UNIVERSITÀ DEGLI STUDI DI TRIESTE

Statistics

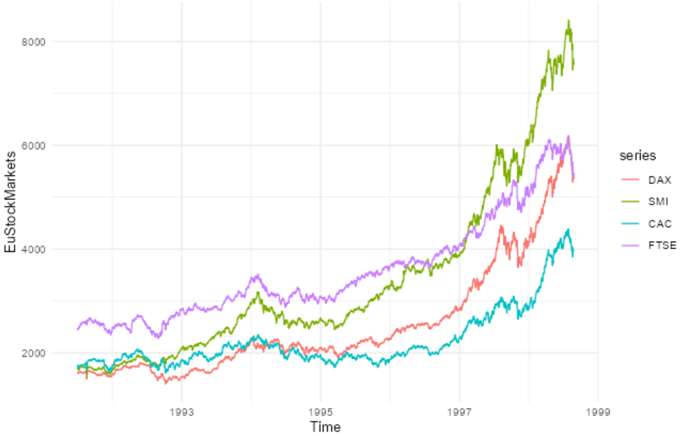
Time Series

Sara Geremia

May 9th, 2024

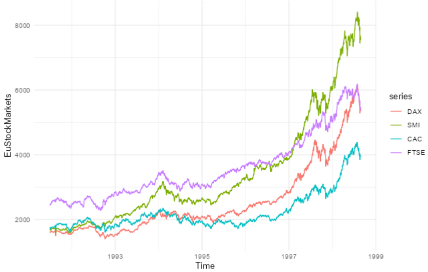
Time Series

- A set of observations indexed by time t: $Y_1, Y_2, \ldots, Y_T$
- Discrete and continuous time series



Time Series

- The interval between observations — that is, the period of time between observation t and observation $t+1$ — is some unit of time such as weeks, months, or years.
- The date $t = 1$ corresponds to the first date in the data set, and $t = T$ corresponds to the final date in the data set.



The data plotted relates to the daily closing prices of major European stock indices from 1991-1998.

Time Series Notation and Terminology

- The change in the value of Y between period $t-1$ and period t is called the first difference in the variable Y_t.
- Δ is used to represent the first difference, so $\Delta Y_t = Y_t - Y_{t-1}$.

- The value of Y in the previous period (relative to the current period, t) is called its first lagged value (or, first lag) and is denoted Y_{t-1}.
- Its jth lagged value (or, jth lag) is its value j periods ago, which is Y_{t-j}.
- Similarly, Y_{t+1} denotes the value of Y one period into the future.

Missing Values

Sometimes there are missing values in time series data.

It is common to replace missing values with the mean of the observed values.

While mean imputation is simple and easy to implement, it often fails in accurately approximating what's really going on due to its inability to account for temporal patterns, variability, relationships between variables, and potential biases. More sophisticated imputation methods are generally required to capture the complexities of such datasets.

Autocorrelation

- In time series data, the value of Y in one period typically is correlated with its value in the next period.
- This correlation of a series with its own lagged values is called autocorrelation or serial correlation.
- The first autocorrelation is the correlation between Y_t and Y_{t-1}—that is, the correlation between the value of Y at time t and its value in the previous period.

Autocorrelation

The autocorrelation coefficient at lag k is given by:

$$\rho_k = \frac{\sum_{t=k+1}^{n}(Y_t - \overline{Y})(Y_{t-k} - \overline{Y})}{\sum_{t=1}^{n}(Y_t - \overline{Y})^2}$$

Autocorrelation

- Autocorrelation can reveal important patterns in time series data, such as trends, cycles, and seasonal effects.
- It is crucial for understanding the underlying structure of the data and for building accurate predictive models.
- High positive autocorrelation indicates that high values of Y tend to be followed by high values, and low values tend to be followed by low values.
- Negative autocorrelation indicates that high values of Y tend to be followed by low values, and vice versa.

Autocovariance

- Autocovariance measures the degree to which a time series is linearly related to a lagged version of itself.
- The autocovariance at lag k quantifies the covariance between Y_t and Y_{t-k}.

