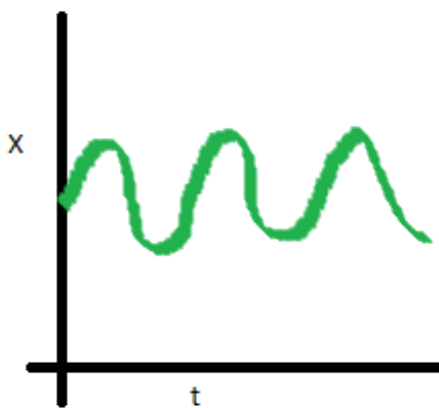# Time Series Analysis Basics
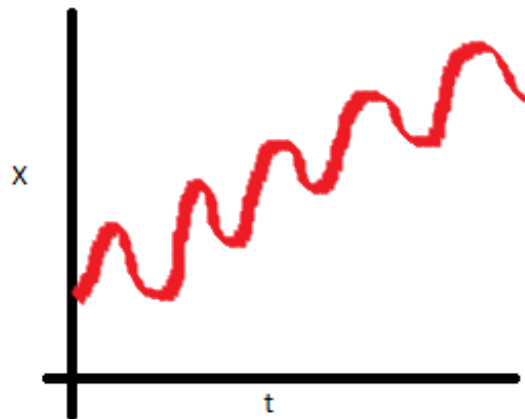
Stationary Series

Three creterias for stationary series:

1. The **mean** of series should be constant but not a function of time.
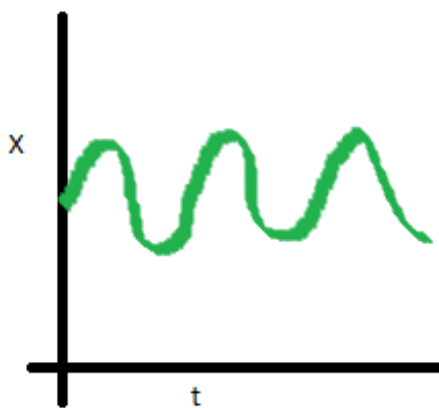

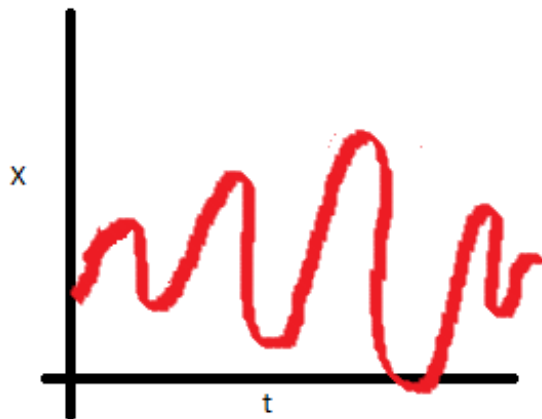Stationary series   Non-Stationary series

2. The **variance** of series should not be a function of time, in economics, it's called homescedasticity.
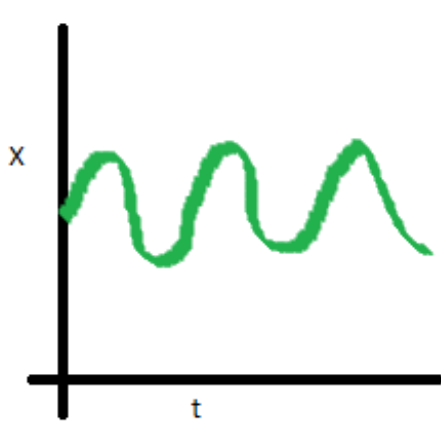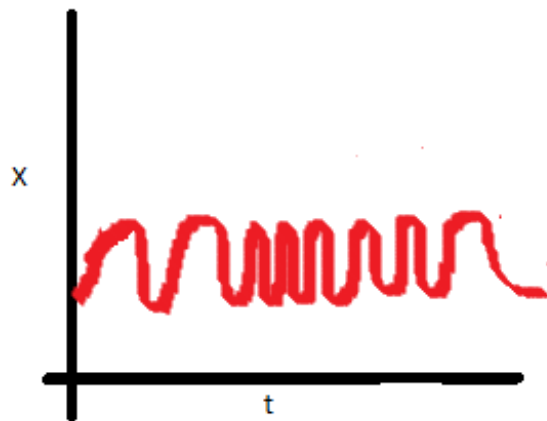

Stationary series   Non-Stationary series

3. The **covariance** of i th term and the (i + m) th term should not be a function of time.

Stationary series   Non-Stationary series

Why stationary series so important?

