Lecture 3: Building Blocks for Time Series

Raul Riva

FGV EPGE

Invalid Date

Intro

Intro

Common data structures in Economics:

- Cross-sectional data: many units, one time period
  - Example: grades of several students in a given exam, GDP growth of several countries in a given year;
- Time series data: one unit, many time periods
  - Example: inflation over time for a given country, amount on rain in a given area, price of a stock over time;
- Panel data: many units, many time periods for the same units
  - Example: GDP growth of several countries over several years, grades of several students in several exams, prices of several products over time...
- Text, spatial data, images, etc.

Intro

- So far: mostly cross-sectional methods (everything about Y_i);
- Next step: time series methods (everything about Y_t);
- Next next step: panel data methods (everything about $Y_{i,t}$) – probably the most prevalent type nowadays;

Intro

- So far: mostly cross-sectional methods (everything about Y_i);
- Next step: time series methods (everything about Y_t);
- Next next step: panel data methods (everything about $Y_{i,t}$) – probably the most prevalent type nowadays;
- Next weeks: focus on understanding the challenges of time series data;
- This lecture: how is a time series different from cross-sectional data? + important definitions;
- Main reference: **Time Series Analysis**, by James Hamilton;
- Hansen's book provides a nice introduction, but it is too short on the topic;

What is a time series?

- A *time series* is a sequence of observations on a variable (or several variables) over time on an equally-spaced interval;
- Example: annual population of a country;
- Unlike cross-sectional data, time series data is **ordered**;
- Also assume in this course that time is discrete (i.e., we observe data at specific time intervals, like days, months, years...);
- Continuous time series (e.g., high-frequency financial data) is a more advanced topic, but with a vast literature as well;

Why should you care?

1. Policy evaluation: every policy takes place over time;
   - What's the impact of a reform? Before vs after? Are effects long-lasting? Fast die-outs?
   - A very important building block for panel data methods;
2. Forecasting: how can the future look like?
   - What will the inflation rate be next month? What is the expected number of COVID cases next week? How many students will enroll next semester?
   - How much inventory should a firm hold? What's the likely path of deforestation in a given area?
3. Nowcasting: high(er)-frequency monitoring of low-frequency variable;
   - What is the current state of the economy? How many people are currently unemployed? How many people are currently infected with COVID?

Why should you care?

1. Policy evaluation: every policy takes place over time;
   - What's the impact of a reform? Before vs after? Are effects long-lasting? Fast die-outs?
   - A very important building block for panel data methods;
2. Forecasting: how can the future look like?
   - What will the inflation rate be next month? What is the expected number of COVID cases next week? How many students will enroll next semester?
   - How much inventory should a firm hold? What's the likely path of deforestation in a given area?
3. Nowcasting: high(er)-frequency monitoring of low-frequency variable;
   - What is the current state of the economy? How many people are currently unemployed? How many people are currently infected with COVID?
4. It has very elegant math behind it!

The Role of Dependence

- In cross-sectional data, we typically assume that observations are i.i.d;
- If we relabel observations, we have no problems;

The Role of Dependence

- In cross-sectional data, we typically assume that observations are i.i.d;
- If we relabel observations, we have no problems;
- In time series, independence is a rare exception;
- Their distribution might also change over time;
- If we relabel observations, we change the meaning of the data;

The Role of Dependence

- In cross-sectional data, we typically assume that observations are i.i.d;
- If we relabel observations, we have no problems;
- In time series, independence is a rare exception;
- Their distribution might also change over time;
- If we relabel observations, we change the meaning of the data;

Examples:

- If inflation is high this month, it will likely be high next month;
- If a student is doing well on every exam, they will likely do well on the next one;
- The major difference w.r.t. cross-sectional data is that the future might depend on the past;

Building Blocks

Building Blocks

Formally, the right way to think about time series is as a **stochastic process**;

First, we define what a *random variable* is:

Definition (Random Variable)

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable is a real function $Y : \Omega \to \mathbb{R}$, such that for all $c \in \mathbb{R}$, $A_c = \{\omega \in \Omega | Y(\omega) \leq c\} \in \mathcal{F}$, $\forall\, c \in \mathbb{R}$.

