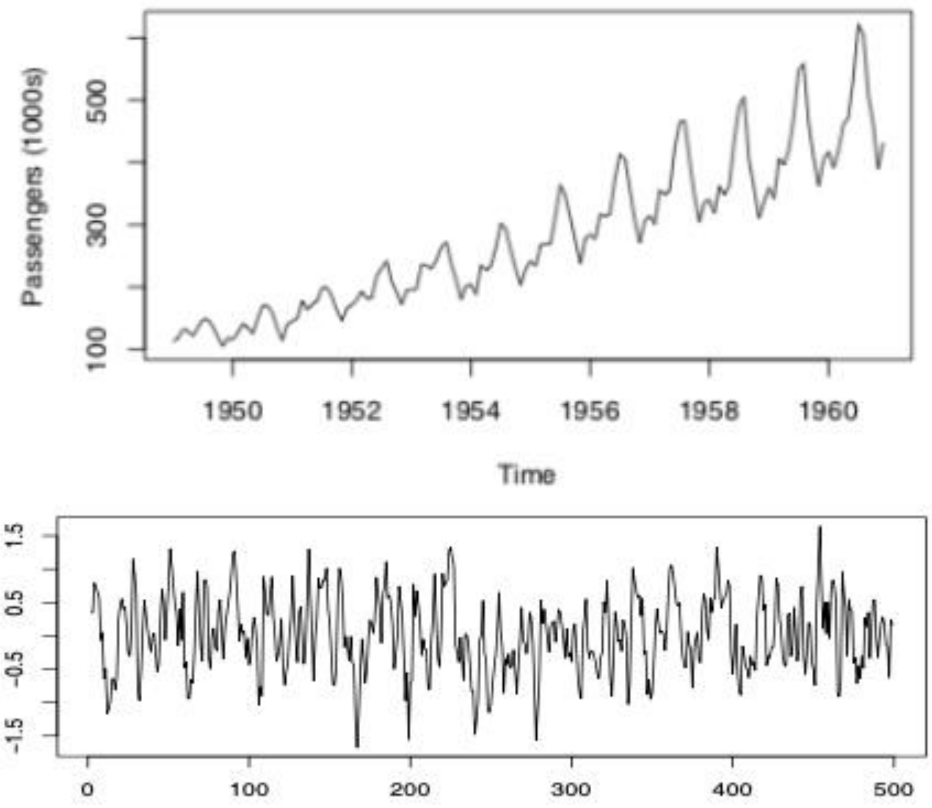
- Autocovariance is a function of the time difference

$$R_X(t_1 - t_2) = R_X(t_2 - t_1)$$

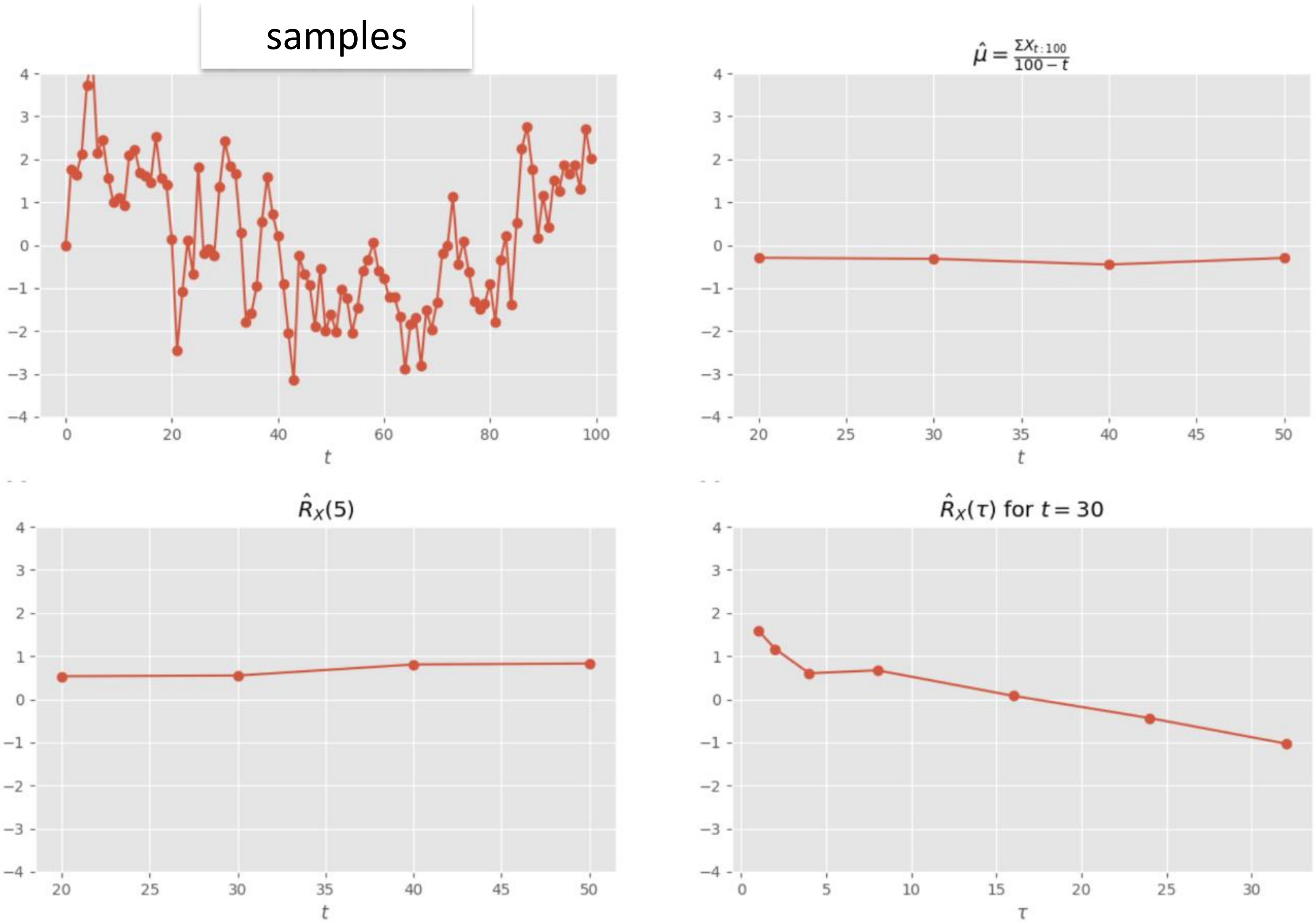
- Sample Autocovariance

$$\hat{R}_X(\tau) = \frac{1}{N-\lambda}\sum_{i=\lambda}^{N-1}(X_i - \hat{\mu})(X_{i+\tau} - \hat{\mu})$$

