A FIRST COURSE

IN

SDE

A FIRST COURSE

IN

SDE

MAT4500 Notebook

Prof. Sang Hu

The Chinese University of Hongkong, Shenzhen

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is from the MAT4001 in fall semester, 2018.

CUHK(SZ)

Notations and Conventions

$\mathbb{R}^n$	n-dimensional real space
$\mathbb{C}^n$	n-dimensional complex space
$\mathbb{R}^{m\times n}$	set of all $m \times n$ real-valued matrices
$\mathbb{C}^{m\times n}$	set of all $m \times n$ complex-valued matrices
x_i	ith entry of column vector $\boldsymbol{x}$
a_{ij}	(i,j)th entry of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	ith column of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i^{\mathrm{T}}$	ith row of matrix $\boldsymbol{A}$
$\mathbb{S}^n$	set of all $n \times n$ real symmetric matrices, i.e., $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $a_{ij} = a_{ji}$ for all i,j
$\mathbb{H}^n$	set of all $n \times n$ complex Hermitian matrices, i.e., $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ and $\bar{a}_{ij} = a_{ji}$ for all i,j
$\boldsymbol{A}^{\mathrm{T}}$	transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ means $b_{ji} = a_{ij}$ for all i,j
$\boldsymbol{A}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$ means $b_{ji} = \bar{a}_{ij}$ for all i,j
$\mathrm{trace}(\boldsymbol{A})$	sum of diagonal entries of square matrix $\boldsymbol{A}$
$\mathbf{1}$	A vector with all 1 entries
$\mathbf{0}$	either a vector of all zeros, or a matrix of all zeros
$\boldsymbol{e}_i$	a unit vector with the nonzero element at the ith entry
$\mathcal{C}(\boldsymbol{A})$	the column space of $\boldsymbol{A}$
$\mathcal{R}(\boldsymbol{A})$	the row space of $\boldsymbol{A}$
$\mathcal{N}(\boldsymbol{A})$	the null space of $\boldsymbol{A}$
$\mathrm{Proj}_{\mathcal{M}}(\boldsymbol{A})$	the projection of $\boldsymbol{A}$ onto the set $\mathcal{M}$

Chapter 1

Week1

1.1. Tuesday

1.1.1. Difference between ODE and SDE

We first discuss the difference between deterministic differential equations and stochastic ones by considering several real-life problems.

Problem 1: Population Growth Model. Consider the first-order ODE

$$\begin{cases} \dfrac{\mathrm{d}N(t)}{\mathrm{d}t} = a(t)N(t) \\ N(0) = N_0 \end{cases}$$

where $N(t)$ denotes the **size** of the population at time t; $a(t)$ is the given (deterministic) function describing the **rate** of growth of population at time t; and N_0 is a given constant.

If $a(t)$ is not completely known, e.g.,

$$a(t) = r(t) \cdot \text{noise, or } r(t) + \text{noise,}$$

with $r(t)$ being a deterministic function of t, and the "noise" term models something random. The question arises: How to *rigorously* describe the "noise" term and solve it?

Problem 2: Electric Circuit. Let $Q(t)$ denote the charge at time t in an electrical circuit, which admits the following ODE:

$$\begin{cases} LQ''(t) + RQ'(t) + \frac{1}{C}Q(t) = F(t), \\ \qquad Q(0) = Q_0, \quad Q'(0) = Q_0' \end{cases}$$

where L denotes the inductance, R denotes the resistance, C denotes the capacity, and $F(t)$ denotes the potential source.

Now consider the scenario where $F(t)$ is not completely known, e.g.,

$$F(t) = G(t) + \text{noise}$$

where $G(t)$ is deterministic. The question is how to solve the problem.

R The differential equations above involving non-deterministic coefficients are called the **stochastic differential equations** (SDEs). Clearly, the solution to an SDE should involve the randomness.

1.1.2. Applications of SDE

Now we discuss some applications of SDE shown in the finance area.

