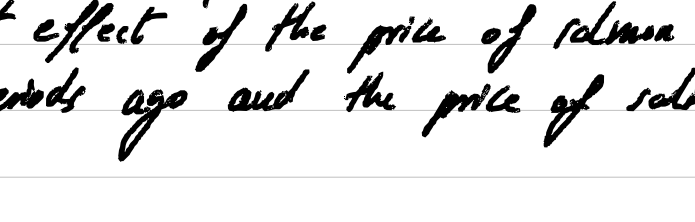
ACF vs PACF

Simply example

- Prediction of the average monthly price of salmon.
  - S_t = avg $ this month
  - S_{t-1} = avg $ last month
  - S_{t-2} = avg $ 2 months ago

There are a lot of external variable that can affect to this month salmon price, but the most intuitive determiner of the of the price of salmon on this month is which was the price of salmon last moth and so on.



So, how do we calculate the Autocorrelation function?

$ACF(S_{t-2}, S_t)$

There are two pieces that build this correlation. S_{t-2} can be correlated with S_t directly and it is also correlated by S_{t-1}:

$ACF(S_{t-2}, S_t)$ → $[S_{t-2} \rightarrow S_t]$
→ $[S_{t-2} \rightarrow S_{t-1} \rightarrow S_t]$

* How is the contrast with PACF?

For PACF we only care about the direct effect.

So, PACF only takes into account the direct correlation, $S_{t-2} \rightarrow S_t$

And AFC takes into account both correlation, so it can happen that the correlation is very high but that correlation comes from INDIRECT effects

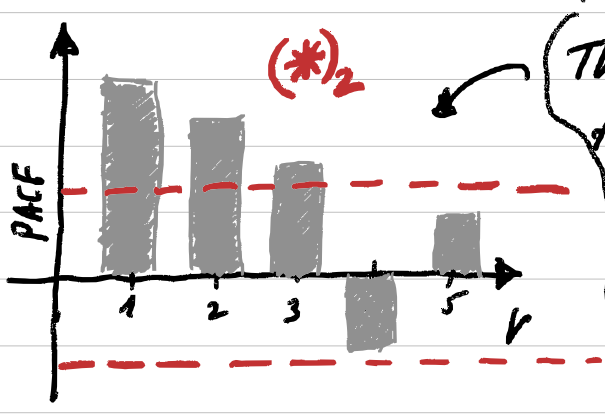
[That is why PACF is so important, because it says, ok, taking all this indirect effects away, just considering the direct ones, what is the direct effect of the price of salmon some number of periods ago and the price of salmon today.]

→ ACF can easily be found using pearson

→ PACF is a little bit more challenging. It must be found using a Regression model.

$$[S_t = \phi_{21} \cdot S_{t-1} + \underbrace{\phi_{22}}_{PACF\,(*)} \cdot S_{t-2} + \varepsilon_t]$$

Let's a if we can find PACF for $k=2$:



This graph indicates us the coef of the price of salmon of t months ago in reference to its price today

(*) We have in red exactly what is PACF, which indicates the effect of S_{t-2} on S_t. We know that this DIRECT EFFECT is being indicated by this coefficient because the other effect is indicated by the first one which is circled by blue.

So the coefficient for a specific term ($S_{t-2}, \ldots, S_{t-n}$) is the PACF for this terms.

$(*)_2$ A good model here would be:

$$[S_t = \beta_0 + \beta_1 \cdot S_{t-1} + \beta_2 S_{t-2} + \ldots + \beta_5 \cdot S_{t-5} + \varepsilon_t]$$

This is called an (AR) or AutoRegressive Model, because it is predicting future values based on past ones.