- Ω is a sample space. Example: the numbers on a die (1, 2, 3, 4, 5, 6);
- $\mathcal{F}$ is a collection of events. Example: even numbers.
- $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure. Example: probability of rolling an even number;

Building Blocks

Formally, the right way to think about time series is as a **stochastic process**;

First, we define what a *random variable* is:

Definition (Random Variable)

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable is a real function $Y : \Omega \to \mathbb{R}$, such that for all $c \in \mathbb{R}$, $A_c = \{\omega \in \Omega | Y(\omega) \leq c\} \in \mathcal{F}$, $\forall\, c \in \mathbb{R}$.

- Ω is a sample space. Example: the numbers on a die (1, 2, 3, 4, 5, 6);
- $\mathcal{F}$ is a collection of events. Example: even numbers.
- $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure. Example: probability of rolling an even number;
- Think about an i.i.d sample as different realizations of ω;
- $Y_1 = Y(\omega_1), Y_2 = Y(\omega_2), \ldots$

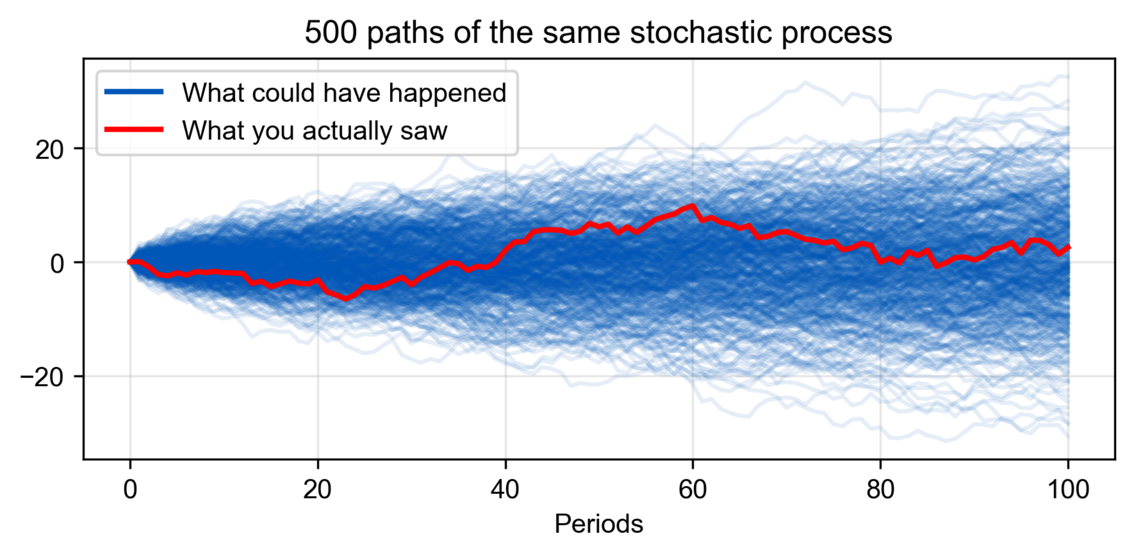
Stochastic Processes

Definition (Stochastic Process)

A *stochastic process* is an ordered sequence (collection) of random variables $\{Y_t(\omega), \omega \in \Omega, t \in \mathcal{T}\}$, such that for all $t \in \mathcal{T}$, $Y_t(\omega)$ **is a random variable** in Ω and $\mathcal{T}$ is an ordering set, for example, $\mathbb{Z} = \{-2, -1, 0, 1, 2, ...\}$.

Loosely speaking:

- A **random variable** is way to model an uncertain number;
- A **stochastic process** is a way to model an uncertain *path*;
- In reality, we only observe one, and only one, realization of the stochastic process;
- Think about the history of the world: it is a single realization of Ω;
- YOLO: you only live once!


500 paths of the same stochastic process
What could have happened
What you actually saw
20
0
−20
0
20
40
60
80
100
Periods

The Challenge Ahead

- You only observe the red, but the blue paths were equally likely to have happened;
- The main challenge is that you see *one* path and you want to make inference about the *whole process*;

The Challenge Ahead

- You only observe the red, but the blue paths were equally likely to have happened;
- The main challenge is that you see *one* path and you want to make inference about the *whole process*;

Example: inflation in Brazil in 1989;

- We know it was super high;
- But we have only one reading of inflation for 1989;
- When can we say something about the “inflation process” in Brazil given **only one** observation?