Testing Stationarity



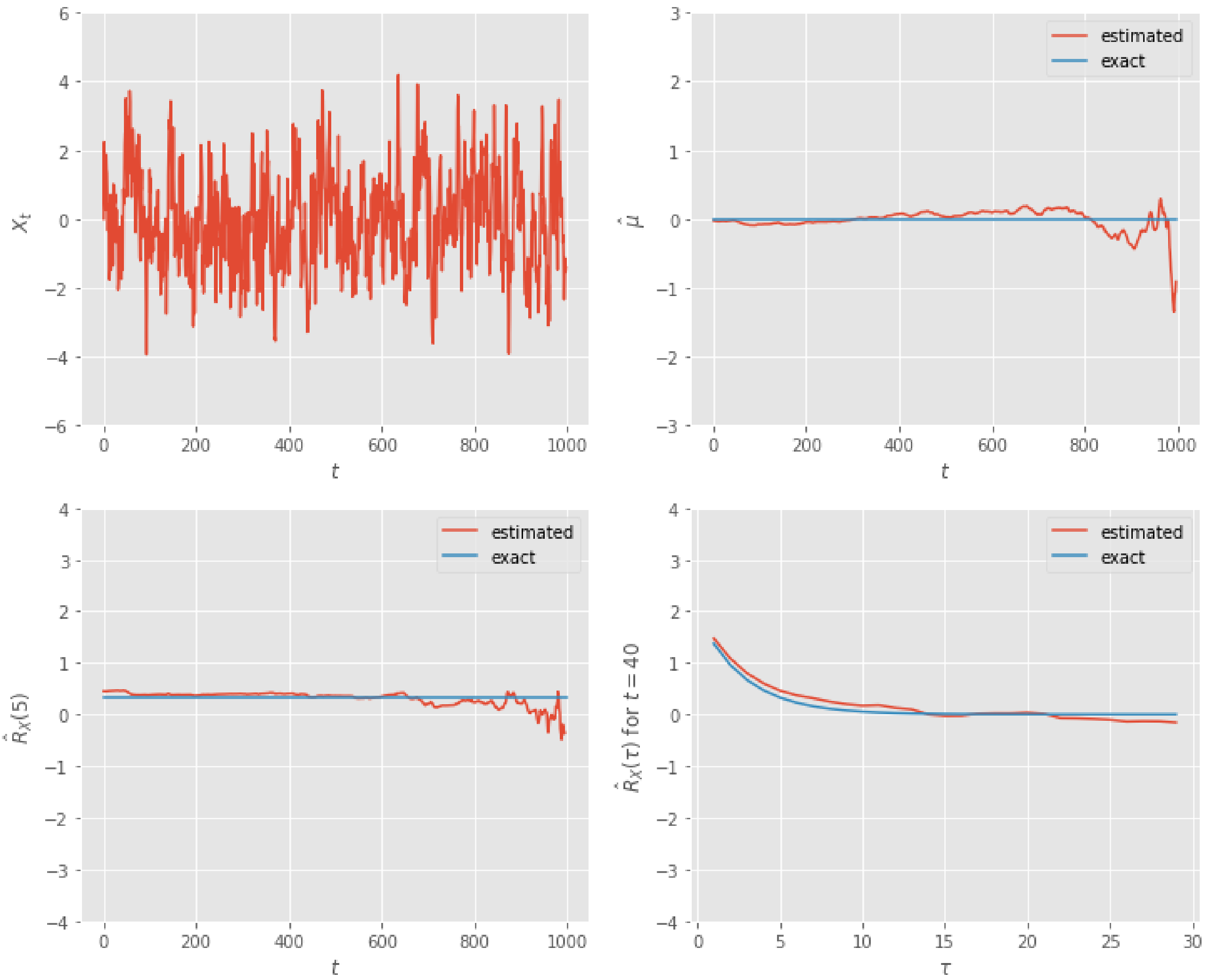
Demonstrate: 100 samples



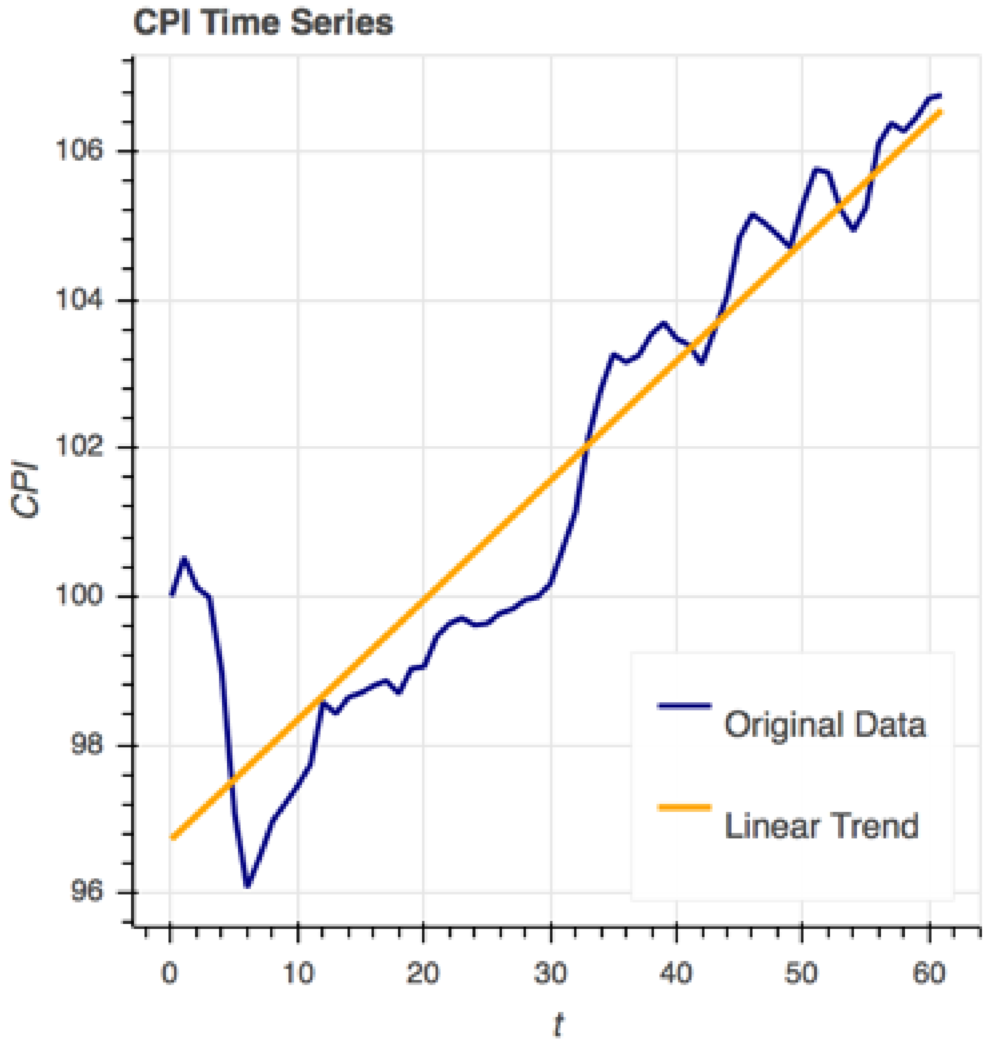
Massachusetts Institute of Technology

MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

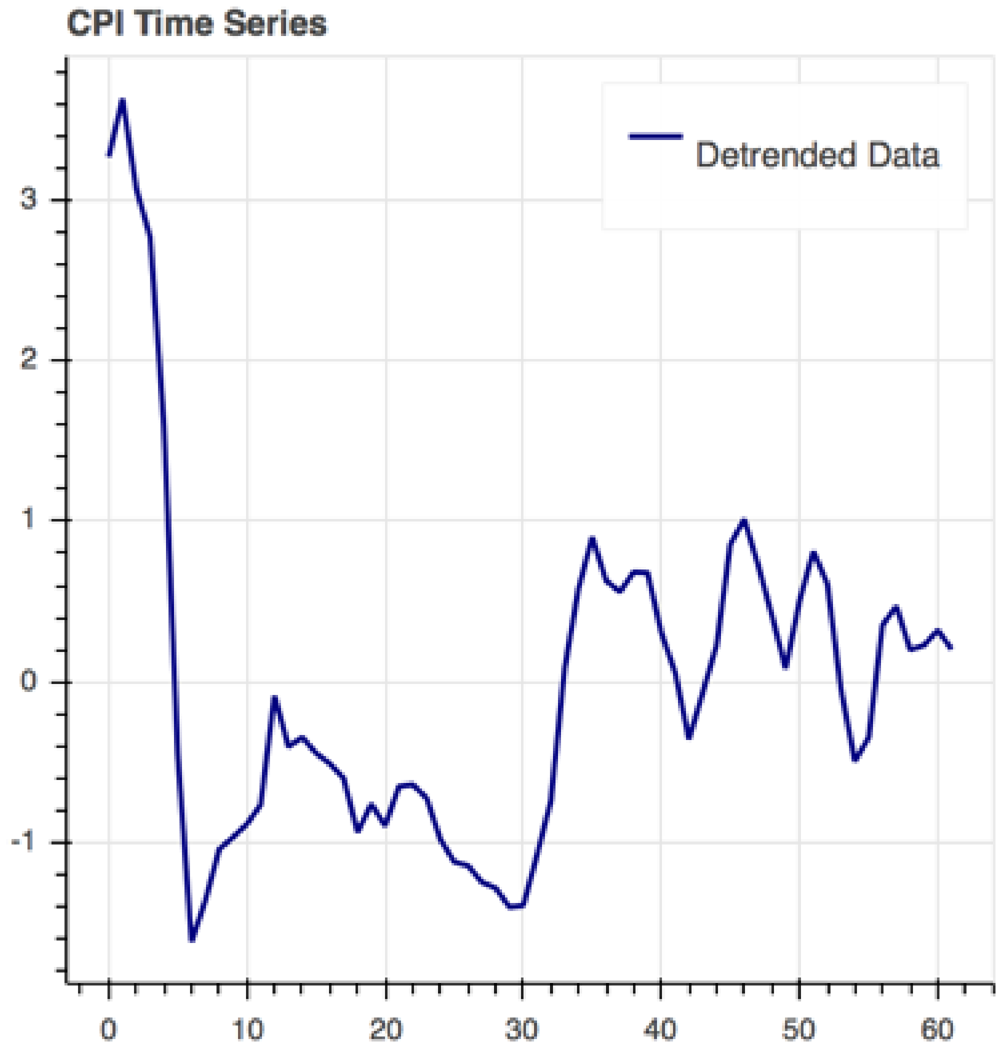
Demonstrate: 1000 samples



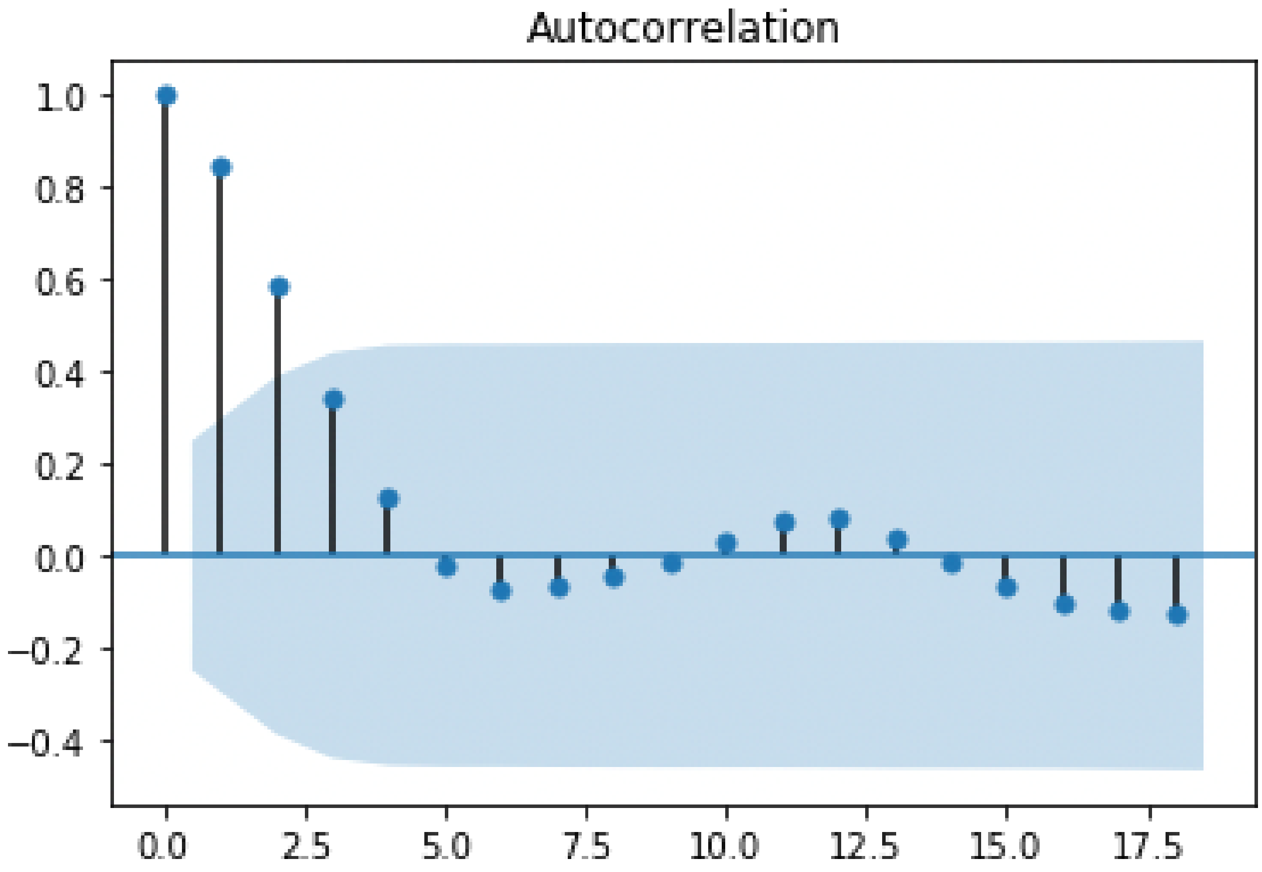
Trends in the Consumer Price Index (CPI) data



De-trended

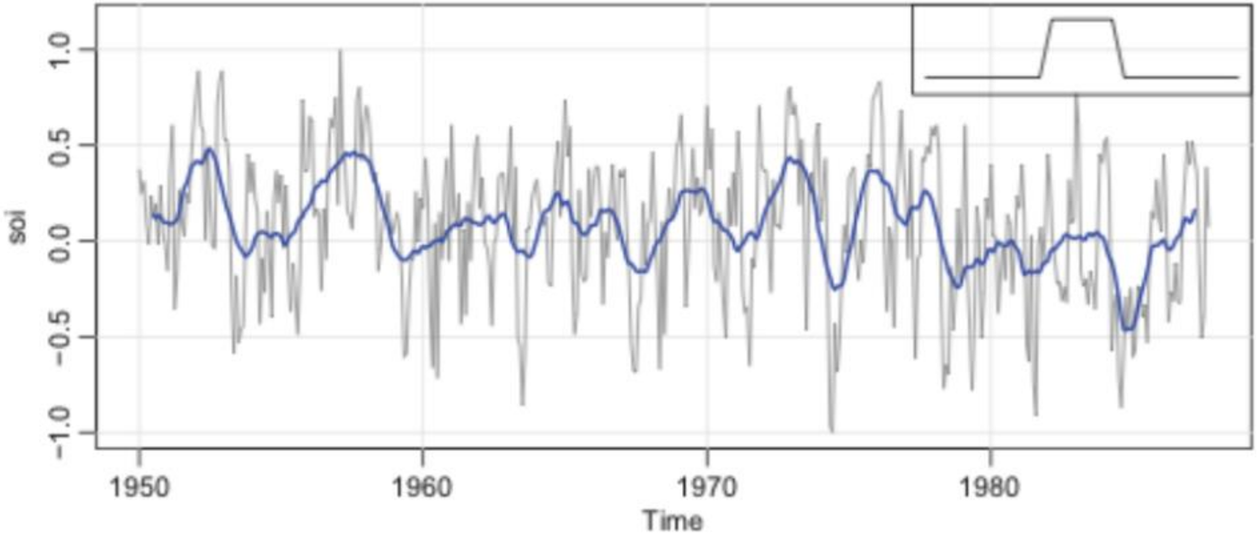


Empirical Autocovariance



Removing Seasonality

- $X_t = S_t + Y_t$
- Estimate S_t using a periodic regression
- Smoothing: $\hat{Y}_t = \sum_{h=-k}^{k} \gamma_h X_{t+h}$.



End

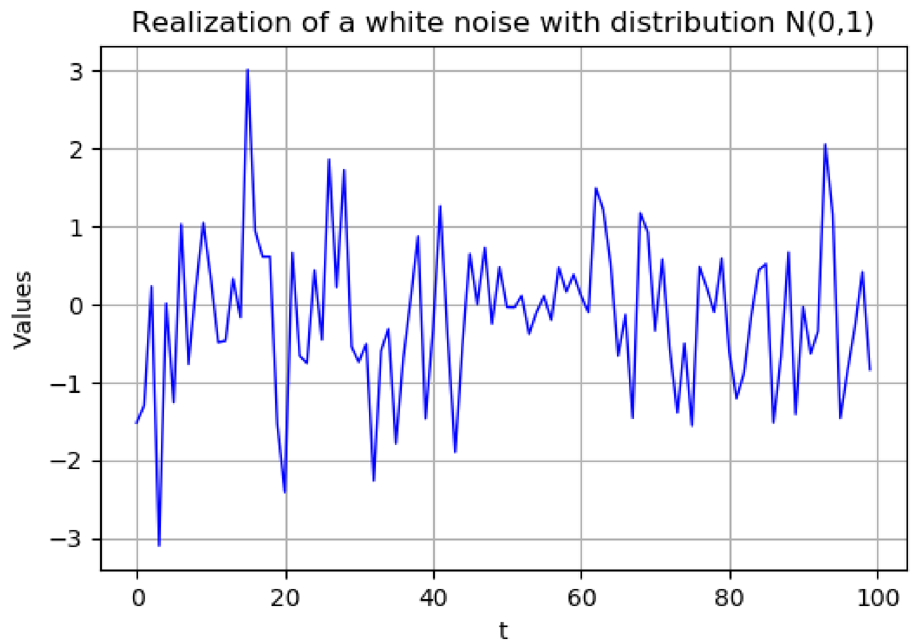
Models of time series

Models of Time-Series

- White noise
- Auto-regressive (AR)
- Moving Average (MA)
- Auto-regressive moving average (ARMA)
- Autocovariance of these examples

White noise

- Process $X_t = w$ $\mathbf{E}(w_t) = 0$

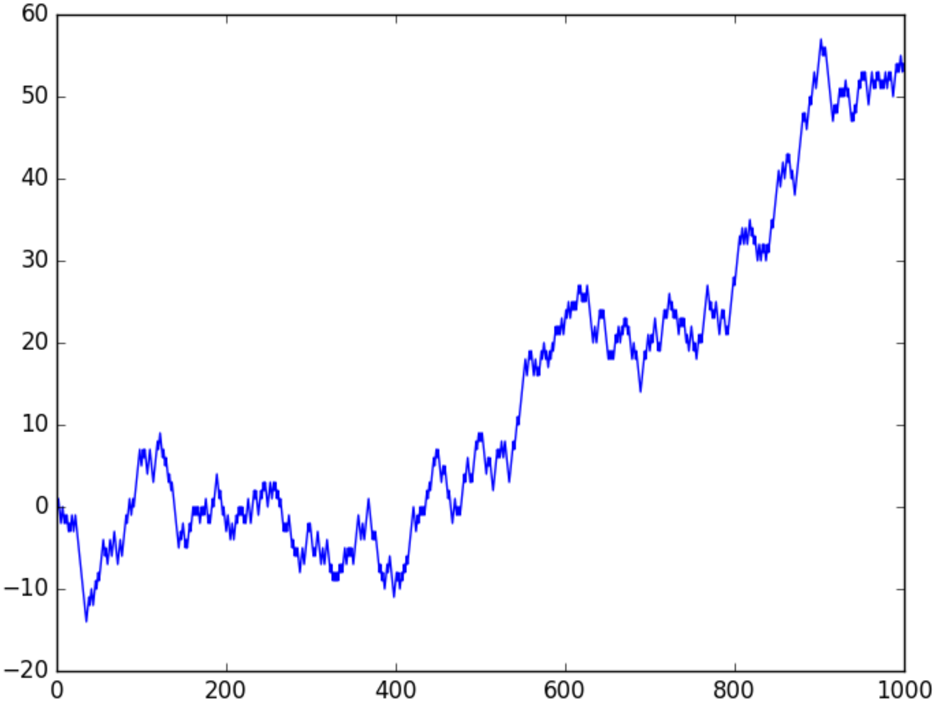


- Autocovariance =

$$\mathbf{E}(w_{t_1} w_{t_2}) = \sigma^2 \, \delta(t_1 - t_2)$$

Random Walk

$$X_t = X_{t-1} + w_t$$



Autoregressive Process AR(p)

