Time Series Data

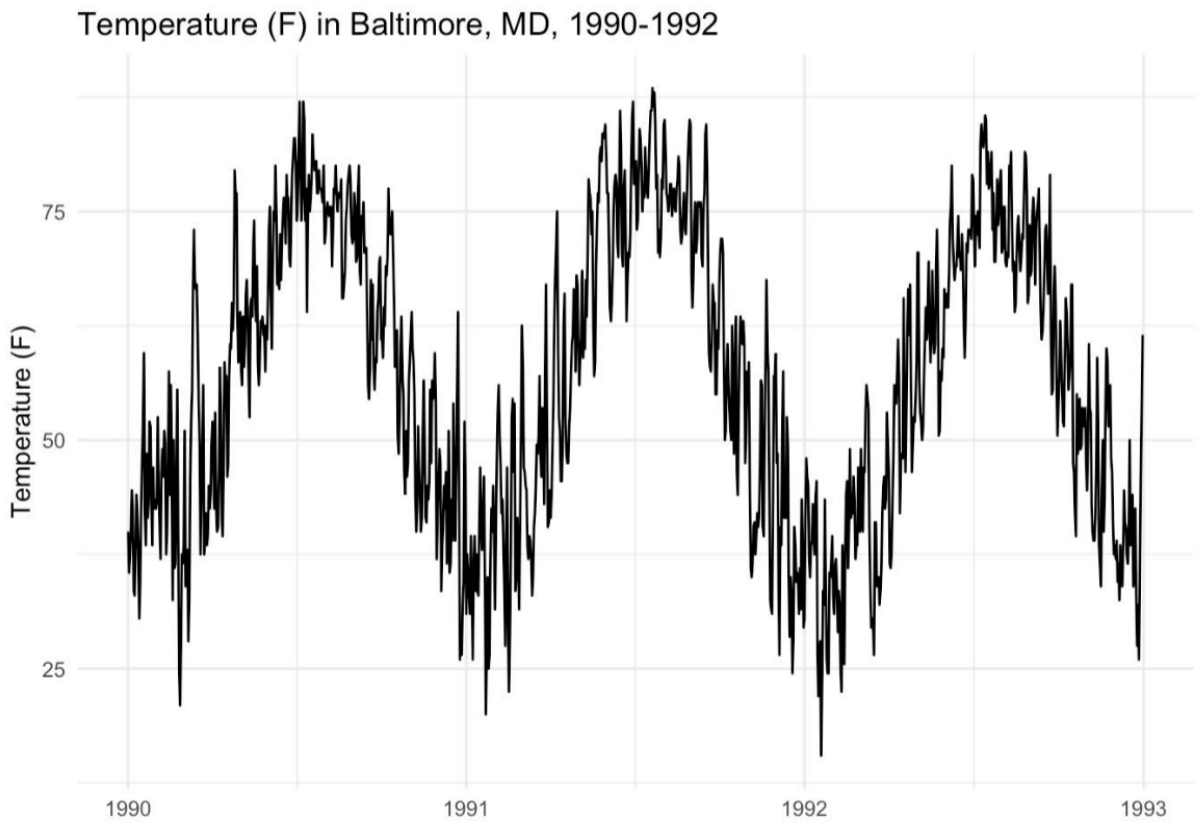
MIE223

Winter 2025

1 Time Series Analysis

1.1 Time Series Data

- Time series data are
- Observations or measurements that are indexed according to time
- Instead of Xi we have Xt where t is specifically a time index
- Why does time make it different?
- The time index t has a special ordering
- Data measured over time are not exchangeable, which is what we often assume when data are indexed by i



1.2 Goals of Time Series Analysis

- Time Scale Analysis: at what scale is data useful to analyze?
- Is weather useful to analyze at the second, hour, or year level? Is it seasonal?
- Smoothing: infer true state of past given a noisy signal (averaging out nearby timepoints)
- Assuming gradual change, use nearby values in a time series to smooth noise
- Filtering: what is the true present state given a noisy signal
- Tracking a rocket trajectory
- Kalman filtering!
- Forecasting: predict future values
- What will the stock price be tomorrow, given the past?
- Regression: give time two series, can we predict the association between them?
- Does public sentiment predict the stock market at a delay?