The Challenge Ahead

- You only observe the red, but the blue paths were equally likely to have happened;
- The main challenge is that you see *one* path and you want to make inference about the *whole process*;

Example: inflation in Brazil in 1989;

- We know it was super high;
- But we have only one reading of inflation for 1989;
- When can we say something about the "inflation process" in Brazil given **only one** observation?

We will need to impose a lot of structure! Without structure, we are lost!

Questions?

Stationarity and Ergodicity

Weak Stationarity

Definition (Weak Stationarity)

A stochastic process $\{Y_t\}$ is said to be *weakly stationary* (or *second-order stationary*, or *covariance stationary*) if, and only if, the first two population unconditional moments of $\{Y_t\}$ exists and are constant:

$$\mathbb{E}[Y_t] = \mu, \ |\mu| < \infty, \ \forall\, t \in \mathcal{T} \text{ and}$$
$$\mathbb{E}[(Y_t - \mu)(Y_{t-h} - \mu)] = \gamma_h, \ |\gamma_h| < \infty, \ \forall\, t \in \mathcal{T} \text{ and} \qquad h = 0, \pm 1, \pm 2, \ldots$$

Weak Stationarity

Definition (Weak Stationarity)

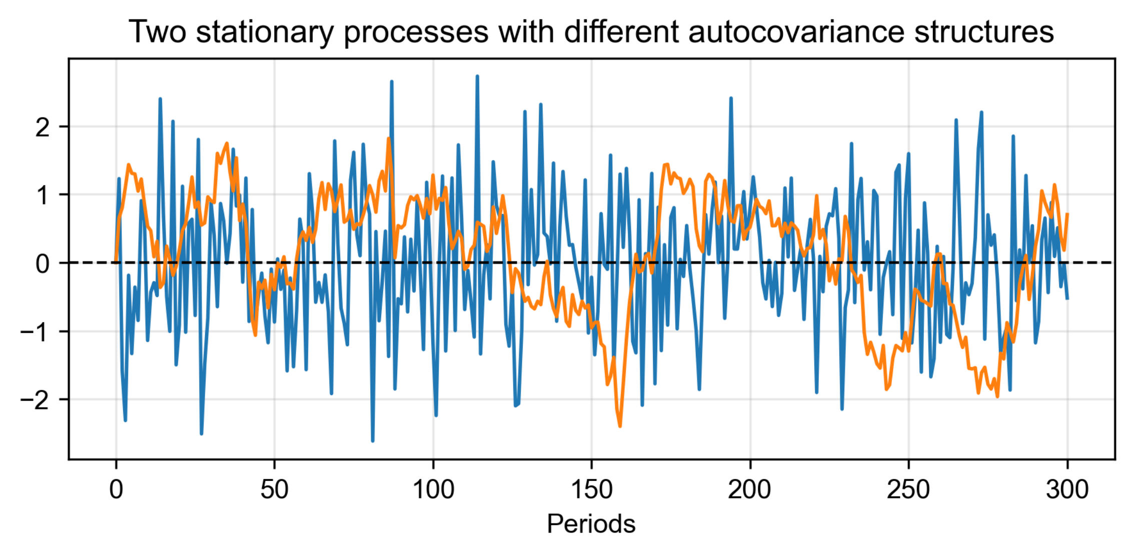
A stochastic process $\{Y_t\}$ is said to be *weakly stationary* (or *second-order stationary*, or *covariance stationary*) if, and only if, the first two population unconditional moments of $\{Y_t\}$ exists and are constant:

$$\mathbb{E}[Y_t] = \mu, \ |\mu| < \infty, \ \forall\, t \in \mathcal{T} \text{ and}$$
$$\mathbb{E}[(Y_t - \mu)(Y_{t-h} - \mu)] = \gamma_h, \ |\gamma_h| < \infty, \ \forall\, t \in \mathcal{T} \text{ and} \qquad h = 0, \pm 1, \pm 2, \ldots$$

True or false?

- $\gamma_0 = \text{Var}(Y_t)$ (True or False?)
- $\gamma_h > \gamma_{-h}$ for $h \neq 0$. (True or False?)
- An i.i.d process with finite variance is weakly stationary. (True of False?)