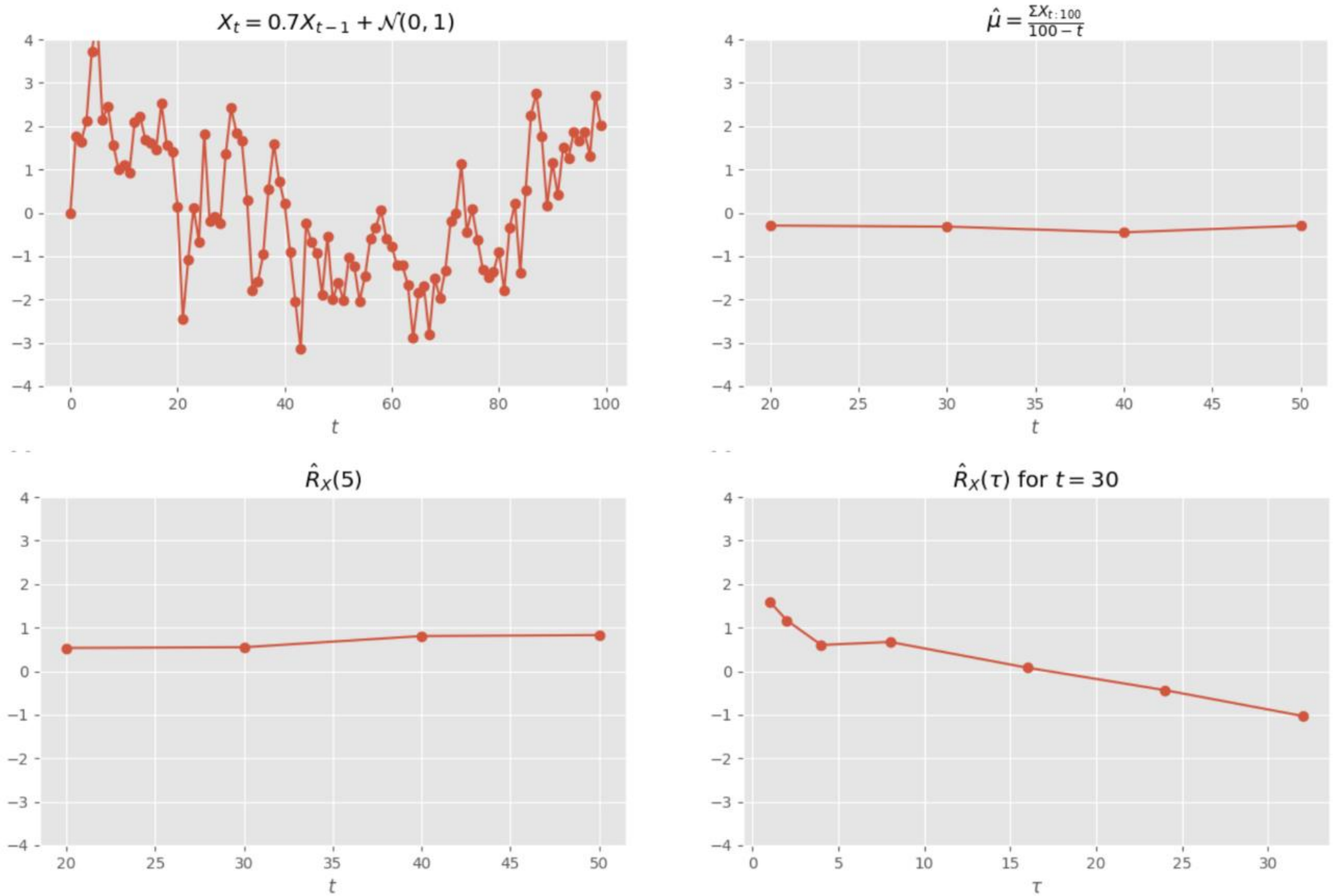
- Process

$$X_t = \sum_{i=1}^{p} a_i X_{t-i} + w_t$$

- w_t is white

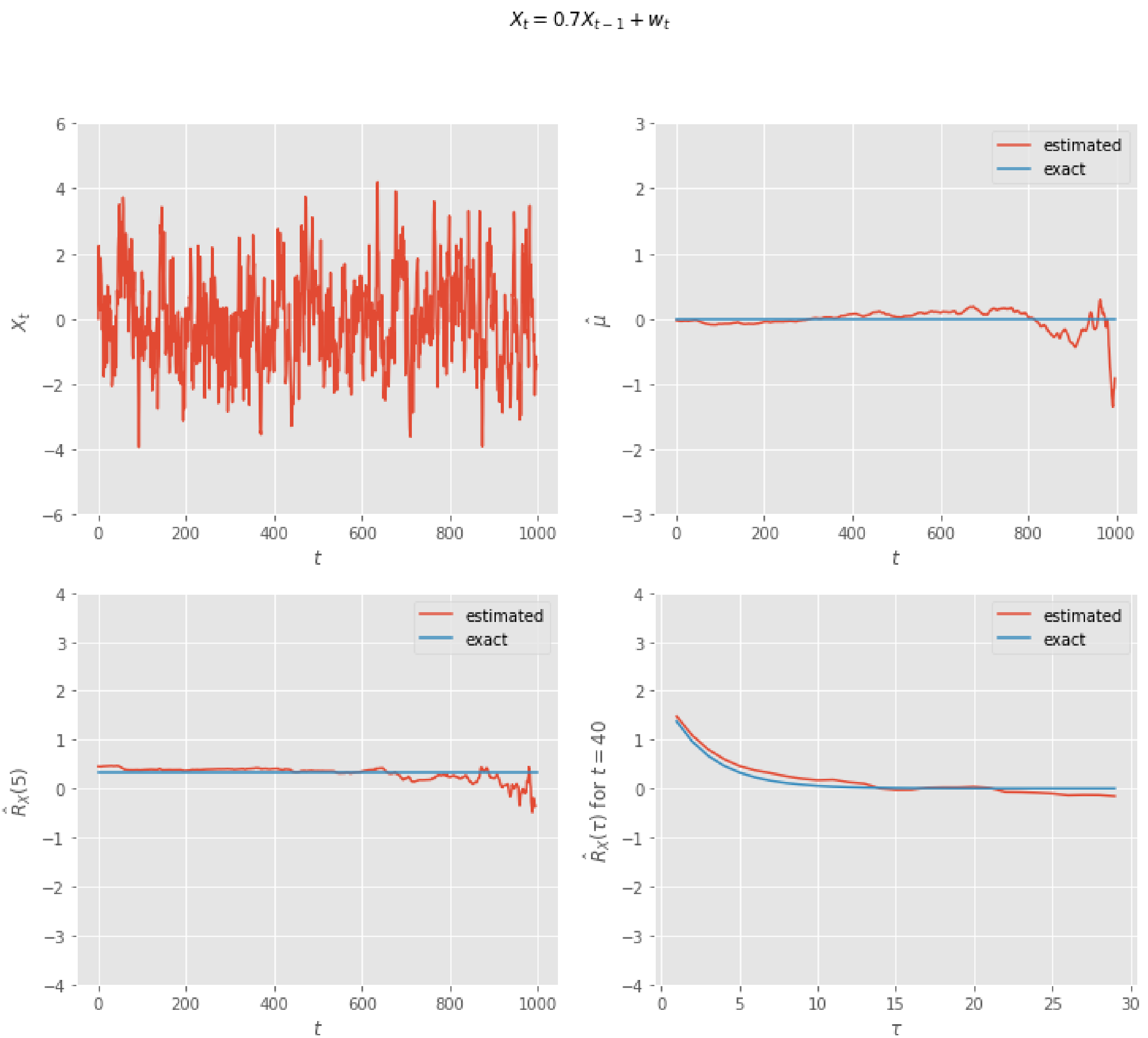
- Interpretation

Example AR(1)



$X_t = 0.7X_{t-1} + \mathcal{N}(0,1)$
$\hat{\mu} = \frac{\Sigma X_{t:100}}{100-t}$
$\hat{R}_X(5)$
$\hat{R}_X(\tau)$ for $t = 30$
t
τ

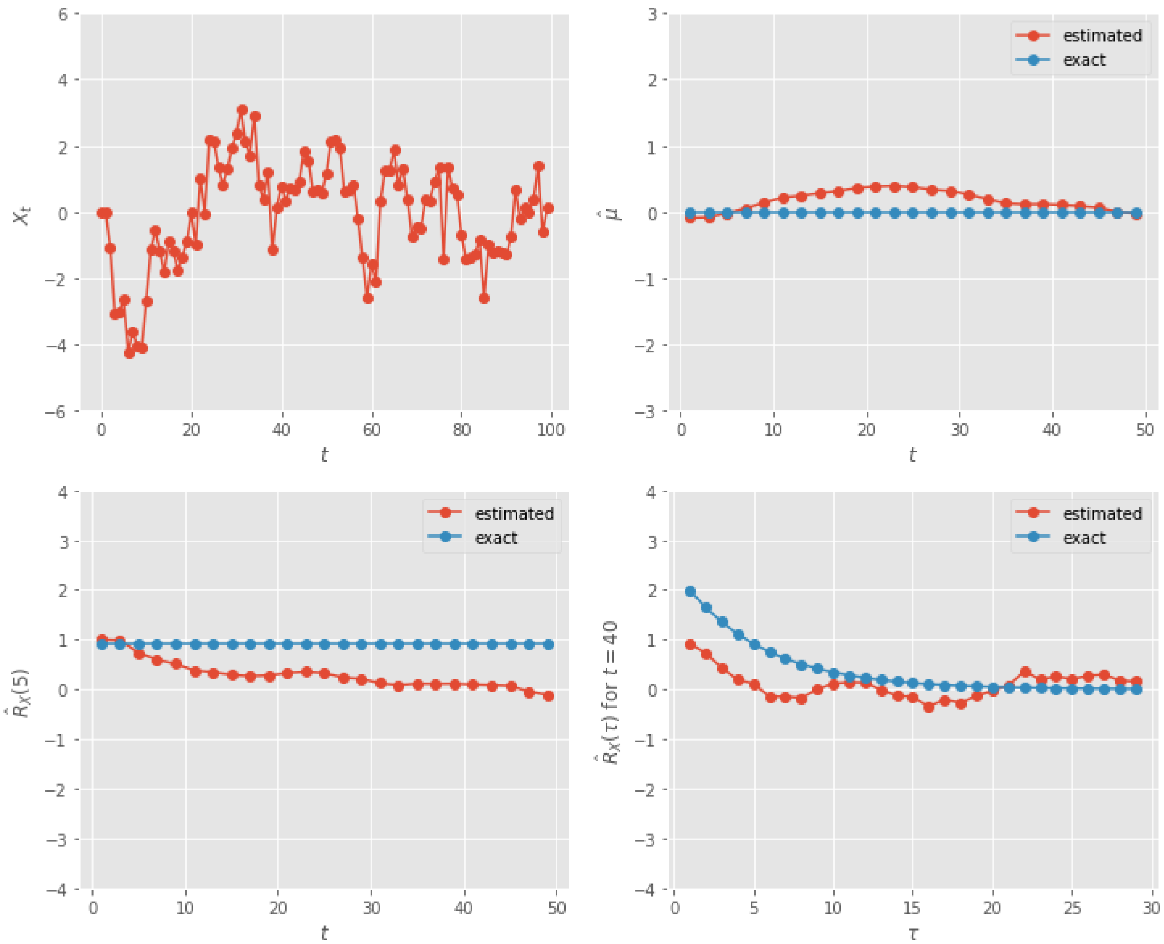
Example: AR(1)



$X_t = 0.7X_{t-1} + w_t$
X_t
t
$\hat{\mu}$
t
estimated
exact
$\hat{R}_X(5)$
t
estimated
exact
$\hat{R}_X(\tau)$ for $t=40$
τ
estimated
exact

Example: AR(2)