Problem 3: Optimal Stopping Problem. Suppose someone holds an asset (e.g., stock, house). He plans to sell it at some future time. Denote $X(t)$ as the price of the asset at time t, satisfying the following dynamics:

$$\frac{\mathrm{d}X(t)}{\mathrm{d}t} = rX(t) + \alpha X(t) \cdot \text{noise}$$

where r, α are given constants. The goal of this person is to maximize the expected selling price:

$$\sup_{\tau \geq 0} \quad \mathbb{E}[X(\tau)]$$

where the optimal solution τ^* is the optimal stopping time.

Problem 4: Portfolio Selection Problem. Suppose a person is interested in two types of assets:

- A risk-free asset which generates a deterministic return ρ, whose price $X_1(t)$ follows a deterministic dynamics

$$\frac{\mathrm{d}X_1(t)}{\mathrm{d}t} = \rho X_1(t),$$

- A risky asset whose price $X_2(t)$ satisfies the following SDE:

$$\frac{\mathrm{d}X_2(t)}{\mathrm{d}t} = \mu X_2(t) + \sigma X_2(t) \cdot \text{noise}$$

  where $\mu, \sigma > 0$ are given constants.

The policy of the investment is as follows. The wealth at time t is denoted as $v(t)$. This person decides to invest the fraction $u(t)$ of his wealth into the risky asset, with the remaining $1 - u(t)$ part to be invested into the safe asset. Suppose that the utility function for this person is $U(\cdot)$, and his goal is to maximize the expected total wealth at the terminal time T:

$$\max_{u(t), 0 \le t \le T} \mathbb{E}[U(v^u(T))]$$

where the decision variable is the portfolio function $u(t)$ along whole horizon $[0, T]$.

Problem 5: Option Pricing Problem. The financial derivates are products in the market whose value depends on the underlying asset. The European call option is a typical financial derivative. Suppose that the underlying asset is stock A, whose price at time t is $X(t)$. Then the call option gives the option holder the right (not the obligation) to buy one unit of stock A at a specified price (strike price) K at maturity date T. The task is to inference the fair price of the option at the current time. The formula for the price of the option is the following:

$$c_0 = \mathbb{E}[(X(T) - K)^+]$$

which is the famous Black-Sholes-Merton Formula.

1.1.3. Reviewing for Probability Space

Firstly, we review some basic concepts in real analysis.

Definition 1.1 [σ-Algebra] A set $\mathcal{F}$ containing subsets of Ω is called a σ-algebra if:

1. $\Omega \in \mathcal{F}$;
2. $\mathcal{F}$ is closed under complement, i.e., $A \in \mathcal{F}$ implies $\Omega \setminus A \in \mathcal{F}$;
3. $\mathcal{F}$ is closed under countably union operation, i.e., $A_i \in \mathcal{F}, i \geq 1$ implies $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition 1.2 [Probability Measure] A function $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ is called a **probability measure** on $(\Omega, \mathcal{F})$ if

- $\mathbb{P}(\Omega) = 1$;
- $\mathbb{P}(A) \geq 0, \forall A \in \mathcal{F}$;
- $\mathbb{P}$ is σ-additive, i.e., when $A_i \in \mathcal{F}, i \geq 1$ and $A_i \cap A_j = \emptyset, \forall i \neq j$,

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i).$$

where $\mathbb{P}(A)$ is called the **probability of the event** A.

Definition 1.3 [Probability Space] A probability space is a triplet $(\Omega, \mathcal{F}, \mathbb{P})$ defined as follows:

1. Ω denotes the **sample space**, and a point $\omega \in \Omega$ is called a sample point;
2. $\mathcal{F}$ is a σ-algebra of Ω, which is a collection of subsets in Ω. The element $A \in \mathcal{F}$ is called an "event"; and
3. $\mathbb{P}$ is a probability measure defined in the space $(\Omega, \mathcal{F})$.

