A little more about Fourier

$$\int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{-\infty}^{\infty} |X(j\omega)|^2 \, d\omega$$

$$x(t) = 1 \xrightarrow{FT} 2\pi\delta(\omega) = X(j\omega)$$

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(j\omega)e^{j\omega t}\,d\omega$$

$$= \frac{1}{\cancel{2\pi}}\int \cancel{2\pi}\,\delta(\omega)e^{j\omega t}\,d\omega$$

$$= \int \delta(\omega)e^{j\omega t}\,d\omega = 1$$

Queen Mary School Hainan
Queen Mary University of London

QHP5701 Exploratory Data Analysis

Time Series Analysis

Nikesh Bajaj, PhD
Lecturer in Data Science,
Queen Mary University of London
nikesh.bajaj@qmul.ac.uk
https://nikeshbajaj.in

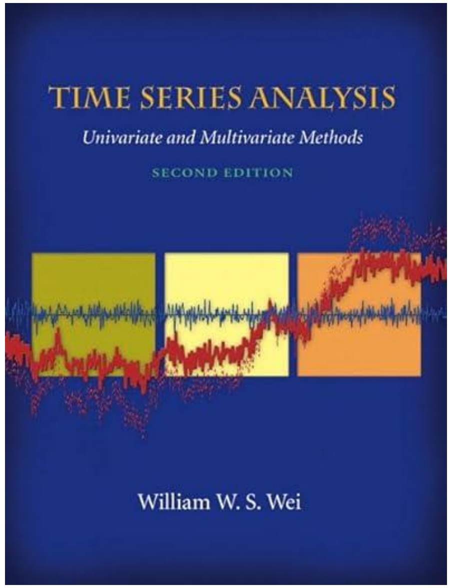
#Ref: Chapter 1 to 4, +
Time Series Analysis by William WS Wei

Time Series Analysis

Time Series Analysis
Sources & Processing

- Stochastics processes
- *Correlation*
- ACF, AVF, PACF, Noise, Visualisation
- MA, AR, ARMA, ARIMA
- Stationary & Non-stationary
- Seasonality

- Time Series Analysis



Queen Mary School Hainan
Queen Mary University of London

QHP5701 Exploratory Data Analysis

Stochastic Process

Nikesh Bajaj, PhD
Lecturer in Data Science,
Queen Mary University of London
nikesh.bajaj@qmul.ac.uk
https://nikeshbajaj.in

#Ref: Chapter 2, TSA, William WS Wei

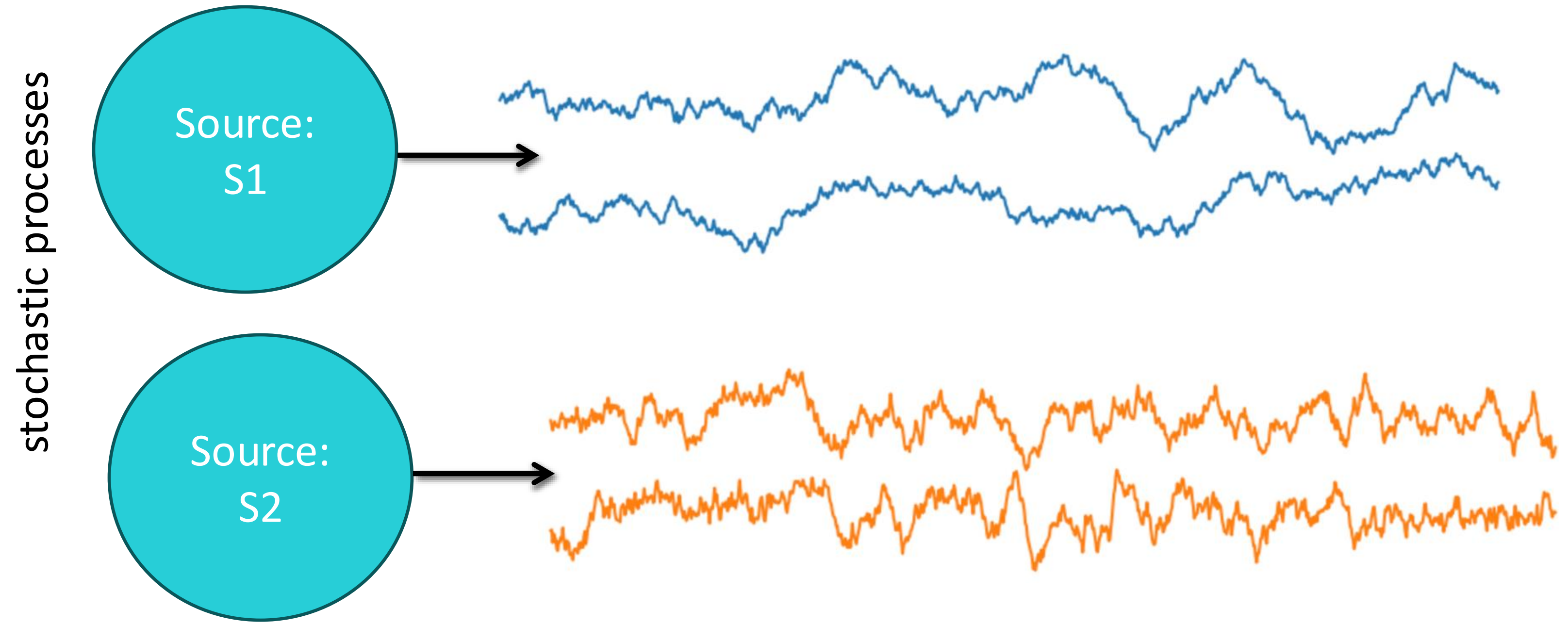
Time Series Analysis: Source

Popul → Sample

$x(t)$ $t \in \mathbb{Z}$

- Time Series - temporal discreet data - time-dependent - ***signal x(n)***
- All time-series are considered to be generated by **a source**
- Objective of TSA is to analyse the source, from given a time-series

"***Time series is a realization of a stochastic process***"

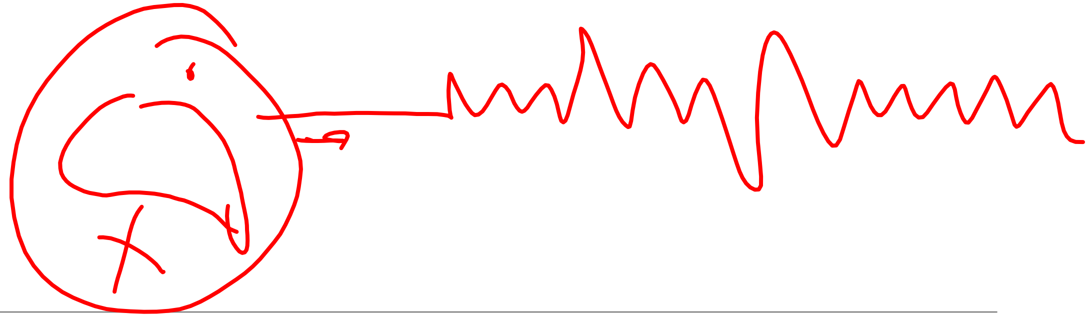


Time Series Analysis: Source

- We analyse time-series, in order to know **more about the source**

- We will model these sources using some mathematical models (stochastic processes)
  - AR (q), MA(p), ARMA, ARIMA

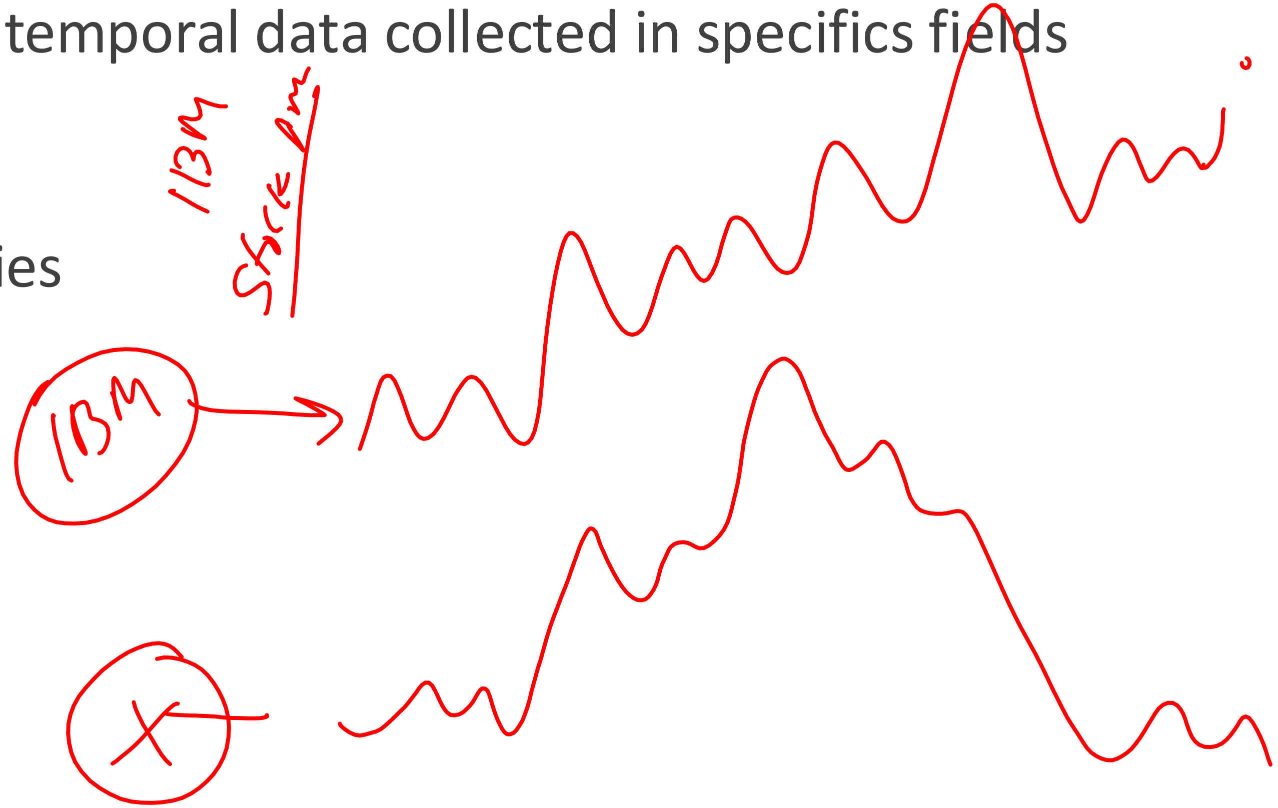
- We will analyse properties of time-series, that help to understand, which model they come from
  - ACF, AVF, PACF



Time Analysis

Time series are typically referred to temporal data collected in specifics fields such as:

- Marketing time series
- Economics and Financial time series
- Demographic time series
- Population time series
- Physical time series



Queen Mary School Hainan
Queen Mary University of London

QHP5701 Exploratory Data Analysis

Stochastic Process: Fundamentals

Nikesh Bajaj, PhD
Lecturer in Data Science,
Queen Mary University of London
nikesh.bajaj@qmul.ac.uk
https://nikeshbajaj.in

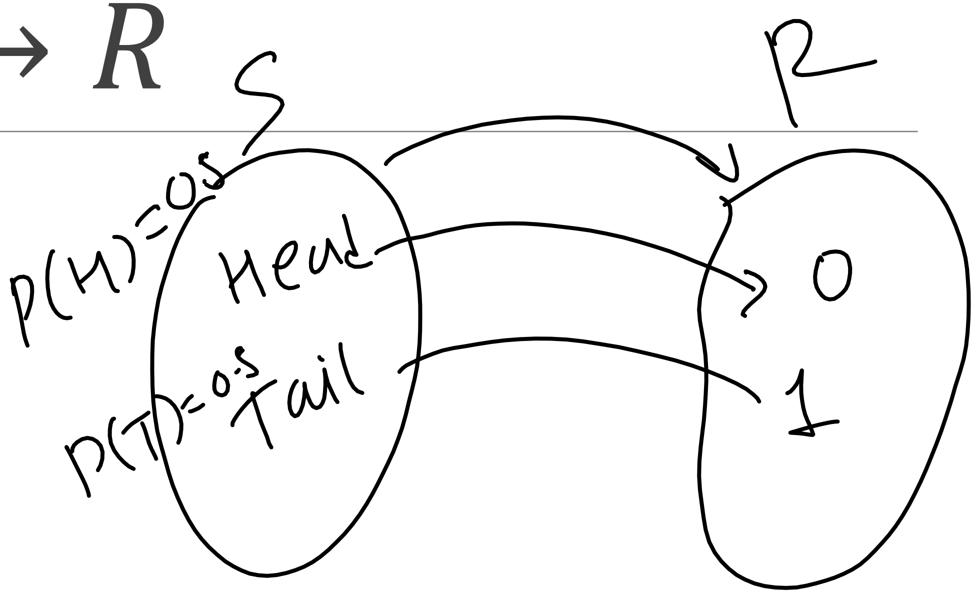
#Ref: Chapter 2, TSA, William WS Wei

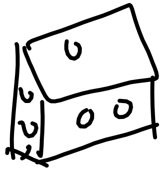
Random Variable $X: S \to R$

Random variable

- Not quite a variable x from algebra
- Ways to map random processes to numbers
- $X: S \to R$