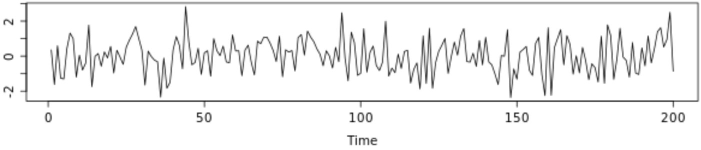
Autocovariance

The autocovariance at lag k is given by:

$$\gamma_k = \frac{1}{n} \sum_{t=k+1}^{n} (Y_t - \overline{Y})(Y_{t-k} - \overline{Y})$$

Stationary Time Series

Some time series do not exhibit any clear trends over time.



A time series is considered stationary if its statistical properties do not change over time.

Stationary Time Series

- Specifically, a time series is stationary if:
  - The mean is constant over time.
  - The variance is constant over time.
  - The autocovariance depends only on the lag between observations, not on the actual time at which the covariance is computed.

Stationary Time Series

A time series $\{Y_t\}$ is stationary if for all t:

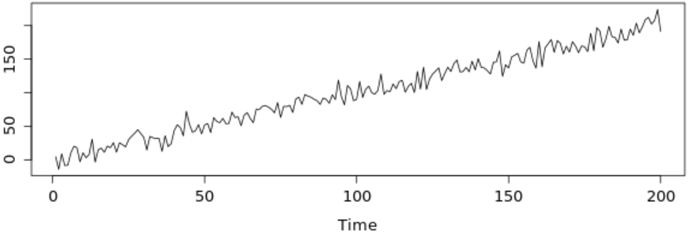
$$E(Y_t) = \mu$$
$$Var(Y_t) = \sigma^2$$
$$Cov(Y_t, Y_{t+k}) = \gamma_k$$

Stationary Time Series

- Stationarity is a key assumption in many time series models and methods, such as ARIMA.
- If a time series is not stationary, it can often be transformed to achieve stationarity through differencing, detrending, or other methods.
- Testing for stationarity can be done using statistical tests.

Linear Trend

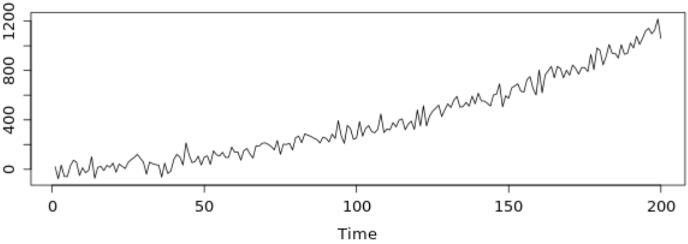
An example of a series with Linear Trend over time.



Data points show a consistent upward or downward movement at a constant rate. This type of trend can be represented by a straight line when plotted on a graph.

Upward Trend

Upward trends may be increasing more quickly than linear.



Rapid decay is also a possibility, but it is not as common in most applications.

Periodic Trend

Examples of series with periodic trends .

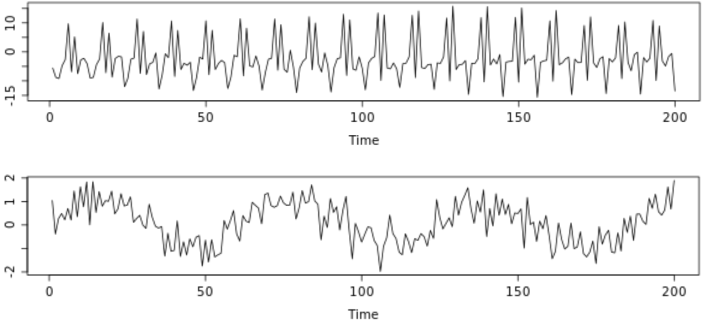
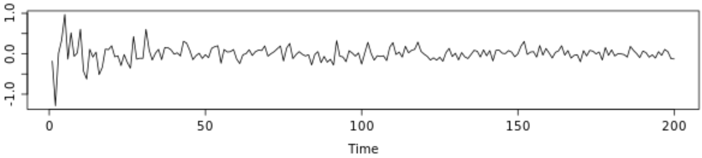


Exhibit patterns that repeat at regular intervals. The key characteristic of such a series is the regularity of peaks and troughs over specific periods.

Trend in Variability

An example of a series with trend in variability .



It refers to changes in the fluctuation or spread of data points over time. This can manifest as either increasing or decreasing variability.

Sample transformations: log()

Data transformations are powerful tools in time series analysis, used to stabilize variance, linearize relationships, and make a dataset more amenable to modeling.