Definition 1.4 [Almost Surely True] A statement S is said to be **almost surely (a.s.) true** or **true with probability 1**, if

- $\mathfrak{B} := \{w : S(w) \text{ is true}\} \in \mathcal{F}$
- $\mathbb{P}(F) = 1$.

■

Definition 1.5 [Topological Space] A **topological space** $(X, \mathcal{T})$ consists of a (non-empty) set X, and a family of subsets of X ("open sets" $\mathcal{T}$) such that

1. $\emptyset, X \in \mathcal{T}$
2. $U, V \in \mathcal{T}$ implies $U \bigcap V \in \mathcal{T}$
3. If $U_\alpha \in \mathcal{T}$ for all $\alpha \in \mathcal{A}$, then $\bigcup_{\alpha \in \mathcal{A}} U_\alpha \in \mathcal{T}$.

When $A \in \mathcal{T}$, A is called the open subset of X. The $\mathcal{T}$ is called a **topology** on X. ■

Definition 1.6 [Borel σ-Algebra] Consider a topological space Ω, with $\mathcal{U}$ being the topology of Ω. The **Borel σ-Algebra** $\mathcal{B}(\Omega)$ on Ω is defined to be the *minimal* σ-algebra containing $\mathcal{U}$:

$$\mathcal{B}(\Omega) \triangleq \sigma(\mathcal{U}).$$

Any element $B \in \mathcal{B}(\Omega)$ is called the **Borel set**. ■

Definition 1.7 [$\mathcal{F}$-Measurable / Random Variable]

1. A function $f : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is called $\mathcal{F}$-**measurable** if

$$f^{-1}(\boldsymbol{B}) = \{w \mid f(w) \in \mathcal{B}\} \in \mathcal{F},$$

   for any $\boldsymbol{B} \in \mathcal{B}(\mathbb{R}^n)$.
2. A random variable X is a function $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ and is $\mathcal{F}$-measurable.

■

Definition 1.8 [Generated σ-Algebra] Suppose X is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. Then the σ-algebra generated by X, say $\mathcal{H}_X$ is defined to be the **minimal σ-algebra** on Ω to make X measurable.

Proposition 1.1 $\mathcal{H}_X = \{X^{-1}(\boldsymbol{B}) : \boldsymbol{B} \in \mathcal{B}(\mathbb{R}^n)\}$.

Proof. Since X is $\mathcal{H}_X$-measurable, for any $\boldsymbol{B} \in \mathcal{B}(\mathbb{R}^n)$, $X^{-1}(\boldsymbol{B}) \in \mathcal{H}_X$. Thus $\mathcal{H}_X \supseteq \{X^{-1}(\boldsymbol{B}) : \boldsymbol{B} \in \mathcal{B}(\mathbb{R}^n)\}$. It suffices to show that $\{X^{-1}(\boldsymbol{B}) : \boldsymbol{B} \in \mathcal{B}(\mathbb{R}^n)\}$ is a σ-algebra to finish the proof, which is true since $\mathcal{B}(\mathbb{R}^n) = \sigma(\mathcal{U})$, with $\mathcal{U}$ being the topology of X. ∎

1.2. Thursday

Reviewing for Probability Space.

- $(\Omega, \mathcal{F}, \mathbb{P})$;
- Random variable;
- Generated σ-algebra;

1.2.1. More on Probability Theory

Definition 1.9 [Distribution] A probability measure μ_X on $\mathbb{R}^n$ induced by the random variabe X isdefined as

$$\mu_X(\boldsymbol{B}) = \mathbb{P}(X^{-1}(\boldsymbol{B})),$$

where $\boldsymbol{B} \in \mathcal{B}(\mathbb{R}^n)$. The μ_X is called the **distribution** of X.

