Time Series

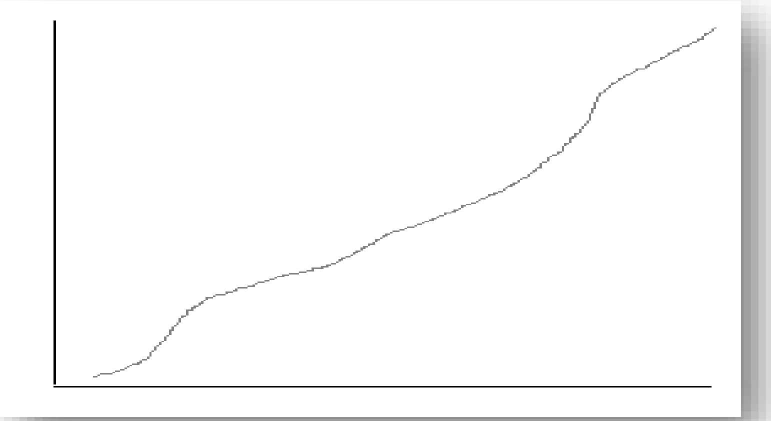
Fundamentals

Assumptions in Time Series Algorithms

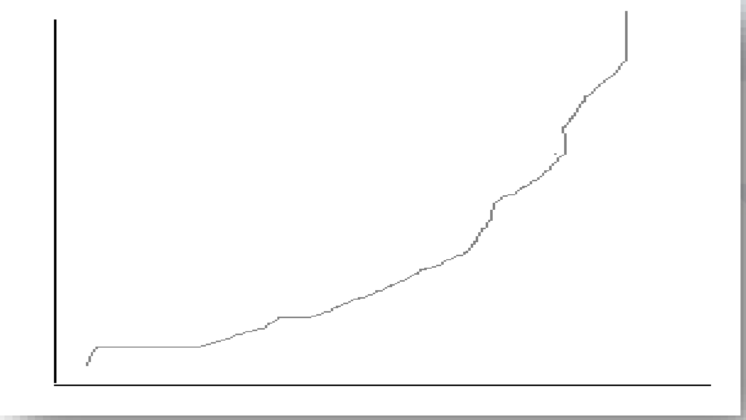
- Consecutive Observations in the series are equally spaced
- Series is indexed on specific period of time. e.g. Weekly, Daily, Yearly etc.
- There aren’t any missing values