Note 1. Smoothing is about the past, filtering is about the present, forecasting is about the future

2 Time Scale Analysis

2.1 Time Scale Analysis

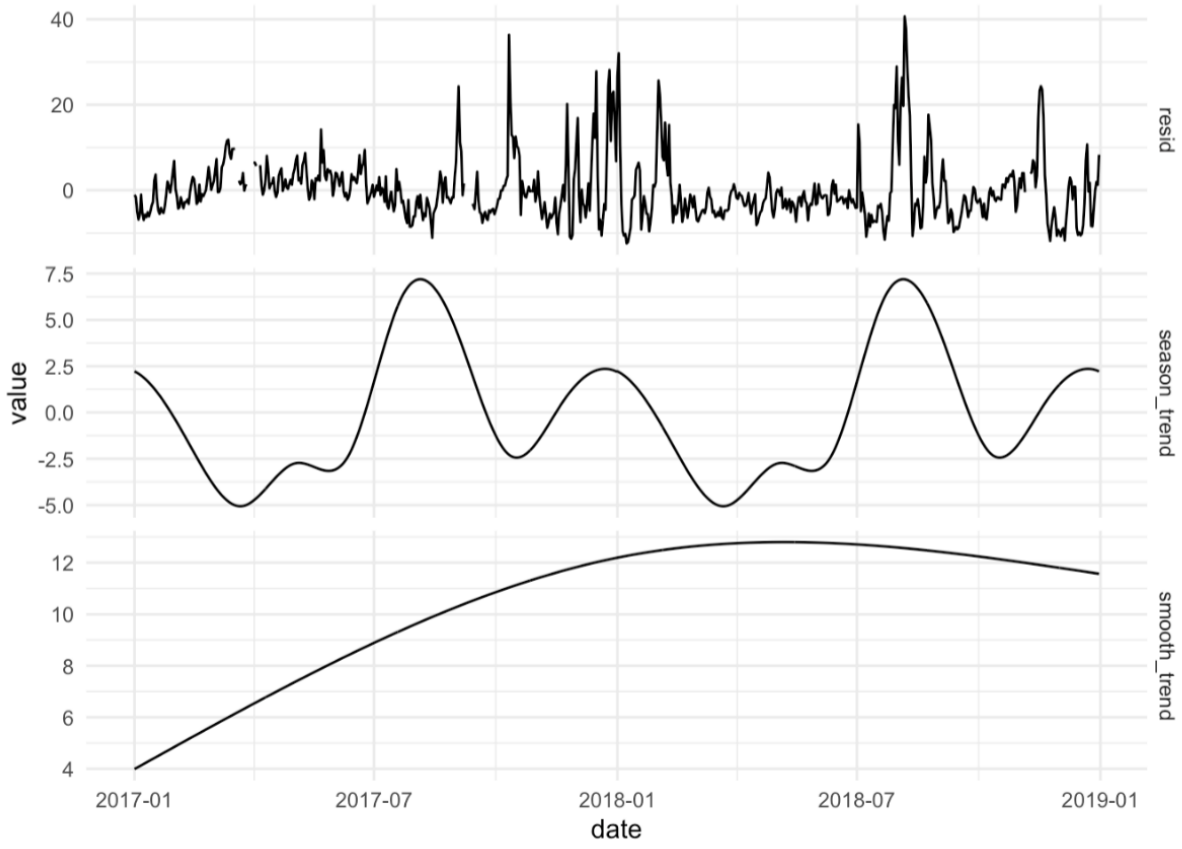
- What time scales of variation dominate or account for the majority of temporal variation in the data? seasonal
- Is there a strong seasonal cycle in the observations of temperature in Baltimore, MD? Hope so
- Is the association between ambient air pollution and mortality in a city primarily driven by large annual changes in pollution levels or by short-term spikes? there can be multiple causes of mortality, and need longer timeframe
- granger causality says that you can predict something if you know the past