Logarithm Transformations

- Purpose:
  - Linearize exponential growth trends.
  - Stabilize variance, particularly for series with increasing variance over time

- How It Works:
  - Reduces the impact of large values, compressing the range of the data and making exponential growth trends appear linear over time.

Sample transformations: diff()

First Difference Transformations

- Purpose:
  - Remove linear trends from a time series.

- How It Works:
  - The first difference transformation involves subtracting each observation from the subsequent observation. This results in a new series of differences, which represents the increments or changes between consecutive observations in the original series.

Seasonal differencing can be applied to remove periodic trends; lag s difference transformation. For monthly or quarterly data, an appropriate value of s would be 12 or 4, respectively. The first difference transformation function has $s = 1$.

White Noise (W-N) Time Series

- The White Noise (W-N) model is the simplest example of a stationary process.
- In a W-N model, the time series consists of a sequence of uncorrelated random variables.
- Each random variable has a mean of zero and a constant variance.

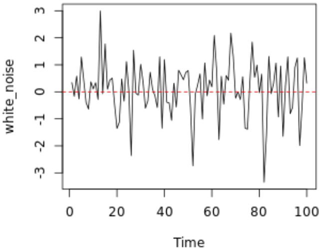
White Noise

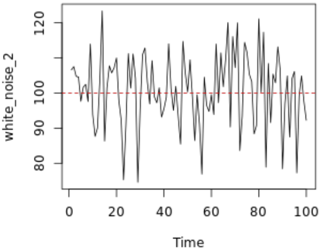
A time series $\{Y_t\}$ is called white noise if:

$$Y_t \sim \text{i.i.d. } (0, \sigma^2)$$

where $E(Y_t) = 0$; $\text{Var}(Y_t) = \sigma^2$ and $Cov(Y_t, Y_{t+k}) = 0$.

White Noise (W-N) Time Series





- **Left Figure**: The series shows no pattern or correlation in the data over time, indicating white noise with zero mean and constant variance.
- **Right Figure**: The series has shifted up, indicating a larger mean and exhibits more vertical variability, signifying a larger variance. However, there are still no patterns or trends over time.

Random Walk (R-W) Time Series

A Random Walk (R-W) is a simple example of an unstable or non-stationary process.

Random Walk

A random walk can be defined as:

$$X_t = X_{t-1} + \epsilon_t$$

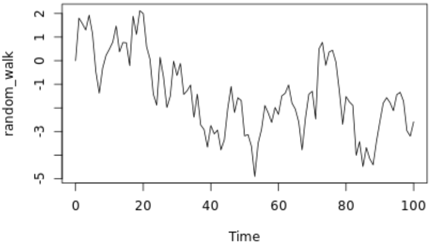
where ϵ_t is a white noise process:

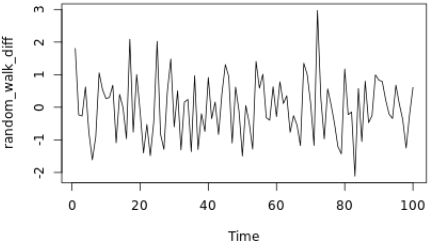
$$\epsilon_t \sim \text{i.i.d. } (0, \sigma^2)$$

Random Walk (R-W) Time Series

- **Characteristics**:
  - **No Specified Mean or Variance**: Unlike stationary processes, a random walk does not have a defined mean or variance.
  - **Strong Dependence Over Time**: Each observation in a random walk is closely related to its immediate predecessor, resulting in a very strong time dependency.
- **White Noise Increments**: The changes or increments ϵ_t follow a white noise process, which is stable and stationary.
- **Implications**:
  - The random walk's path can drift widely over time, lacking a central tendency or predictable long-term behavior.
  - Despite the non-stationarity of the overall process, the increments ϵ_t remain stationary and stable.

Random Walk (R-W) Time Series





- **Left Figure**: In the series, each observation is strongly dependent on its immediate predecessor, resulting in a path that curves over time without a fixed mean or variance.
- **Right Figure**: The first difference of the random walk series is white noise. This is because a random walk is simply recursive white noise data. By removing the long-term trend, you end up with simple white noise.

Back on Stationary Time Series

To obtain parsimony in a time series model you often assume some form of distributional invariance over time, or stationarity.

For an observed time series: Fluctuations appear random. However, the same type of random behaviour often holds from one time period to the next.

Why focus on stationary models?