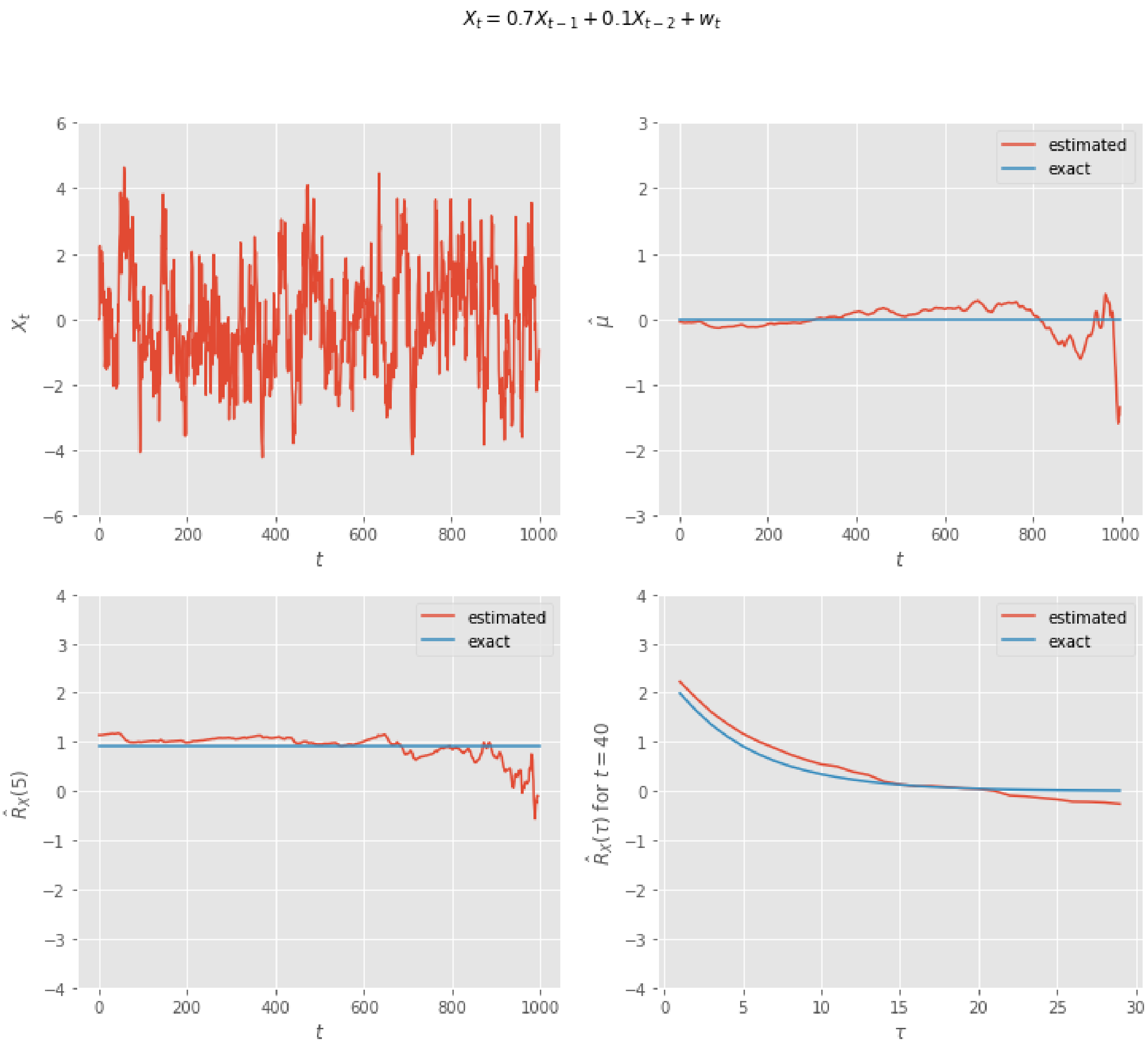
$$X_t = 0.7X_{t-1} + 0.1X_{t-2} + w_t$$



Massachusetts Institute of Technology

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

Example: AR(2)



$X_t = 0.7X_{t-1} + 0.1X_{t-2} + w_t$
X_t
t
$\hat{\mu}$
estimated
exact
$\hat{R}_X(5)$
$\hat{R}_X(\tau)$ for $t = 40$
τ

Example: AR(2)—argue it is an exponential

- Recall: $R_X(\tau) = R_X(-\tau)$

- Solve for $R_X(0), R_X(1), R_X(2)$ from

$$R_X(0) = a_1 R_X(1) + a_2 R_X(2) + \sigma^2$$

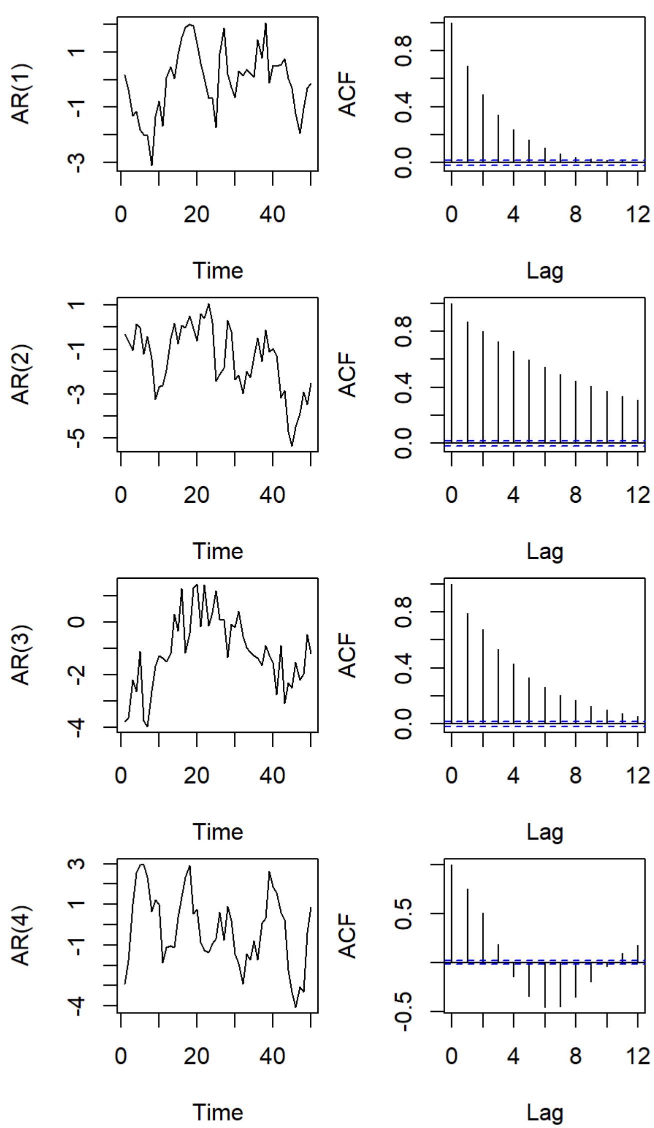
$$R_X(1) = a_1 R_X(0) + a_2 R_X(1)$$

$$R_X(2) = a_1 R_X(1) + a_2 R_X(0)$$

- Compute the rest

$$R_X(\tau) = \sum_{i=1}^{p} a_i R_X(\tau - i), \tau \geq 3$$

Geometric Shape of ACF for AR(p)



Moving Average MA(q)

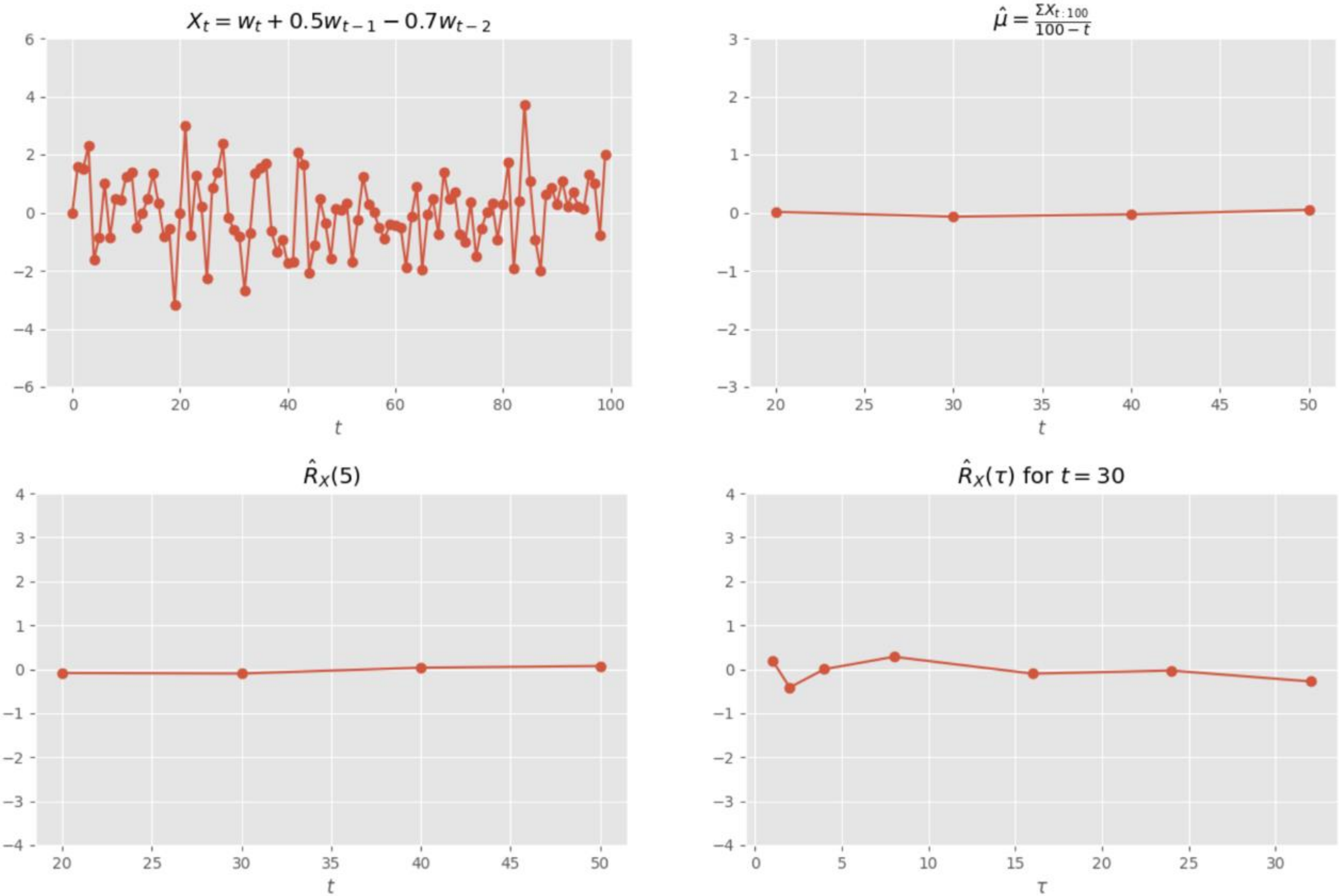
- Process

$$X_t = \sum_{i=0}^{q} b_i w_{t-i}$$

- w_t is white

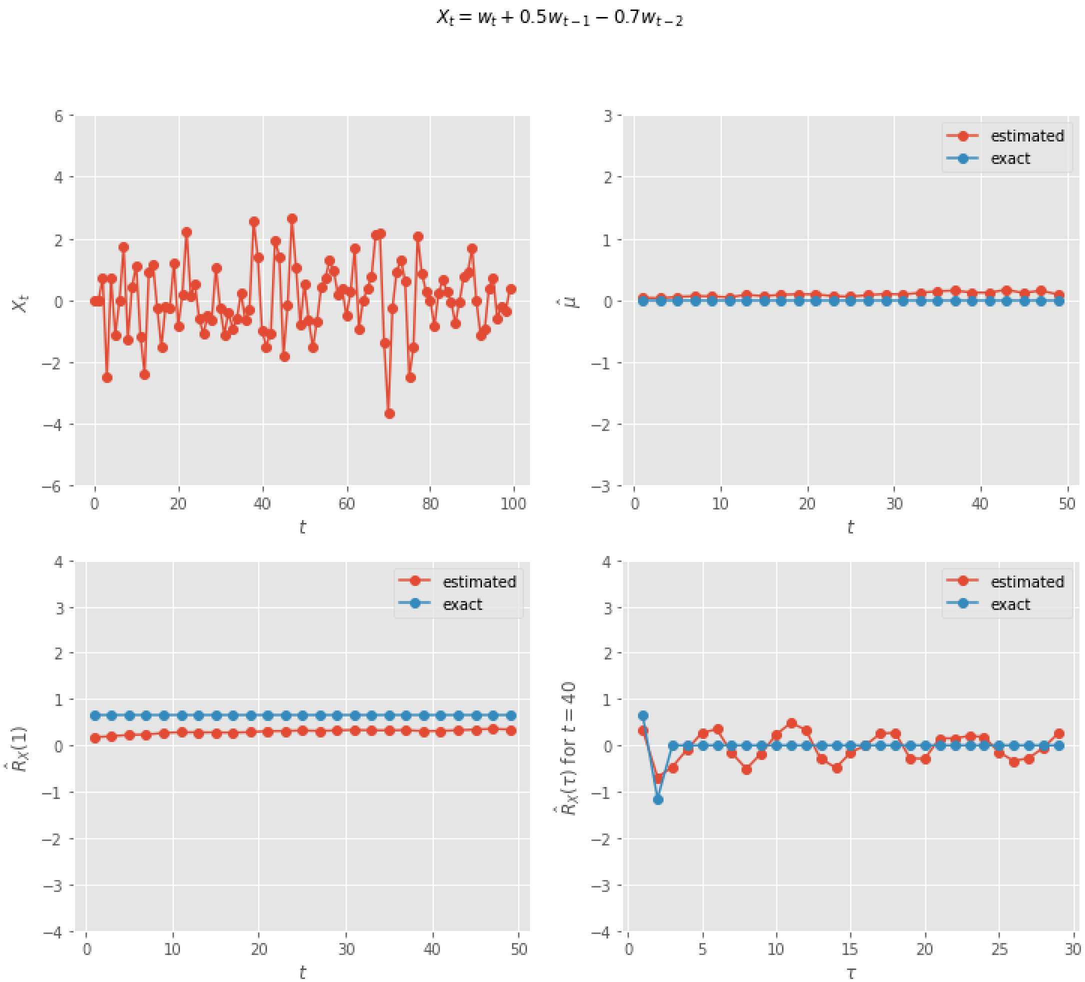
- Interpretation

Example MA(2)



