PMDS504L: Introduction to Time Series

Generalized Linear Models

Dr. Jisha Francis

Department of Mathematics
School of Advanced Sciences
Vellore Institute of Technology
Vellore Campus, Vellore - 632 014
India

Table of Contents

What is a Time Series?

- A **time series** is a sequence of observations recorded at specific time intervals.
- Each observation is indexed by time t, denoted as x_t.
- Time series data occurs in many fields:
  - **Finance:** Stock market prices
  - **Economics:** GDP growth rates
  - **Weather Forecasting:** Temperature variations
  - **Medicine:** ECG signal recordings

Example: Stock Market Trends

- Stock prices fluctuate over time based on market conditions.
- Key factors affecting stock prices:
    - Market demand and supply
    - Economic news and policies
    - Investor sentiment

Types of Time Series

- **Discrete-Time Series:** Observations are recorded at fixed time intervals (e.g., daily, monthly, yearly).
- **Continuous-Time Series:** Observations occur continuously over an interval (e.g., temperature readings, ECG signals).

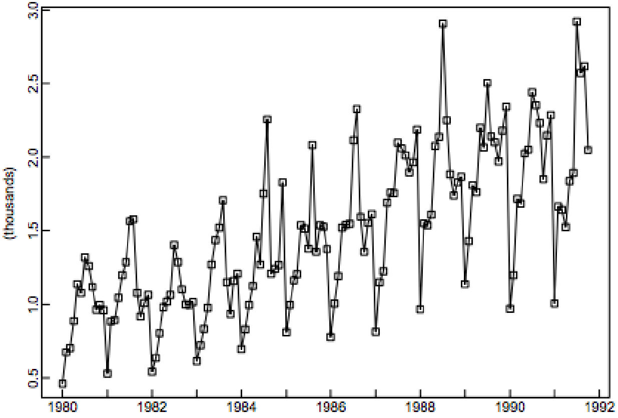
Graphical Representation

- Time series can be visualized using line plots, bar charts, or scatter plots.
- The figure below shows an example of a time series plot for Australian red wine sales.

Example: Australian Red Wine Sales (WINE.TSM)

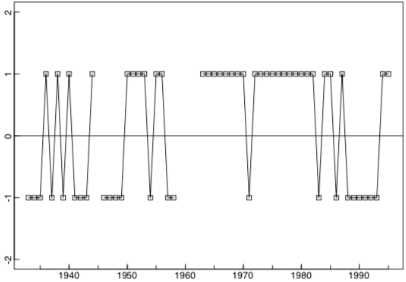
- Monthly sales (in kiloliters) of red wine by Australian winemakers from January 1980 to October 1991.
- The set T_0 consists of 142 time points: $\{$(Jan. 1980), (Feb. 1980), ..., (Oct. 1991)$\}$.
- Often, the time axis is rescaled so that T_0 becomes $\{1, 2, ..., 142\}$, where January 1980 is month 1.
- The data exhibits an **upward trend** and a **seasonal pattern** with peaks in July and troughs in January.

Example: Australian Red Wine Sales (WINE.TSM)



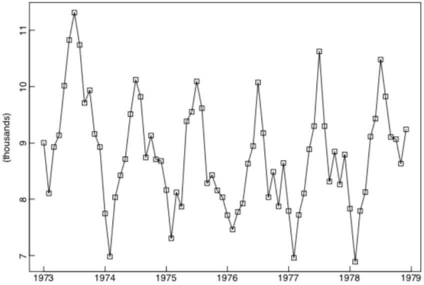
Example: All-Star Baseball Games (1933?1995)

- Results encoded as +1 if the National League won and -1 if the American League won.
- Missing values exist: No game in 1945; two games played in 1959?1962.
- Useful for analyzing win streaks and league dominance over time.



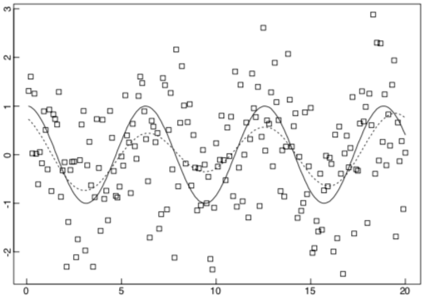
Example: Accidental Deaths in the U.S. (1973?1978)

- Monthly data showing strong seasonal variation.
- Peaks in July, troughs in February.
- Can be decomposed into trend, seasonality, and residual components.



