JOHNS HOPKINS
CAREY BUSINESS SCHOOL

BU.610.740: Forecasting Models for Business Intelligence
Chapter 1: Characteristics of Time Series

Ali Eshragh

Spring II, 2024

Topics Covered in These Slides

1. **Time Series Statistical Models**
   - Introduction
   - White Noise Series
2. **Measures of Dependence**
   - Mean Function
   - Autocorrelation Function
3. **Stationary Time Series**
   - Definition
   - Characteristics

Time Series Analysis

Definition (Time Series Analysis)

The **analysis** of experimental data collected at various points in **time** introduces unique challenges in statistical **modeling** and **inference** commonly referred to as **time series analysis**.

Conventional Statistical Methods vs. Time Series Analysis

Sampling adjacent time points creates a strong **correlation**, challenging the use of conventional **statistical methods** that assume observations are **independent** and **identically distributed** (**iid**).

Time Series

Definition (Time Series)

A collection of **random variables** $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ is referred to as a **time series**, where X_t represents the value of the series at **time** t.

- We should **distinguish** X_t, a **random variable**, from x_t, its **observed value**, referred to as a **realization** of X_t
- Here, a collection of **observed values** $\{x_t : t = 0, \pm 1, \pm 2, \ldots\}$ is also referred to as a **time series**.

White Noise

Definition (White Noise)

A sequence of **uncorrelated** random variables $\{W_t : t = 0, \pm 1, \pm 2, \ldots\}$, all with a mean of 0 and a finite variance σ_W^2, is known as **white noise**, denoted by $W_t \sim \mathtt{WN}[0, \sigma_W^2]$.

- The designation **white** originates from the analogy with **white light** which has the **same** intensity at every **frequency**.
- A particularly **useful** type of white noise series is **Gaussian white noise**, wherein the W_t are independent **Normal** random variables, $W_t \overset{\text{iid}}{\sim} \mathtt{N}[0, \sigma_W^2]$

Outline

Mean Function

Definition (Mean Function)

Let $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ be a **time series**. The **mean function** of this time series at time t is denoted by $\mu_X(t)$ and defined as

$$\mu_X(t) := \mathbb{E}[X_t] = \int_{-\infty}^{\infty} x f_{X_t}(x) dx$$

- Throughout, if **no confusion** arises regarding which time series is being discussed, we will **simplify** our notation by **removing** the subscript X across all instances, such as simplifying $\mu_X(t)$ to $\mu(t)$

Example

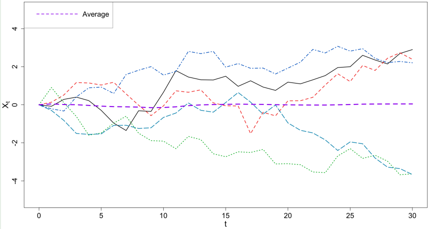
Example (Simple Ransom Walk)

A simple **random walk** model can explain **stock price** dynamics, where today's price equals yesterday's price plus random noise.

$$\begin{cases} X_t = X_{t-1} + W_t & \text{for } t = 1, 2, \dots \\ X_0 = 0 & \text{for } t = 0 \end{cases}$$

where $W_t \overset{\text{iid}}{\sim} \mathtt{N}[0, \sigma_W^2]$

1. **Show** that $X_t = \sum_{j=1}^{t} W_j$ for $t = 1, 2, \dots$
2. Find the **mean** function at time t



Autocovariance Function

Definition (Autocovariance Function)

Let $\{X_t : t = 0, \pm 1, \pm 2, \dots\}$ be a **time series**. The **autocovariance function** of this time series at times s and t is denoted by $\gamma_X(s,t)$ and defined as

$$\gamma_X(s,t) := \mathsf{Cov}(X_s, X_t) = \mathbb{E}[(X_s - \mu_X(s))\,(X_t - \mu_X(t))]$$

- It can be easily **observed** that $\gamma_X(t,t) = \mathsf{Var}(X_t)$

Autocorrelation Function

Definition (Autocorrelation Function)