2.2 Trend-Season-Residual Decomposition

A common tool to use is to decompose a time series into

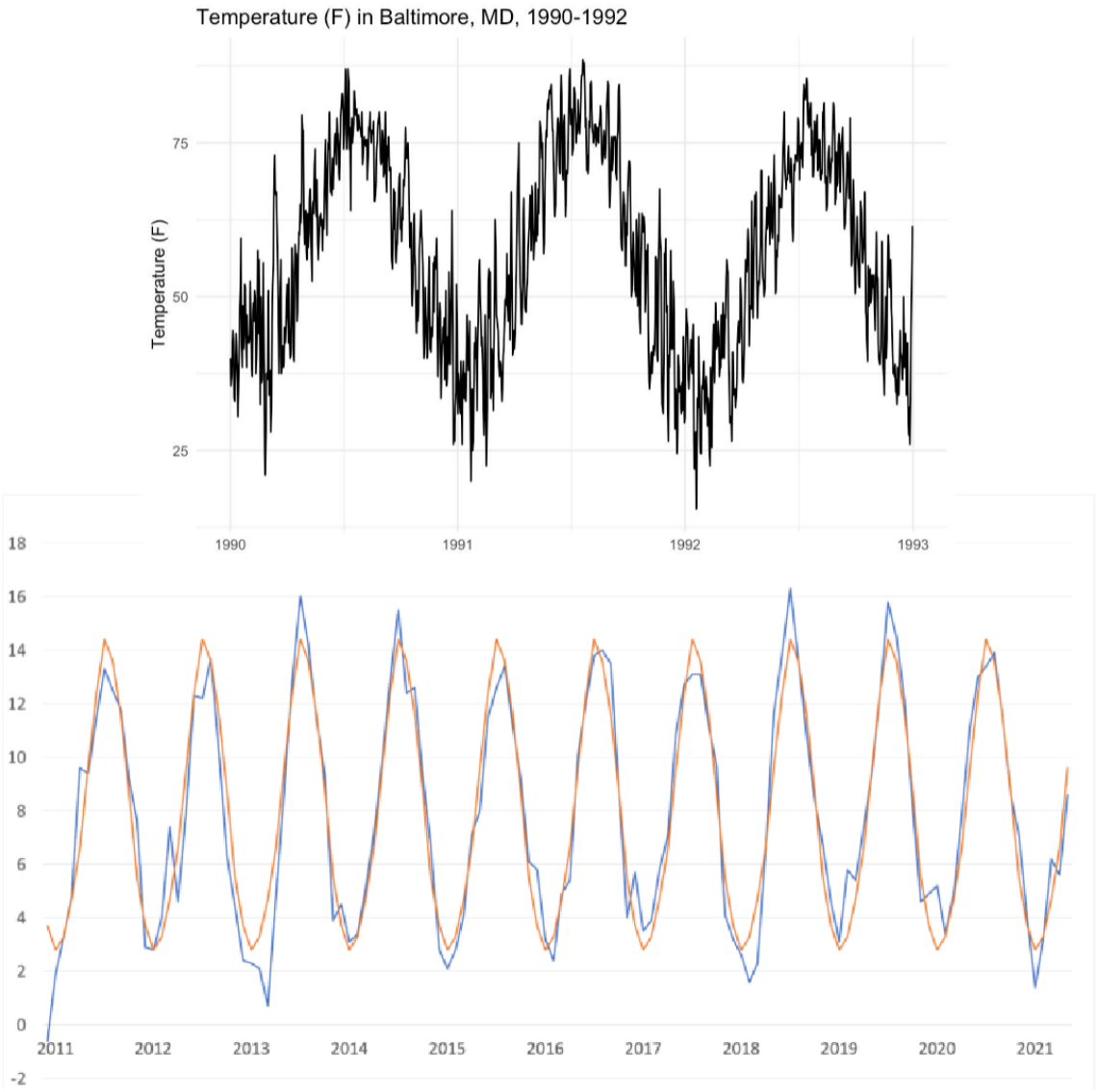
1. Smooth long-term trend (trend)
2. Seasonal variation (seasonal)
3. Residual variation (day to day)

The primary benefit is of looking at long-term trend and seasonal variation is that they are highly interpretable and are very common.
These graphs can be made by smoothing over different intervals.



2.3 Trend-Season-Residual Example 1

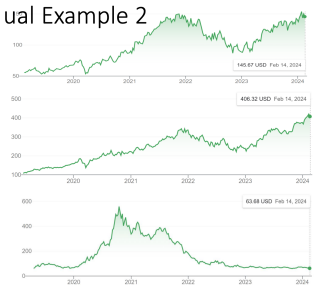
- Daily and monthly temperature vs. year • What time scales are informative?
- What do you think could be
  - Long-term trend: climate change
  - Is there seasonality? weather seasons
  - Cause of variation? day to day fluctuations in the weather



2.4 Trend-Season-Residual Example 2

Three stocks: GOOG, MSFT, ZOOM • Which is which?