- Continues/ Discreet random Variable





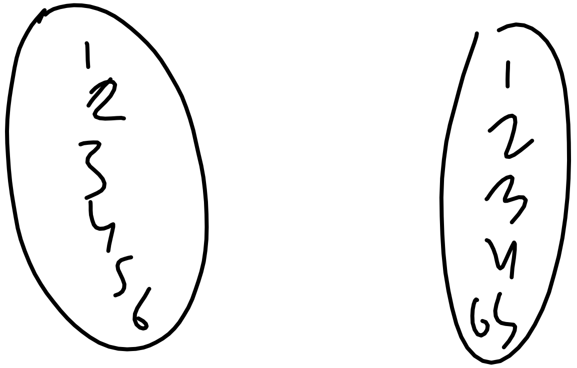
Examples

- $X = \begin{cases} 1, & for\ head \\ 0, & for\ tail \end{cases}$

$$\frac{P(X = 1)}{P(X=0)}$$

- $Y = sum\ of\ two\ dice$

$P(Y < 3)$



$P(X=4) = 1/6$

Discreet RV

$S : \{H, T\}$

$\{0, 1\}$

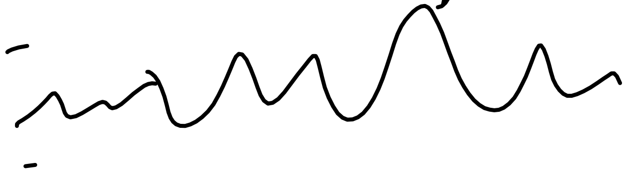
$P(H) = 0.5$

$P(T) = 0.5$

Cond RV.

$S : (-1, 1)$

$x(n) =$

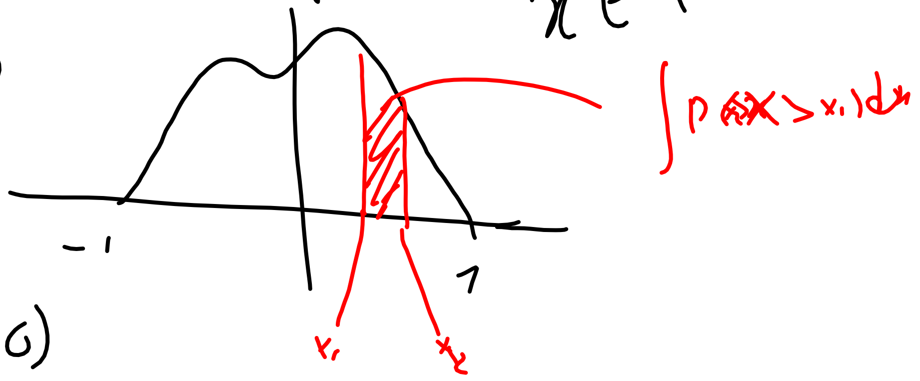


0 0.0001

0 12345

$x \in R$

$P(x)$



$P(x=0)$

$P(X \geq 0)$

Expected value

$E(X)$

$X \rightarrow 0\ 0\ 1\ 1\ 0\ 1\ 1$

Expected value $E[X]$ is a mean of random variable X

X

$$\mu_X = E[X] = \sum x p(X = x)$$

$$\int_{-\infty}^{\infty} x\, p(x)\, dx$$

$$\mu_X = E[X] \equiv \text{mean of source X}$$

$E(X)$

Expected value: Example

Head / Tail | Coin Toss

$X = \begin{cases} H = 0 \\ T = 1 \end{cases}$

$\cancel{P(H)}$ $p(X=H) = 0.5$

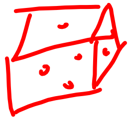
$p(X=T) = 0.5$

$X \rightarrow 0,1,1,0,1,0$ $E[X] = \sum x\, p(X=x) = \mu_X = 0.5$

$= 0 \times p(0) + 1\, p(1)$

$= 0 + 0.5 = 0.5$

Expected value: Example

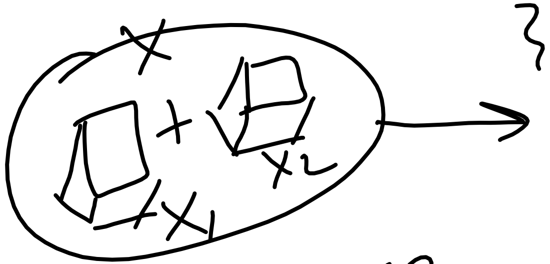
Dice  $X \rightarrow 1, 2, 1, 1, 6, 3, 4, 5,$

$$X = \begin{cases} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{cases} \qquad p(X=x) = 1/6$$

$$E(X) = \mu_X = \sum x\,p(x) = 1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6$$

$$= 21/6 = 7/2 = 3.5$$

$$\mu_X = 3.5$$



$\{2.$

123

Dice + Dice

$\{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$

$E(X) = \cancel{6}\ \cancel{5}\ 7 = \mu_X$

$p(x)$

$X = \begin{cases} 2 \text{ — } 1/36 \\ 3 \text{ — } 2+1,\ 1+2 \quad 2/36 \\ 4 \text{ — } 3/36 \text{ — } \{1+3,\ 3+1,\ 2+2,\ \cancel{2+2} \\ 5 \text{ — } 4/36 \text{ — } 1+4,\ 2+3,\ 3+2,\ 4+1 \\ 6 \text{ — } 5/36 \text{ — } 1+5,\ 2+4,\ 3+3,\ 4+2,\ 5+1 \\ 7 \rightarrow 6/36 \end{cases}$

$12 \rightarrow 1/36$

Estimation of Expected value: mean

Expected value $E[X]$ is a mean of random variable X

$$\mu_X = E[X] = \sum xp(X = x)$$

- Which is estimated from realization x, as sample mean,
  We don't know the TRUE probability distribution P(X)

$$\boxed{\hat{\mu}_X = \frac{1}{N}\sum x_i}$$

X

1
2
3
4
5
6
7

$\hat{\mu}_X$

Example

$X \longrightarrow 1, 2, 3, 4, 1, 1, 1, 3, 4, 5,$

iid

$E[X] = ?$

$\hat{\mu}_X = 25/10 = 2.5$

Example

$$X \sim N(0,1) \quad + \quad Y \sim N(0,1) \quad = \quad Z$$

$X \sim N(0,1)$

$E[X] = 0$

$Y \sim N(0,1)$

$E[Y] = 0$

$$E(Z) = E[Z] = E[X+Y]$$
$$= E[X] + E[Y]$$
$$= 0$$

$$E[(X-\mu_X)^2] = E(X^2) - E(\mu_X^2)$$

Variance & Standard Deviation

Variance: $V[X] = \underline{E[(X-\mu_X)^2]} = \sigma_X^2$ $= \cancel{\sum (x-\mu_X)^2 P(X=x)}$

Standard Deviation = σ_X $= \cancel{\sum (x^2+\mu_X^2-2\mu_X x)\, P(X=x)}$

$$E[(X-\mu_X)^2] = E[(X-\mu_X)(X-\mu_X)]$$

$$=$$

$$= \sum (x-\mu_X)^2 P(X=x)$$

Estimation of Variance & Standard Deviation

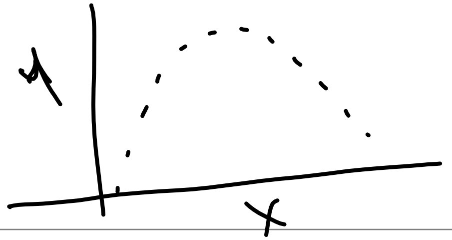
Variance: $V[X] = E[(X-\mu_X)^2] = \sigma_X^2$

$Standard\ Deviation = \sigma_X$

- This is estimated from realization x,

$$\boxed{\hat{\sigma}_X^2 = \frac{1}{N-1}\sum(x_i - \hat{\mu}_X)^2}$$

Covariance



Variance of x

$$\sigma_X^2 = E[(X-\mu_X)^2] = E[(X-\mu_X)(X-\mu_X)] = \sum (x-\mu_X)^2 P(X=x) = V(X)$$

Covariance of X and Y

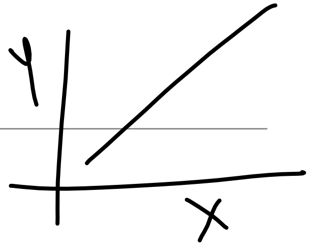
$$\sigma_{XY}^2 = E[(X-\mu_X)(Y-\mu_Y)] = \sigma_{YX}^2$$

$$E[(X-\mu_X)(Y-\mu_Y)] = \sum (x-\mu_X)(y-\mu_Y)P(X=x, Y=y)$$

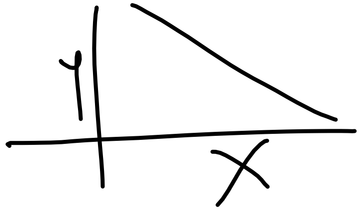
linear relationship betw two variable

Example

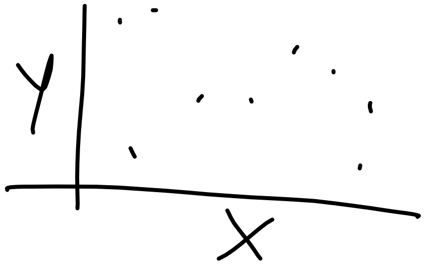
$\sigma_{xy}^2 \rightarrow +ve$ $\quad X\uparrow \; Y\uparrow \quad X\downarrow \; Y\downarrow$



$\sigma_{xy}^2 \longrightarrow -ve$ $\quad X\uparrow \; Y\downarrow \quad X\downarrow, \; Y\uparrow$



$\sigma_{xy}^2 = 0$ $\quad X\uparrow\downarrow \perp Y\downarrow\uparrow$



$X \rightarrow$

$Y \rightarrow$

1, ..., 6

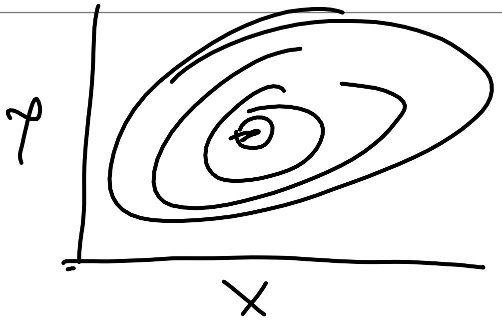
Dice

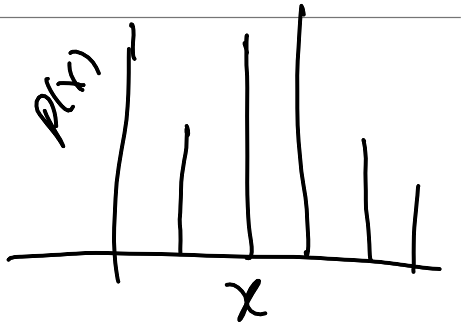
$X \rightarrow$ 1 4 ,

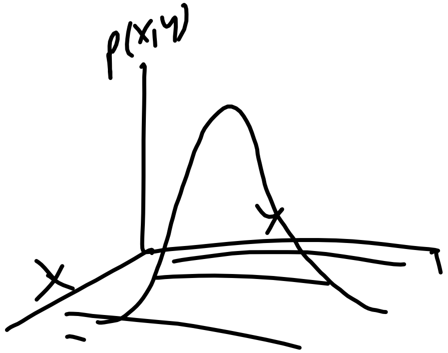
$Y \rightarrow$ 3 1

(1 , 6)

$\sigma^2_{XY} \approx 0$







$X \to 1, 2, 3, 4, 5, 6$

$Y \to 0, 0, 1, 0, 3, 4, 6$

Estimation of Covariance

Variance of x

$$\sigma_X^2 = E[(X-\mu_X)^2] = E[(X-\mu_X)(X-\mu_X)] = \sum (x-\mu_X)^2 P(X=x)$$

Covariance of x and y (cross-covariance)

$$\sigma_{XY}^2 = E[(X-\mu_X)(Y-\mu_Y)] = \sigma_{YX}^2$$

$$E[(X-\mu_X)(Y-\mu_Y)] = \sum (x-\mu_X)(y-\mu_Y)\boxed{P(X=x, Y=y)}$$

Covariance of x and y can be estimated as

$$\hat{\sigma}_{XY}^2 = Cov(x,y) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\hat{\mu}_X)(y_i-\hat{\mu}_Y)$$

Example

Correlation

Correlation is closely related to covariance, so it can be computed as:

ρ_{XY}

$X \rightarrow$

$Y \rightarrow$

$$\boxed{\rho_{XY} = \frac{\sigma^2_{XY}}{\sigma_X \sigma_Y}}$$

$$-1 \leq \rho_{XY} \leq 1$$

$$\boxed{Corr(x,y) = \frac{Cov(x,y)}{\hat{\sigma}_x \hat{\sigma}_y}} = \frac{\hat{\sigma}^2_{xy}}{\hat{\sigma}_x \hat{\sigma}_y}$$

Cor

$r_{xy} =$

$$= \frac{\sum(x-\bar{x})(y-\bar{y})}{\left(\sum(x-\bar{x})^2 \sum(y-\bar{y})^2\right)^{1/2}}$$

Example

Independently and Identically Distributed: IID

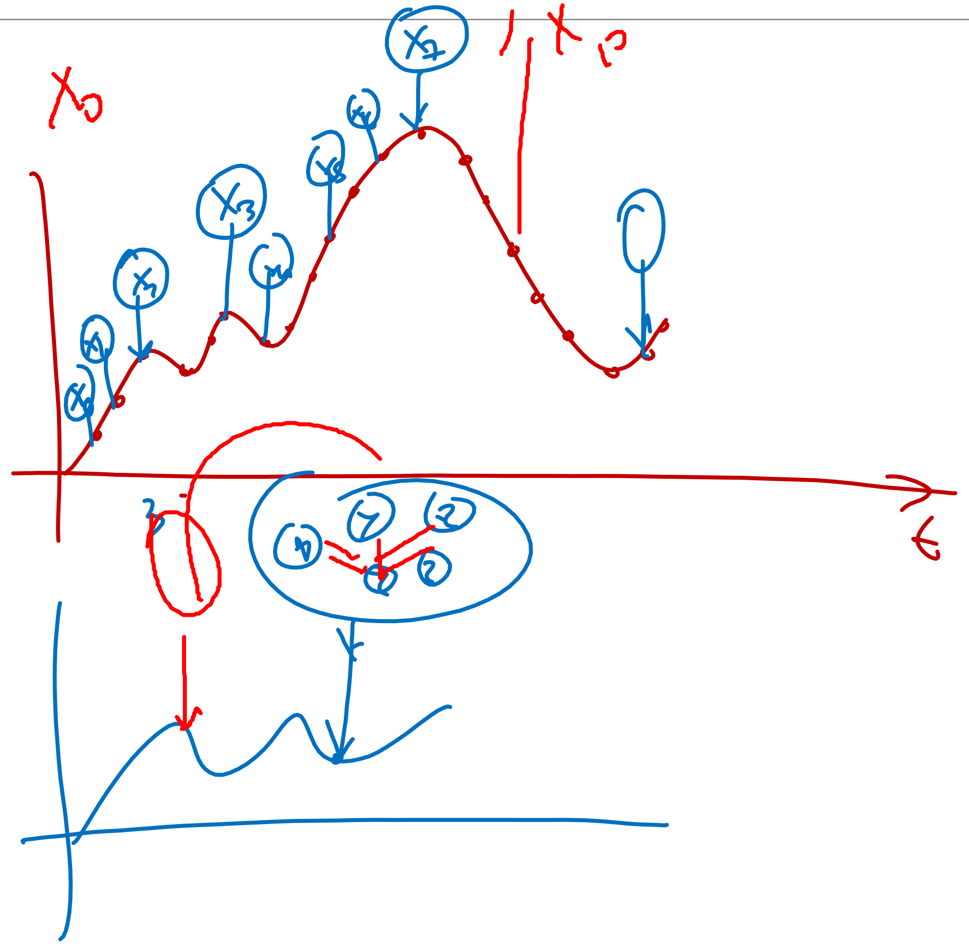
ii iid

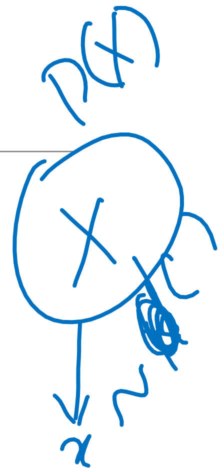
$$z_i \sim N(0, 1) \quad \text{iid}$$

Stochastic Processes: Time series

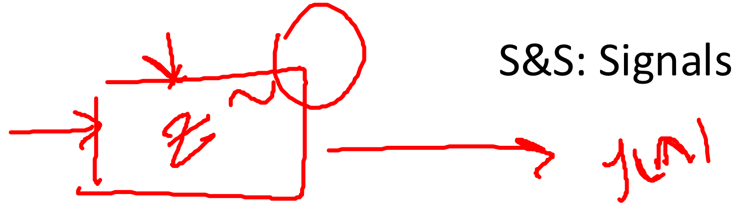
Collection of Random variable

- $X_1 \quad X_2 \quad X_3 \quad \dots$ $\qquad X_n \to x_n$
- 20 32 22
- $X_n \sim distribution(\mu_n, \sigma_n^2, \ \dots),$
- Time series $\qquad X_t \to x_t$

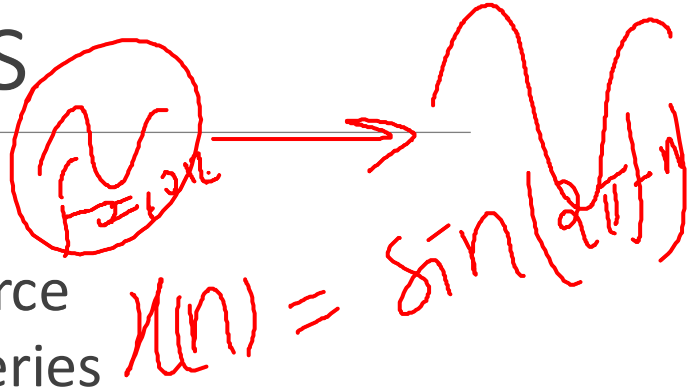




S&S: Signals & Systems

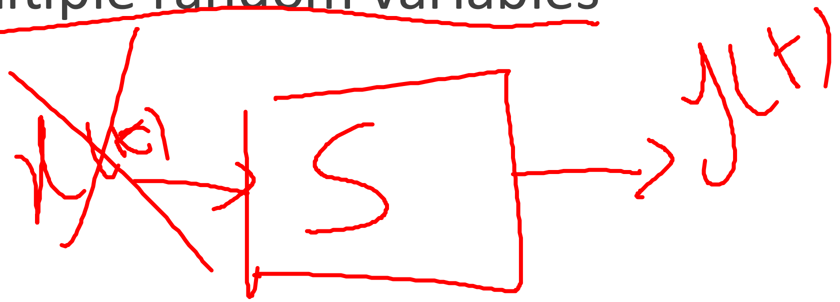


Stochastic Processes: Time series



Stochastic process is a collection of Random variable

- Opposite of deterministic process, deterministic system/source
- We are more interested to learn about the **source** of time-series
- S&S: A source can be seen as a system, with no or a little input
- S&S: A time-series is a signal, output of the system
- Every sample produced at time n, can be seen as a random variable.
- All the samples of a time-series generated from multiple random variables with each having different distribution



- $X_n \sim distribution(\mu_n, \sigma_n^2, \ \dots),$

Example

Types of characteristics

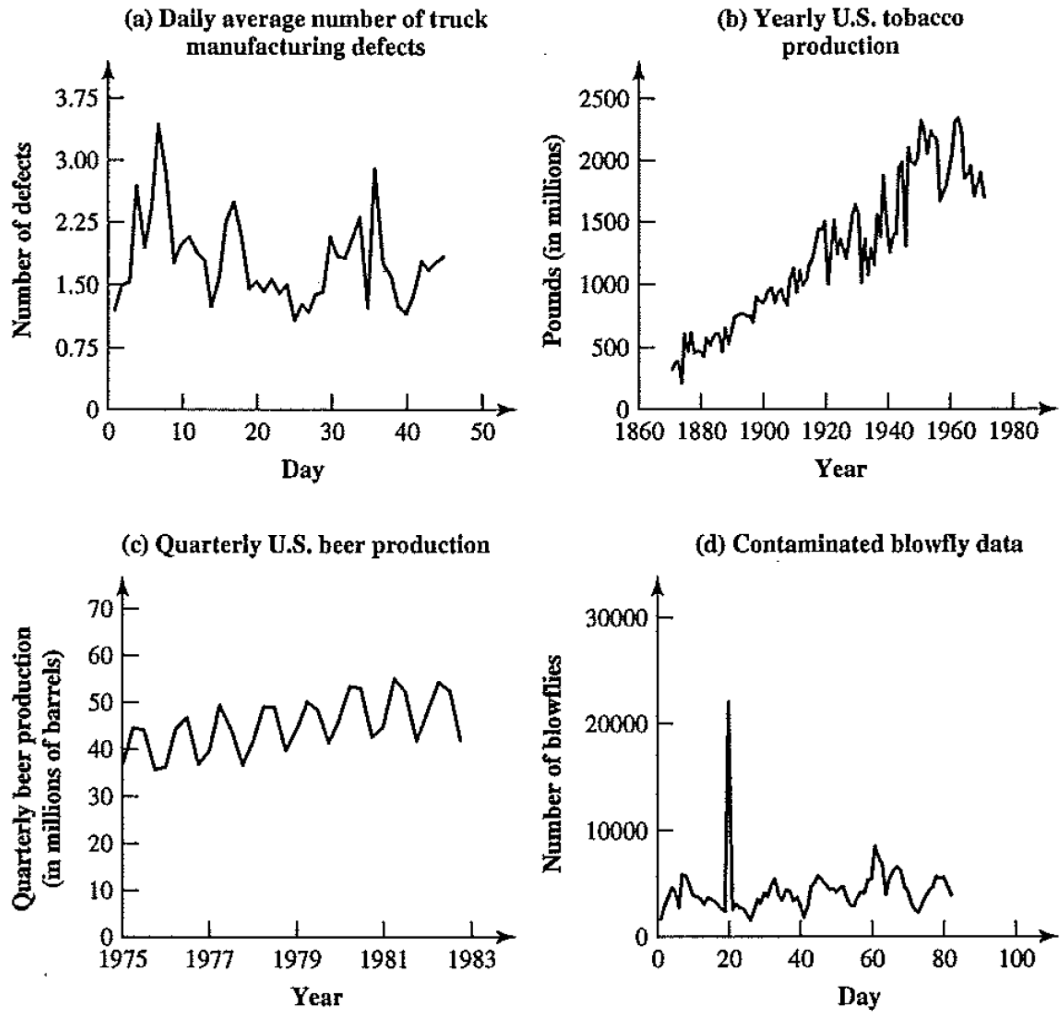
Time-series ≡ signal

(a) (weak) Stationary

(b) Non-stationary

(c) Season variation: seasonality

(d) Non-stationary due to external disturbance



Strict Stationary

detrend

trend

Stationarity

Stationary: plot (a):

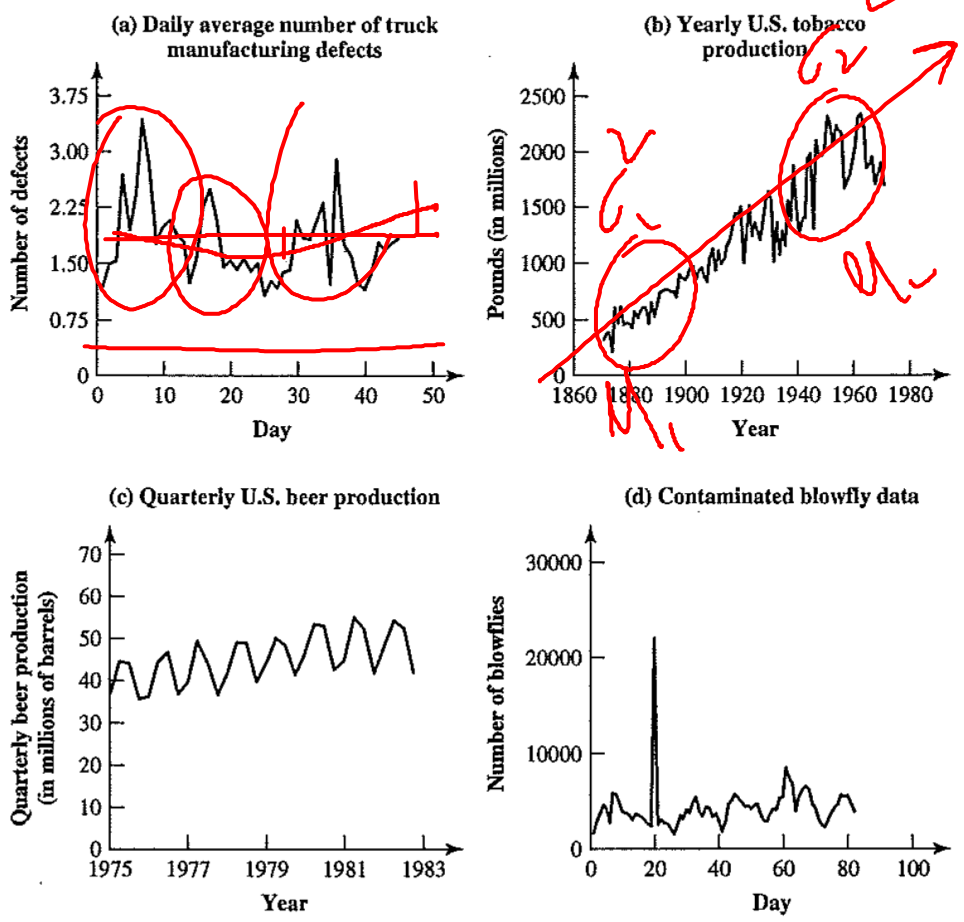
- No systematic change in time-series,
- No trends,
- No systematic change in mean
- No systematic change in variation
- No period fluctuation

μ

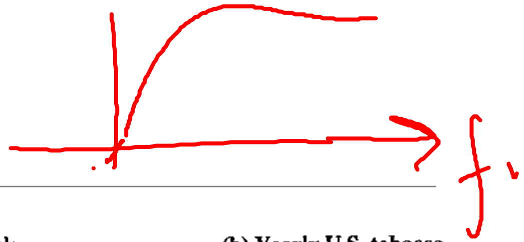
Properties of one part (section) of signal, is same as another part of signal

It is property of the stochastic process of the model, not the property of time-series, but often we say - a (weak) stationary time series

Why weak?



