ACF, PACF and Autoregression

ACF → Auto correlation function
PACF → Partial Auto Correlation function.

ACF → Auto + Correlation

Auto ⇓ Correlation itself in the feature.

Correlation ↓ Relationship b/w two feature.

Date	y
Jan	10
Feb	20
Mar	30
Apr	40

↓ time series data

* ACF measures the correlation between time series & its lag value

Month	y_t	1st lag	2nd lag	3rd lag.
Jan	10	NA	Na	Na
Feb	25	10 (Jan)	Na	Na
Mar	35	25 (Feb)	10 (Jan)	Na
Apr	42	35 (Mar)	25 (Feb	10 (Jan)
May	50	42	35 (Mar)	25 (Feb)
June	55	50	42	35
July	62	55	50	42
	$\boxed{y_t}$	$\boxed{y_{t-1}}$	$\boxed{y_{t-2}}$	$\boxed{y_{t-3}}$

Auto correlation

{Pearson, Spearman rank, Kendall}

$Corr(y_t, y_{t-1})$

$Cor(y_t, y_{t-2})$

$Corr(y_t, y_{t-3})$

Partial Autocorrelation Function (PACF)

$t-2 \leftarrow \cancel{t-1} \leftarrow t$

remove the Intermediate affect

Monday ~~Tuesday~~ Wed

Sunny ~~Storm~~ rain

	Month	y_t	$y_t - y_{t-2}$
t-6	Jan	10	
t-5	Feb	18 X	
t-4	Mar	$\boxed{25}$	15
t-3	April	29 X	
t-2	May	$\boxed{35}$	10
t-1	June	45 X	
t	July	$\boxed{50}$	15

25	15
35	10
50	15

PACF

$y_t - y_{t-3}$

$y_t - y_{t-4}$

* Auto regression

Auto → regression with itself.

regression

regression

$X \rightarrow y$

X — IV

y — DV

$y = mx + c$

slope dp intercept.

MLR ⇒ $y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \ldots \theta_n x_n$

$\theta_0 \theta_1 \theta_2 =$ Coeff

$$y_t = \psi y_{t-1} + c$$

$$y_t = \psi_1 y_{t-1} + \psi_2 y_{t-2} + c$$

$$y_t = \psi_1 y_{t-1} + \psi_2 y_{t-2} + \psi_3 y_{t-3} + c$$

y_t — value at current timestamp

ψ — Coeff

c — Constant.

$\varepsilon \rightarrow$ Error.

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \varphi_3 y_{t-3} + \cdots \varphi_n y_{t-n} + C$$

How to decide if you have to go till $t-1$, or $t-2$ or $t-3$??

$\Downarrow$

ACF | PACF

ARIMA