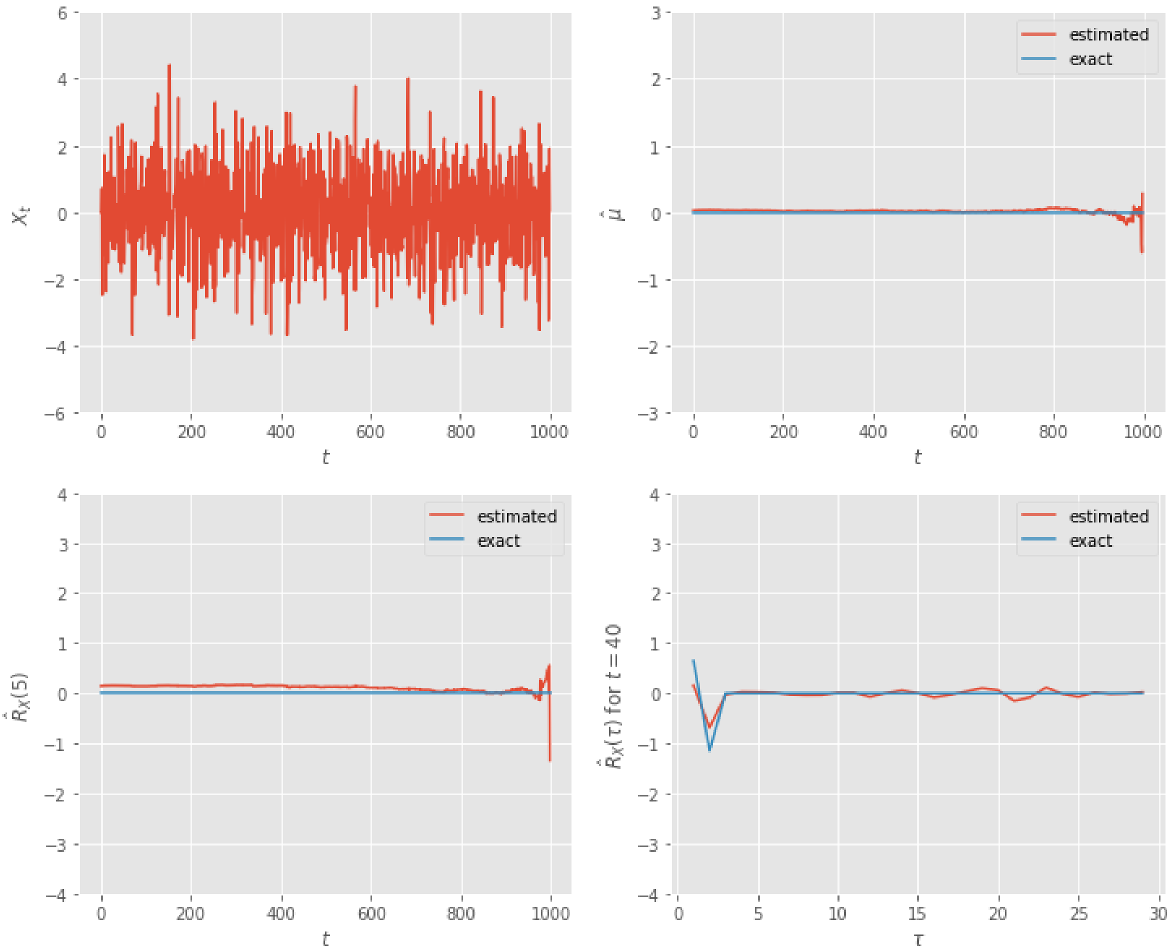
$X_t = w_t + 0.5w_{t-1} - 0.7w_{t-2}$
$\hat{\mu} = \frac{\Sigma X_{t:100}}{100-t}$
$\hat{R}_X(5)$
$\hat{R}_X(\tau)$ for $t = 30$
t
τ

Sample vs Actual



More Samples

$$X_t = w_t + 0.5w_{t-1} - 0.7w_{t-2}$$



Autocovariance of MA—finite (remove proof)

- How do you compute the exact Autocovariance of

$$X_t = \sum_{i=0}^{q} b_i w_{t-i}$$

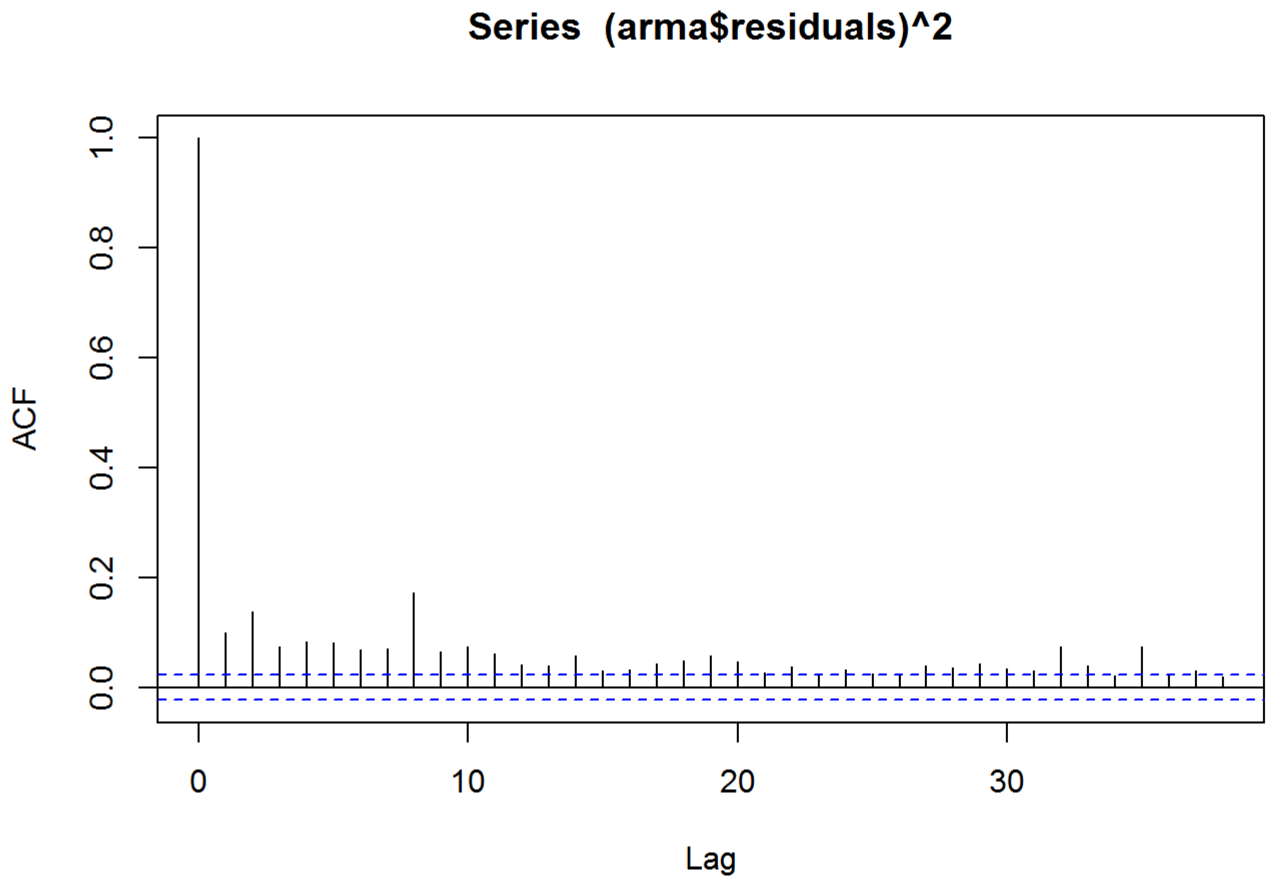
- It follows that:

$$\mathbf{E}(X_t X_{t+\tau}) = \sum_{i=0}^{q} \sum_{j=0}^{q} b_i b_j \mathbf{E}(w_{t-i} W_{t+\tau-j})$$

- Conclusion: simple convolution with finite support

$$R_X(\tau) = \sigma^2 \sum_{j=0}^{q} b_j b_{j-\tau}$$

Finite lag Shape of ACF for MA



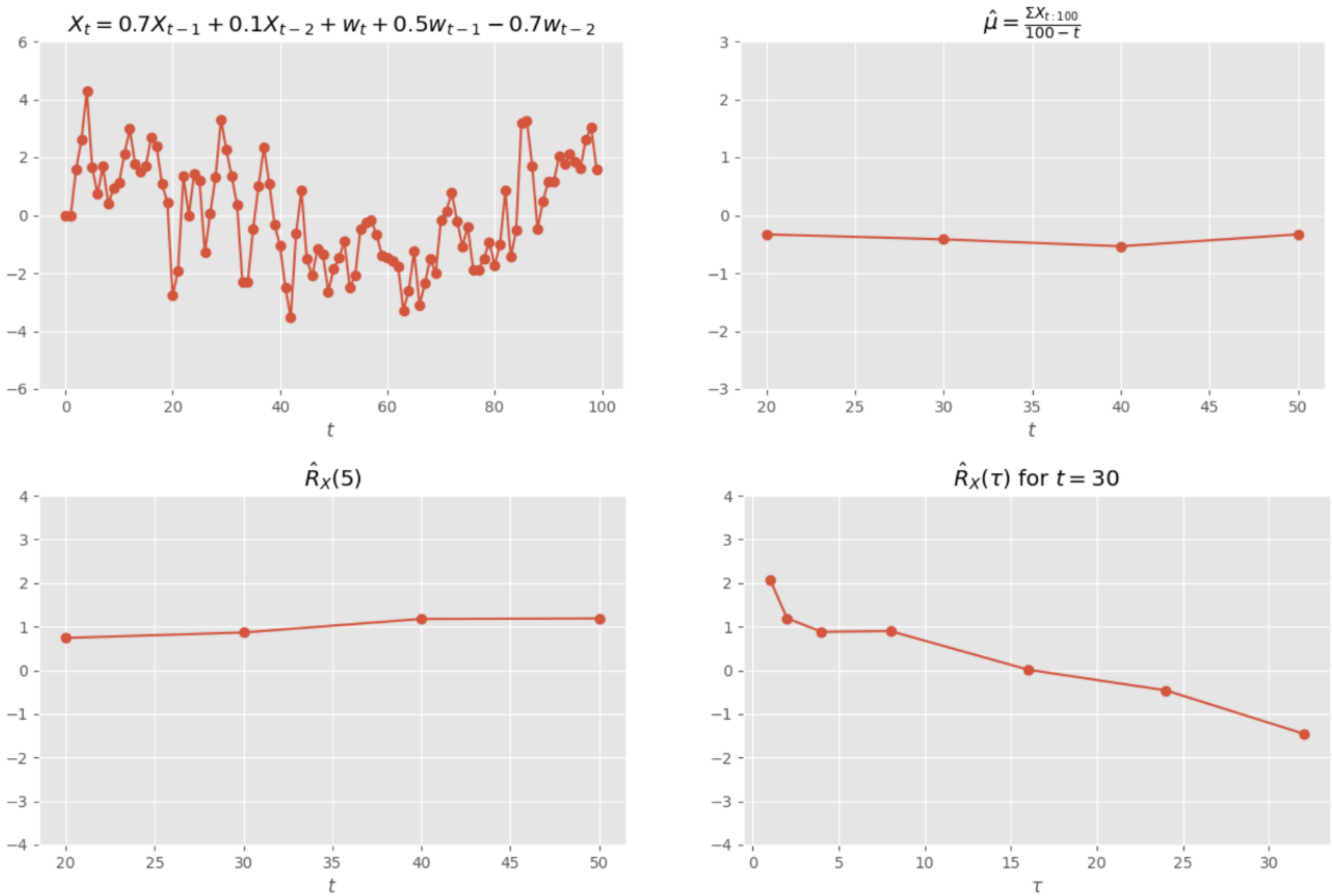
Autoregressive Moving Average Process (ARMA)