It cannot be modeled unless it's stationary series. Typically, the first step is stationarizing the time series. The methods include Detrending, Differencing etc.

Random Walk

Suppose X represent a stochastic number (also can be viewed as time series data) and has following feature:

$$dX = dW$$

Which means the movement of X is purely random and X is random walk.

Assume at time 0, the value of X is known as X_0. We can write down the expression of X_t.

$$X_t = X_0 + \int_0^t dW_t$$

As a result, X_t has following chracteristics:

1. **Mean Constant**
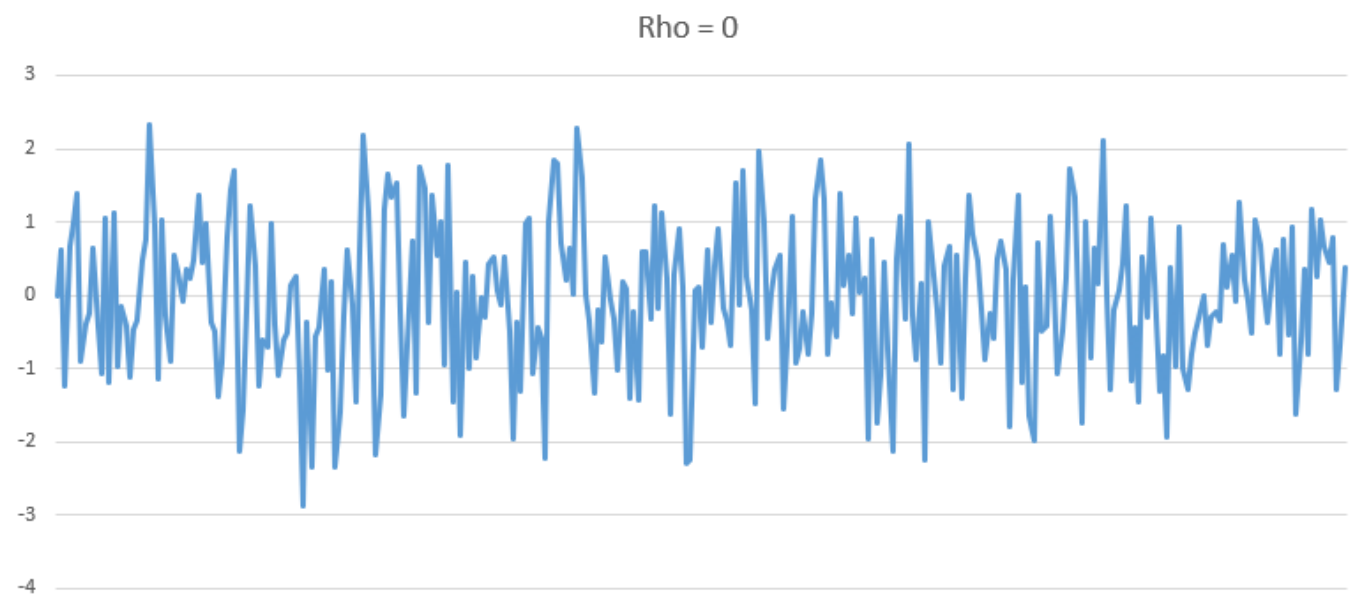
$$E(X_t) = X_0$$