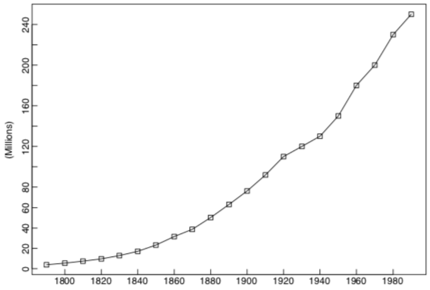
Example: Signal Detection Problem

- Data modeled as signal plus noise: $X_t = \cos(t/10) + N_t$, where N_t is normally distributed noise.
- Can be analyzed using frequency domain methods.
- Smoothing techniques help extract the true signal.



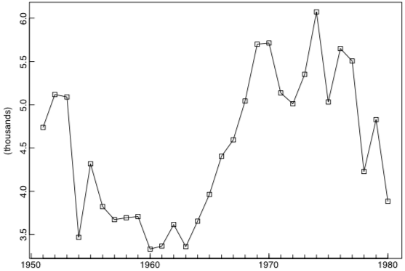
Example: U.S. Population (1790?1990)

- Population recorded at ten-year intervals.
- Possible quadratic or exponential trend.
- Useful for demographic and economic forecasting.



Example: Strikes in the U.S. (1951?1980)

- Annual data on the number of strikes.
- Fluctuations around a slowly changing level.
- Can be analyzed using trend detection techniques.



Time Series

- Time series analysis helps in understanding trends, seasonal effects, and forecasting.
- Discrete and continuous time series are used in various fields, including economics, engineering, and medicine.

Objectives of Time Series Analysis

- To develop a probabilistic model representing the data.
- To estimate parameters and check for goodness of fit.
- To decompose data into trend, seasonal, and random components.
- To apply models for forecasting and hypothesis testing.
- To use models for filtering noise from signals and improving interpretations.
- To support decision-making in fields such as economics, engineering, and science.

Time series model

- An important part of the analysis of a time series is the selection of a suitable probability model (or class of models) for the data.
- To allow for the possibly unpredictable nature of future observations, it is natural to suppose that each observation x_t is a realized value of a certain random variable X_t.

Definition of a Time Series Model

Definition

A **time series model** for the observed data $\{x_t\}$ is a mathematical specification that describes the joint distributions (or at least the means and covariances) of a sequence of random variables $\{X_t\}$.

Time Series Model Selection

Time Series Model Selection

- An important part of time series analysis is selecting a suitable probability model for the data.
- Since future observations may be unpredictable, we assume each observation x_t is a realization of a random variable X_t.

Remark on Time Series

We shall frequently use the term **time series** to mean both the data and the process of which it is a realization.

- **Time Series as Data (Observed Sequence)**
  - Suppose we collect daily temperature readings in a city over 10 days:
    $(x_1 = 30^\circ C, \quad x_2 = 32^\circ C, \quad x_3 = 31^\circ C, \quad \ldots, \quad x_{10} = 29^\circ C$
  - This sequence $\{x_t\}$ is a **realization** of the underlying stochastic process governing temperature variations.
- **Time Series as a Stochastic Process**
  - The temperature readings are not deterministic; they depend on various factors such as weather patterns, atmospheric pressure, and seasonality.
  - We model these observations as a sequence of **random variables** $\{X_t\}$, where each X_t represents the temperature on day t with an associated probability distribution.

Example: Time Series as Both Data and a Process

- **Example: Stock Prices**
  - Consider the closing prices of a company's stock over five days:

$$x_1 = 100, \quad x_2 = 102, \quad x_3 = 98, \quad x_4 = 105, \quad x_5 = 101$$

  - This sequence represents one realization of the **stochastic process** governing stock prices.
  - The process itself $\{X_t\}$ is influenced by multiple random factors like market news, investor sentiment, and economic conditions.

Probabilistic Model of Time Series

A complete probabilistic time series model specifies the joint distributions of the sequence $\{X_1, X_2, \ldots\}$.

- It defines all joint distributions of the random vectors:

$$(X_1, X_2, ..., X_n)'$$

- Alternatively, it provides probabilities:

$$P[X_1 \leq x_1, ..., X_n \leq x_n], \quad -\infty < x_1, ..., x_n < \infty, \quad n = 1, 2, \ldots$$

- Also, in practice, we rarely use full probabilistic specifications because they contain too many parameters to estimate.