- What do you think could be
  - Long-term trend
  - Is there seasonality?
  - Cause of variation?



bottom is zoom, middle is google, top is microsoft

3 Time Series Regression with Autocorrelation and Cross-correlation

3.1 Regression (in general)

Given a time series of two phenomena, what is the association between them?

- What is the association between daily levels of air pollution and daily values of cardiac hospitalizations?
- What is the lag (in months) between a change in a country's unemployment rate and a change in the gross domestic product?
- What is the cumulative number of excess deaths that occurs in the 2 weeks following a major hurricane?

3.2 Recap of Covariance

- Covariance recap:
  - Two random variables X, Y with means $E(X), E(Y)$
  - Covariance $Cov_{X,Y} = E((X - E(X)) \cdot (Y - E(Y)))$
  - If X and Y are independent:
    - $Cov_{X,Y} = E(XY - E(X)Y - E(Y)X + E(X)E(Y)) = 0$
  - If $Cov_{X,Y} > 0$:
    - High values for X go with high values for Y
  - If $Cov_{X,Y} < 0$:
    - High values for X go with low values for Y

probability is model based, stats is data based.

$$E[XY] = E[X]E[Y] \text{if X and Y are independent} \tag{1}$$

3.3 Pearson Correlation Coefficient

Covariance depends on the units in which X and Y are given
Therefore, we look at the Pearson correlation coefficient
If x and y have standard deviations σx and σy respectively:

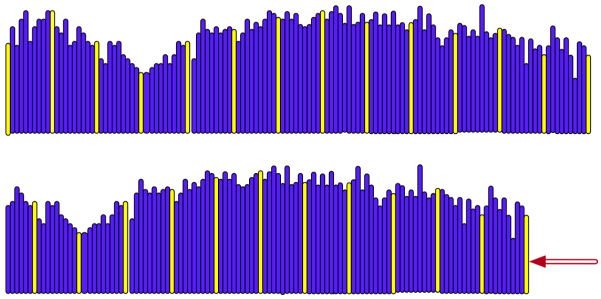
- $\rho_{X,Y} = \frac{Cov_{X,Y}}{\sigma_X \sigma_Y} \in [-1,1]$

$$= \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

independence implies correlation should be 0 but not vice versa

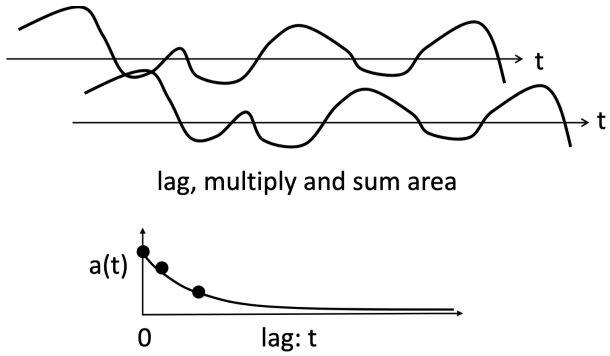
3.4 Autocorrelation

Auto-correlation: Compare the correlation of a time series with itself moved (lagged) by k positions



shift the data by k and see how correlated it is with itself. highest peak represents the most correlated point. tracing out the correlation as a vector. have to get 1 at lag = 0

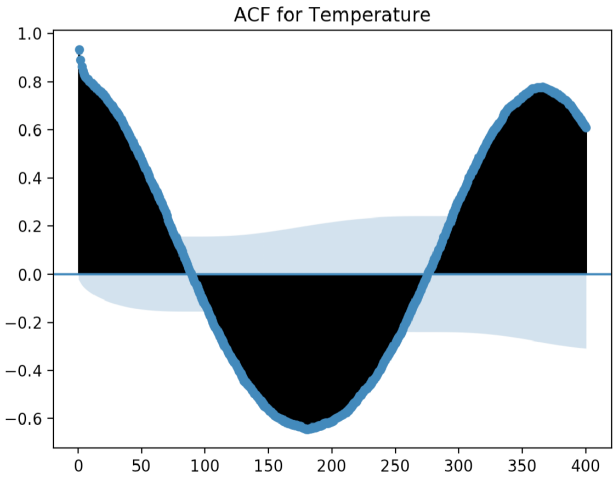
3.5 Autocorrelation: Measure of correlation in time series at different lags