What process has a higher γ_1?



Strong (or Strict) Stationarity

Definition (Strong Stationarity)

A stochastic process $\{Y_t\}$ is said to be *strongly stationary* (ou *strictly stationary*) if, and only if, the joint distribution of $(Y_1, Y_2, \dots, Y_T)$ is invariant with respect to time shifts:

$$F_Y(Y_1, Y_2, \dots, Y_n) = F_Y(Y_{1+\tau}, Y_{2+\tau}, \dots, Y_{n+\tau}), \ \forall\, \tau$$

where $F_Y(.)$ is the joint CDF of the random vector $(Y_1, Y_2, \dots, Y_n)$.

Strong (or Strict) Stationarity

Definition (Strong Stationarity)

A stochastic process $\{Y_t\}$ is said to be *strongly stationary* (ou *strictly stationary*) if, and only if, the joint distribution of $(Y_1, Y_2, \ldots, Y_T)$ is invariant with respect to time shifts:

$$F_Y(Y_1, Y_2, \ldots, Y_n) = F_Y(Y_{1+\tau}, Y_{2+\tau}, \ldots, Y_{n+\tau}), \ \forall \tau$$

where $F_Y(.)$ is the joint CDF of the random vector $(Y_1, Y_2, \ldots, Y_n)$.

True or false?

- If a process is weakly stationary, it is also strongly stationary. (True or False?)
- If a process is strongly stationary, it is also weakly stationary. (True or False?)
- An i.i.d process with finite variance is strongly stationary. (True of False?)

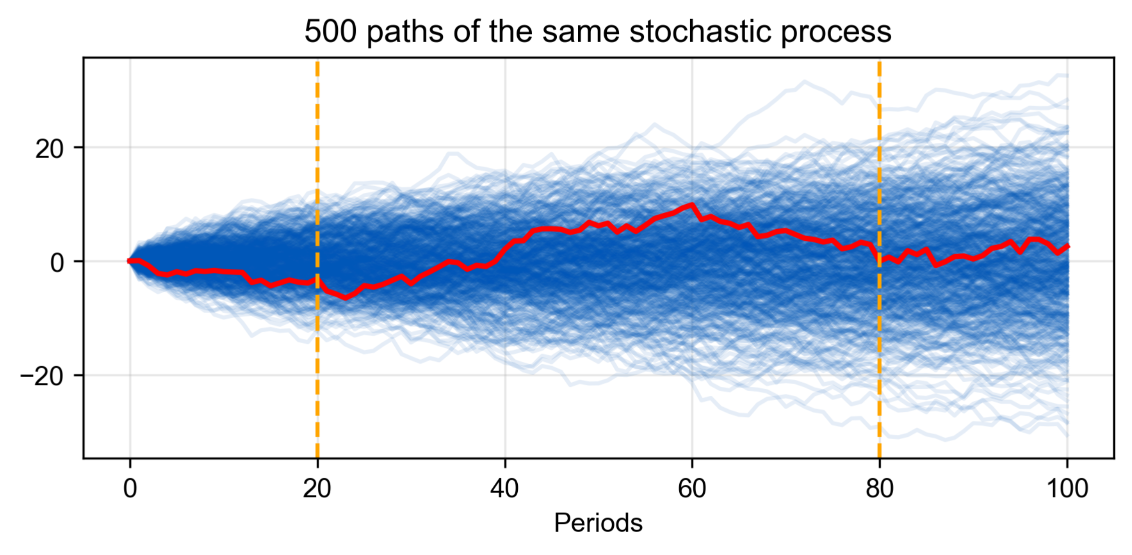
Ergodicity - Motivation

- Let's call the mysterious process from the YOLO simulation m_t;
- Denote by $m_t^{(i)}$ as the realization of m at time t for the i-th path;
- The paths are independent (because I chose so!);