Let $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ be a **time series**. The **autocorrelation function** (ACF) of this time series at times s and t is denoted by $\rho_X(s,t)$ and defined as

$$\rho_X(s,t) := \frac{\gamma_X(s,t)}{\sqrt{\gamma_X(s,s)\gamma_X(t,t)}}$$

- It can be easily **observed** that $\rho_X(t,t) = 1$
- The **autocorrelation** function is similar to **correlation**, but it measures the **linear relationship** of the same random variable at different **time points**, instead of between two different **random variables**.
- If the **magnitude** of $\rho_X(s,t)$ is nearly one for some time points $s < t$, then X_t can be approximately **predicted** from X_s through a **linear** regression model $X_t = \beta_0 + \beta_1 X_s + W_t$, where W_t is a **Gaussian white noise** series.

Cross-correlation Function

Definition (Cross-correlation Function)

Let $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ and $\{Y_t : t = 0, \pm 1, \pm 2, \ldots\}$ be two **time series**. The **cross-correlation function** (CCF) of these two time series at times s and t is denoted by $\rho_{X,Y}(s,t)$ and defined as

$$\rho_{X,Y}(s,t) := \frac{\gamma_{X,Y}(s,t)}{\sqrt{\gamma_X(s,s)\gamma_Y(t,t)}}$$

where $\gamma_{X,Y}(s,t) := \text{Cov}(X_s, Y_t)$ is the **cross-covariance function** between the two time series at times s and t.

- If the **magnitude** of $\rho_{X,Y}(s,t)$ is nearly one for some time points $s < t$, then Y_t can be approximately **predicted** from X_s through a **linear** regression model $Y_t = \beta_0 + \beta_1 X_s + W_t$, where W_t is a **Gaussian white noise** series.

Outline

Stationary Time Series

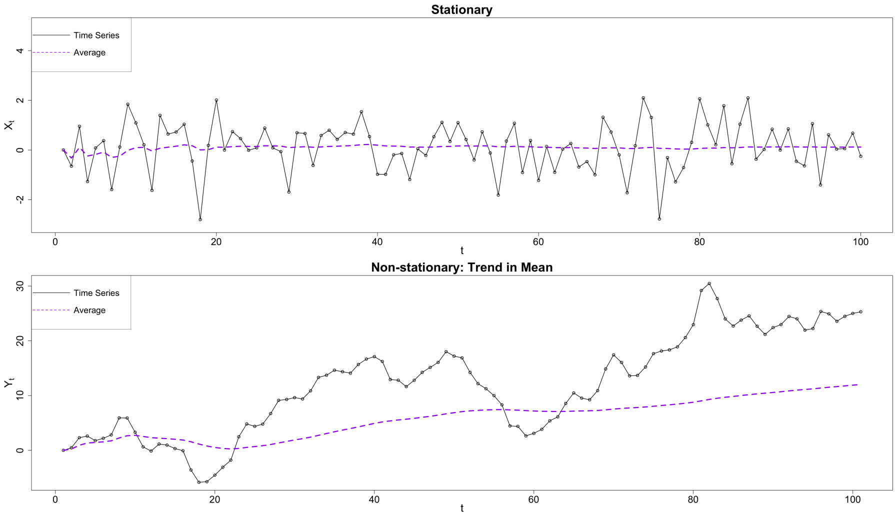
Definition (Stationary Time Series)

A time series $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ is **stationary** if:

1. its **mean** value function, $\mu_X(t)$, remains **constant** and independent of time, for all time t, and
2. its **autocovariance** function, $\gamma_X(s,t)$, is **solely** determined by the **lag** $|s-t|$ between time points, for all times s and t.

- In some textbooks, including your **prescribed** textbook, this type of stationarity is also called **weakly stationary**.