- A stationary process can be modelled with fewer parameters.
- All Y_t share a common mean μ, which can be estimated accurately by the sample mean $\bar{y}$.

Introduction to Autoregressive Models

- An autoregressive (AR) model is useful for capturing long-term dependencies and trends in time series data.
- AR models assume that the current value of the series is a linear combination of its previous values and a random error term.
- The stationarity assumption ensures that the parameters and the error term have consistent statistical properties over time.
- It is one of the simplest and most widely used models for time series forecasting.

Mathematical Formulation of AR Models

- The AR model of order p (AR(p)) is defined as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t$$

- Where:
  - $\phi_1, \phi_2, \ldots, \phi_p$ are parameters of the model.
  - ϵ_t is white noise error term.

The AR model's core concept is "autoregression," implying a variable's regression against its own past values. It's an intuitive method for capturing time dependence in time series data.

AR(1) Model

- The simplest AR model is the AR(1) model:

$$Y_t = \phi_1 Y_{t-1} + \epsilon_t$$

- Key properties:
  - Mean: $\mu = 0$
  - Variance: $\sigma^2 = \frac{\sigma_\epsilon^2}{1-\phi_1^2}$
  - Autocorrelation: $\rho_k = \phi_1^k$

Estimation of AR Model Parameters

- Parameters $\phi_1, \phi_2, \ldots, \phi_p$ can be estimated using methods such as:
  - Ordinary Least Squares (OLS)
  - Maximum Likelihood Estimation (MLE)
- Once parameters are estimated, the model can be used for forecasting future values.
- Forecast future values by iterating the AR equation.
- Example: Forecast Y_{t+1} using:

$$\hat{Y}_{t+1} = \hat{\phi}_1 Y_t + \hat{\phi}_2 Y_{t-1} + \cdots + \hat{\phi}_p Y_{t-p+1}$$

Model Diagnostics

- After fitting an AR model, it's crucial to check the model diagnostics to ensure the adequacy of the model.
- Key diagnostics include:
  - **Residual Analysis:**
    - Residuals are the differences between the observed values and the values predicted by the model.
    - For a well-fitted AR model, the residuals should resemble white noise.
  - **Autocorrelation Function (ACF) of Residuals:**
    - The ACF measures the correlation between residuals at different lags.
    - Plotting the ACF of residuals helps to check for remaining autocorrelation.
    - For a good model, the ACF should show no significant autocorrelations beyond lag zero.
    - If significant autocorrelations are present, it suggests that the model may be missing some pattern in the data.

Introduction to Moving Average (MA) Models

- Moving Average (MA) models are particularly useful for capturing short-term shocks or noise in time series data.
- An MA model expresses the current value of the series as a linear combination of past white noise error terms.
- It is used to capture the short-term dependencies in the data.

Mathematical Formulation of MA Models

- The MA model of order q (MA(q)) is defined as:

$$Y_t = \mu + \epsilon_t + \theta_1\epsilon_{t-1} + \theta_2\epsilon_{t-2} + \cdots + \theta_q\epsilon_{t-q}$$

- Where:
  - μ is the mean of the series.
  - $\theta_1, \theta_2, \ldots, \theta_q$ are parameters of the model.
  - ϵ_t is white noise error term.

MA(1) Model

- The simplest MA model is the MA(1) model:

$$Y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1}$$

- Key properties:
  - Mean: $\mathbb{E}(Y_t) = \mu$
  - Variance: $\mathrm{Var}(Y_t) = \sigma^2(1 + \theta_1^2)$
  - Autocorrelation: $\rho_k = 0$ for $k > 1$

Estimation of MA Model Parameters

- Parameters $\theta_1, \theta_2, \ldots, \theta_q$ can be estimated using methods such as:
  - **Method of Moments:**
    - Matching sample autocorrelations to theoretical autocorrelations.
  - **Maximum Likelihood Estimation (MLE):**
    - Maximizing the likelihood function under the assumption of normally distributed errors.
- Once parameters are estimated, the model can be used for forecasting future values.
- Example: Forecast Y_{t+1} using:

$$\hat{Y}_{t+1} = \mu + \epsilon_t + \hat{\theta}_1 \epsilon_{t-1} + \hat{\theta}_2 \epsilon_{t-2} + \cdots + \hat{\theta}_q \epsilon_{t-q}$$

  In most cases, the mean μ is assumed to be zero, and the model is centred around this assumption.