2. **Time Varing Variance**

$$Var(X_t) = \int_0^t dt = t$$

So X is not a stationary process

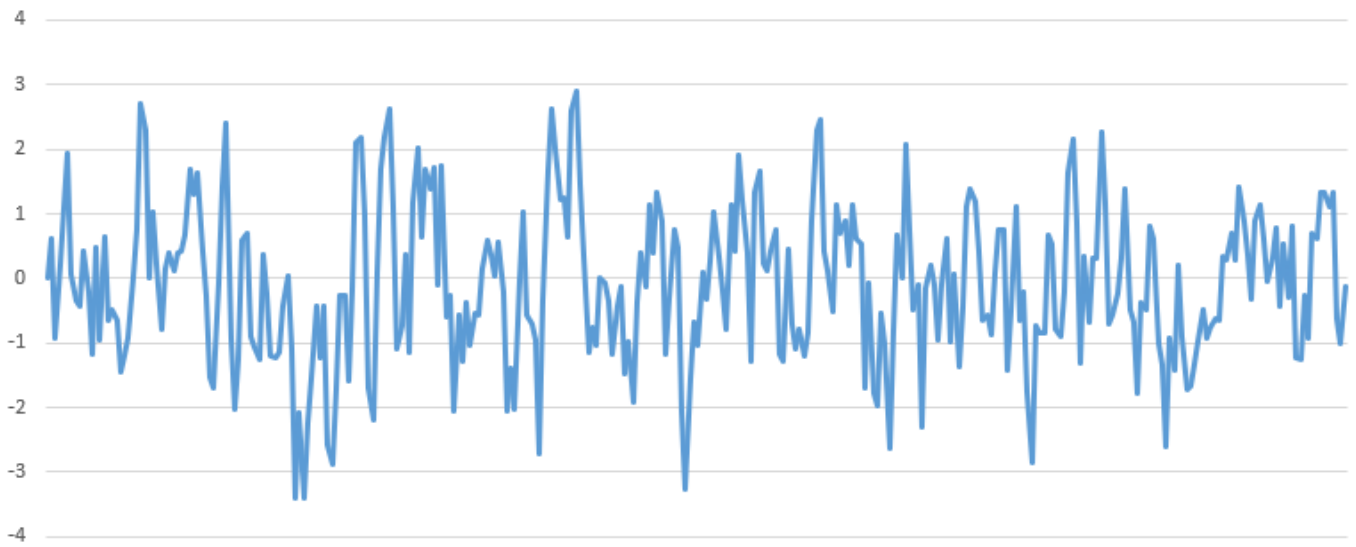# Rho Coefficient and ADF Test

$$X_t = Rho * X_{t-1} + W_{t-(t-1)}$$

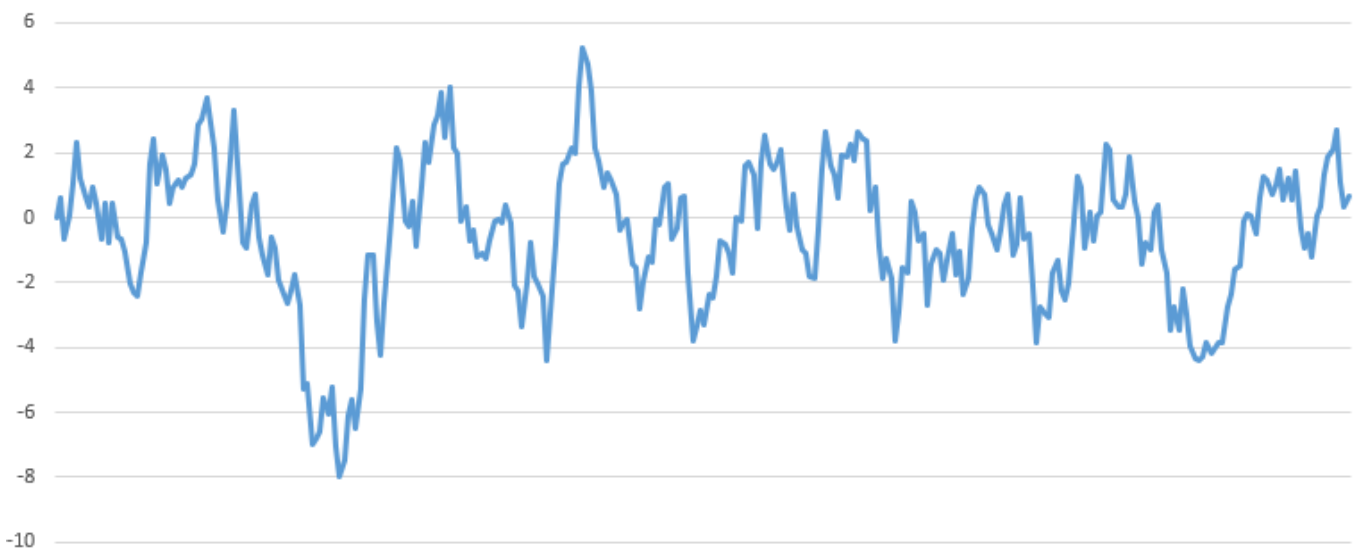If Rho = 0,here is the plot for the time series:

Rho = 0

Rho = 0.5

Rho = 0.5

Rho = 0.9

Rho = 0.9

Rho = 1

As we can see, as Rho increase, the series becomes more and more non-stationary.

Take the expectation:

$$E(X_t) = Rho * E(X_{t-1})$$

If Rho is less than 1, then X is pulled back to 0. For example, $X_{t-1} = 2$ and $Rho = 0.5$, then $E(X_t) = 1$.