Types of Trends

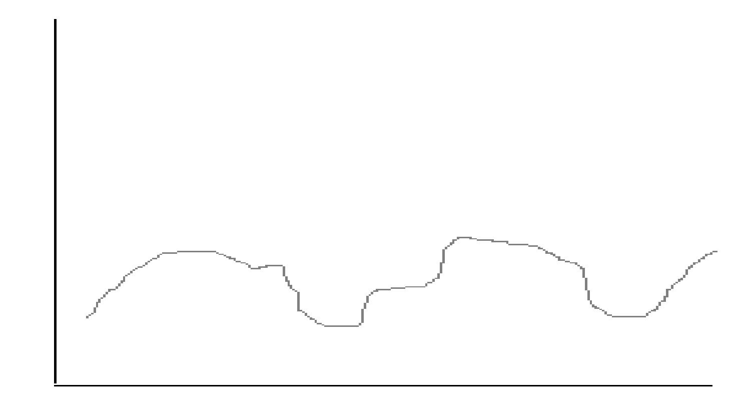
Linear



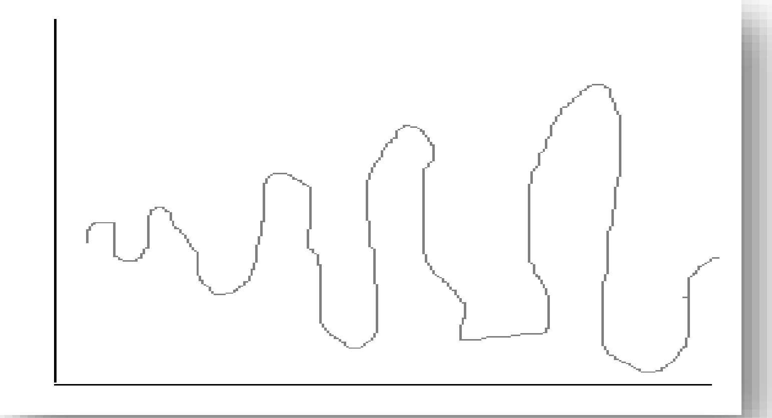
Rapid Growth



Periodic



Varying Variance



Some Transformations

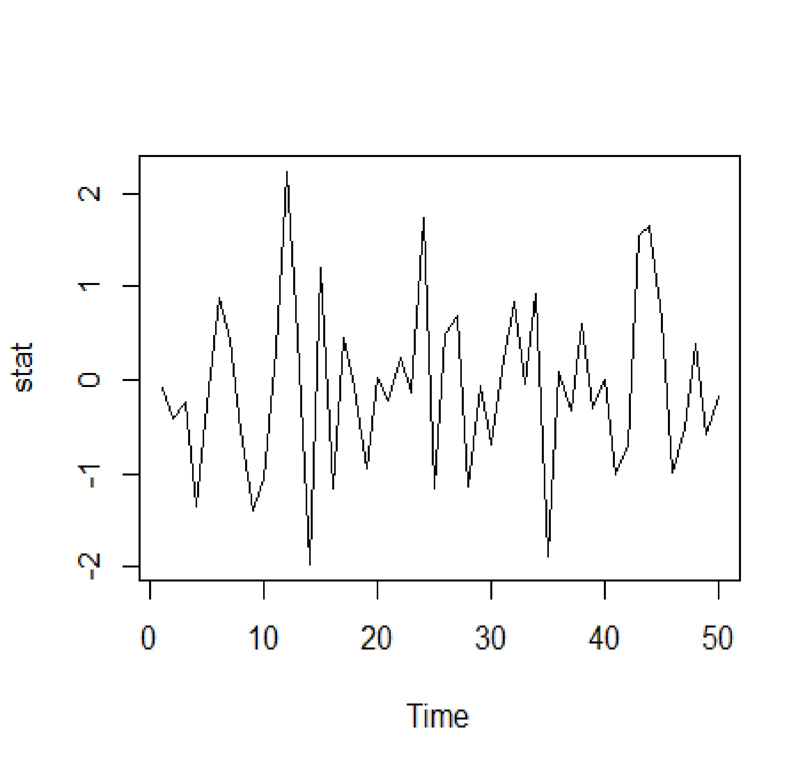
- log: The log() function can linearize the rapid growth trend. It can also stabilize the varying variance series. It is only for positive values.
- diff: The diff() function can remove the linear trends. It can also remove periodic trends.

Stationary Process

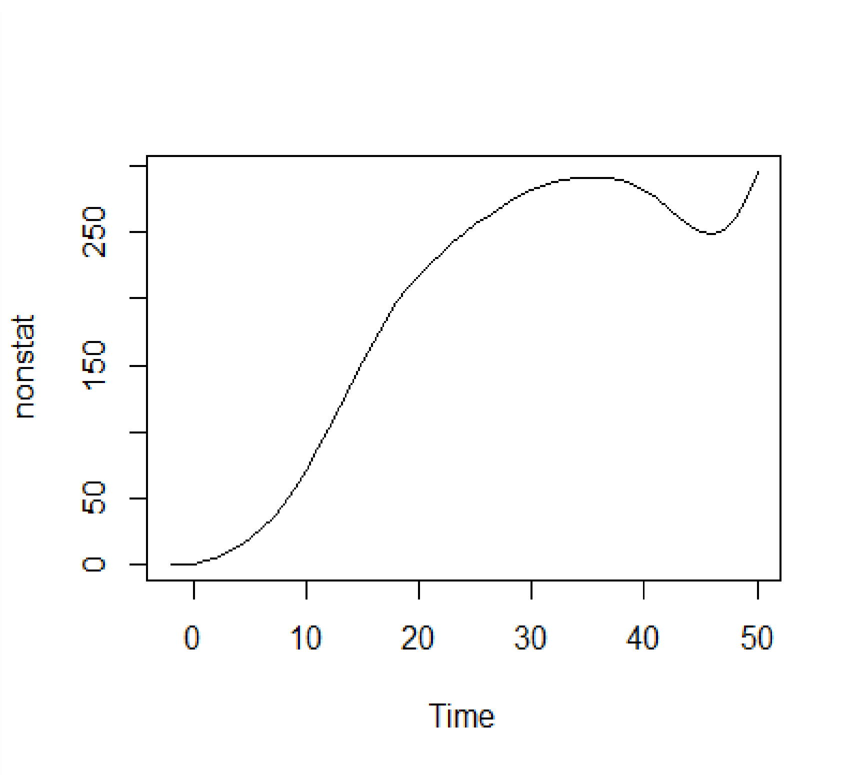
- Stationary process is that stochastic (probabilistic) process whose joint probability distribution does not change when shifted in time.
- In our context of time series, it is that time series whose mean and variance do not change over time.
- White Noise Model is the simplest example of Stationary series.
- For weak stationarity, covariance of y_t and y_s is constant for all $|t-s|=h$, for all h. e.g. $Cov(y_3, y_7) = Cov(y_{22}, y_{26})$