3.6 Autocorrelation Example

Autocorrelation for Chenai temperatures

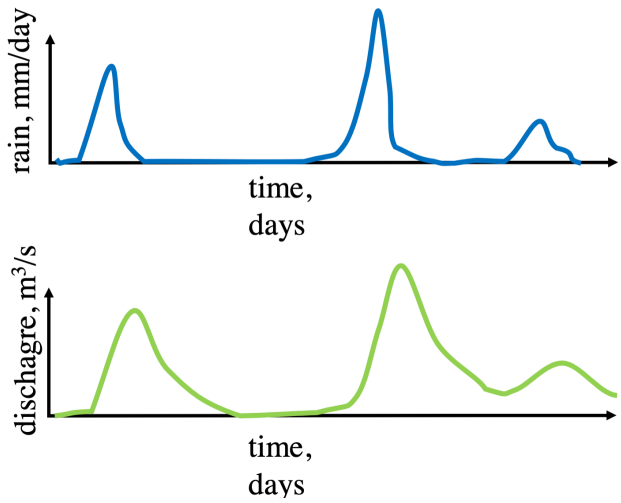
- High auto-correlation for the next day and for a year afterwards
- The light blue denotes significance



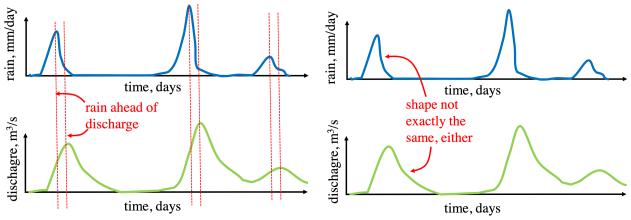
daily measurements. negative correlation with the seasonal changes. there is autocorrelation in the air temperature.

3.7 Cross-correlation

- Consider how one time series "predicts another"
  - Example: how does the amount of rainfall affect (predict) the discharge water flow rate of a river measured at a given point?
    - Discharge correlated with rain
    - Discharge is delayed behind rain because rain takes time to drain from the land

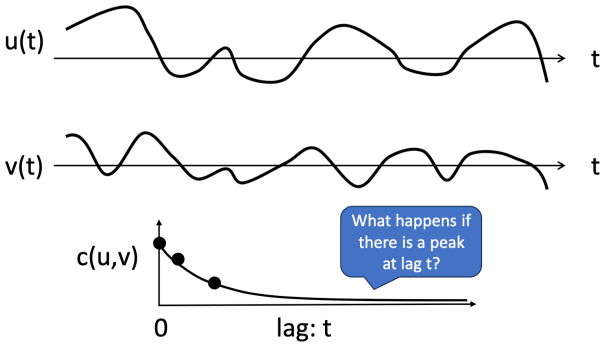


takes time for water to flow. the peak of the discharge is delayed behind the peak of the rain.
Sample midterm question: Would you use VADER to predict the stock market?
Answer: No, VADER is a sentiment analysis tool that is used to analyze text data. It is not used to predict the stock market. Simple polarity sentiment is not complex enough to analyze the stock market



how do we find the lag automatically? shift at which the correlation peaks is the lag.

3.8 Cross-correlation Measure of correlation between two time series at different lags



the cross-correlation does not necessarily peak at 0. the peak is the lag at which the two series are most correlated.

4 Smoothing

4.1 Smoothing

- Given a complete (noisy) dataset, what can I infer about the true state of nature in the past?
- Given a noisily measured signal, can I reconstruct the true signal from the data?
- Now that my spacecraft has flown around the moon, what the distance of its closest approach to the moon?

4.2 Smoothing (example of 3-point weighted average smoothing)

smoothed data = weighted average of observed data

or

$$d_i^{smooth} = \tfrac{1}{4} d_{i-1}^{obs} + \tfrac{1}{2} d_i^{obs} + \tfrac{1}{4} d_{i+1}^{obs}$$

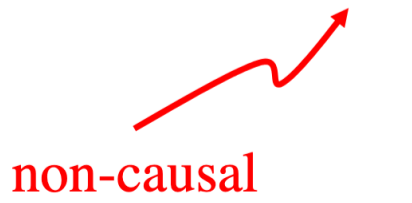
i is the index of the data point, and t is the time index.