Stationary vs. Non-stationary Time Series



ACF of a Stationary Time Series

Definition (Autocovariance, Variance, and Autocorrelation)

Let $\{X_t : t = 0, \pm 1, \pm 2, \ldots\}$ be a **stationary** time series. Then:

- The **autocovariance function** at lag $h = 0, 1, 2, \ldots$ is defined as

$$\gamma_X(h) := \mathrm{Cov}(X_{t+h}, X_t) \quad \text{for any time } t$$

- The **variance** of the time series equals

$$\mathrm{Var}(X_t) = \gamma_X(0)$$

- The **autocorrelation function** at lag $h = 0, 1, 2, \ldots$ is defined as

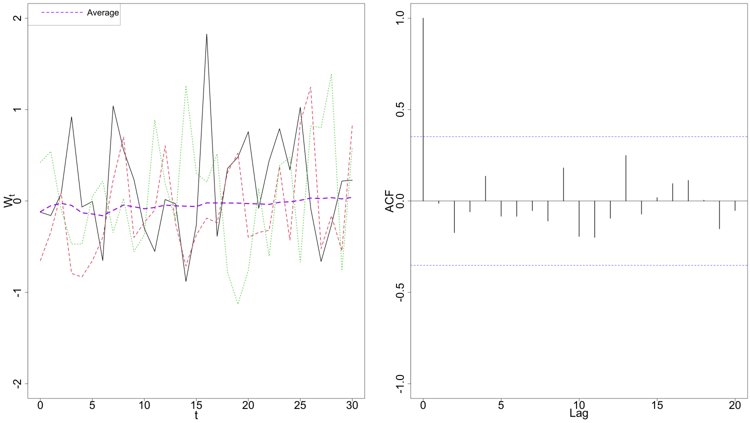
$$\rho_X(h) := \frac{\gamma_X(h)}{\gamma_X(0)}$$

  and $\rho_X(h) = \rho_X(-h)$ for $h = -1, -2, \ldots$

Example

Example (Gaussian White Noise Series)

Is the **Gaussian white noise** $\{W_t : t = 0, \pm 1, \pm 2, \ldots\}$ **stationary**?

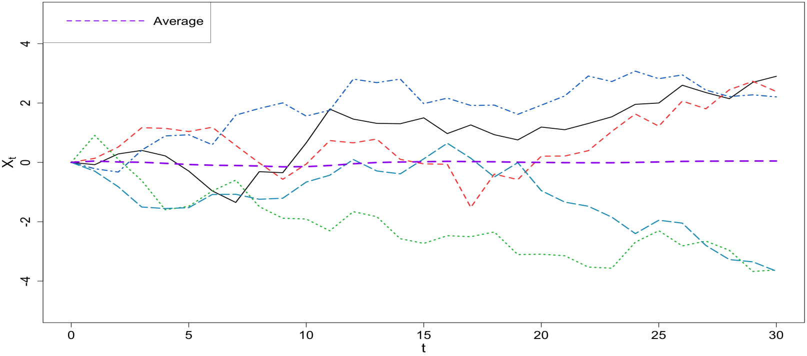


Example

Example (Simple Random Walk)

Is the simple **random walk** model, where $W_t \overset{\text{iid}}{\sim} \mathtt{N}[0, \sigma_W^2]$, **stationary**?

$$\begin{cases} X_t = X_{t-1} + W_t & \text{for } t = 1, 2, \ldots \\ X_0 = 0 & \text{for } t = 0 \end{cases}$$



Assignment 2

Question 1

Is the following time series model **stationary**? Justify your answer.

$$\begin{cases} X_t = X_{t-1} + W_t & \text{for } t = 1, 2, \ldots \\ X_0 = W_0 & \text{for } t = 0 \end{cases}$$

where $\{W_t : t = 0, 1, 2, \ldots\}$ is a **Gaussian white noise** series with mean zero and variance one.