- We can combine the two processes

$$X_t = \sum_{i=1}^{p} a_i X_{t-i} + \sum_{j=0}^{q} b_j w_{t-j}$$

- Superposition of different processes

Example: ARMA

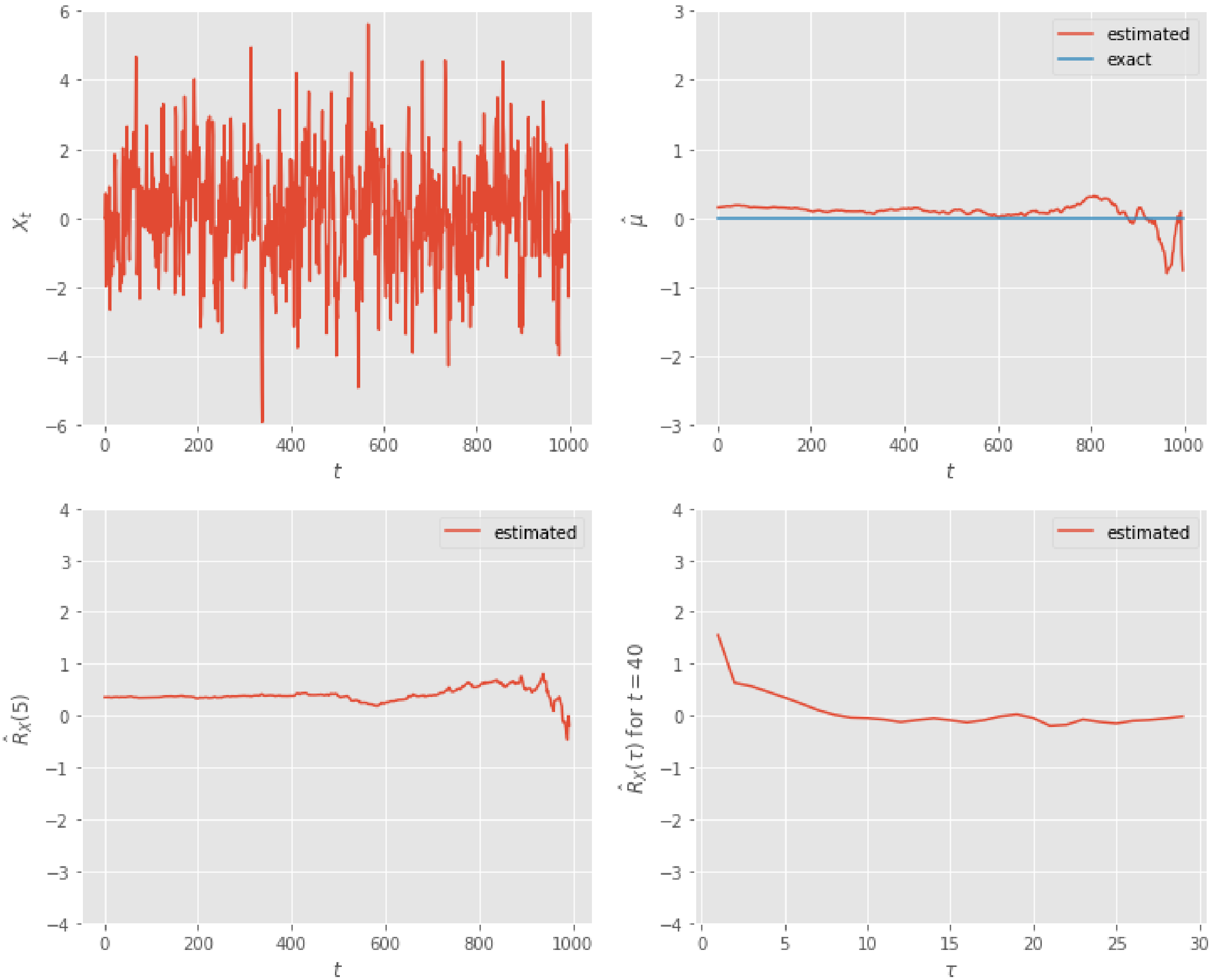


Massachusetts Institute of Technology

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

ARMA: More Samples

$$X_t = 0.7X_{t-1} + 0.1X_{t-2} + w_t + 0.5w_{t-1} - 0.7w_{t-2}$$



End

Learning Time Series

Learning a Time Series Model

- Data to Models
  - Autoregressive (AR) models
  - Moving Average (MA)
- AR learning Looks like a standard Least Squares
  - What's the catch?
  - What can we learn?
- How about other models
- Example: Consumer Price Data

Example

- Generate 100 data points according to

$$x_t = .7\, x_{t-1} + .1\, x_{t-2} + \mathrm{w_t}$$

- Fit an AR(1) model
  – Answer: $a = .8704$

- Fit an AR(2) model
  – Answer $a_1 = 0.7448 \;\; and \; a_2 = \; .1742$

- Observations?

Order Estimation

- Example highlights the importance of order estimation
- Derive multiple estimates and choose based on error on cross validation data
- Add penalty to model complexity (MDL, AIC)
- Use the Autocovariance function (ACF) as a guidance!
  - Can we do better than ACF?

Partial Autocovariance Function (PACF)

- $X_t \;\; | X_{t+1}, X_{t+2}, \ldots., X_{t+k-1} | \; X_{t+k}$

- Project X_{t+1} and X_{t+k} on the variables in between:

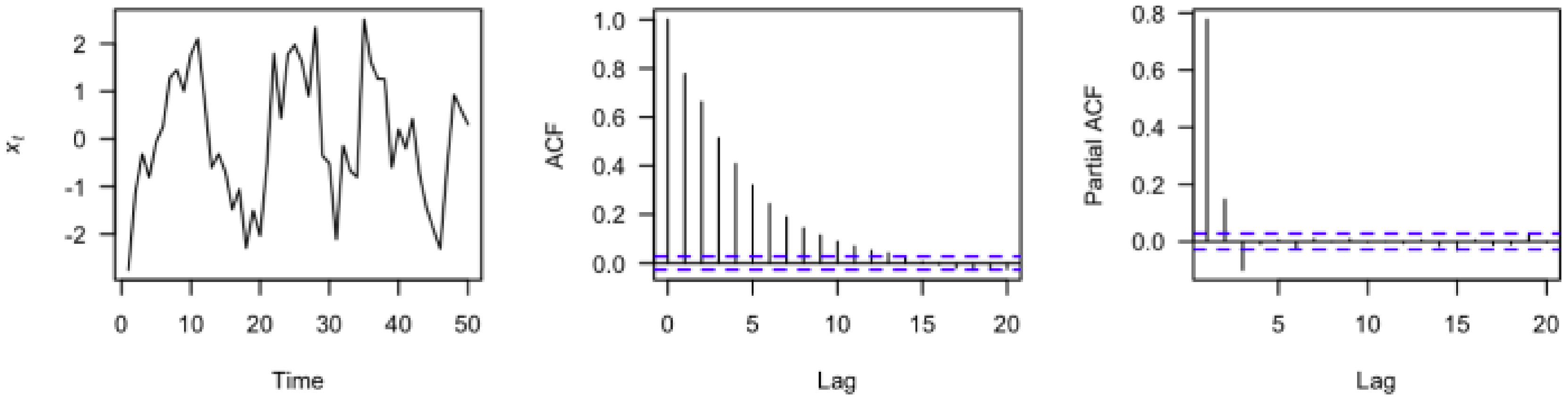
$$P_t \quad and \quad P_{t+k}$$

- In the case of ARX, the new PACF is zero outside p

- *Use to estimate order!*

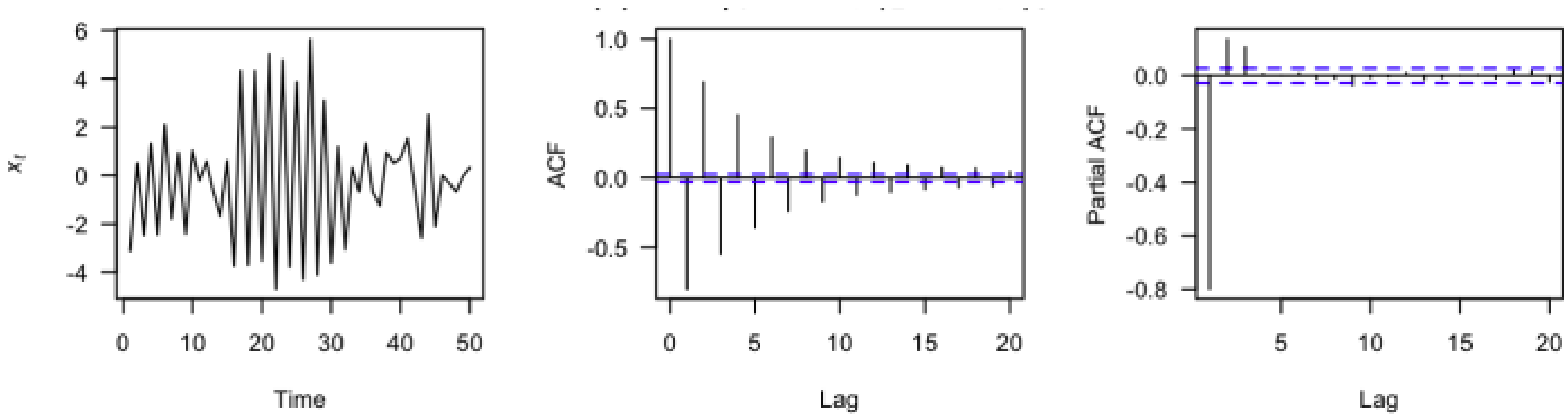
Example: AR(3)

$$X_t = .7\,X_{t-1} + .2\,X_{t-2} - .1\,X_{t-3}$$

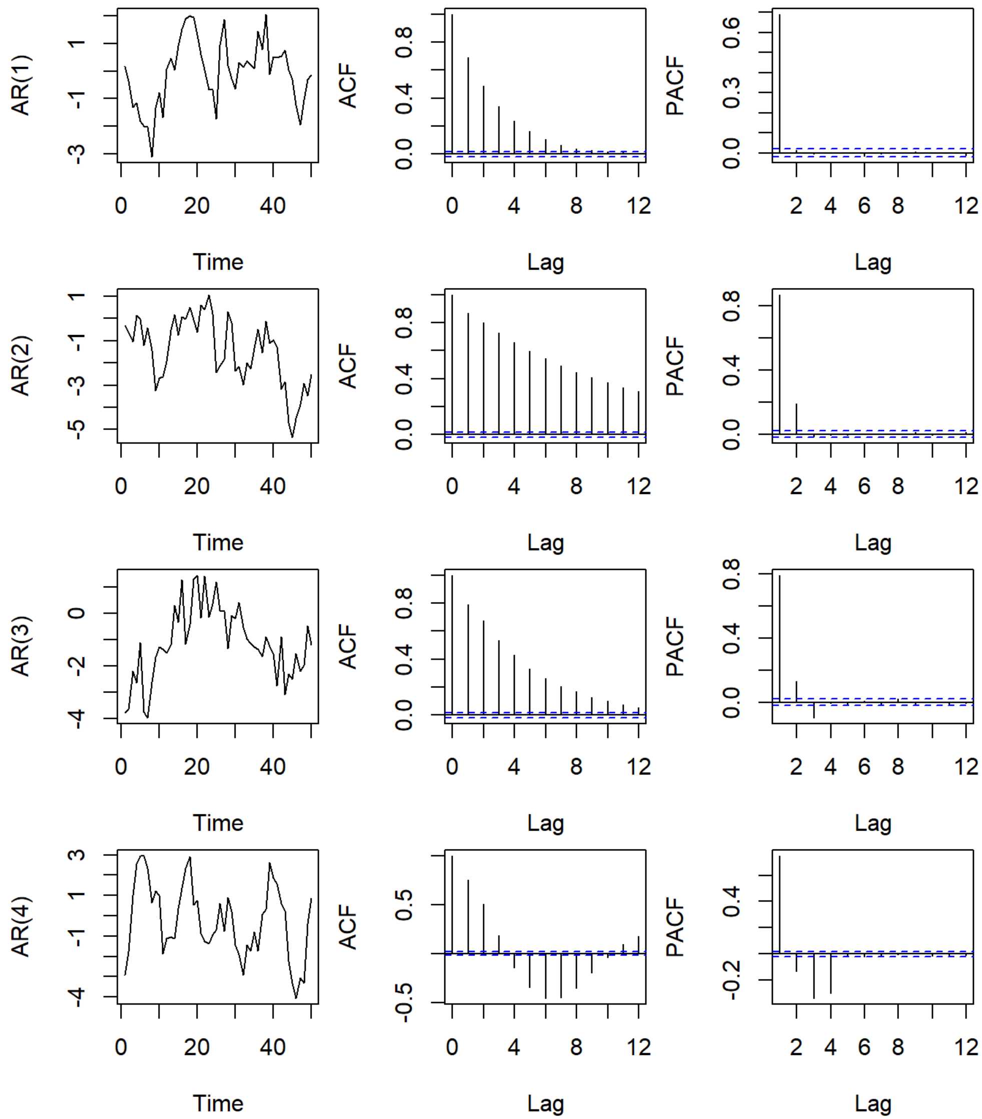


Example: AR(3)

$$X_t = -.7\,X_{t-1} + \;.2\,X_{t-2} + .1\,X_{t-3}$$



Geometric Shape of ACF for AR(p)