Second-Order Properties

- Instead of full joint distributions, we specify only first- and second-order moments:
    - Expected values: $E[X_t]$
    - Expected products: $E[X_{t+h}X_t]$, for $t = 1, 2, \ldots$ and $h = 0, 1, 2, \ldots$
- These properties provide a practical way to analyze time series models.

Multivariate Normal Case

- If all joint distributions are multivariate normal, second-order properties fully determine the distributions.
- This gives a complete probabilistic characterization of the time series sequence.

Information Loss and Justification

- **Information Loss:** Using only second-order properties (mean and covariance) may lead to a loss of higher-order statistical information.
  - Full joint distributions capture all dependencies in a time series.
  - Second-order properties do not capture skewness, kurtosis, or tail dependencies.
  - Two different time series may have the same mean and covariance but different higher-order moments.
- **Minimum Mean Squared Error (MMSE) Linear Prediction:**
  - MMSE linear predictors rely only on second-order properties.
  - Given a stationary time series, the best linear predictor is determined by its mean and covariance.
  - Even for non-Gaussian time series, second-order properties play a dominant role in forecasting.

Components of a Time Series Model

A time series consists of four main components:

- **Trend ($\mathbf{T_t}$)**: The long-term movement in the series.
  - Represents the overall direction (increasing, decreasing, or stable).
  - Example: The gradual increase in global temperature over decades.
- **Seasonality ($\mathbf{S_t}$)**: Regular periodic fluctuations.
  - Recurring patterns over fixed intervals (e.g., daily, monthly, yearly).
  - Example: Higher sales of ice cream in summer.

Components of a Time Series Model

A time series consists of four main components:

- **Cyclic Component (C_t)**: Long-term oscillations.
  - Unlike seasonality, cycles have varying durations.
  - Example: Economic cycles influencing stock markets.
- **Irregular Component (I_t)**: Random fluctuations.
  - Caused by unexpected events (e.g., natural disasters, political crises).
  - Example: A sudden drop in air travel due to a pandemic.

Mathematical Representation

The time series model can be expressed as:

- **Additive Model:**

$$Y_t = T_t + S_t + C_t + I_t$$

- **Multiplicative Model:**

$$Y_t = T_t \times S_t \times C_t \times I_t$$

Example: Monthly Ice Cream Sales

- **Trend:** Overall increase in sales due to population growth.
- **Seasonality:** Higher sales in summer, lower in winter.
- **Cyclic Component:** Economic trends affecting spending habits.
- **Irregular Component:** Unexpected supply chain disruptions.

Zero-Mean Models: IID Noise

Definition:

- A time series model with no trend or seasonal component.
- Observations are independent and identically distributed (iid) random variables with zero mean.
- Represented as:

$$X_1, X_2, X_3, \cdots \sim \text{iid}$$

Mathematical Representation

Probability Distribution:

- For any positive integer n and real numbers $x_1, x_2, \ldots, x_n$,

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P(X_1 \leq x_1) \cdots P(X_n \leq x_n)$$

- Since the variables are identically distributed,

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) = F(x_1) \cdots F(x_n)$$

where $F(\cdot)$ is the cumulative distribution function.

Independence Property

Key Characteristic: No dependence between observations.

- For all $h \geq 1$ and all $x, x_1, \ldots, x_n$,

$$P(X_{n+h} \leq x | X_1 = x_1, \ldots, X_n = x_n) = P(X_{n+h} \leq x)$$

- This shows that past values of X_t do not provide any useful information for predicting future values.

Binary Process: An Example of IID Noise

Definition:

- Consider a sequence of iid random variables $\{X_t, t = 1, 2, \dots\}$.
- Each X_t takes values $+1$ or -1 with probabilities:

$$P(X_t = 1) = p, \quad P(X_t = -1) = 1 - p.$$

- A common example is a fair coin toss where $p = \frac{1}{2}$.

Example: Coin Tossing Model

- Toss a fair coin repeatedly:
  - Assign $+1$ for heads.
  - Assign -1 for tails.
- This forms an iid sequence $\{X_t\}$.
- Used to model independent events, such as sports outcomes.