Stationary and Non-Stationary

Stationary



Non-Stationary



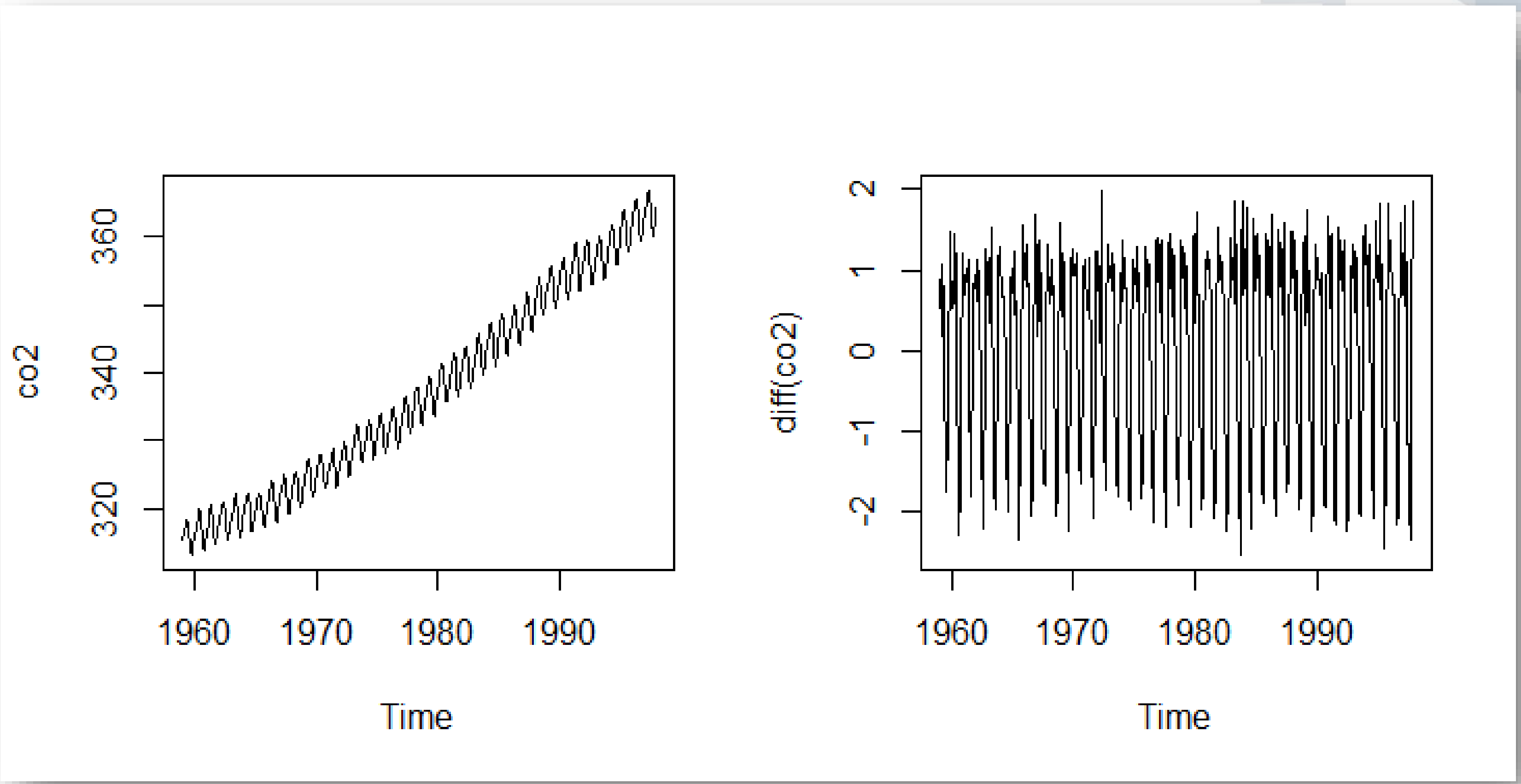
White Noise Model (WN Model)