4.3 Non-causal Smoothing

smoothed data = weighted average of observed data

or

$$d_i^{smooth} = \tfrac{1}{4}\, d_{i-1}^{obs} + \tfrac{1}{2}\, d_i^{obs} + \tfrac{1}{4}\, d_{i+1}^{obs}$$



non-causal

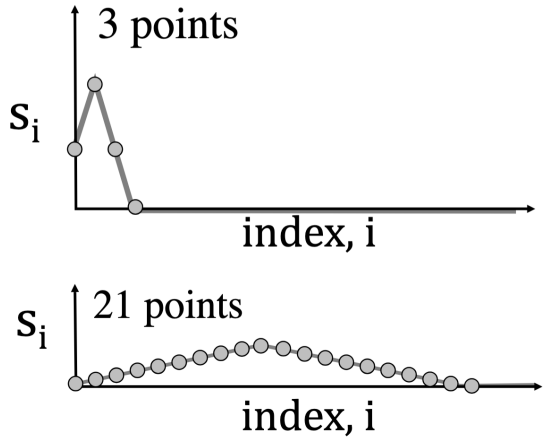
you cannot use this form of smoothing for forecasting. the smoothed value depends on future values. causal smoothing is used for forecasting. causal smoothing is used for filtering and forecasting. non-causal smoothing is used for smoothing. causal smoothing is used for filtering and forecasting.
non-causal is using data from the future to predict the past
filtering and forecasting not tested

4.4 Causal Smoothing (smoothed value does not depend on future)

$$d_i^{smoothed\ and\ delayed} = \tfrac{1}{4}\, d_i^{obs} + \tfrac{1}{2}\, d_{i-1}^{obs} + \tfrac{1}{4}\, d_{i-2}^{obs}$$

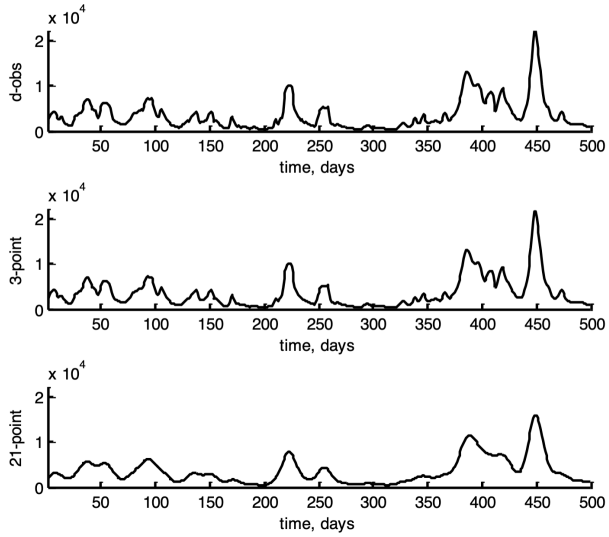
only use lagged values to predict the current value.

4.5 Triangular-weighted Smoothing Filters



the weights form a triangle. the weights are highest at the center and decrease as you move away from the center.

4.6 Smoothing of Neuse River Hydrograph



4.7 Smoothing Methods

- Box (windowed average) smoothing
  - Uniform weights within a fixed window size
- Triangular-weighted and Gaussian smoothing (make it normal)
  - See previous slide for triangular weighting
  - Gaussian just uses Gaussian-shaped weights
- Exponential smoothing
  - Exponential decay to weights, efficiently computed over all history

- compute the next time point prediction as the observed timepoint + (1-alpha) * previous prediction

$$\hat{y}_{T+1|T} = \alpha y_T + \alpha(1-\alpha)y_{T-1} + \alpha(1-\alpha)^2 y_{T-2} + \cdots$$

s1 = alpha * x1 + (1-alpha) * s0 = x0
s2 = alpha * x2 + (1-alpha) * alpha * x1 + $(1 - alpha)^2$ * s0
ARIMA (not covered) - An autoregressive (trained) predictive regression model