Week 1 — Introduction to Forecasting

Time Series Components, Stationarity, and Naive Methods

Dr Juan Zurita

Economic & Business Forecasting
The University of Edinburgh · School of Economics

What Is Forecasting?

Why Forecast?

- Monetary policy: central banks forecast inflation to set interest rates
- Fiscal planning: governments forecast GDP to plan budgets
- Business: firms forecast demand for inventory and staffing
- Finance: risk models rely on forecasts of returns and volatility

Key question: given observations $y_1, y_2, \ldots, y_T$, what is our best prediction of y_{T+h} for horizon $h \geq 1$?

Types of Forecasting

Qualitative

- Expert judgement
- Delphi method
- Surveys

Quantitative

- Time-series models
- Causal / econometric models
- Machine learning

This course focuses on quantitative methods.

Time Series Components

Decomposition

Any time series can be decomposed:

$$y_t = T_t + S_t + C_t + \varepsilon_t$$

T_t **Trend** — long-run increase or decrease

S_t **Seasonality** — regular periodic pattern (fixed period)

C_t **Cycle** — fluctuations without a fixed period

ε_t **Irregular** — random noise

Additive vs Multiplicative

Additive:

$$y_t = T_t + S_t + \varepsilon_t$$

Seasonal amplitude is *constant*.

Multiplicative:

$$y_t = T_t \times S_t \times \varepsilon_t$$

Seasonal amplitude *grows* with the level.

Stationarity

Covariance Stationarity

A time series $\{y_t\}$ is covariance stationary if:

1. $E[y_t] = \mu$ for all t (constant mean)
2. $\text{Var}(y_t) = \sigma^2$ for all t (constant variance)
3. $\text{Cov}(y_t, y_{t-k}) = \gamma_k$ depends only on k, not t

Why it matters: most forecasting models assume stationarity, or that we can transform the data to achieve it.

Testing for Stationarity

- **Visual inspection**: does the series wander or mean-revert?
- **ACF**: slow decay $\Rightarrow$ likely non-stationary
- **Augmented Dickey-Fuller (ADF) test**: H_0: unit root (non-stationary)
- **KPSS test**: H_0: stationary

Common remedy: differencing — $\Delta y_t = y_t - y_{t-1}$.

Autocorrelation

ACF and PACF

Autocorrelation Function (ACF):

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\mathrm{Cov}(y_t, y_{t-k})}{\mathrm{Var}(y_t)}$$

Measures *total* correlation at lag k.

Partial ACF (PACF):

Correlation between y_t and y_{t-k} *after* removing the effect of lags $1, \ldots, k-1$.

ACF and PACF are the main tools for model identification.

A First Forecast

The Naive Method

Forecast: $\hat{y}_{T+h|T} = y_T$ (repeat the last observation)

- Surprisingly hard to beat for random walks
- Serves as a **benchmark** — any useful model should beat it
- Seasonal variant: $\hat{y}_{T+h|T} = y_{T+h-m}$ (same season last year)

Forecast evaluation:

$$\text{RMSE} = \sqrt{\frac{1}{H}\sum_{h=1}^{H}\left(y_{T+h} - \hat{y}_{T+h|T}\right)^2}$$

Summary

- Time series have **trend**, **seasonal**, **cyclical**, and **irregular** components
- **Stationarity** is the key assumption behind most models
- **ACF/PACF** reveal the dependence structure
- The **naive method** is a useful benchmark

Next week: probability and statistics foundations for forecasting.