Direct current
DC filter
of freq component
filter



Non-Stationarity

Non-Stationary: plot (b), (d):

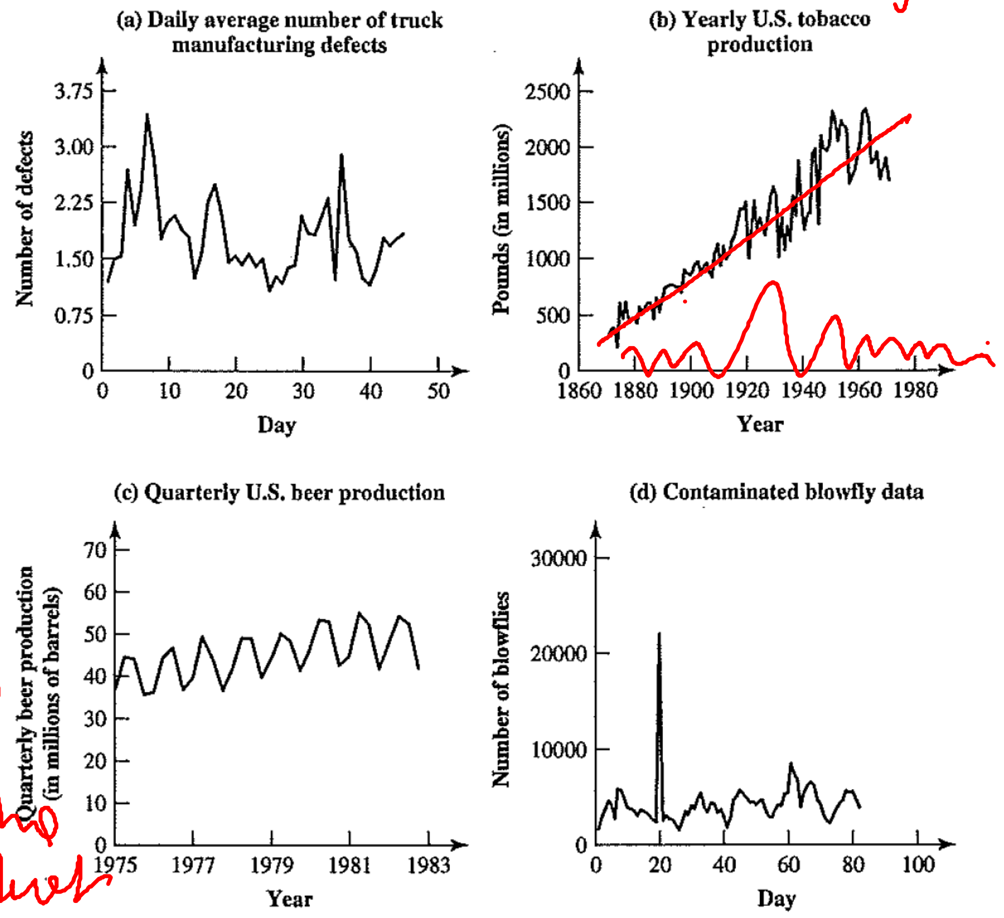
- trends,

*Properties of one part (section) of signal, is **NOT** same as another part of signal*

Transform into Stationary time series

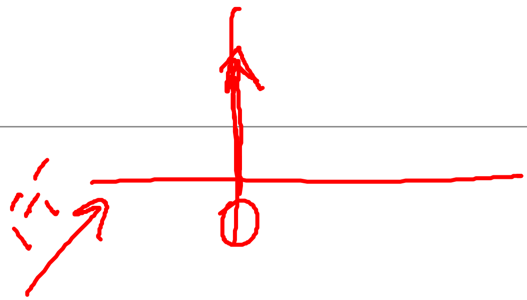
detrend

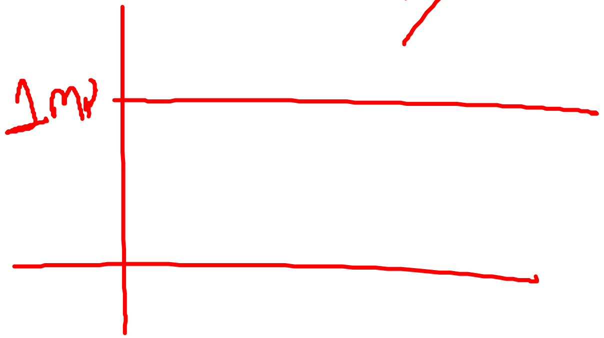
Removing trend
Removing Butterworth
DC Removal