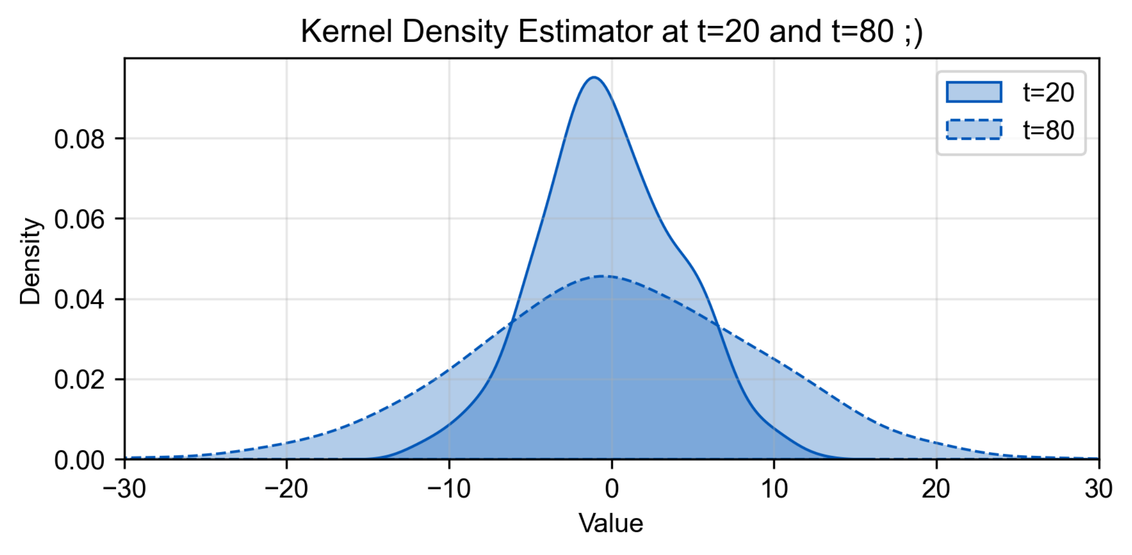
Ergodicity - Motivation

- Let's call the mysterious process from the YOLO simulation m_t;
- Denote by $m_t^{(i)}$ as the realization of m at time t for the i-th path;
- The paths are independent (because I chose so!);
- Let's say I want to estimate two (potentially different) quantities:
  - $\mathbb{E}[m_{20}]$
  - $\mathbb{E}[m_{80}]$;
- How can I do that?

Ergodicity - Motivation



Ergodicity - Motivation



Ergodicity - Motivation

Proposal 1: Estimate $\mathbb{E}[m_{20}]$ by averaging the values at $t = 20$ across all paths;

- Formally: $\hat{m}_{20} = \frac{1}{\text{number of paths (n)}} \sum_{i=1}^{n} m_{20}^{(i)}$;
- Do the same for $t = 80$;
- Would this yield a consistent estimator?

Ergodicity - Motivation

Proposal 1: Estimate $\mathbb{E}[m_{20}]$ by averaging the values at $t = 20$ across all paths;

- Formally: $\hat{m}_{20} = \frac{1}{\text{number of paths (n)}} \sum_{i=1}^{n} m_{20}^{(i)}$;
- Do the same for $t = 80$;
- Would this yield a consistent estimator?
- What's the main problem with this approach?

Ergodicity - Motivation

Proposal 1: Estimate $\mathbb{E}[m_{20}]$ by averaging the values at $t = 20$ across all paths;

- Formally: $\hat{m}_{20} = \frac{1}{\text{number of paths (n)}} \sum_{i=1}^{n} m_{20}^{(i)}$;
- Do the same for $t = 80$;
- Would this yield a consistent estimator?
- What's the main problem with this approach?

Proposal 2: Since I do not have access to the whole process, I will use the one path I have;

- Estimate $\mathbb{E}[m_{20}]$ by averaging the values at $t = 20$ across all time periods in the path;
- Formally: $\widetilde{m}_{20} = \frac{1}{T} \sum_{t=1}^{T} m_t^{(1)}$;
- Would I need stationarity for this to yield a consistent estimator?
- Intuitively, would that be enough?

Ergodicity - Definition

- Ergodicity is a property that some stochastic processes have;
- Intuitive definition (don't quote me on this):

Definition (Ergodicity - Intuitive)

A stochastic process $\{Y_t\}$ is said to be *ergodic* if its realized paths are "rich enough" with probability 1. By "rich enough", we mean that it **will not** get stuck in a subset of the state space or will get into cyclic trajectories with probability 1.