Random Walk Process

Definition:

- A random walk is formed by cumulatively summing iid random variables.
- Defined as:

$$S_0 = 0, \quad S_t = X_1 + X_2 + \cdots + X_t, \quad t = 1, 2, \ldots$$

where $\{X_t\}$ is iid noise.

Simple Symmetric Random Walk

- If X_t follows a binary process with $p = \frac{1}{2}$, the random walk is called a **simple symmetric random walk**.
- Example:
    - A pedestrian starts at position 0.
    - Tosses a fair coin at each step:
        - Moves right if heads.
        - Moves left if tails.

Reconstructing the Original Process

- The outcomes of coin tosses can be recovered from the random walk.
- By differencing:

$$X_t = S_t - S_{t-1}.$$

- This shows that differencing a random walk retrieves the underlying iid noise process.

Models with Trend and Seasonality

Definition:

- Many real-world time series exhibit trend and/or seasonality .
- A trend represents a long-term increase or decrease in the data.
- Seasonality refers to repeating patterns at regular intervals.

Examples of Trend in Time Series

- Australian Red Wine Sales:
  - Shows an increasing trend over time.
  - Suggests the need for a non-zero mean model .
- U.S. Population (1790-1990):
  - Data follows a clear upward trend.
  - No periodic pattern observed.
  - Can be modeled using a polynomial regression.

Mathematical Representation

A general model with trend:

$$X_t = m_t + Y_t$$

where:

- m_t is the trend component , a slowly changing function.
- Y_t is the zero-mean residual component .

Estimating the Trend Component

Least Squares Method:

- A common approach to estimate m_t.
- Fits a function of the form:

$$m_t = a_0 + a_1 t + a_2 t^2$$

- Parameters a_0, a_1, a_2 are chosen to minimize:

$$\sum_{t=1}^{n} (X_t - m_t)^2$$

Applying Least Squares Regression

- To fit a polynomial trend to the U.S. Population Data :
  1. Relabel time axis ($t = 1$ for 1790, $t = 21$ for 1990).
  2. Use a polynomial regression of order 2 .
  3. Estimate parameters using the Least Squares Method .
- This can be done using statistical software.

Trend Component

- Many time series require models with trend and seasonality .
- The Least Squares Method is a key tool for estimating trends.
- Choosing an appropriate model depends on the structure of the data .

A General Approach to Time Series Modeling

Key Steps in Time Series Analysis:

1. **Plot and Examine the Series**
    - Identify major features:
        - Presence of a **trend** (increasing/decreasing behavior).
        - Existence of a **seasonal component** (repeating patterns).
        - Any **sharp changes** in behavior.
        - Presence of **outliers** (unusual values).

Removing Trend and Seasonality

Steps for Achieving Stationarity:

- **Transform the Data** if fluctuations vary with level.
  - Example: Apply logarithm: $Y_t = \ln X_t$.
  - Ensure all values are positive before taking logarithms.
- **Trend and Seasonality Removal Techniques:**
  - Estimate and subtract trend and seasonal components.
  - Use **differencing**:

$$Y_t = X_t - X_{t-d}, \quad d \text{ is a chosen lag}$$

  - Aim: Obtain a stationary series (residuals).

Modeling the Residuals

Choosing a Model:

- Analyze the **sample autocorrelation function** (ACF).
- Identify the appropriate model:
    - AR (AutoRegressive).
    - MA (Moving Average).
    - ARMA / ARIMA (Integrated models).

Forecasting

Steps to Forecast Future Values:

- Forecast residuals using the chosen model.
- Invert transformations to obtain final forecasts:
  - Add back trend and seasonal components.
  - Reverse any transformations (e.g., exponentiate if logarithms were applied).

Fourier Transform in Time Series

Alternative Approach: Spectral Analysis

- Express series in terms of **Fourier components** (sinusoidal waves).
- Important for:
    - **Signal processing** (engineering applications).
    - **Structural design** (avoiding resonance issues).
- Helps analyze periodic behavior in data.

A General Approach to Time Series Modeling

- Time series modeling follows a structured approach.
- Removing trend and seasonality is key to obtaining stationary data.
- Residuals can be analyzed using statistical models.
- Fourier methods provide an alternative approach for frequency-based analysis.