Rearrange the equation we get:

$$X_t - X_{t-1} = (Rho - 1) * X_{t-1} + W_{t-(t-1)}$$

Then we can do **Dickey Fuller Test of Stationarity**. Run the regression and do hypothesis test.

Null hypothesis: (Rho-1) equals zero.
Alternative hypythesis: (Rho-1) is significantly different than zero.

If reject null hypothesis, which means rho is not 1, the series is stationary.

Stric Stationary and Weak Stationary

Suppose X_t is a time series. For any n, m and k, if the joint distribution of $Z_n, Z_{n-1} \ldots Z_m$ is the same as $Z_{n-k}, Z_{n-k-1} \ldots Z_{m-k}$. ($n > m$) Then it's defined as stric stationary.

For weak stationary, it has to satisfy two requirements:

1. The **mean** of time series is constant over time..
2. **Covariance stationary**. The autocovariance is the same for all times and lags k. Simpliy speaking, if lags k is decided, the autocov should be the same.

Make time series stationary

There are 2 major reasons behind non-stationary time series:

1. **Trend** - varying mean over time.
2. **Seasonality** - variations at specific time-frames.

Tricks 1 - Transformation

Take a log on time series data which penalize higher values more than smaller values.

Tricks 2 - Elimilate Trend

The first step is modeling the trend, there are several technics:
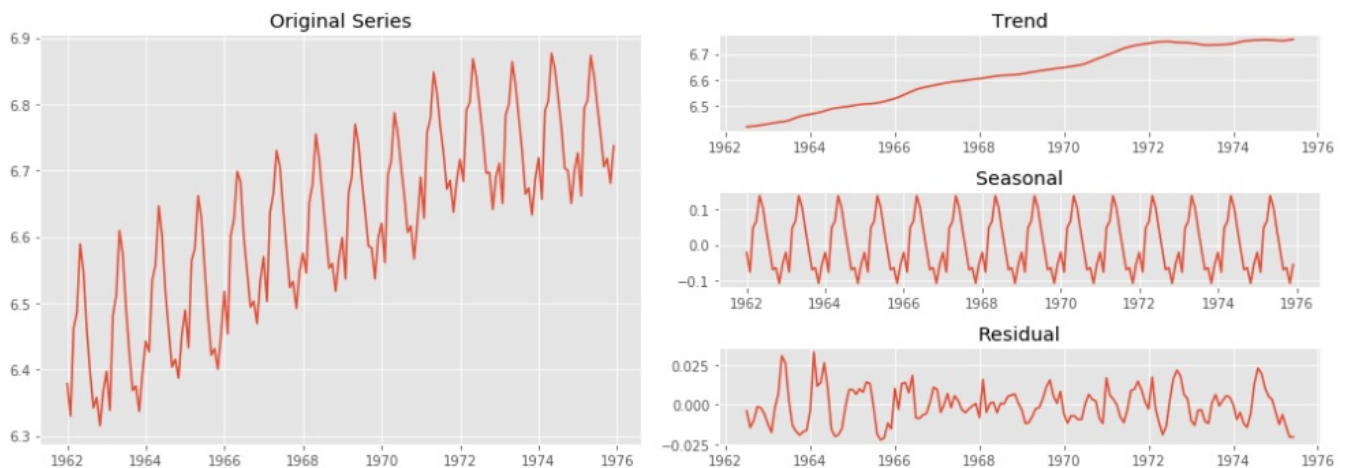
1. **Aggregation** - taking average for a time period
2. **Smoothing** - taking rolling averages
3. **Polynomial** Fitting - fit a regression model

Tricks 3 - Handling Trend and Seasonality

There are mainly two approach:

1. **Differencing**

2. **Decomposition**

Time Series Modeling

ARIMA Model Parameters Determining

1. ACF measures the correlation between the time series and a lagged version of itself.
2. PACF measures the correlation between the TS with a lagged version of itself but after eliminating the variations already explained by the intervening comparisons. For example, ACF between X_t and X_{t-k} is not their pure correlation, since X_t are also affected by $X_{t-1}, X_{t-2} \ldots X_{t-k+1}$. PACF is the correlation which elimilate these affections.tions.
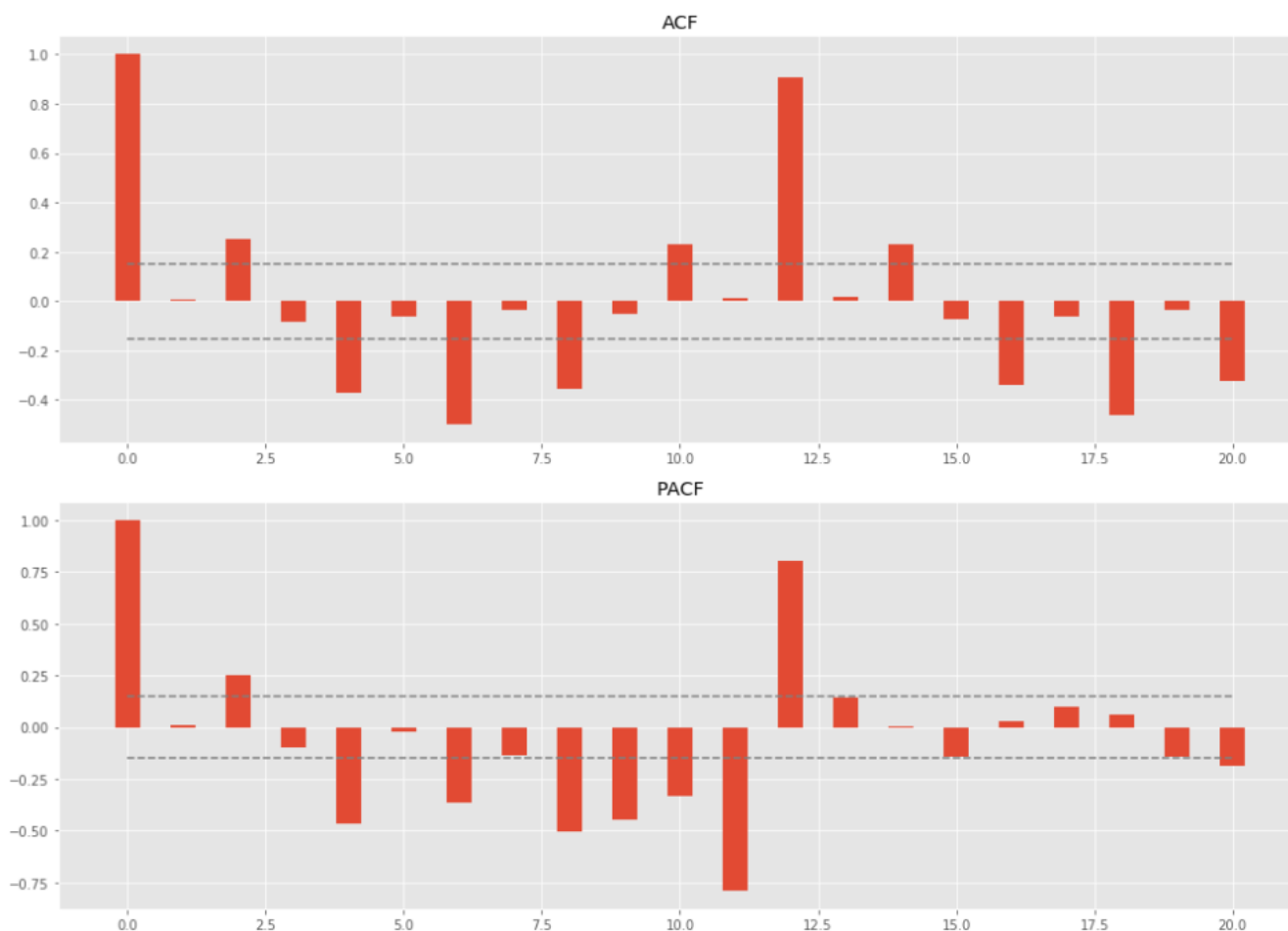
Creteria

Table 3.1. Behavior of the ACF and PACF for ARMA Models

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

Quantitative Model Selection - Information Creteria

As we increase the parameters in the model, the performance tends to be better, but we need to penalize the new parameters. There are two common infomation creteria

AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion).

Suppose k is the number of parameters, n is the sample size and L is likelihood function. For AIC:

$$AIC = 2k - 2Ln(L)$$

For BIC:

$$BIC = kLn(n) - 2Ln(L)$$

The less AIC/BIC the better the model.

References

A Complete Tutorial on Time Series Modeling in R

A comprehensive beginner's guide to create a Time Series Forecast (with Codes in Python)

Analytics Vidhya - Time Series

时间序列建模完整教程（R语言)

时间序列预测初学者指南（Python)