- WN Model is a simple example of stationary process
- A weak White Noise has
  - A fixed constant mean
  - A fixed constant variance
  - No correlation of any time point value with any time point value

Random Walk (RW) Model

- RW Model is a simple example of non-stationary time series
- A random walk series has
  - No specific mean and variance
  - Strong dependence over time
- Changes or increments in RW series are white noise
- Random Walk Recursion: Today's value = Yesterday's Value + Noise
- In other words, $y_t = y_{t-1} + \in_t$, where $\in_t$ is white noise with mean zero
- RW Model has only one parameter i.e. variance of the white noise σ_ϵ^2

Example of RW Model



Random Walk with Drift

- Random Walk Recursion:
  $$Today's\ Value = Constant + Yesterday's\ Value + Noise$$
- In other words, $y_t = c + y_{t-1} + \in_t$, where $\in_t$ is white noise with mean zero
- This has two parameters, drift constant c and σ_ϵ^2

Autocorrelation

What is Autocorrelation?

- Autocorrelation is correlation between the elements of a series and others from the same series separated from them by a given interval.
- Lag 1 Autocorrelation: Correlation of today’s value with yesterday’s value
- Lag 2 Autocorrelation: Correlation between today’s and day before yesterday’s values
- Lag k Autocorrelation: Correlation between Day 1 with Day k values

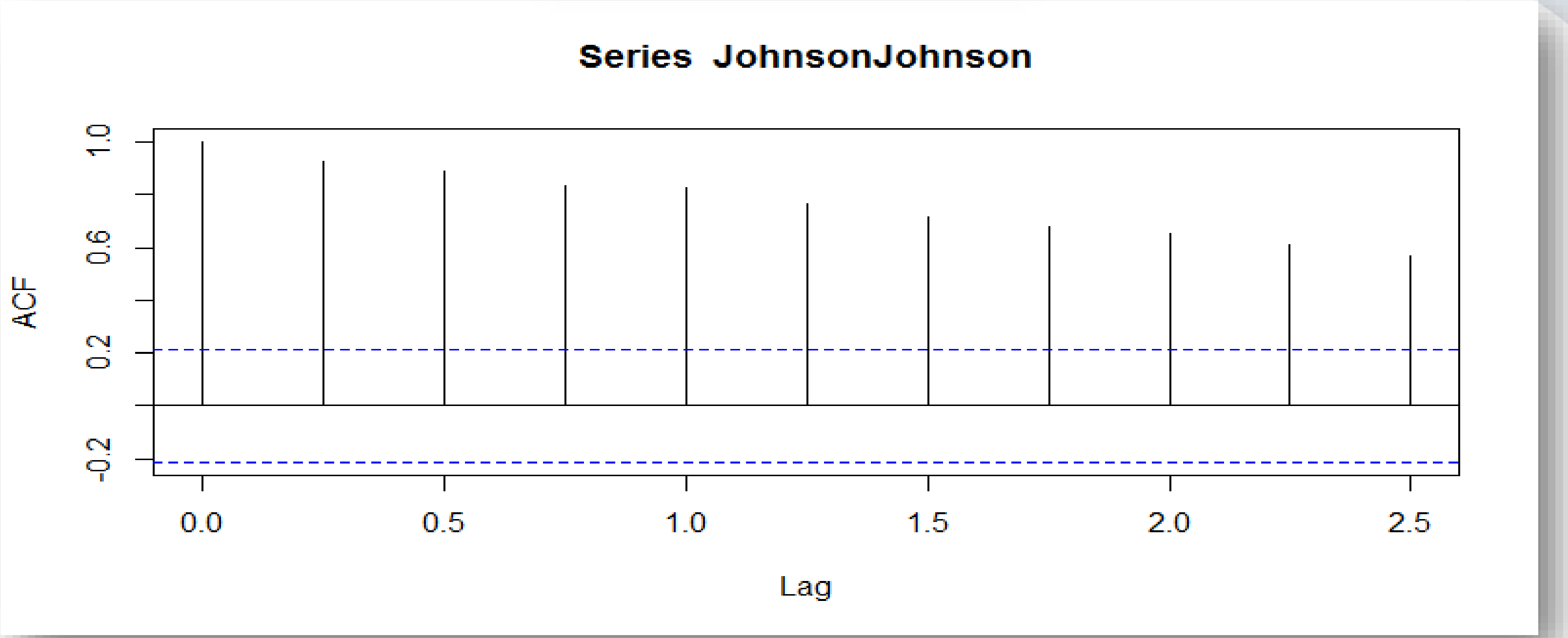
Calculating acf

```
> acf(JohnsonJohnson,10, plot = F)

Autocorrelations of series 'JohnsonJohnson', by lag

 0.00  0.25  0.50  0.75  1.00  1.25  1.50  1.75  2.00  2.25  2.50
1.000 0.925 0.888 0.833 0.824 0.764 0.718 0.675 0.654 0.608 0.564
```