Stationary Models and the Autocorrelation Function

Definition of Stationarity:

- A time series $\{X_t\}$ is said to be **stationary** if its statistical properties remain unchanged over time. That is, shifting the time series forward or backward does not alter its overall behavior.
- A time series $\{X_t\}$ is stationary if its statistical properties remain unchanged under time shifts.

Mean and Covariance Functions of a Time Series

Let $\{X_t\}$ be a time series with $E(X_t^2) < \infty$.

Mean Function

The **mean function** of $\{X_t\}$ is given by:

$$\mu_X(t) = E(X_t).$$

Covariance Function

The **covariance function** of $\{X_t\}$ is defined as:

$$\gamma_X(r, s) = \text{Cov}(X_r, X_s) = E[(X_r - \mu_X(r))(X_s - \mu_X(s))],$$

for all integers r and s.

Definition of Weak Stationarity

A time series $\{X_t\}$ is **weakly stationary** if:

1. **Constant Mean:** The expected value of the series is independent of time:

$$E(X_t) = \mu, \quad \forall t.$$

2. **Constant Covariance Structure:** The covariance between two points in time depends only on their time difference (lag h), not on their absolute positions:

$$\gamma_X(t, t+h) = \text{Cov}(X_t, X_{t+h}) = E[(X_t - \mu)(X_{t+h} - \mu)] = \gamma_X(h).$$

Importance of Stationarity

- Many statistical methods in time series analysis assume stationarity.
- Ensures that relationships within the data remain consistent over time.
- Helps in reliable forecasting and model interpretation.
- If a series is not stationary, transformations (e.g., differencing, detrending) are often applied.

Remark on Strict and Weak Stationarity

Remark 1

Strict stationarity of a time series $\{X_t\}$, where $t = 0, \pm 1, \pm 2, \ldots$, is defined by the condition that:

$$(X_1, X_2, \ldots, X_n) \quad \text{and} \quad (X_{1+h}, X_{2+h}, \ldots, X_{n+h})$$

have the **same joint distributions** for all integers h and $n > 0$.

Relation to Weak Stationarity

It is easy to check that if $\{X_t\}$ is **strictly stationary** and $E(X_t^2) < \infty$ for all t, then $\{X_t\}$ is also **weakly stationary**.

Unless stated otherwise, whenever we use the term **stationary**, we shall mean **weakly stationary**.

Strict vs. Weak Stationarity

Strict Stationarity:

- A time series $\{X_t\}$ is **strictly stationary** if the joint distribution of $(X_1, ..., X_n)$ is the same as $(X_{1+h}, ..., X_{n+h})$ for all h, n.
- This implies weak stationarity if $E[X_t^2] < \infty$.

Weak Stationarity:

- Only first- and second-order moments are invariant over time.
- Most practical applications assume weak stationarity.

Autocovariance Function (ACVF)

Definition:

- The autocovariance function of a stationary time series $\{X_t\}$ at lag h is:

$$\gamma_X(h) = \text{Cov}(X_{t+h}, X_t) = E[(X_{t+h} - \mu_X)(X_t - \mu_X)]$$

- $\gamma_X(h)$ depends only on h, not t.

Autocorrelation Function (ACF)

Definition:

- The autocorrelation function (ACF) at lag h is:

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$$

- Measures the strength of linear dependence between X_t and X_{t+h}.
- $-1 \leq \rho_X(h) \leq 1$.

Properties of Autocovariance and ACF

- $\gamma_X(0)$ is the variance of X_t, i.e., $\gamma_X(0) = \text{Var}(X_t)$.
- $\gamma_X(h)$ is symmetric:

$$\gamma_X(h) = \gamma_X(-h)$$

- ACF values closer to 1 indicate strong correlation.
- ACF helps in identifying patterns such as trends and seasonality.

Example: Linearity of Covariances

Property: For any real constants a, b, c, and random variables X, Y, Z,

$$\text{Cov}(aX + bY + c, Z) = a\text{Cov}(X, Z) + b\text{Cov}(Y, Z).$$

Example Application:

- If X_t follows a stationary process, then a linear combination of past values also follows a predictable pattern.

Autocorrelation Function

- Stationary models are essential for time series analysis.
- Weak stationarity ensures constant mean and time-invariant covariance.
- ACF provides insights into dependency structures.
- Understanding these concepts aids in model selection for forecasting.

