STA457S 2023 Summer

Anton S.

- **Professor:** Esam Mahdi, `e.mahdi@mail.utoronto.ca`
- **Lectures:** MW 6-9 BA1160

Contents

1. Introduction

Time series can be defined as a collection of random variables indexed according to the order they are obtained in time. We can consider time series as a sequence of random variables

$$x_1, x_2, \dots, x_t, \dots$$

where x_t is obtained at t-th time point. In this course, the indexing variable t will typically be discrete and not continuous. I.e. $t \in \mathbb{N}$ or $t \in \mathbb{Z}$. A *time series* is a series of observed values (x_t), we call the unrealized model a *process* in this course.

Definition 1. A series is *stationary* if it remains around a mean value over time.

Examples: Daily temperature, stock prices, generally measurements

Box-Jenkins Methodology

1. **Identification:** Examine graphs and identify patterns and dependency in an observed time series. We look for: *trend, periodic trend, outliers, irregular change*
2. **Estimation:** Select a suitable fitted model for predicting future values.
3. **Diagnostic checking:** Goodness of fit tests and residual scores to estimate adequacy of the model, determine unaccounted for patterns.
4. **Forecasting:** Use model to forecast the future values.

We say forecasting instead of prediction to indicate foretelling closely into the future.

Financial Time Series

We motivate a lot of this course with financial data, so we define terminology for financial time series.

Definition 2. The *net return* from the holding period $t-1$ to t is

$$R_t = \frac{x_t - x_{t-1}}{x_{t-1}} = \frac{x_t}{x_{t-1}} - 1$$

i.e. relative percent increase of (x_k) from $t-1$ to t.

Definition 3. The *simple gross return* from the holding period $t-1$ to t is

$$\frac{x_t}{x_{t-1}} = 1 + R_t$$

Definition 4. The *gross return over the most recent k periods* is defined as

$$1 + R_t(k) = \frac{x_t}{x_{t-k}} = \prod_{i=0}^{i=k} \frac{x_{t-i}}{x_{t-i-1}} = (1 + R_t) \dots (1 + R_{t-k})$$

Definition 5. The *log returns* or *continuously compounded returns* are denoted r_t and defined as

$$r_t = \log(1 + R_t) = \log(x_t) - \log(x_{t-1})$$

Returns are scale-free but not unitless since they depend on t.

Definition 6. The *volatility* is the conditional standard deviation of underlying asset return.

In most financial time series data, the scale of the volatility appears to be the same. Highly volatile periods tend to be clustered together.

We may decompose a financial time series as

$$x_t = \underbrace{T_t}_{\text{trend}} + \underbrace{s_t}_{\text{season}} + \underbrace{c_t}_{\text{cycle}} + \underbrace{I_t}_{\text{irregularity}}$$

If these components are corelated, use a multiplicative decomposition $x_t = T_t s_t c_t I_t$. If only some are corelated, use a mixed model, i.e. $x_t = s_t T_t + c_t + I_t$.

Time Series Models

Definition 7 (Moving average)**.** The k-th (odd) moving average of a time series (x_t) is defined as the sum of the k values of the time series around x_t. For example, the third moving average series for (x_t) is

$$y_t = \frac{1}{3}(x_{t-1} + x_t + x_{t+1})$$

If k is even, we reindex and define the time of the moving average to be at the middle of the times we evaluate. For example the 4-th moving average of (x_t) is

$$y_t = \frac{1}{4}(x_{t-2} + x_{t-1} + x_{t+1} + x_{t+2})$$

Moving averages allow us to 'smooth' a time series by reducing the noise while maintaining the trend in the series.

Definition 8 (White noise)**.** A *white noise process* is a collection of uncorrelated and identically distributed random variables (w_t), each with 0 mean and finite variance σ_w^2 for every t. If the white noise follows a normal distribution, i.e.

$$w_t \sim N(0, \sigma_w^2)$$

then it is *Gaussian white noise*. In the Gaussian case, independent and uncorrelated are the same, so w_t are i.i.d.

Definition 9 (Random walk)**.** A *random walk with drift* (x_t) is a series

$$x_t = \delta + x_{t-1} + w_t$$

where $w_t \sim \mathrm{wn}(0, \sigma^2)$. For $t \geqslant 1$, δ is the *drift*. When $\delta = 0$, the series is simply a random walk:

$$x_t = x_{t-1} + w_t$$

The series is the same as in the previous time step plus a white noise shock. Therefore we may write

$$x_t = \delta t + \sum_{j=1}^{t} w_j, \quad t \geqslant 1$$

If $\delta \neq 0$, the series is not stationary.

Definition 10 (Signal in noise)**.** Many realistic models for generating time series assume an underlying sinusoidal signal:

$$x_t = A\sin(\omega t + \phi) + \omega_t$$

Definition 11. The *mean function* is defined as

$$\mu_t = E(x_t)$$

μ_t is the expectation of the process at the given t.

The goal of time series analysis is to apply a series of transformations in order to reduce the remaining model to a white noise series. Through these transformations we address trends in the series.