

Forecasting: Problem Formulation

- Forecasting: predicting the **future values** of the series using **current information set**
- **Current information set** consists of current and past values of the series and other “exogenous” series

Time Series Forecasting Requires a Model

Time Series Forecasting Requires Models

$$\begin{aligned}\hat{y}_{y+H|t} &= f(\text{current information set}) \\ &= f(\underbrace{y_t, y_{t-1}, y_{t-2}, \dots, y_1, \mathbf{X}_t, \mathbf{X}_{t-1}, \dots, \mathbf{X}_1}_{\text{Information Set}})\end{aligned}$$

Forecast
horizon: H

A statistical model or a
machine learning
algorithm

Information Set:

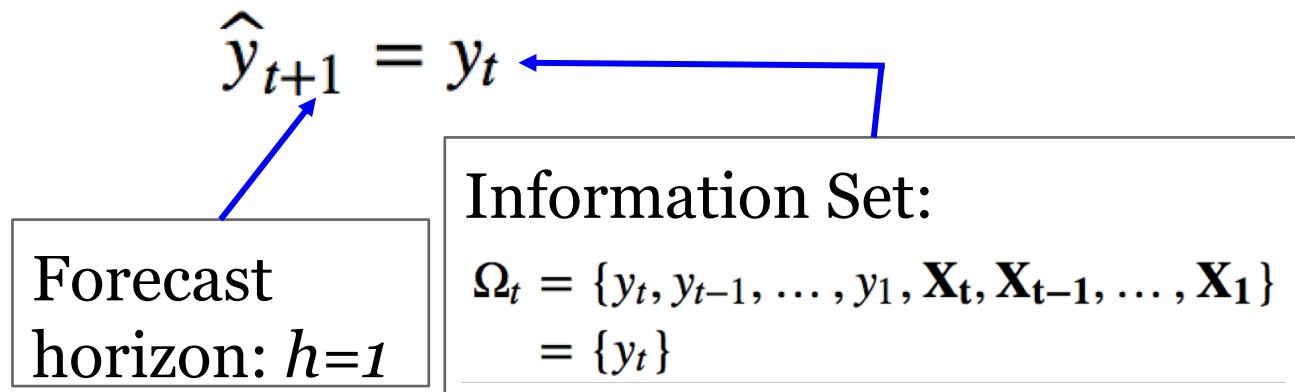
$$\Omega_t = \{y_t, y_{t-1}, y_{t-2}, \dots, y_1, \mathbf{X}_t, \mathbf{X}_{t-1}, \dots, \mathbf{X}_1\}$$

But ... What models?

A Naïve, Rule-based Model:

A model, $f()$, could be as simple as “a rule” - naive model:

The forecast for tomorrow is the observed value today



However ...

**$f()$ could be a slightly more complicated
“model” that utilizes more information
from the current information set**

“Rolling” Average Model

The forecast for time $t+1$ is an average of the observed values from a predefined, past k time periods

$$\hat{y}_{t+1} = \frac{1}{k} \sum_{s=t-k}^t y_s$$

Forecast horizon: $h=1$

Information Set:

$$\Omega_t = \{y_t, y_{t-1}, \dots, y_1, \mathbf{X}_t, \mathbf{X}_{t-1}, \dots, \mathbf{X}_1\}$$

$$= \{y_t, \dots, y_{t-k}\}$$

More Sophisticated Models

... there are other, much more sophisticated models that can utilize the information set much more efficiently

A Preview

Autoregressive Moving Average (ARIMA)

Model: 1-Minute Recap

$$y_t = a + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \dots + \theta_q \omega_{t-q}$$

values from own series

shocks / “error” terms

$$\omega_t \sim N(0, \sigma_\omega^2) \quad \forall t$$

It models the dynamics of the series