Ergodicity - Definition

- Ergodicity is a property that some stochastic processes have;
- Intuitive definition (don't quote me on this):

Definition (Ergodicity - Intuitive)

A stochastic process $\{Y_t\}$ is said to be *ergodic* if its realized paths are "rich enough" with probability 1. By "rich enough", we mean that it **will not** get stuck in a subset of the state space or will get into cyclic trajectories with probability 1.

Said in another way: if you observe a process for long enough T, you will be able to learn everything about the process;

Ergodicity vs Stationarity

- There are ways to test whether a process is stationary. We will get there;
- There is no way to test for ergodicity! You have to assume it!

Ergodicity vs Stationarity

- There are ways to test whether a process is stationary. We will get there;
- There is no way to test for ergodicity! You have to assume it!

Important:

- When a process is stationary, there are simple conditions that imply ergodicity;
- Under ergodicity and strict stationarity, $\widetilde{m}_{20}$ is consistent!
- This result is called the **Ergodic Theorem** (Theorem 14.9 on Hansen's book).

Ergodicity vs Stationarity

- There are ways to test whether a process is stationary. We will get there;
- There is no way to test for ergodicity! You have to assume it!

Important:

- When a process is stationary, there are simple conditions that imply ergodicity;
- Under ergodicity and strict stationarity, $\widetilde{m}_{20}$ is consistent!
- This result is called the **Ergodic Theorem** (Theorem 14.9 on Hansen's book).

Also important:

- Stationarity does not imply ergodicity: one example in the problem set;

Formal Definition - Ergodicity (Time Allowing)

Trivial Invariant Events

- Let $G \subset \mathbb{R}^{\infty}$;
- An event A is $A = \{\omega \in \Omega | \tilde{Y}_t \in G\}$, where $\tilde{Y}_t = (..., Y_{t-1}, Y_t, Y_{t+1}, ...)$ is the history of the process;
- The l-th time shift of A is $A_l = \{\omega \in \Omega | \tilde{Y}_{t+l} \in G\}$, where $\tilde{Y}_{t+l} = (..., Y_{t-1+l}, Y_{t+l}, Y_{t+1+l}, ...)$ is the history of the process shifted by l periods;

Trivial Invariant Events

- Let $G \subset \mathbb{R}^\infty$;
- An event A is $A = \{\omega \in \Omega | \tilde{Y}_t \in G\}$, where $\tilde{Y}_t = (..., Y_{t-1}, Y_t, Y_{t+1}, ...)$ is the history of the process;
- The l-th time shift of A is $A_l = \{\omega \in \Omega | \tilde{Y}_{t+l} \in G\}$, where $\tilde{Y}_{t+l} = (..., Y_{t-1+l}, Y_{t+l}, Y_{t+1+l}, ...)$ is the history of the process shifted by l periods;
- An event A is called **invariant** if $A_l = A$ for all $l \in \mathbb{Z}$;
- An event A is called **trivial** is $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.

Trivial Invariant Events

- Let $G \subset \mathbb{R}^\infty$;
- An event A is $A = \{\omega \in \Omega | \tilde{Y}_t \in G\}$, where $\tilde{Y}_t = (..., Y_{t-1}, Y_t, Y_{t+1}, ...)$ is the history of the process;
- The l-th time shift of A is $A_l = \{\omega \in \Omega | \tilde{Y}_{t+l} \in G\}$, where $\tilde{Y}_{t+l} = (..., Y_{t-1+l}, Y_{t+l}, Y_{t+1+l}, ...)$ is the history of the process shifted by l periods;
- An event A is called **invariant** if $A_l = A$ for all $l \in \mathbb{Z}$;
- An event A is called **trivial** is $\mathbb{P}(A) = 0$ or $\mathbb{P}(A) = 1$.

A process Y_t is called ***ergodic*** if every invariant event is trivial.

Example

- Consider the process $A = \{\max_{t \in \mathbb{Z}} Y_t \leq 0\}$;
- In words: the maximum value of the process is never negative;

Example

- Consider the process $A = \{\max_{t \in \mathbb{Z}} Y_t \leq 0\}$;
- In words: the maximum value of the process is never negative;
- For any l, $A_l = \{\max_{t \in \mathbb{Z}} Y_{t+l} \leq 0\}$;
- Is this event invariant?