Direct current

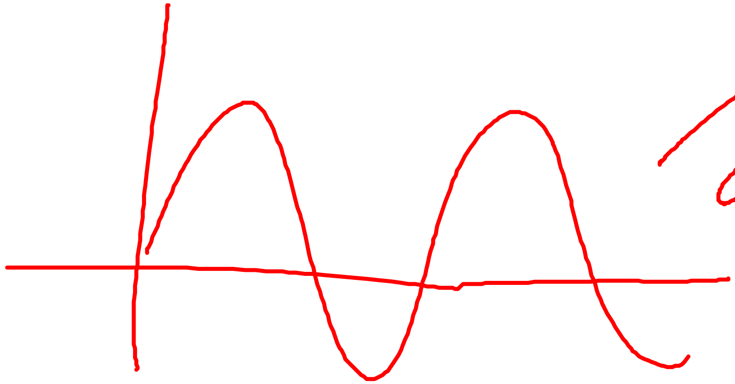
DC





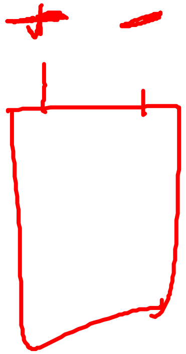
No

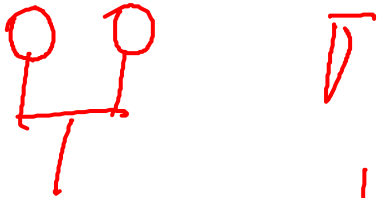
AC Alternating current

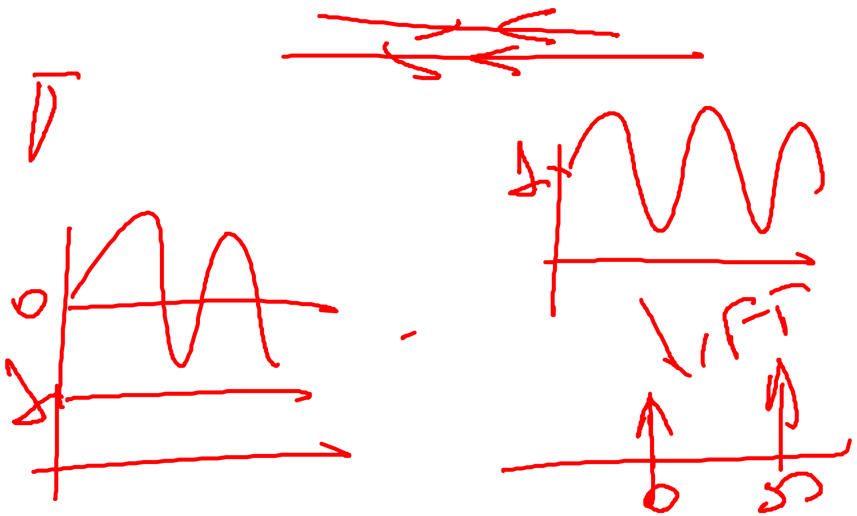


50Hz

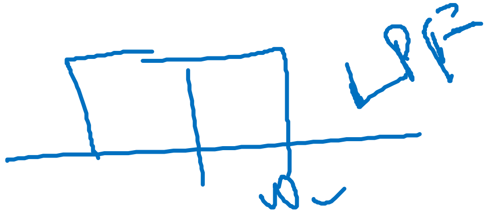
No =0





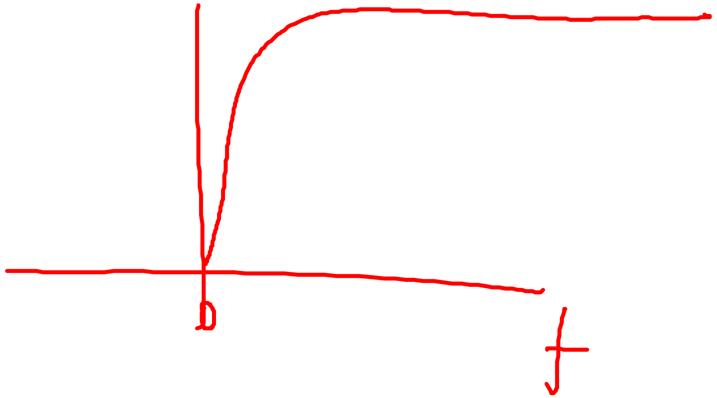


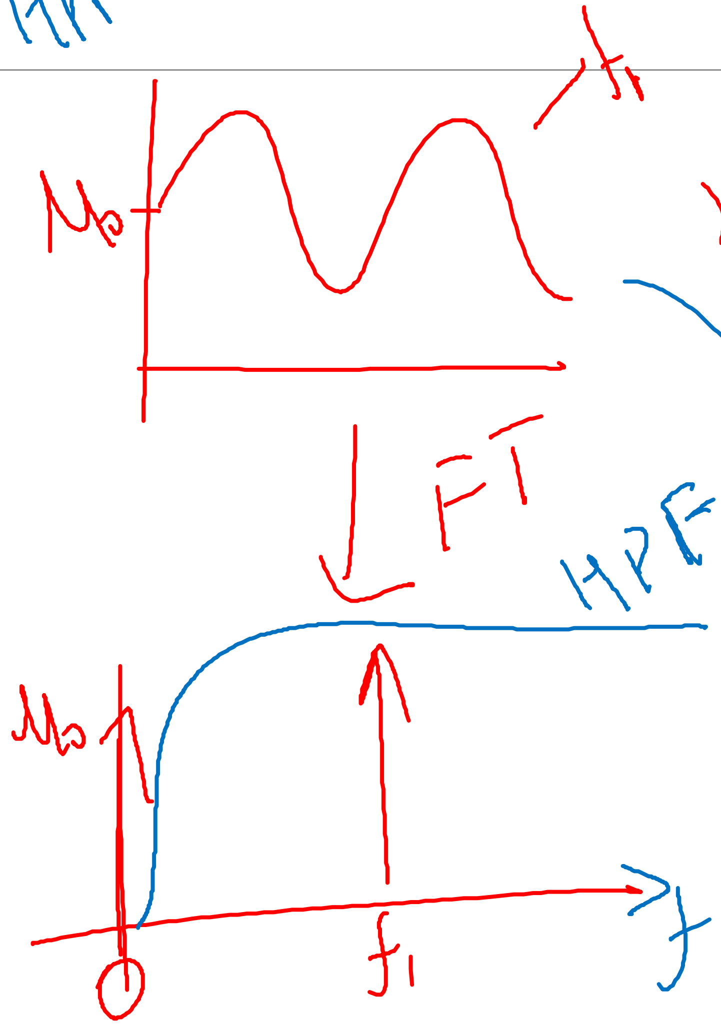
DC Removal

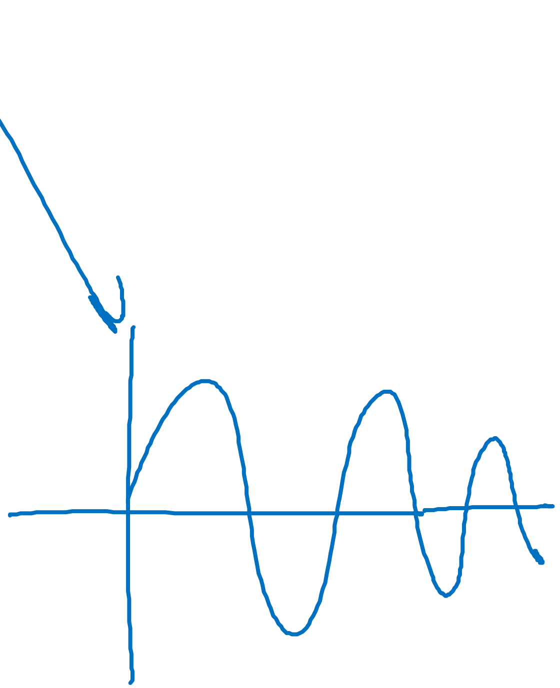


DC Filtering

$x(t) = M_0 + \sin(f_1 t)$







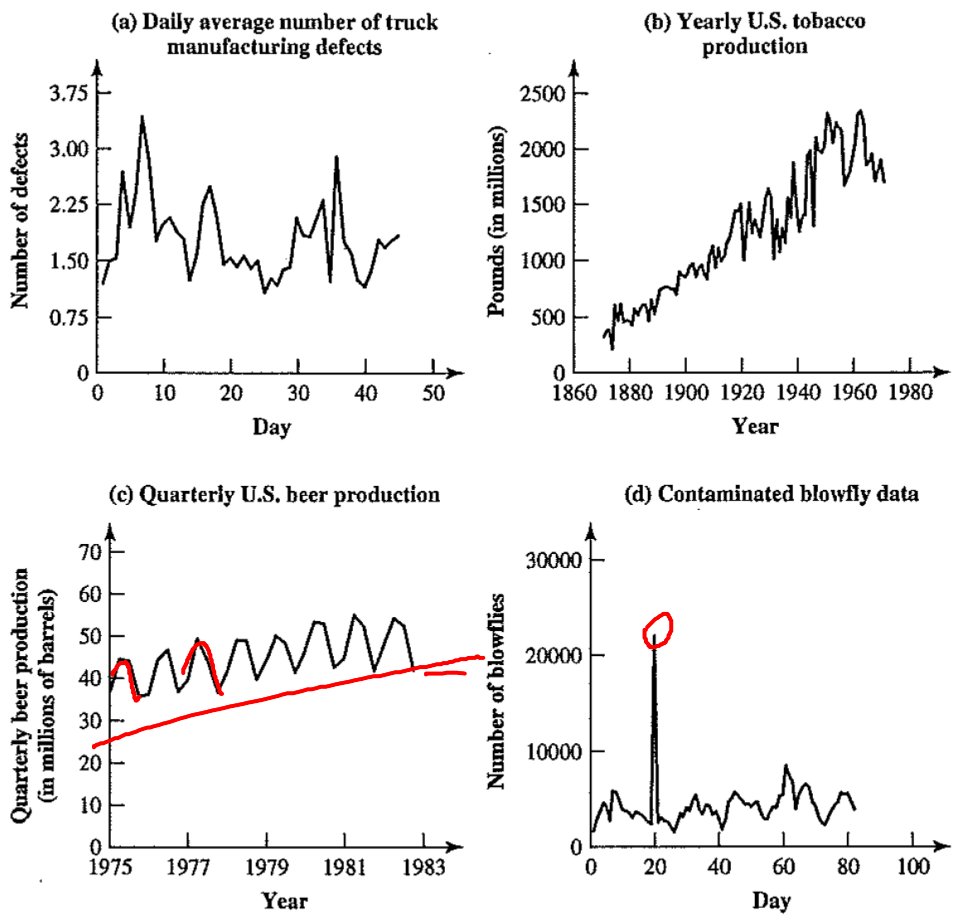
~*Periodicity*

Seasonality: Season variation:

Seasonality: plot (c):

- Periodic fluctuations,

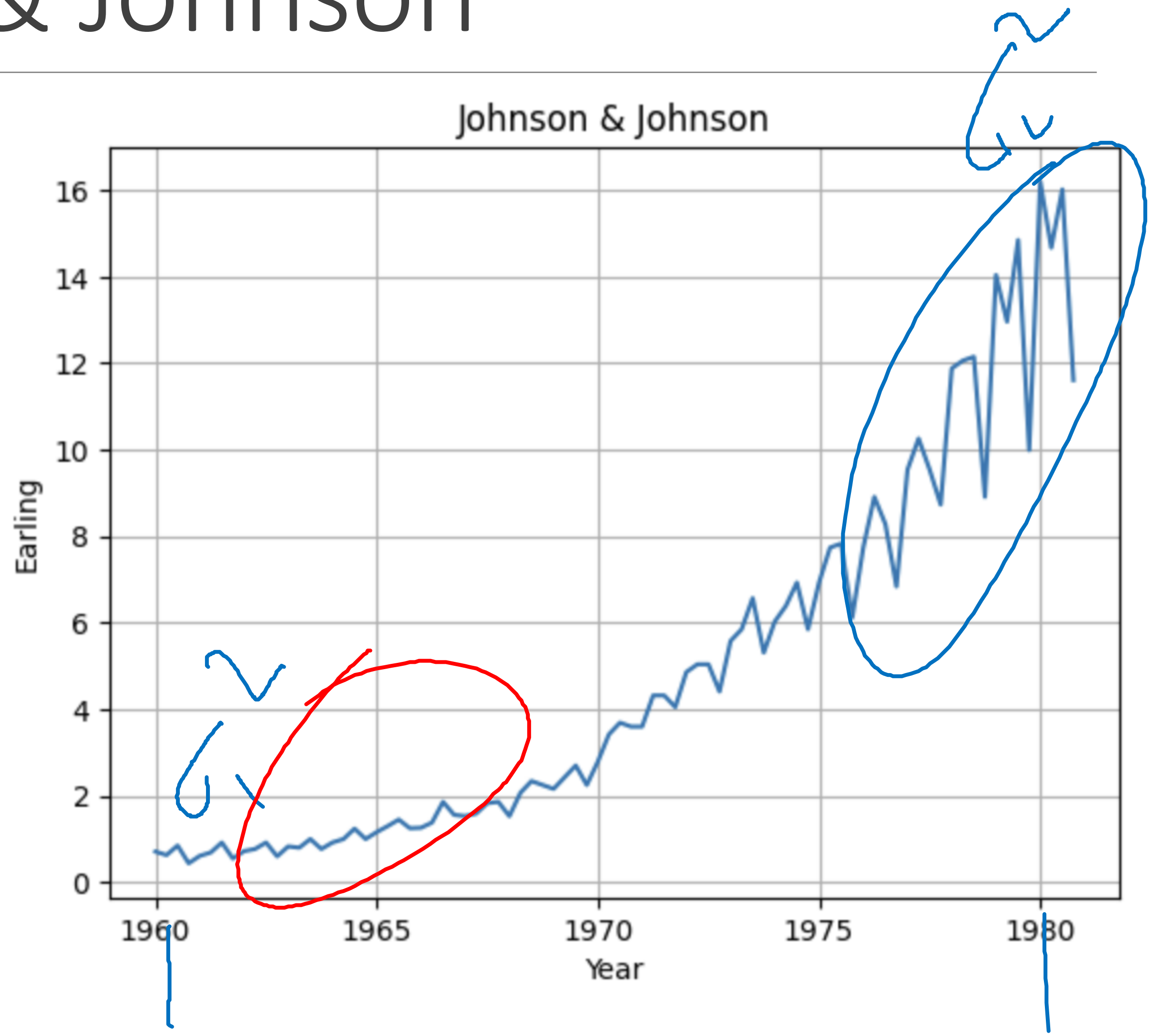
Time-series have periodicity, something that repeats like seasons



Example: Johnson & Johnson

Johnson & Johnson, quarterly Earning per share

- Collected from 21 years
- see anything meaningful?