Definition 1.10 [Expectation] The expectation of X is given by

$$\mathbb{E}[X] = \int_\Omega X(\omega)\,\mathrm{d}\mathbb{P}(\omega)$$

When $\Omega = \mathbb{R}^n$, the expectation can be written in terms of distribution function:

$$\mathbb{E}[X] = \int_{\mathbb{R}^n} y\,\mathrm{d}\mu_X(y)$$

Note that the expectation of the random variable X is well-defined when X is integrable:

Definition 1.11 [Integrable] The random variable X is **integrable**, if

$$\int_\Omega |X(w)|\,\mathrm{d}\mathbb{P}(w) < \infty.$$

In other words, X is said to be $\mathcal{L}^1$-integrable, denoted as $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$.

Example 1.1 If $f : \mathbb{R}^n \to \mathbb{R}$ is Borel measurable, and $\int_\Omega |f(X(\omega))| \, d\mathbb{P}(\omega) < \infty$, then

$$\mathbb{E}[f(x)] = \int_\Omega f(X(\omega)) \, d\mathbb{P}(\omega) = \int_{\mathbb{R}^n} f(y) \, d\mu_X(y).$$

Definition 1.12 [L^p space] Suppose $X : \Omega \to \mathbb{R}$ is a random variable and $p \geq 1$.

- Define L^p-norm of X as

$$\|X\|_p = \left(\int_\Omega |X(\omega)|^p \, d\mathbb{P} \right)^{1/p}$$

If $p = \infty$, define

$$\|X\|_\infty = \inf\{N \in \mathbb{R} \mid |X(w)| \leq N, \text{ a.s.}\}$$

- A random variable X is said to be in the L^p space (p-th integrable) if

$$\int_\Omega |X(\omega)|^p \, d\mathbb{P}(\omega) < \infty,$$

denoted as $X \in \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$.

Proposition 1.2 If $p \geq q$, then $\|X\|_p \geq \|X\|_q$. Thus $\mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P}) \subseteq \mathcal{L}^q(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. The inequality is shown by using Holder's inequality:

$$\|X\|_q^q = \int_\Omega |X|^q \, d\mathbb{P} \leq \left(\int_\Omega (|X|^q)^{p/q} \, d\mathbb{P} \right)^{q/p} = \left(\int_\Omega |X|^p \, d\mathbb{P} \right)^{\frac{1}{p} \cdot q} = \|X\|_p^q.$$

∎

Then we discuss how to define independence between two random variables, by the following three steps:

Definition 1.13 [Independence]

1. Two events $A_1, A_2 \in \mathcal{F}$ are said to be **independent** if $\mathbb{P}(A_1 \bigcap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2)$.
2. Two σ-algebras $\mathcal{F}_1, \mathcal{F}_2$ are said to be **independent** if F_1, F_2 are independent events for $\forall F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2$
3. Two random variables X, Y are said to be **independent** if $\mathcal{H}_X, \mathcal{H}_Y$, the σ-algebra generated by X and Y, respectively, are independent.

R The independence defined above can be generalized from two events into finite number of events.

Proposition 1.3 If X and Y are two independent random variables, and $\mathbb{E}[|X|] < \infty, \mathbb{E}[|Y|] < \infty$, then

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] < \infty.$$

Proof. The first step is to simplify the probability distribution for the product random variable (X, Y), i.e., $\mu_{X,Y}$.

R From now on, we also write the event $\{X^{-1}(\boldsymbol{B})\}$ as $\{X \in \boldsymbol{B}\}$ for $\boldsymbol{B} \in \mathcal{B}(\mathbb{R}^n)$.