Example

- Consider the process $A = \{\max_{t \in \mathbb{Z}} Y_t \leq 0\}$;
- In words: the maximum value of the process is never negative;
- For any l, $A_l = \{\max_{t \in \mathbb{Z}} Y_{t+l} \leq 0\}$;
- Is this event invariant?
- Suppose $Y_t = Z$, where $Z \sim U[-1, 1]$, for all $t \in \mathbb{Z}$;
- This means that the process is constant after Z is drawn;
- What's $\mathbb{P}(A)$ here?

Example

- Consider the process $A = \{\max_{t\in\mathbb{Z}} Y_t \leq 0\}$;
- In words: the maximum value of the process is never negative;
- For any l, $A_l = \{\max_{t\in\mathbb{Z}} Y_{t+l} \leq 0\}$;
- Is this event invariant?
- Suppose $Y_t = Z$, where $Z \sim U[-1, 1]$, for all $t \in \mathbb{Z}$;
- This means that the process is constant after Z is drawn;
- What's $\mathbb{P}(A)$ here?
- Is this process ergodic?

One useful characterizations of Ergodicity

Theorem (Ergodicity - Characterization)

A strictly stationary series $Y_t \in \mathbb{R}^m$ *is ergodic if, and only if, for all events* A *and* B

$$\lim_{n \to \infty} \frac{1}{n} \sum_{l=1}^{n} \mathbb{P}(A_l \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$$

- Intuition: as we go back in time, on average, events become nearly independent;

Two Useful Theorems

Stationarity + Ergodicity = LLN

- You will be interested in approximating means and seconds moments of a process;
- Stationarity + ergodicity is the "right set of assumptions" to do that ;

Theorem (Theorem 2 (page 203) in Hannan (1970))

If Y_t is strictly stationary and ergodic with $\mathbb{E}[|Y_t|] < \infty$, then

$$\bar{Y}_T = \frac{1}{T}\sum_{t=1}^{T} Y_t \xrightarrow{a.s.} \mathbb{E}[Y_t] \quad \text{as} \quad T \to \infty$$

Also, if $\mathbb{E}[Y_t^2] < \infty$, then

$$\hat{\gamma}_h = \frac{1}{T}\sum_{t=1}^{T-h} (Y_t - \bar{Y}_T)(Y_{t+h} - \bar{Y}_T) \xrightarrow{a.s.} \gamma_h \quad \text{as} \quad T \to \infty$$

What to do in practice?

- The last theorem is not ideal: it assumes something we cannot test;
- If we are only concerned with estimation, there is an easier way out:

Theorem (Theorem 6 (page 210) in Hannan (1970))

If Y_t is weakly stationary and $\sum_{i=1}^{\infty} |\gamma_i| < \infty$, then we have that

$$\bar{Y}_T = \frac{1}{T}\sum_{t=1}^{T} Y_t \xrightarrow{L_2} \mathbb{E}[Y_t] \quad \text{as} \quad T \to \infty$$

$$\hat{\gamma}_h = \frac{1}{T}\sum_{t=1}^{T-h} (Y_t - \bar{Y}_T)(Y_{t+h} - \bar{Y}_T) \xrightarrow{L_2} \gamma_h \quad \text{as} \quad T \to \infty$$

- Recall that L_2-convergence is stronger than convergence in probability, but weaker than almost-sure convergence;

Questions?

White Noise and Martingale Difference Sequences

White Noise

- In the cross-section context, it's commom to write $y_i = \alpha + \beta x_i + u_i$ where u_i is i.i.d;
- We will need more flexible assumptions on "error terms" now;
- Two very important weaker notions of "error term" are **white noise** processes and **martingale difference sequences**.

White Noise

- In the cross-section context, it's commom to write $y_i = \alpha + \beta x_i + u_i$ where u_i is i.i.d;
- We will need more flexible assumptions on "error terms" now;
- Two very important weaker notions of "error term" are **white noise** processes and **martingale difference sequences**.

Definition (White Noise Process)

A sequence of random variables $(\varepsilon_t)_{t\in\mathbb{Z}}$ is called a **white noise process** if:

1. $\mathbb{E}[\varepsilon_t] = 0$ for all t;
2. $\mathrm{Var}(\varepsilon_t) = \sigma^2 < \infty$ for all t;
3. $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for all $t \neq s$.

In words, a white noise process is a sequence of uncorrelated random variables with mean zero and constant variance.