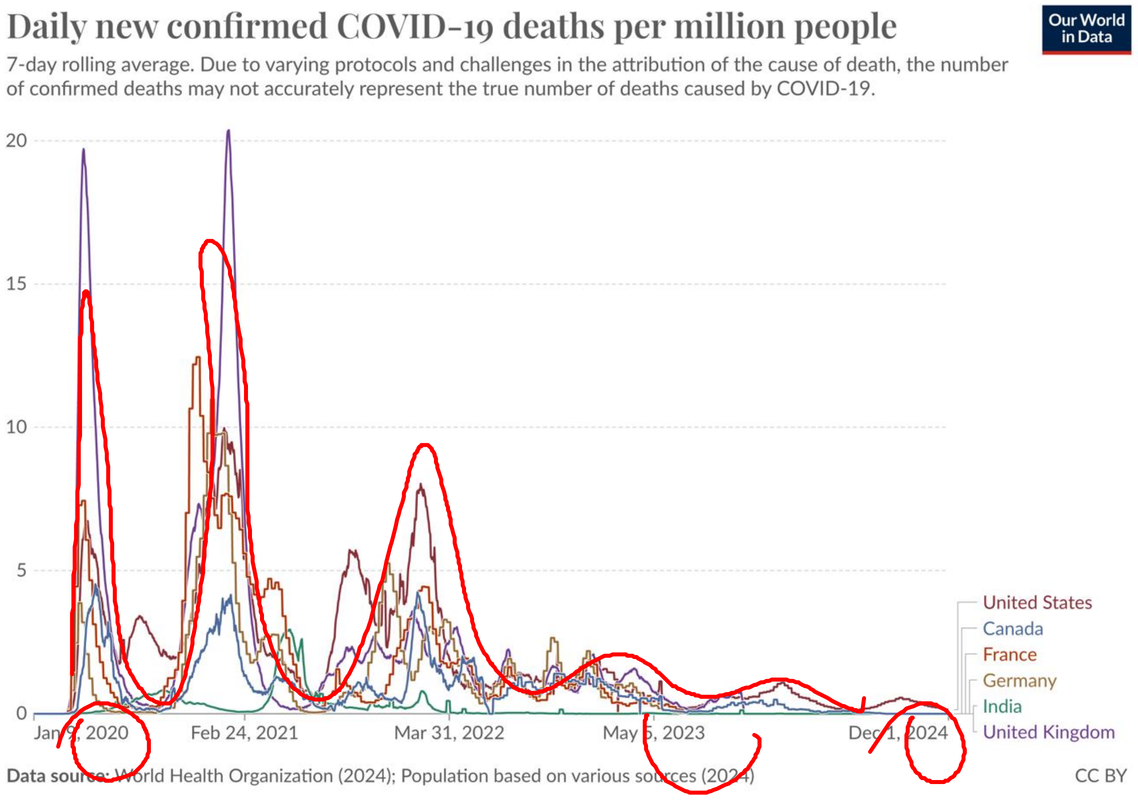
Example: Covid-19 Deaths

Covid-19 deaths per millions

- Check online

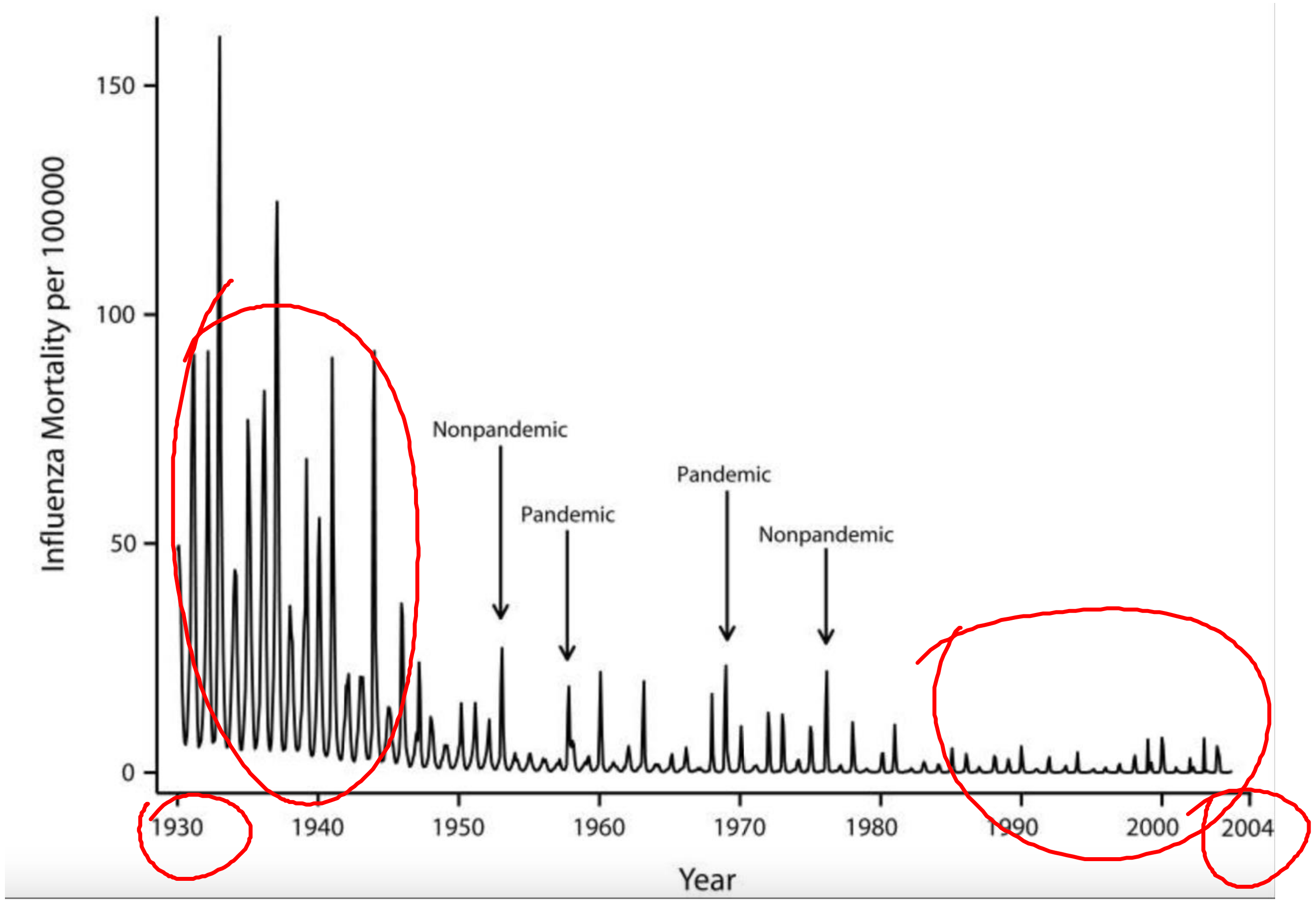
- see anything meaningful?

[Source: https://ourworldindata.org/covid-deaths](https://ourworldindata.org/covid-deaths)



Example: Influenza

- see anything meaningful?



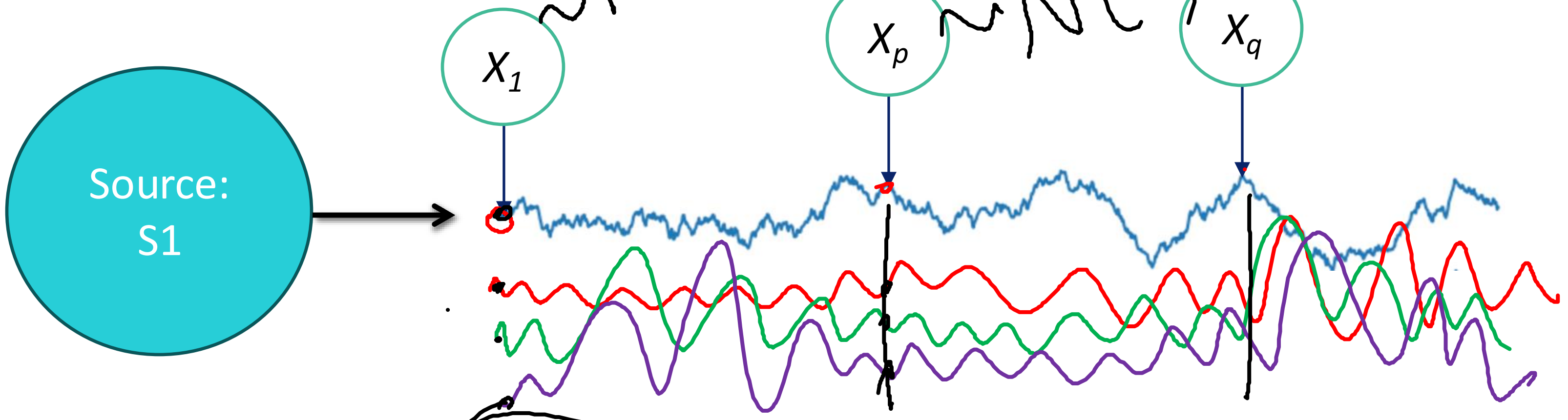
Source:
https://pmc.ncbi.nlm.nih.gov/articles/PMC2374803/

Autocovariance Function: AVF

Recall Covariance of two random variables X and Y: $\sigma^2_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = \sigma^2_{YX}$

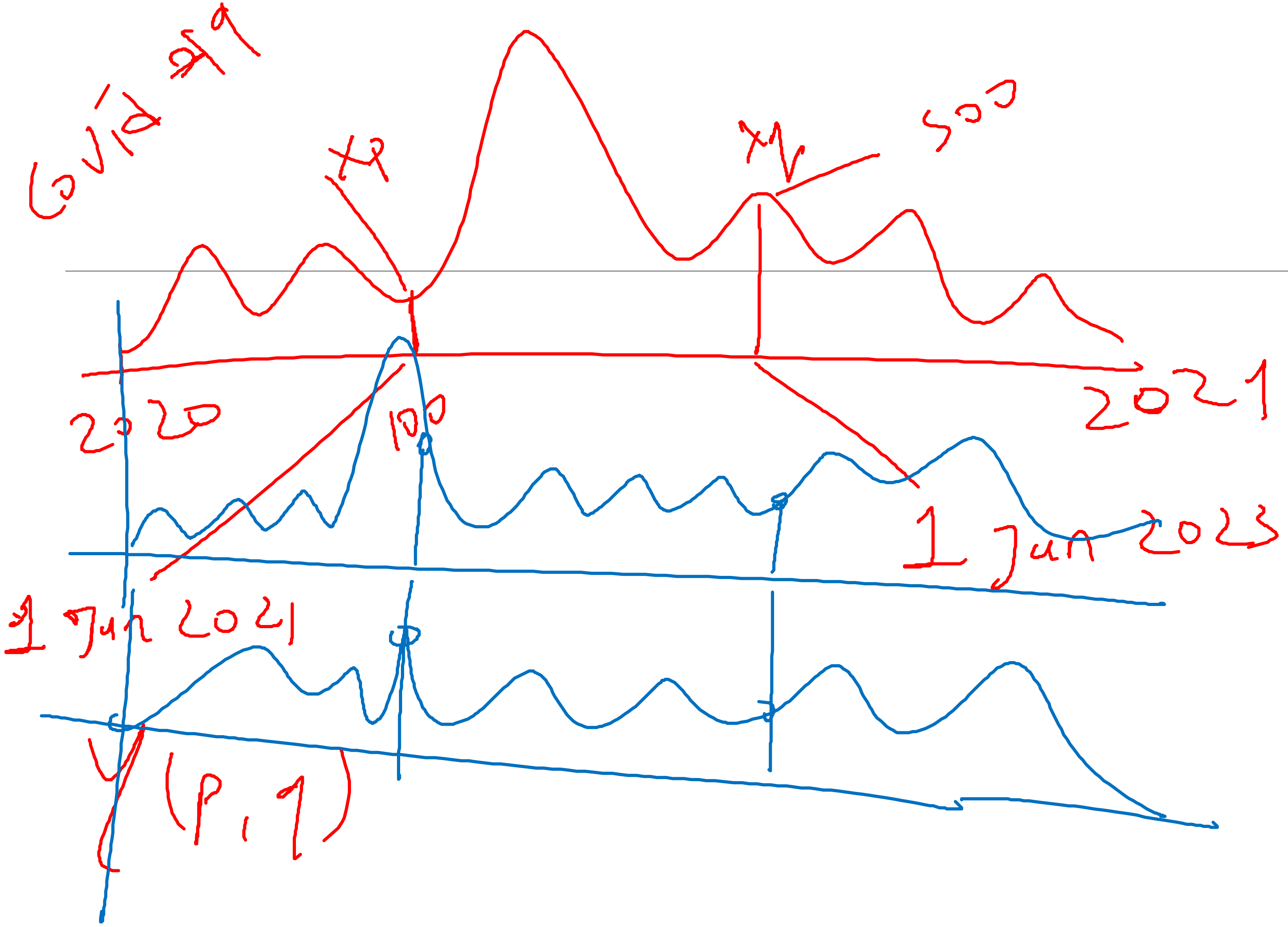
$\sim N(\ )$ $\sim N(\ )$

Time-series



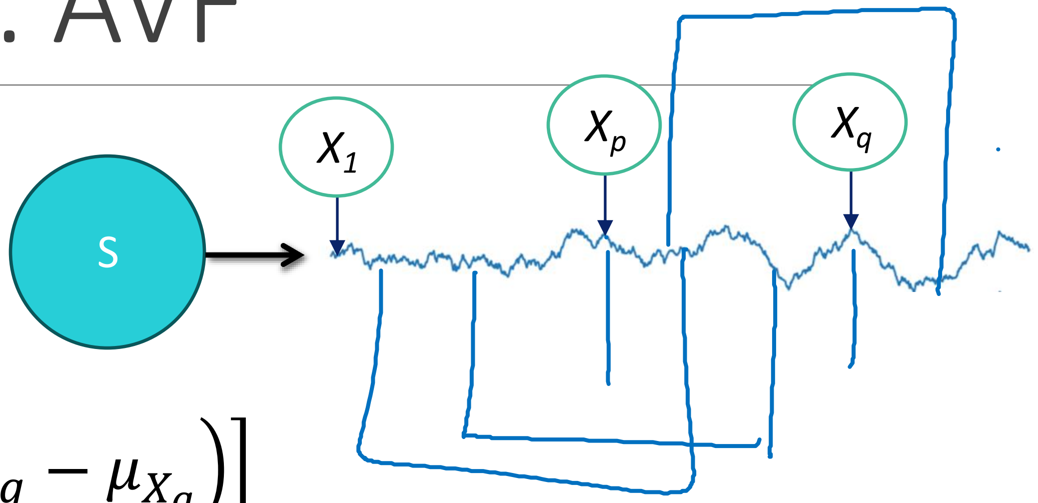
- Autocovariance Function: $\gamma(p,q)$

$$\gamma(p,q) = \sigma^2_{X_pX_q} = E\left[\left(X_p - \mu_{X_p}\right)\left(X_q - \mu_{X_q}\right)\right]$$


Covid 19
2020
2021
100
500
1 Jun 2023
1 Jun 2021
(P, 1)

Autocovariance Function: AVF

- Autocovariance Function: $\gamma(p, q)$



$$\gamma(p, q) = \sigma^2_{X_p X_q} = E\left[\left(X_p - \mu_{X_p}\right)\left(X_q - \mu_{X_q}\right)\right]$$

$$\gamma(p, p) = \sigma^2_{X_p} = E\left[\left(X_p - \mu_{X_p}\right)^2\right]$$

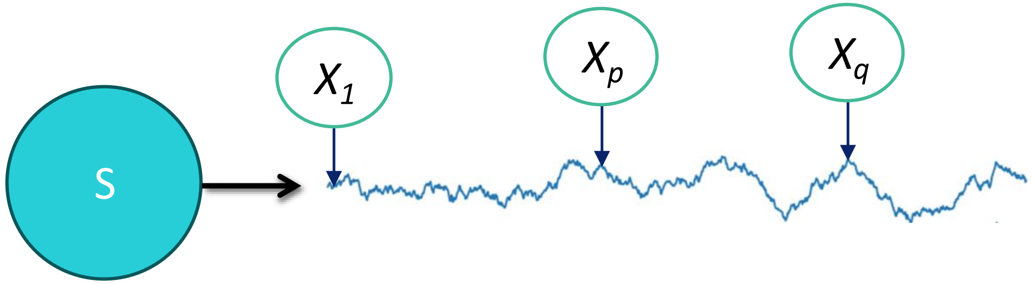
- Assuming the stationary time series: *k - time difference*

$$\boxed{\gamma_k = \gamma(p, p+k) = \sigma^2_{X_p X_{p+k}}}$$

Autocovariance Function: AVF

AVF

$$\boxed{\gamma_k = \gamma(p, p+k) = \sigma^2_{X_p X_{p+k}}}$$



γ_2

X_0

X

$P: 1, 2, 3, 4, 5, 1, 1, 1, 7, 8, 2$

X_2

γ_3 X_0

X_3

$$\hat{\sigma}^2_{XY} = Cov(x, y) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{\mu}_X)(y_i - \hat{\mu}_Y)$$

$$\hat{\sigma}^2_{XY} = Cov(x,y) = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \hat{\mu}_X)(y_i - \hat{\mu}_Y)$$

Autocovariance Coefficients: γ_k

Autocovariance function for a stationary time-series

$$\gamma_k = \gamma(p, p+k) = \sigma^2_{X_p X_{p+k}} = E\left[\left(X_p - \mu_{X_p}\right)\left(X_{p+k} - \mu_{X_{p+k}}\right)\right]$$