Simple Time Series Models

- The terms **IID Noise** and **White Noise** are fundamental in time series analysis.
- While they share similarities, they have important differences in assumptions and properties.

IID Noise (Independent and Identically Distributed Noise)

The simplest model for a time series is one in which there is no trend or seasonal component and in which the observations are independent and identically distributed (IID) random variables with zero mean. We refer to such a sequence of random variables $X_1, X_2, \ldots$ as IID noise.

Mathematical Representation:

$P[X_1 \leq x_1, \ldots, X_n \leq x_n] = P[X_1 \leq x_1] \cdots P[X_n \leq x_n] = F(x_1) \cdots F(x_n)$, where $F(\cdot)$ is the cumulative distribution function of each identically distributed random variable.

Forecasting Property: $P[X_{n+h} \leq x | X_1 = x_1, \ldots, X_n = x_n] = P[X_{n+h} \leq x]$, showing that knowledge of past values does not help predict future values.

Notation: $\{X_t\} \sim \text{IID}(0, \sigma^2)$

IID Noise (Independent and Identically Distributed Noise)

A sequence $\{X_t\}$ is called **IID Noise** if:

1. **Independence:** The random variables X_t are independent for all $t \neq s$.
2. **Identically Distributed:** Each X_t follows the same probability distribution.
3. $E(X_t) = 0$
4. **Finite Second Moment:** $E(X_t^2) = \sigma^2 < \infty$.

Mathematical Representation:

$$\gamma_X(t+h,t) = \begin{cases} \sigma^2, & \text{if } h = 0, \\ 0, & \text{if } h \neq 0. \end{cases}$$

Notation:

$$\{X_t\} \sim \text{IID}(0, \sigma^2)$$

White Noise

A sequence $\{X_t\}$ is called **White Noise** if:

1. **Zero Mean:** $E(X_t) = 0$.
2. **Constant Variance:** $Var(X_t) = \sigma^2$.
3. **No Autocorrelation:** $E(X_t X_s) = 0$ for $t \neq s$ (uncorrelated but not necessarily independent).

Notation:

$$\{X_t\} \sim \text{WN}(0, \sigma^2)$$

Random Walk

A random walk is a stochastic process where the value at each step is determined by summing independent, identically distributed (IID) random variables.

Definition:

$$S_0 = 0, \quad S_t = X_1 + X_2 + \cdots + X_t, \quad \text{for } t = 1, 2, 3, \ldots$$

where X_t is IID noise.

Expected Value and Variance:

- $E(S_t) = 0$ if X_t has zero mean.
- $\mathrm{Var}(S_t) = t\sigma^2$, meaning the variance increases with time.

Autocovariance Function of a Random Walk

The autocovariance function of a random walk $\{S_t\}$ is given by:

Consider the autocovariance function:

$$\begin{aligned}\gamma_S(t+h,t) &= \text{Cov}(S_{t+h}, S_t) \\ &= \text{Cov}(S_t + X_{t+1} + \cdots + X_{t+h}, S_t)\end{aligned}$$

By the linearity of covariance,

$$\gamma_S(t+h,t) = \text{Cov}(S_t, S_t) + \sum_{i=1}^{h} \text{Cov}(X_{t+i}, S_t)$$

Autocovariance Function of a Random Walk

Since the increments $X_{t+1}, X_{t+2}, \ldots, X_{t+h}$ are independent of S_t, we have:

$$\text{Cov}(X_{t+i}, S_t) = 0 \quad \forall i.$$

Thus,

$$\gamma_S(t+h, t) = \text{Var}(S_t) = t\sigma^2.$$

Conclusion: Since $\gamma_S(t+h, t)$ depends on t, the process $\{S_t\}$ is **not stationary**.

Simple Symmetric Random Walk

A simple symmetric random walk is a special case where:

- X_t follows a Bernoulli distribution:

$$X_t = \begin{cases} +1, & \text{with probability } 0.5, \\ -1, & \text{with probability } 0.5. \end{cases}$$

- The position at time t is given by:

$$S_t = X_1 + X_2 + \cdots + X_t.$$

- This process models a pedestrian who starts at zero and moves left or right with equal probability at each step.