By the definition of independence, we have the following:

$$\begin{aligned}
\mu_{X,Y}(A_1 \times A_2) &\triangleq \mathbb{P}(\{(X,Y) \in (A_1 \times A_2)\}) = \mathbb{P}(\{X \in A_1, Y \in A_2\}) \\
&= \mathbb{P}(\{X \in A_1\})\mathbb{P}(\{Y \in A_2\}) = \mu_X(A_1)\mu_Y(A_2).
\end{aligned}$$

Now we begin to simplify the expectation of product:

$$\begin{aligned}
\mathbb{E}[XY] &= \int xy \, \mathrm{d}\mu_{X,Y}(x,y) = \iint xy \, \mathrm{d}\mu_X(x)\mu_Y(y) \\
&= \int y \left[\int x \, \mathrm{d}\mu_X(x) \right] \mu_Y(y) = \int \mathbb{E}[X] y \mu_Y(y) = \mathbb{E}[X]\mathbb{E}[Y].
\end{aligned}$$

■

1.2.2. Stochastic Process

Consider a set T of time index, e.g., a non-negative integer set or a time interval $[0,\infty)$. We will discuss a discrete/continuous time stochastic process.

Definition 1.14 [Stochastic Process] A collection of random variables $\{X_t\}_{t\in T}$, defined on $(\Omega,\mathcal{F},\mathbb{P})$ and taking values in $\mathbb{R}^n$, is called a **stochastic process**. ■

R A stochastic process $\{X_t\}_{t\in T}$ can also be viewed as a random function, since it is a mapping $\Omega\times T\to\mathbb{R}^n$. Sometimes we omit the subscript to denote a stochastic process $\{X_t\}$.

Definition 1.15 [Sample Path] Fixing $\omega\in\Omega$, then $\{X_t(\omega)\}_{t\in T}$ (denoted as $X.(\omega)$) is called a **sample path**, or **trajectory**. ■

Definition 1.16 [Continuous] A stochastic process $\{X_t\}$ is said to be **continuous** (right-cot, left-cot, resp.) a.s., if $t\to X_t(\omega)$ is **continuous** (right-cot, left-cot, resp.) a.s., i.e.,

$$\mathbb{P}\Big(\{\omega : t\to X_t(\omega)\text{ is continuous (right-cot, left-cot, resp.)}\}\Big)=1.$$

■

■ **Example 1.2** [Poisson Process] Consider $(\xi_j, j=1,2,\dots)$ a sequence of i.i.d. random variables with Possion distribution with intensity $\lambda>0$. Let $T_0=0$, and $T_n=\sum_{j=1}^n\xi_j$. Define $X_t=n$ if $T_n\le t<T_{n+1}$. Verify that $\{X_t\}$ is a stochastic process with right-continuity and left-limit exists. Instead of giving a mathematical proof, we provide a numerical simulation of $\{X_t\}$ plotted in Figure. 1.1. [a] ■

[a] The corresponding matlab code can be found in

https://github.com/WalterBabyRudin/Courseware/tree/master/MAT4500/week1

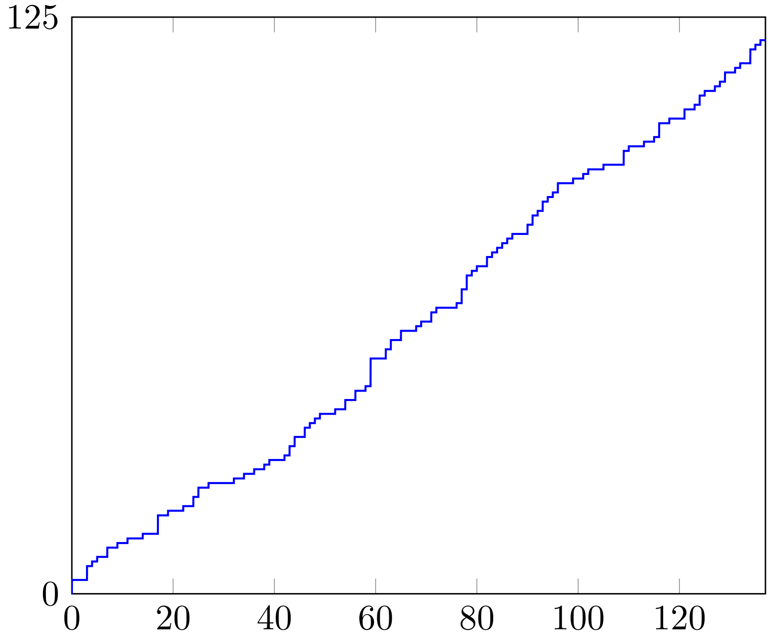

Figure 1.1: One simulation of $\{X_t\}$ with intensity $\lambda = 1.2$ and 500 samples

Chapter 2

Week2

2.1. Tuesday

R One can generalize the "independence" of two events / σ-algebras / random variables to a countable collection.