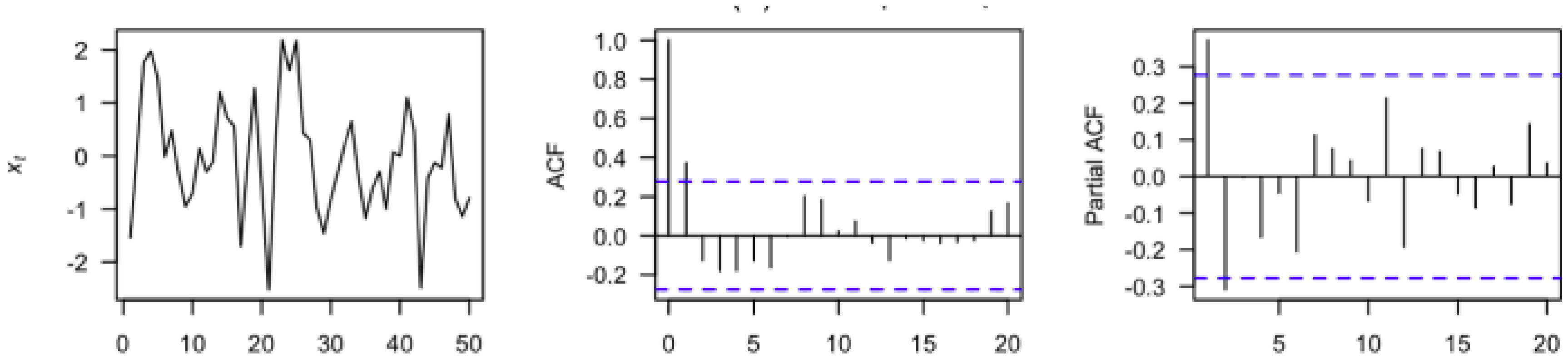


Massachusetts Institute of Technology

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY

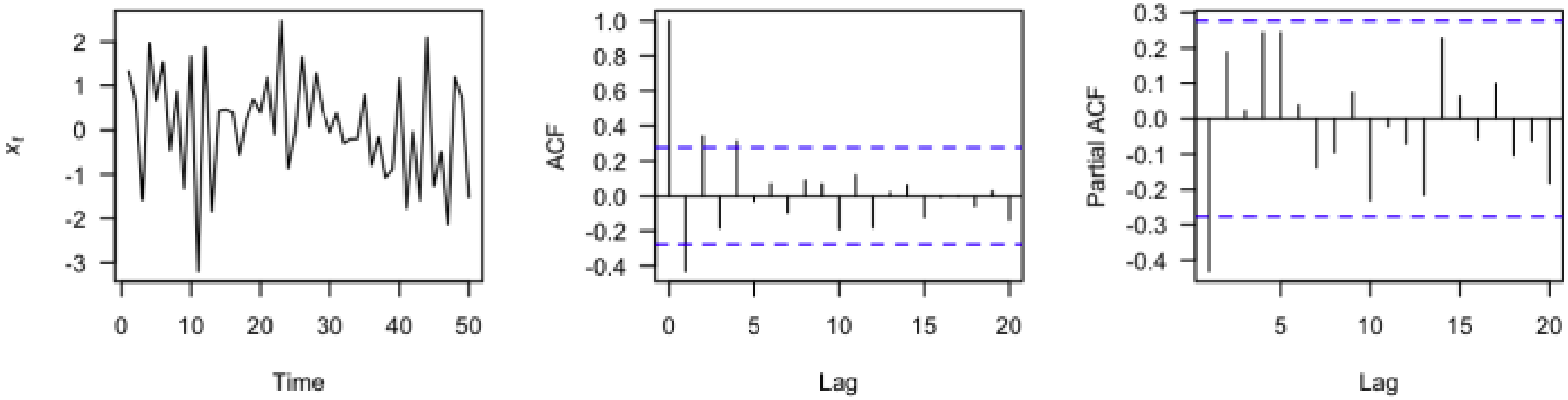
Example: MA(2)

$$X_t = .7\,W_{t-1} + W_t$$



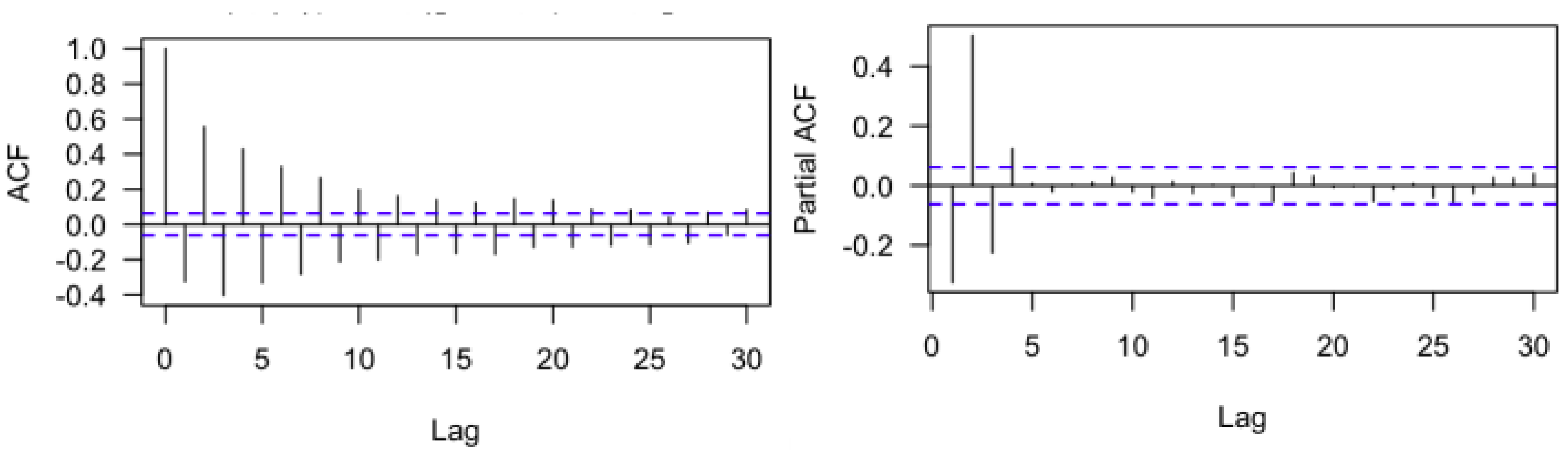
Example: MA(3)

$$X_t = W_t - .7\, W_{t-1} + \; .2\, W_{t-2} + .1 \;\; W_{t-3}$$



Example: ARMA(3)

$$X_t = .7\,X_{t-1} + \;.2\,X_{t-2} + .7\,W_t + \;.2\,W_{t-1}$$



ACF vs PACF

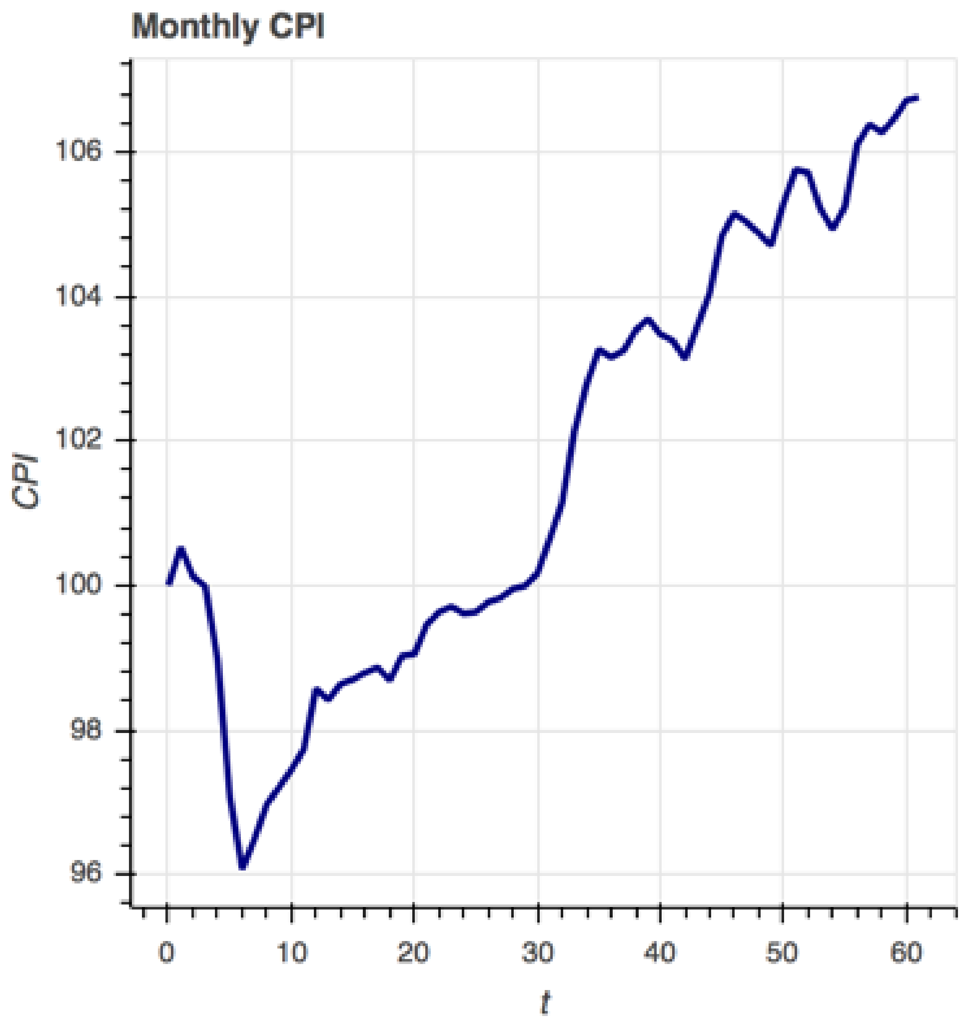
	ACF	PACF
AR(p)	decays	zero for $h > p$
MA(q)	zero for $h > q$	decays
ARMA(p,q)	decays	decays

Billion Price Data

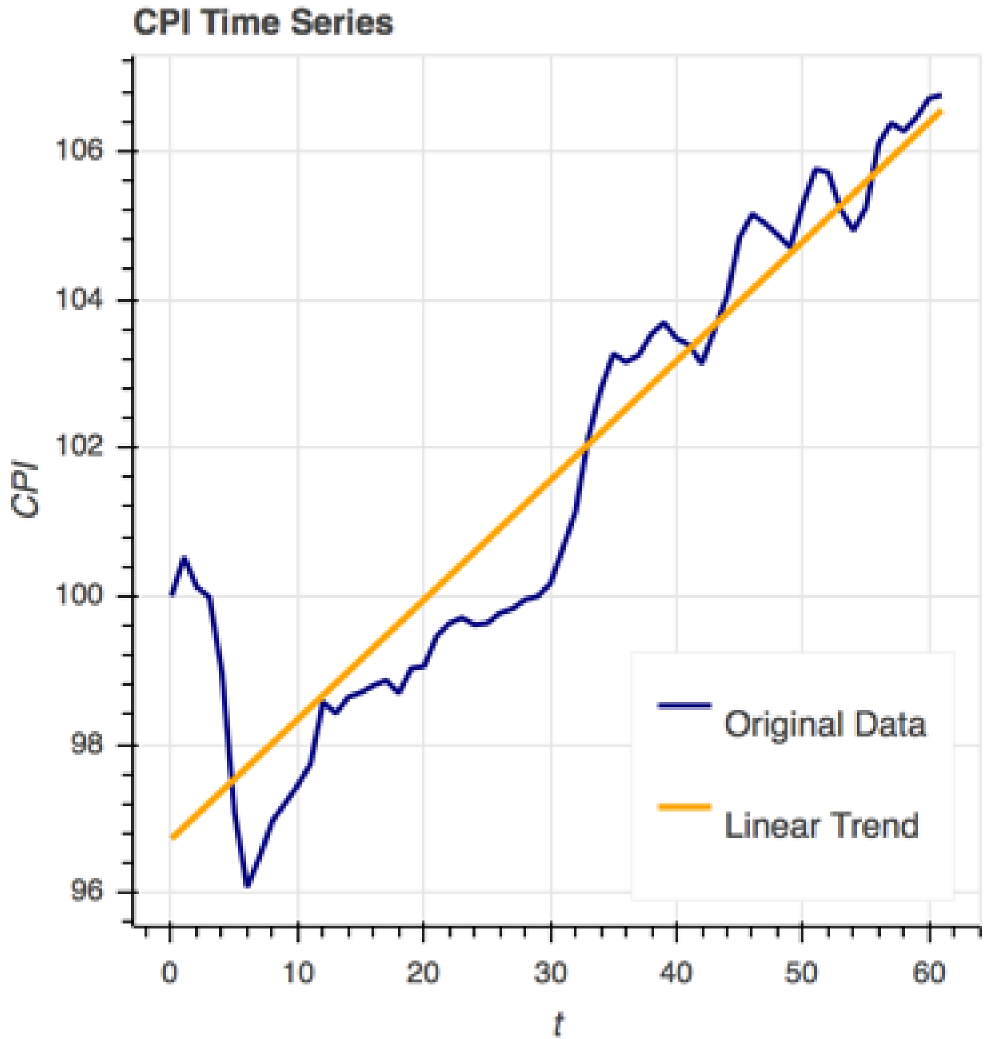
- The goal of this problem is to analyze the PriceStats data from the MIT Billion Prices Project, provided by Professor Rigobon.