Autocovariance Coefficients $\hat{\gamma}_k$ is estimation of γ_k

For given a realization x (time-series) of process

$$\boxed{\hat{\gamma}_k = \frac{1}{N-k}\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}$$

$$\boxed{\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i}$$

Autocovariance Coefficients: c_k

$$\hat{\gamma}_k = \frac{1}{N} \sum_{i=1}^{N-k} (x_i - \bar{x})(x_{i+k} - \bar{x})$$

bias

bias

$$\hat{\hat{\gamma}}_k = \frac{1}{N - k} \sum_{i=1}^{N-k} (x_i - \bar{x})(x_{i+k} - \bar{x}) \quad = c_k$$

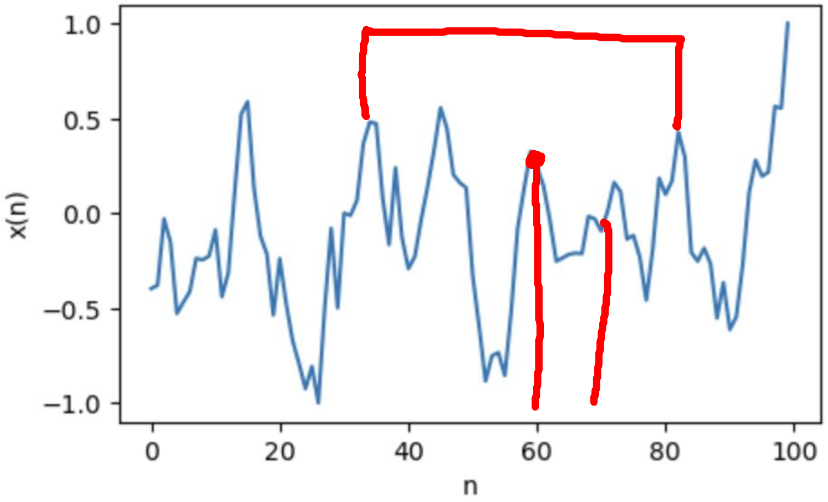
In general, $\hat{\gamma}_k$ has larger bias than $\hat{\hat{\gamma}}_k$

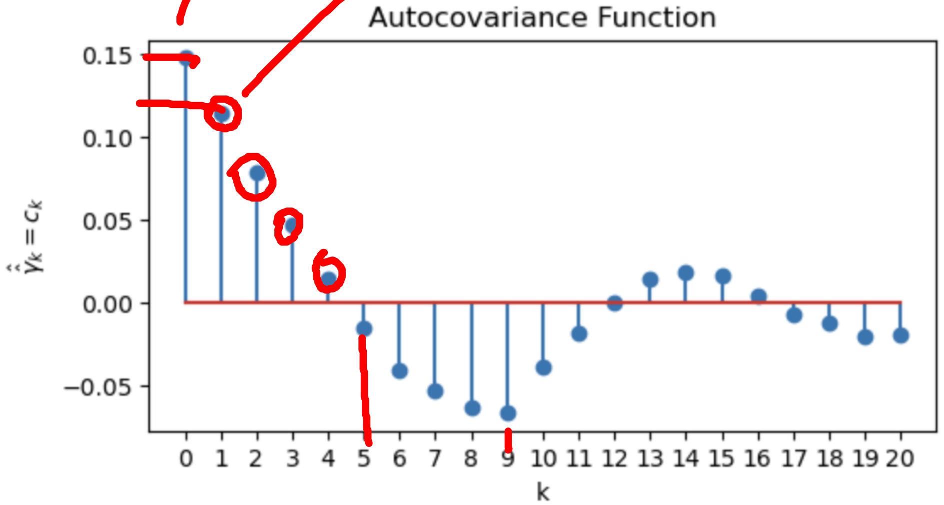
{proof: Book: 2.5.2}

Example: AVF

y_0 x_0

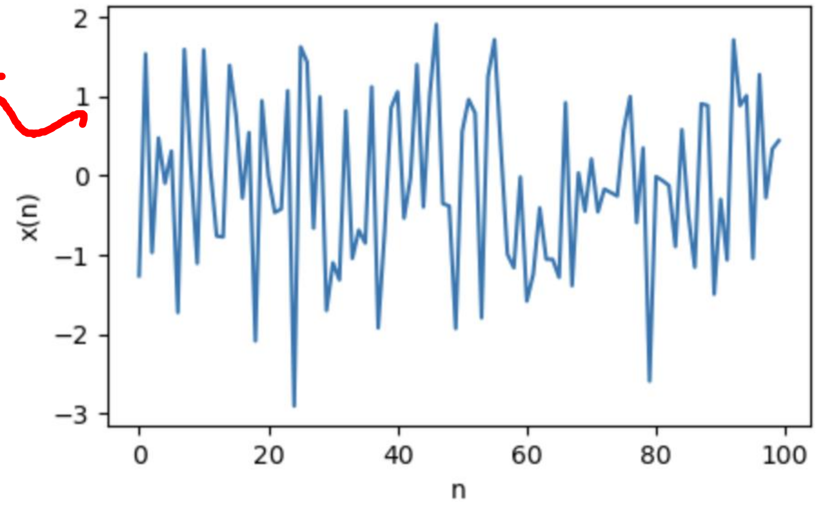
$y_1 \sim x_{0+1}$

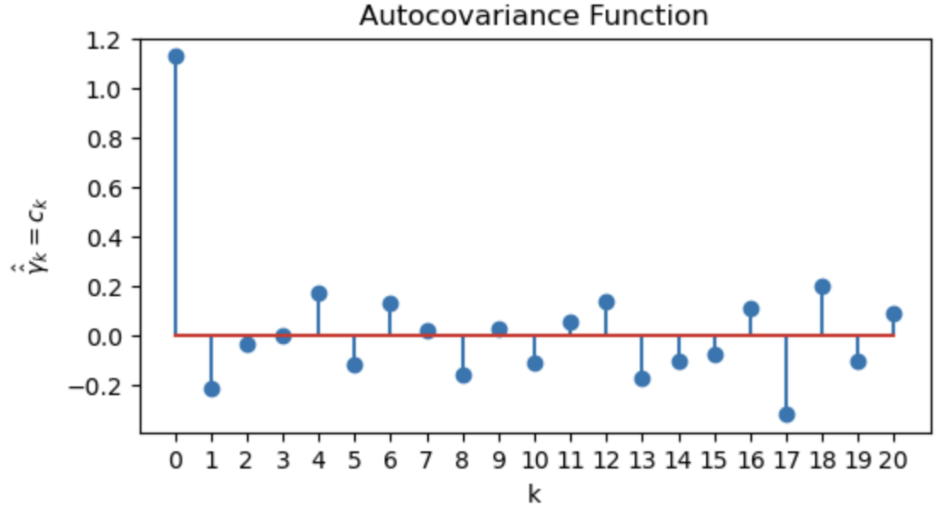




iid

$N(0,1)$





Example

Example

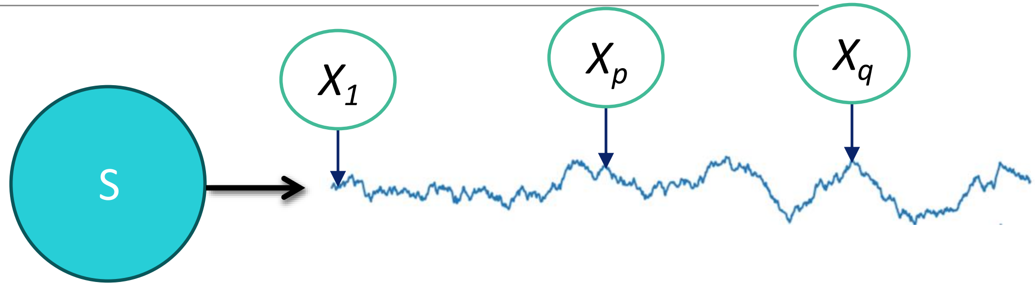
$$\rho_{XY} = \frac{\sigma_{XY}^2}{\sigma_X \sigma_Y}$$

Autocorrelation Function: ACF

Autocorrelation Function: ρ_k

- Assume weak stationarity
- Computed for X_p and X_{p+k}
-

$$-1 \le \rho_k = \frac{\gamma_k}{\gamma_0} \le 1$$



- Estimation: *lag k*

$$\boxed{\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = r_k}$$

$$\hat{\rho}_k = \frac{\hat{\hat{\gamma}}_k}{\hat{\hat{\gamma}}_0} = r_k$$

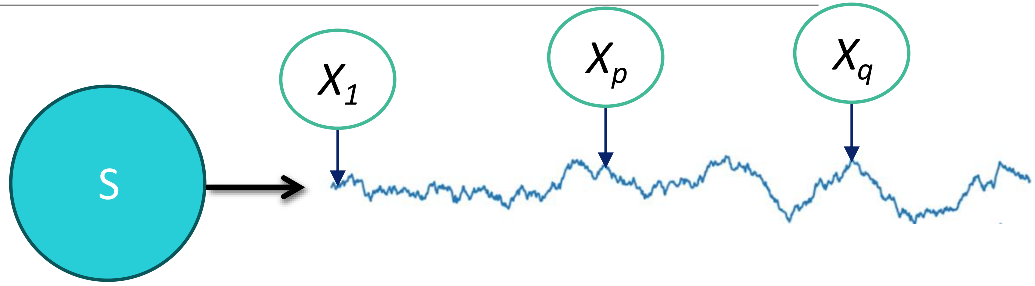
$$\hat{\rho}_0 = r_0 = 1$$

$$\rho_{XY} = \frac{\sigma_{XY}^2}{\sigma_X \sigma_Y}$$

Autocorrelation Function: ACF

Autocorrelation Function: ρ_k

- Estimation: *lag k*



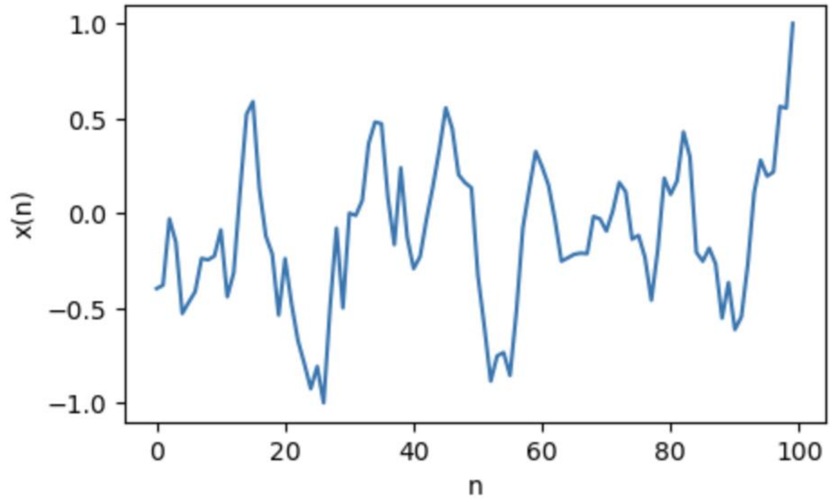
$$\hat{\rho}_k = \frac{\sum_{i=1}^{N-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{N}(x_i - \bar{x})^2} = r_k$$