■ **Example 2.1** If two random variables $X, Y : \Omega \to \mathbb{R}$ are independent, provided that X and Y are integrable, i.e.,

$$\mathbb{E}|X| < \infty, \quad \mathbb{E}|Y| < \infty,$$

then we have $\mathbb{E}[XY] = \mathbb{E}|X|\mathbb{E}|Y|$ ■

2.1.1. Stochastic Process

Stochastic processes are mathematical models that are used to describe random phenomena evolving in the time.

Therefore, we need to have a time set T as parameters. T can be **non-negative** integers $\mathbb{Z}_+$ (i.e., discrete process), or be $[0, \infty)$ (i.e., continuous process).

Definition 2.1 [Stochastic Process] A stochastic process $\{x_t\}_{t \in T}$ is a collection of parameterized random variables, which is defined on $(\Omega, \mathcal{F}, \mathbb{P})$, and taking values in $\mathbb{R}^n$. ■

R A stochastic process $\{X_t\}$ can be viewed as a function: $T \times \Omega \to \mathbb{R}^n$.

Definition 2.2 1. For fixed $t \in T$, the function $\omega \to X_t(\omega)$ is a random variable.

2. For fixed $\omega \in \Omega$,

$$t \mapsto X_t(\omega)$$

is called a sample path.

3. The stochastic process $\{X_t\}_{t\in T}$ is **continuous** (resp. right-continuous, right-continuous with left limit) if the sample path $t \mapsto X_t(\omega)$ are continuous (resp. right-continuous, right-continuous with left limit), almost surely.

R A function $f : (a,b) \to \mathbb{R}$ is called **right-continuous** at $t_0 \in (a,b)$ if the right limit of f exists at t_0 and equal to $f(t_0)$.

The function f is called **right-continuous with left limit exists** at $t_0 \in (a,b)$ if the limit of f exists at t_0 , and the left limit of f exists at t_0.

Example 2.2 [Possion Process] Let $\{\xi_j\}_{j\geq 1}$ be i.i.d. random variables with possion distribution with intensity λ. Let $T_0 = 0$ and $T_n = \sum_{j=1}^{n} \xi_j$. For each $t \geq 0$, define $x_t = n$ if $T_n \leq t < T_{n+1}$.

Then for fixed ω, the function $t \to X_t(\omega)$ is a step function, with jump possibly at random time T_n, and is right-continuous with left-limit exists at T_n.

Definition 2.3 If $\{X_t\}_{t\in T}$ is a stochastic process taking values in $\mathbb{R}^n$, the joint distribution of $x_{t_1}, x_{t_2}, \ldots, x_{t_k}$ for given $t_1, t_2, \ldots, t_k \in T$,

$$\mu_{t_1,t_2,\ldots,t_k}(F_1 \times \cdots \times F_k) = \mathbb{P}(x_{t_1} \in F_1, \cdots, x_{t_k} \in F_k),$$

where $F_1, \ldots, F_k$ are Borel sets in $\mathbb{R}^n$. This $\mu_{t_1,t_2,\ldots,t_k}$ is also called **finite-dimensional distribution**, which is a probability measure on $\mathbb{R}^n \times \cdots \times \mathbb{R}^n$

■ **Example 2.3** [Brownian Motion] Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, for fixed $x \in \mathbb{R}$, define

$$\mathbb{P}(y; t, x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(y-x)^2}{2t}, y \in \mathbb{R}, t > 0\right)$$