- Consumer Price Index Data: (consumer price index, the price of a ”market basket of consumer goods and services” - a proxy for inflation)

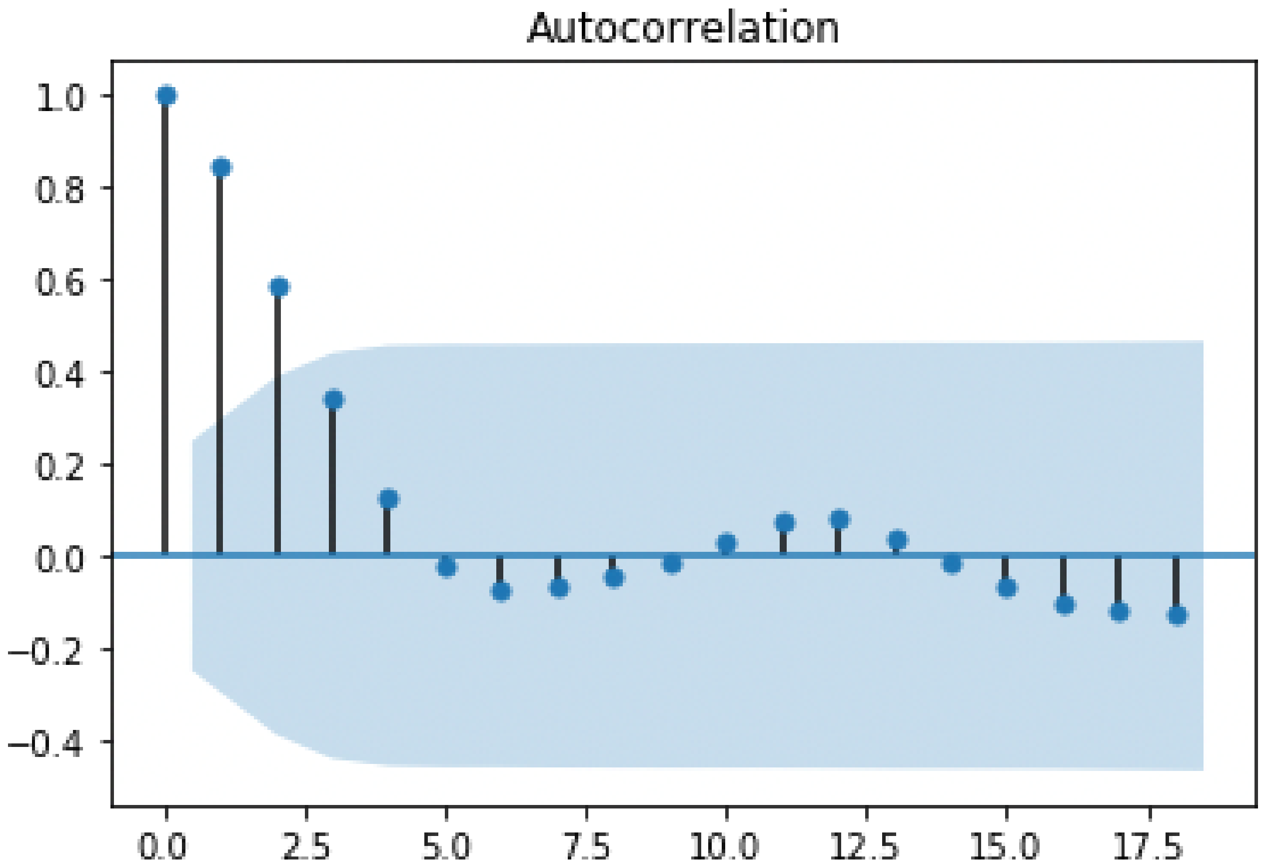
Consumer Price Index



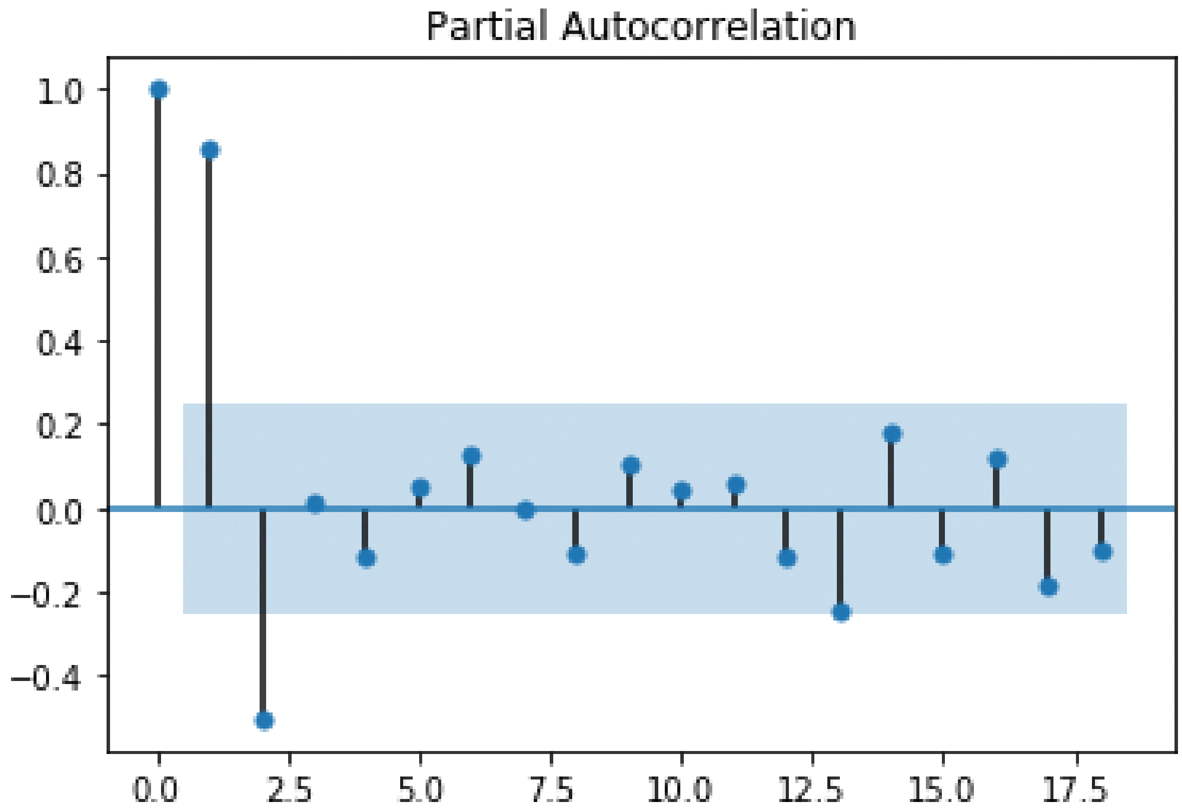
Identify Trends in the CPI data



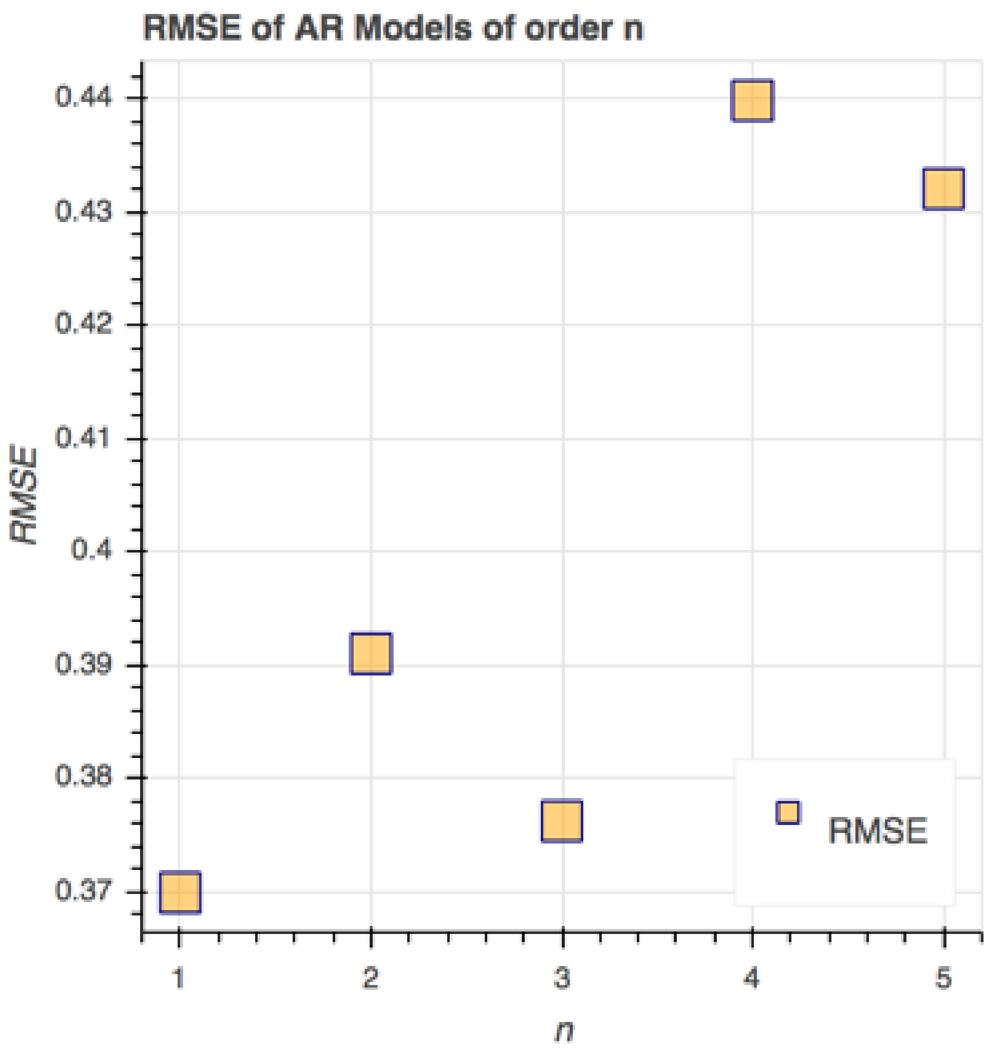
Empirical Autocovariance



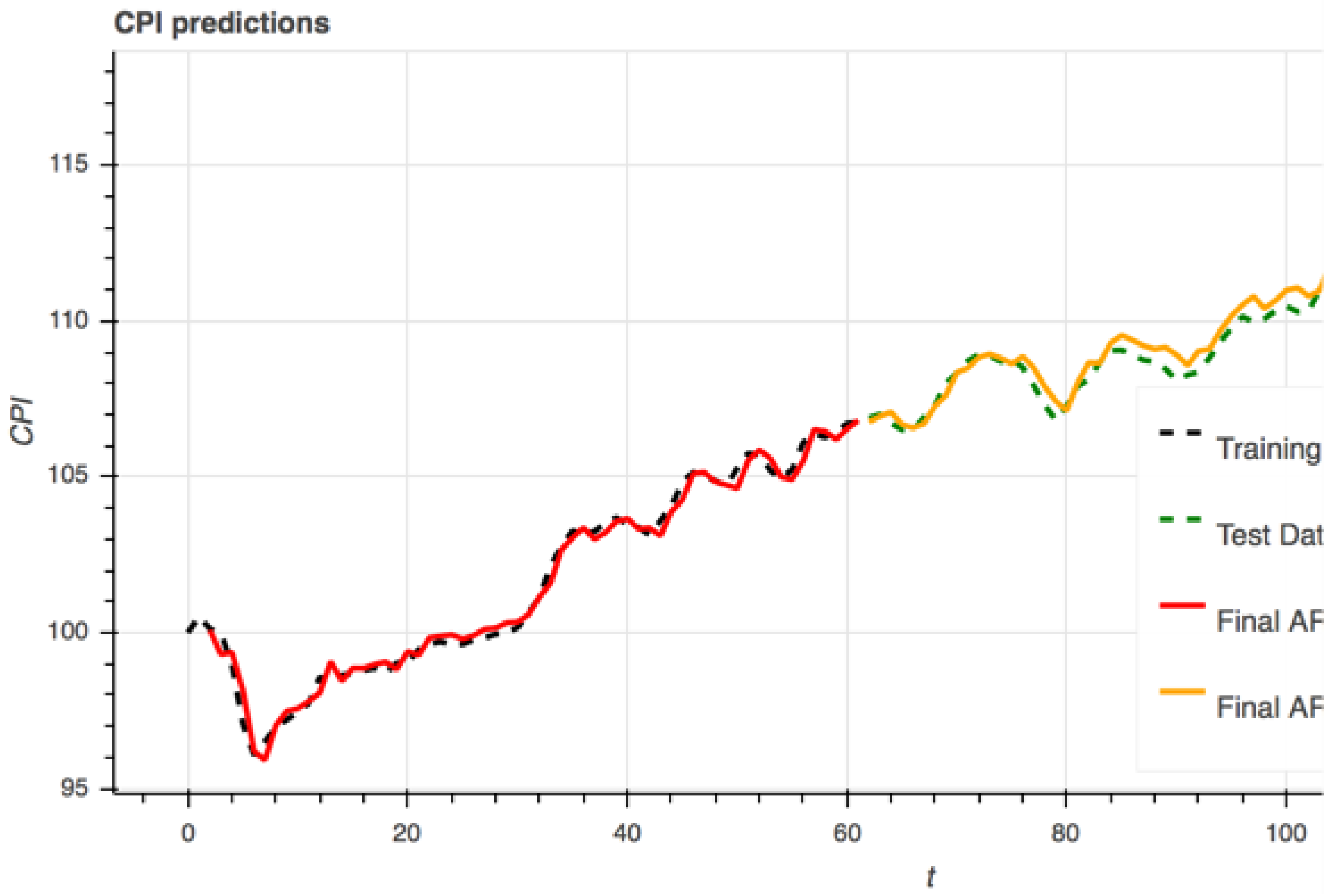
Empirical Partial Autocovariance



AR(1) or AR(2)


RMSE of AR Models of order n
RMSE
0.44
0.43
0.42
0.41
0.4
0.39
0.38
0.37
1
2
3
4
5
n
RMSE

Prediction of AR(2)



Thank You

Massachusetts Institute of Technology

IDSS
MIT INSTITUTE FOR DATA, SYSTEMS, AND SOCIETY