Key Differences Between IID Noise and White Noise

Property	IID Noise	White Noise
Independence	Yes, completely independent	No, only uncorrelated
Identically Distributed	Yes	Not necessarily
Autocorrelation	Zero for all lags	Zero for all lags
Finite Second Moment	Required	Not necessarily required
Example	Random samples from $N(0, \sigma^2)$	Residuals in a time series

- IID noise is a stronger condition than white noise because it implies independence.
- White noise only requires uncorrelated values, meaning there could be some hidden dependencies.
- If a white noise process is also independent and identically distributed, then it is IID noise.

Definition of MA(1) Process

The first-order moving average $MA(1)$ process is given by:

$$X_t = Z_t + \theta Z_{t-1}, \quad t = 0, \pm 1, \pm 2, \ldots \tag{1}$$

where:

- $\{Z_t\} \sim WN(0, \sigma^2)$ (white noise with mean zero and variance σ^2)
- θ is a real-valued constant.

This process is stationary if the second moment is finite.

Expectation and Variance of MA(1)

Expectation:

$$E(X_t) = E(Z_t) + \theta E(Z_{t-1}) = 0. \tag{2}$$

Variance:

$$\begin{aligned}
\mathrm{Var}(X_t) &= E(X_t^2) - (E(X_t))^2 && (3)\\
&= E[(Z_t + \theta Z_{t-1})^2] && (4)\\
&= E[Z_t^2 + 2\theta Z_t Z_{t-1} + \theta^2 Z_{t-1}^2] && (5)\\
&= E[Z_t^2] + 2\theta E[Z_t Z_{t-1}] + \theta^2 E[Z_{t-1}^2]. && (6)
\end{aligned}$$

Expectation and Variance of MA(1)

Since Z_t is white noise,

$$E[Z_t Z_{t-1}] = 0, \quad E[Z_t^2] = E[Z_{t-1}^2] = \sigma^2. \tag{7}$$

Therefore,

$$\text{Var}(X_t) = \sigma^2(1 + \theta^2) < \infty. \tag{8}$$

Autocovariance Function of MA(1)

Definition: The autocovariance function is given by

$$\gamma_X(h) = \mathrm{Cov}(X_t, X_{t+h}). \tag{9}$$

Calculation: For $h = 0$:

$$\gamma_X(0) = \mathrm{Var}(X_t) = \sigma^2(1 + \theta^2). \tag{10}$$

Autocovariance Function of MA(1)

For $h = \pm 1$:

$$\gamma_X(1) = E[(Z_t + \theta Z_{t-1})(Z_{t+1} + \theta Z_t)] \tag{11}$$

$$= E[Z_t Z_{t+1}] + \theta E[Z_{t-1} Z_{t+1}] + \theta E[Z_t^2] + \theta^2 E[Z_{t-1} Z_t]. \tag{12}$$

Since white noise has zero correlation except at lag zero,

$$\gamma_X(1) = \theta\sigma^2. \tag{13}$$

By symmetry:

$$\gamma_X(-1) = \theta\sigma^2. \tag{14}$$

Why $\gamma_X(h) = 0$ for $|h| > 1$?

- The MA(1) process depends only on Z_t and Z_{t-1}.
- For $|h| > 1$, the terms involved in X_t and X_{t+h} are independent.
- Since white noise terms are uncorrelated:

$$\gamma_X(h) = 0, \quad \text{for } |h| > 1. \tag{15}$$

Autocorrelation Function

The autocorrelation function (ACF) is:

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}. \tag{16}$$

For $h = 0$:

$$\rho_X(0) = \frac{\sigma^2(1+\theta^2)}{\sigma^2(1+\theta^2)} = 1. \tag{17}$$

For $h = \pm 1$:

$$\rho_X(1) = \frac{\gamma_X(1)}{\gamma_X(0)} = \frac{\theta\sigma^2}{\sigma^2(1+\theta^2)} = \frac{\theta}{1+\theta^2}. \tag{18}$$

For $|h| > 1$:

$$\rho_X(h) = 0. \tag{19}$$

MA(1) process

- The MA(1) process has a short memory.
- Its autocovariance function is nonzero only for $h = 0, \pm 1$.
- Beyond lag 1, there is no correlation because past values do not influence the present.
- This is a key property of moving average processes.

References

This presentation is adapted from:

- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to Linear Regression Analysis, Fifth Edition. Wiley.
- MTH 416 : Regression Analysis — Shalabh, IIT Kanpur

Thank You!

Thank you for your attention!