The Brownian motion $\{B_t\}_{t \geq 0}$ is a continuous stochastic process, and the distribution of $\{B_t\}$ with $t_0 = 0, B_0 = x$ is given by:

$$\begin{aligned}
&\mathbb{P}(B_{t_1} \in F_1, \ldots, B_{t_k} \in F_k) = \\
&\quad \int_{F_1 \times \cdots \times F_k} p(t_1, x, x_1) p(t_2 - t_1, x_1, x_2) \cdots p(t_k - t_{k-1}, x_{k-1}, x_k) \, \mathrm{d}(x_1, \ldots, x_k)
\end{aligned}$$

■

R If $\{x_t\}_{t \geq 0}$ is a continuous stochastic process and we need to deal with such kind of set below

$$F = \{\omega \mid x_t(\omega) \in [0, 1), \forall t \leq 1\}$$

Such set F may be not measurable, i.e., F may not be an event. Then $\mathbb{P}(F)$ does not make sense. Therefore, we need the additional conditions.

Exercise. Let $\{x_t\}_{t \geq 0}$ be a stochastic process on $(\Omega, \mathcal{F}, \mathbb{P})$ and take values in $\mathbb{R}^n$. Let $\boldsymbol{B}$ be a Borel subset. If T is a countable set (or can be finite), then

$$\{\omega \mid x_t(\omega) \in \boldsymbol{B}, \forall t \in T\} \text{ is measurable} = \bigcap_{t \in T} \{\omega : X_t(\omega) \in \boldsymbol{B}\}.$$

We can also show that $\sup_{t \in T} |x_t(\omega)|$ is measurable.

Definition 2.4 Let $\{x_t\}_{t \geq 0}$ and $\{Y_t\}_{t \geq 0}$ be two stochastic process. We call $\{Y_t\}$ be a version of $\{x_t\}$ if for every $t \geq 0$,

$$\mathbb{P}\{w \mid x_t(\omega) = Y_t(\omega)\} = 1,$$

then $\{x_t\}$ and $\{Y_t\}$ are also called to be equivalent. ■

R If $\{Y_t\}_{t \geq 0}$ is a version of $\{x_t\}_{t \geq 0}$, then they have the same joint distribution.

Chapter 3

Week3

3.1. Tuesday

3.1.1. Uniform Integrability

Definition 3.1 [L_1-convergence] We say $f_n \to f$ in L^1 if

$$\lim_{n\to\infty} \int_S |f_n - f| \, \mathrm{d}\mu = 0$$

The **uniform integrability** for a family of integrable random variables is used to handle the convergence of random variables in L^1.

Proposition 3.1 If a random variable X is integrable, i.e., $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, then for any given $\varepsilon > 0$, there exists $\delta > 0$ such that for any $F \in \mathcal{F}$ with $\mathbb{P}(F) < \delta$, we have

$$\mathbb{E}[|X|; F] := \mathbb{E}[|X| 1_F] = \int_F |X| \, \mathrm{d}\mathbb{P} < \varepsilon$$

Proof. Suppose the conclusion is false, then there exists some $\varepsilon_0 > 0$, and a sequence of $\{F_n\}$ with each $F_n \in \mathcal{F}$ such that

$$\mathbb{P}(F_n) < \frac{1}{2^n}, \qquad \mathbb{E}[|X|; F_n] \geq \varepsilon_0.$$

Let $H := \lim_{n\to\infty} \sup F_n$. Note that $\sum_n \mathbb{P}(F_n) < \sum \frac{1}{2^n} < \infty$.

By applying the Borel-Centelli lemma, we have $\mathbb{P}(H) = 0$.

However, with the reverse fatou's lemma, since $1_H(w) = \lim_{n\to\infty} \sup 1_{F_n}(w)$,

$$\int |X| 1_H \, d\mathbb{P} \geq \limsup \int |X| 1_{F_n} \, d\mathbb{P}$$

since $\{|X| 1_{F_n}\}$ is dominated by the integrable random variable $|X|$.