- Output shows quarterly autocorrelations. Plot is rendered FALSE so that the function doesn't produce graph. Here 10 is for maximum lags to produce.
- We observe that, the correlation goes on decreasing with the increase in the lag. This is not the case with every time series.

Plotting acf

```
acf(JohnsonJohnson,10)
```



- We observe here that as the lag goes on increasing, the correlation goes on decreasing

Autoregressive Models

AR Process

Autoregressive Model

- In this model, we consider that today's observation is regressed on yesterday's observation or any of the previous day's observation.
- Model:

$$Today's\ Value = Constant + Slope * Yesterday's\ Value + Noise$$

- Software may use mean centered version of this model as

$$(Today's\ Value - Mean) = Slope * (Yesterday's\ Value - Mean) + Noise$$

- By notations, $y_t - \mu = \phi(y_{t-1} - \mu) + \epsilon_t$, where ϵ_t is a white noise with mean 0 with variance σ_ϵ^2 and ϕ and μ are the slope and mean respectively

AR Process

$$y_t - \mu = \phi(y_{t-1} - \mu) + \epsilon_t$$

- If slope $\phi = 0$ then $y_t = \mu + \epsilon_t$ and y_t will be white noise with mean μ and variance σ_ϵ^2
- If slope $\phi \neq 0$ then the process of $\{y_t\}$ is autocorrelated
- Large value of Ø implies greater dependency of current values with previous values
- Negative value of Ø implies oscillatory time series
- If $\mu = 0$ and slope $\phi = 1$, then $y_t = y_{t-1} + \epsilon_t$, which is a random walk process

Simple Moving Average Model

MA Process

Simple Moving Average Model

- Simple MA model:
$$Today's\ \text{Value} = \text{Mean} + \text{Noise} + \text{Slope} * (\text{Yesterday's Noise})$$
- In mathematical notations,
$$y_t = \mu + \epsilon_t + \theta\epsilon_{t-1}$$

Where

μ: Mean of the series

θ: Slope

ϵ_t: Error or Noise at time t which has mean 0 and some variance σ_ϵ^2

- At $\theta = 0$, the model will be a white noise with mean μ and variance σ_ϵ^2

Simple Moving Average Model

$$y_t = \mu + \epsilon_t + \theta\epsilon_{t-1}$$

- If θ is non-zero then y_t depends on both ϵ_t and ϵ_{t-1} and the process is auto correlated
- Larger values of θ imply greater autocorrelation
- Negative values of θ imply oscillatory time series