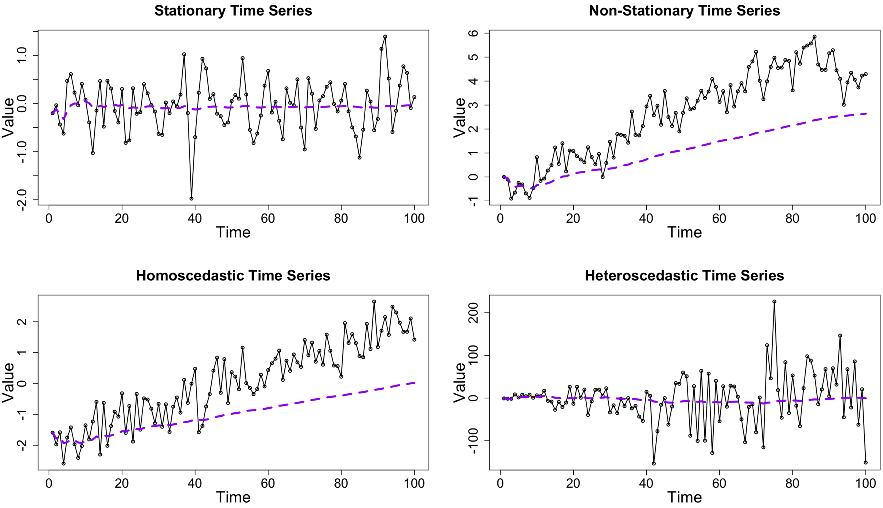
Homoskedastic vs. Heteroskedastic Time Series

Definition (Homoskedasticity and Heteroskedasticity)

A time series is called **homoskedastic** if its variance is **constant** over time, and **heteroskedastic** if its variance **changes** over time.

- Every **stationary** time series is homoskedastic, and any heteroskedastic time series is **non-stationary**.
- **Differentiating** between heteroskedasticity and homoskedasticity is **critical** for **model** selection and **method** application in time series analysis.
- Heteroskedasticity is common in **financial** time series data with **changing** volatility.

Trends in Mean and Variance



Statistical Hypothesis Test

- **Plotting** the time series data offers **insightful** information about the data's **nature**.
- However, making **inferences** should rely on conducting an appropriate **statistical test**.
- **KPSS** (Kwiatkowski-Phillips-Schmidt-Shin) test for checking **stationarity**:

$$\begin{cases} H_0 : \text{The time series is } \textbf{stationary} \\ H_a : \text{The time series is } \textbf{not} \text{ stationary} \end{cases}$$

Jointly Stationary Time Series

Definition (Jointly Stationary)

Two time series $\{X_t : t = 0, \pm 1, \pm 2, \dots\}$ and $\{Y_t : t = 0, \pm 1, \pm 2, \dots\}$ are **jointly stationary** if:

1. they are each **stationary**, and
2. their **cross-covariance** function, $\gamma_{X,Y}(h) := \mathrm{Cov}(X_{t+h}, Y_t)$, is **solely** determined by the **lag** h, for all time points t.

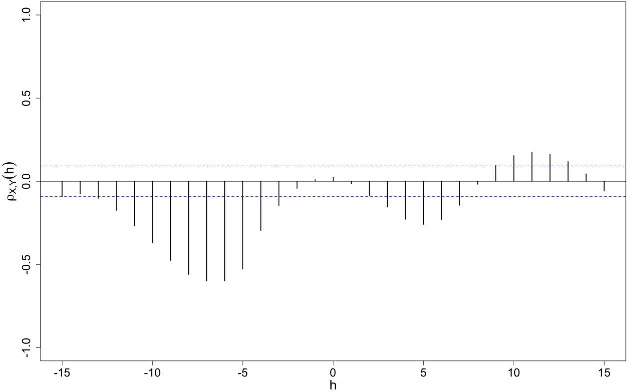
Definition (Cross-correlation Function)

The **cross-correlation function** of **jointly stationary** time series X_t and Y_t at lag $h = 0, 1, 2, \dots$ is defined as

$$\rho_{X,Y}(h) := \frac{\gamma_{X,Y}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}$$

and $\rho_{X,Y}(h) = \rho_{Y,X}(-h)$ for $h = -1, -2, \dots$

Case Study: SOI (X_t) vs. Recruitment (Y_t) Cross-correlation Analysis



References

1. R.H. Shumway and D.S. Stoffer, *Time Series Analysis and Its Applications With R Examples*, Springer, New York, 2017.