White Noise

- In the cross-section context, it's commom to write $y_i = \alpha + \beta x_i + u_i$ where u_i is i.i.d;
- We will need more flexible assumptions on "error terms" now;
- Two very important weaker notions of "error term" are **white noise** processes and **martingale difference sequences**.

Definition (White Noise Process)

A sequence of random variables $(\varepsilon_t)_{t \in \mathbb{Z}}$ is called a **white noise process** if:

1. $\mathbb{E}[\varepsilon_t] = 0$ for all t;
2. $\mathrm{Var}(\varepsilon_t) = \sigma^2 < \infty$ for all t;
3. $\mathrm{Cov}(\varepsilon_t, \varepsilon_s) = 0$ for all $t \neq s$.

In words, a white noise process is a sequence of uncorrelated random variables with mean zero and constant variance.

Any i.i.d. sequence of variables with mean zero and finite variance is a white noise process.

Information Sets

- Conditioning on past information is common in economic models and in Econometrics;
- The correct mathemtical object to use here are σ-**fields** and **filtrations**;
- We will avoid measure-theoretic formalism right now;
- Intuitively: the σ-field (or -algebra) generated by $(Y_t, Y_{t-1}, Y_{t-2}, ...)$ is the collection of possible histories for this process up to t;
- We denote this in two equivalent ways: $\mathscr{F}_t = \sigma(Y_t, Y_{t-1}, Y_{t-2}, ...)$;
- In other areas of Economics, we tipically refer to σ-fields as *information sets*;
- $\mathscr{F}_t$ contains all information available up to time t;
- Example: $\mathbb{E}[Y_t|\mathscr{F}_{t-1}]$ is our best guess for Y_t conditional on information from this process, and only that, up to $t-1$;

Martingale Difference Sequences

Definition (Martingale Difference Sequence)

A sequence of random variables $(\varepsilon_t)_{t \in \mathbb{Z}}$ is called a **martingale difference sequence** (MDS) with respect to the information set $\mathcal{F}_{t-1}$ if:

1. ε_t is adapted to $\mathcal{F}_t$;
2. $\mathbb{E}[|\varepsilon_t|] < \infty$ for all t;
3. $\mathbb{E}[\varepsilon_t | \mathcal{F}_{t-1}] = 0$ almost surely for all t.

Martingale Difference Sequences

Definition (Martingale Difference Sequence)

A sequence of random variables $(\varepsilon_t)_{t\in\mathbb{Z}}$ is called a **martingale difference sequence** (MDS) with respect to the information set $\mathcal{F}_{t-1}$ if:

1. ε_t is adapted to $\mathcal{F}_t$;
2. $\mathbb{E}[|\varepsilon_t|] < \infty$ for all t;
3. $\mathbb{E}[\varepsilon_t|\mathcal{F}_{t-1}] = 0$ almost surely for all t.

True or false?

- Every MDS sequence has (unconditional) mean zero; (True of false?)
- $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t+h}) = 0$ for all $h \neq 0$. (True of false?)
- $\mathbb{E}[\varepsilon_{t+h}|\mathcal{F}_t] = 0$ for all $h > 0$. (True of false?)
- $\mathbb{E}[\varepsilon_{t-h}|\mathcal{F}_t] = 0$ for all $h > 0$. (True of false?)

Example

- Let u_t be an i.i.d sequence with mean zero and variance $\sigma^2 < \infty$.
- Define $v_t = u_t \cdot u_{t-1}$.
  - Is this process i.i.d? Is it a MDS with respect its natural filtration (information set)? Is it white noise?

Example

- Let u_t be an i.i.d sequence with mean zero and variance $\sigma^2 < \infty$.
- Define $v_t = u_t \cdot u_{t-1}$.
  - Is this process i.i.d? Is it a MDS with respect its natural filtration (information set)? Is it white noise?
- Now, define $m_t = u_t + u_{t-1} \cdot u_{t-2}$;
  - Is this process i.i.d? Is it a MDS with respect its natural filtration (information set)? Is it white noise?

The End

References

- Hannan, E. J. (1970). *Multiple Time Series*. Wiley.
- Chapters 3 and 4 from Hamilton's book;
- MDS sequences are covered in Chapter 7;
- Chapter 14 from Hansen's book;