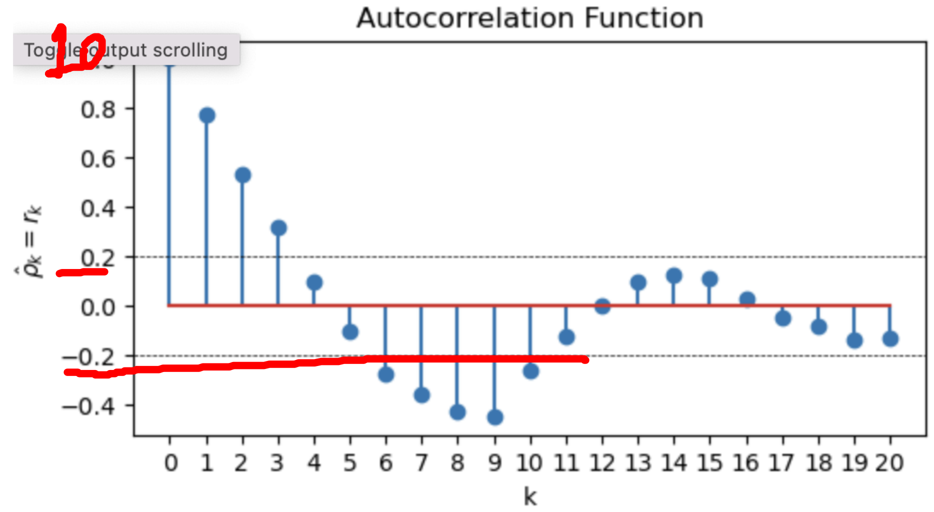
ρ rho

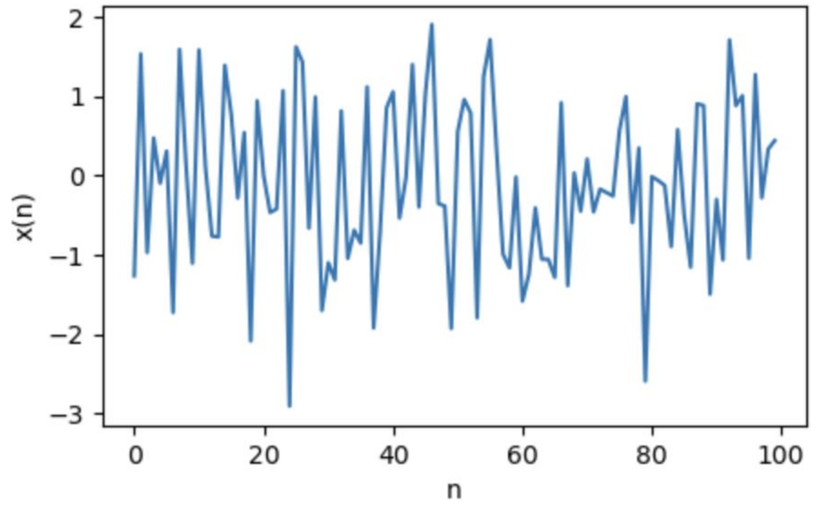
$$\hat{\rho}_0 = r_0 = 1$$

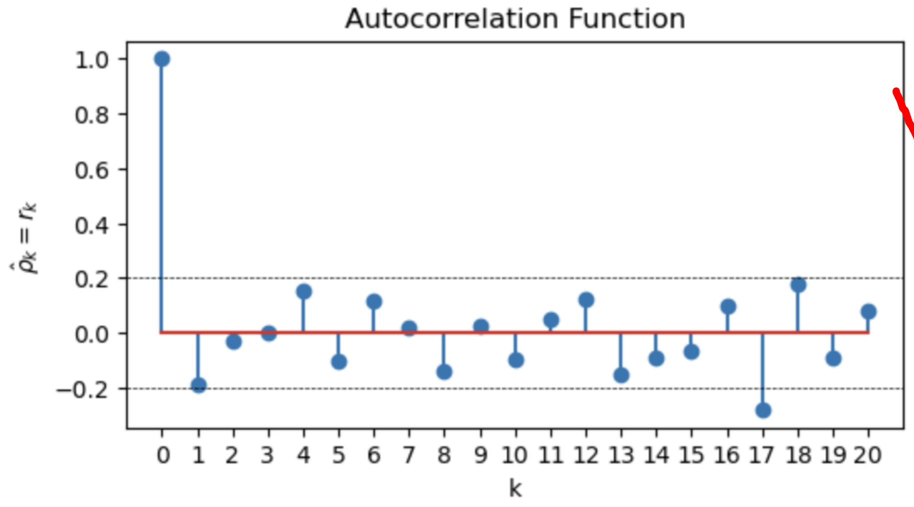
Example: ACF

<u>Correlogram</u>









iid

Example

Summary so far

$x(n) = [1\ 0\ 2\ 1\ 4\ 1\ 3\ 4\ 5\ 1\ 7\ 8]$

$\hat{\rho}_k = ?$

$\hat{\gamma}_k = ?$

So far, we have learned:

- Basic concepts of Stochastic Processes, Random Variables
- Time series as Stochastic Process
- Characteristics of time series: Stationarity, Seasonality, Non-stationarity

- Autocovariance Function AVF: γ_k and estimations
- Autocorrelation Function ACF: ρ_k and estimations

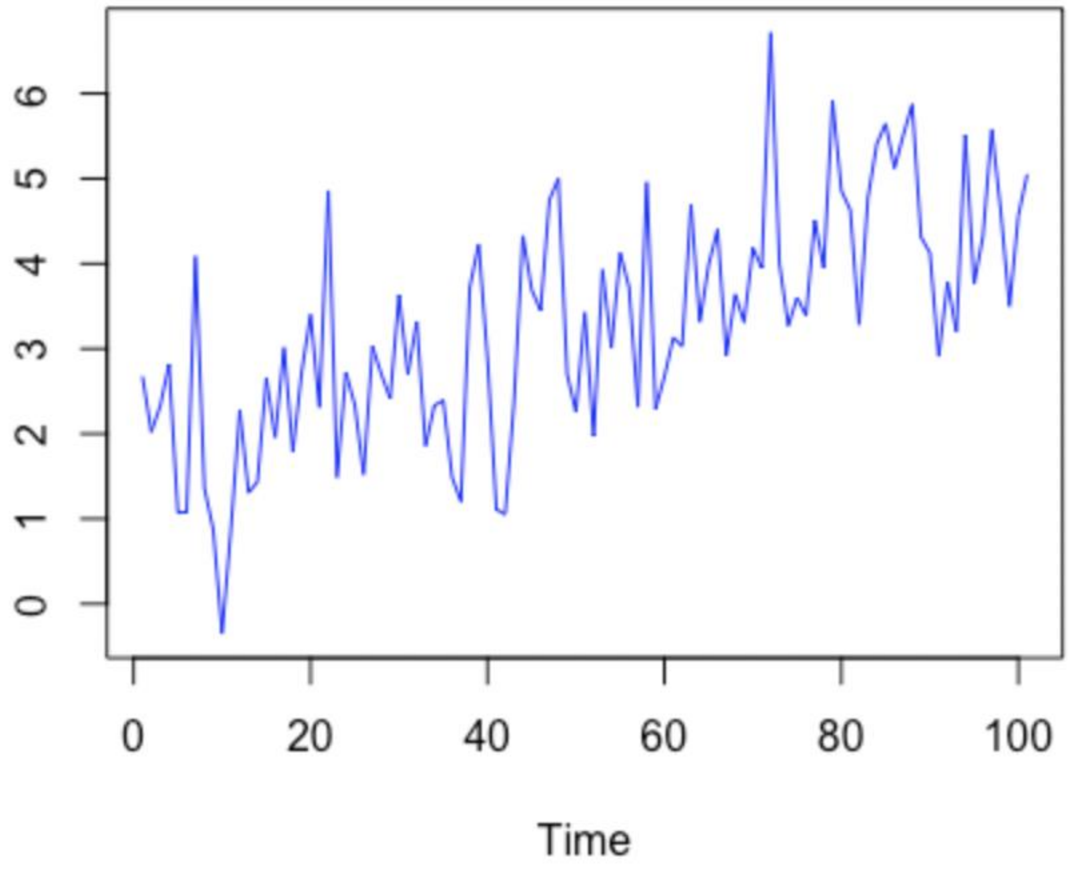
γ_5

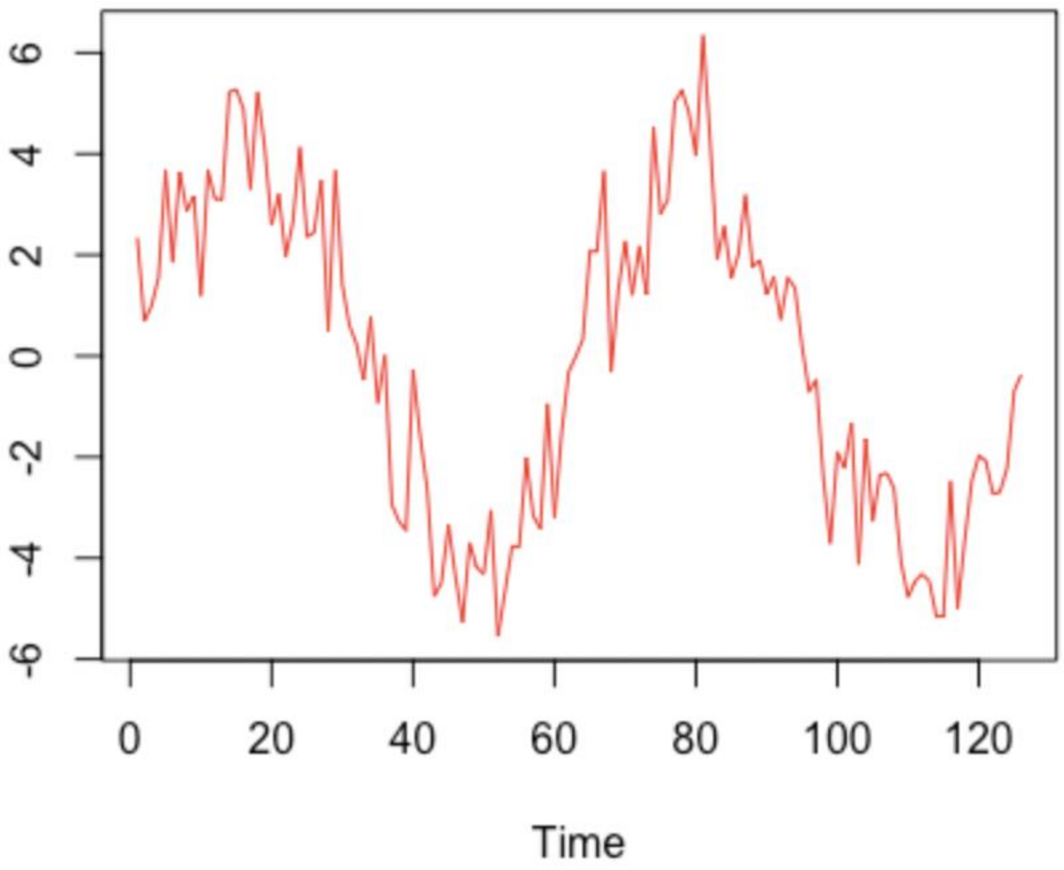
ρ_4

$1 \le k \le 9$

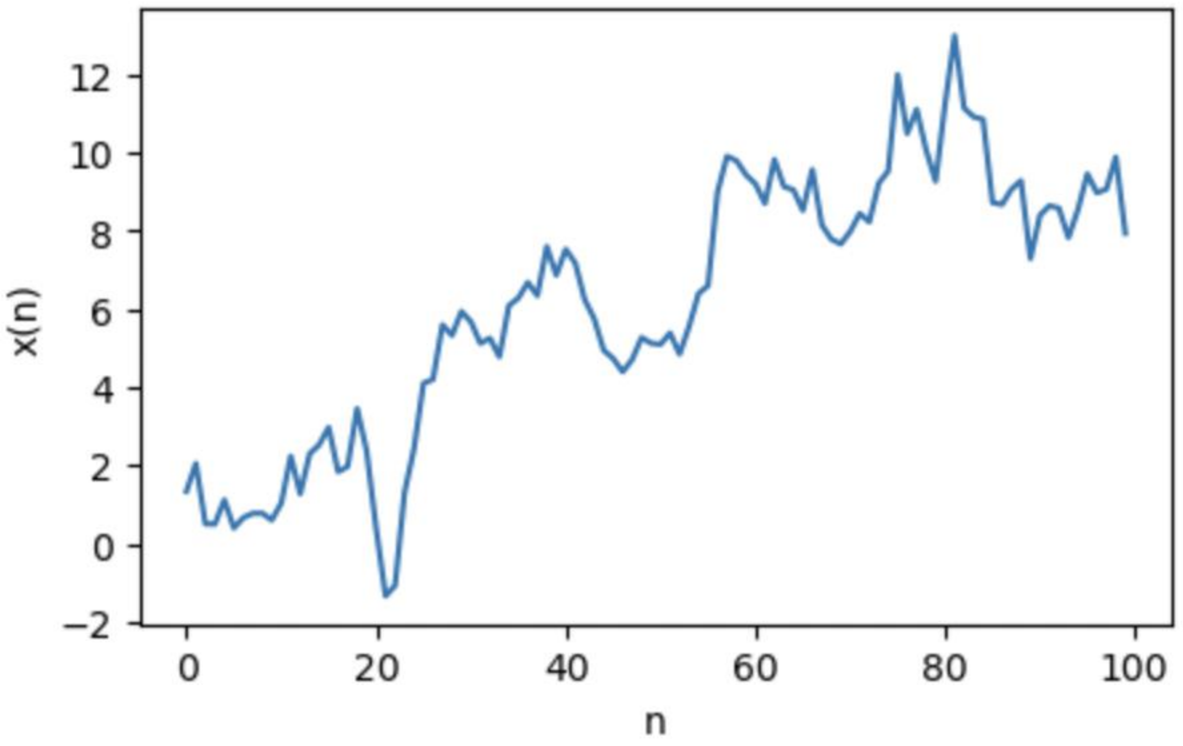
- Partial Autocorrelation Function PACF - will be covered after models

Let’s see some examples





6
4
2
0
-2
-4
-6
0
20
40
60
80
100
120
Time


x(n)
n
12
10
8
6
4
2
0
−2
0
20
40
60
80
100

Queen Mary
University of London