Therefore,

$$\mathbb{E}[|X|; H] \geq \limsup \mathbb{E}[|X|; F_n] \geq \varepsilon_0$$

which contradicts with $\mathbb{P}(H) = 0$. ■

Corollary 3.1 Suppose $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then for any given $\varepsilon > 0$, there exists $K \geq 0$, such that $\mathbb{E}[|X|; |X| > K] := \int_{|X|>K} |X| \, d\mathbb{P} < \varepsilon$.

Proof. Note that

$$\begin{aligned}
\mathbb{E}[|X|] &= \mathbb{E}[|X|; |X| > K] + \mathbb{E}[|X|; |X| \leq K] \\
&\geq \mathbb{E}[K; |X| > K] = K\mathbb{E}[1_{|X|>K}] \\
&= K\mathbb{P}(|X| > K)
\end{aligned}$$

Therefore, we imply

$$\mathbb{P}(|X| > K) \leq \frac{\mathbb{E}|X|}{K}$$

Applying Proposition (3.1), we choose K large enough such that $\frac{\mathbb{E}|X|}{K} < \delta$.

Therefore, $\mathbb{P}(|X| > K) < \delta$, which implies

$$\int_{|X|>K} |X| \, d\mathbb{P} < \varepsilon.$$

■

Definition 3.2 A class $\mathcal{C}$ of random variables are called **uniform integrable** if and only if for any given $\varepsilon > 0$, there exists $K \geq 0$ such that

$$\mathbb{E}[|X|;|X| > K] < \varepsilon, \qquad \forall X \in \mathcal{C}$$

Note that for such uniform integrable class $\mathcal{C}$, we choose $\varepsilon_1 = 1$, then there exists $K_1 \geq 0$ such that

$$\begin{aligned}\forall X \in \mathcal{C},\ \mathbb{E}[|X|] &= \mathbb{E}[|X|;|X| > K_1] + \mathbb{E}[|X|;X \leq K_1] \\ &\leq \varepsilon_1 + K_1 = 1 + K_1,\end{aligned}$$

i.e., class $\mathcal{C}$ is uniformly bounded in L^1.

The reverse is not true:

Example 3.1 Take $(\Omega, \mathcal{F}, \mathbb{P}) = ([0,1], \mathcal{B}[0,1], \text{Leb})$

Let $E_n := (0, \frac{1}{n})$, and set

$$X_n(\omega) = n1_{E_n}(\omega) = \begin{cases} n, & \text{if } \omega \in E_n \\ 0, & \text{if } \omega \notin E_n \end{cases}$$

Then $\mathbb{E}[X_n] = 1, \forall n$, which implies that $\{X_n\}$ are uniformly bounded in L^1.

However, for any $K \geq 0$, as long as $n > K$,

$$\mathbb{E}[|X_n|;|X_n| > K] = 1$$

Therefore, X_n's are not uniformly integrable.

Ovserve that $X_n \to 0$ a.s., but $1 = \mathbb{E}|X_n|$ not converging to 0.

Question: what about L^p-boundness for $p > 1$?

Theorem 3.1 Suppose a class $\mathcal{C}$ of random variables are uniformly bounded in L^p

$(p > 1)$:

$$\exists A > 0, \text{ s.t. } \mathbb{E}[|X|^p] < A, \forall x \in \mathcal{C}$$

Then the class $\mathcal{C}$ is uniformly integrable (UI).

Proof. Note that

$$\begin{aligned}
\mathbb{E}[|X|; |X| > K] = \int_{|X|>K} |X| \, d\mathbb{P} \leq \int_{|X|>K} \frac{|X|^p}{K^{p-1}} \, d\mathbb{P} = \frac{1}{K^{p-1}} \int_{|X|>K} |X|^p \, d\mathbb{P} \\
\leq \frac{1}{K^{p-1}} \int_{\Omega} |X|^p \, d\mathbb{P} \\
\leq \frac{1}{K^{p-1}} A, \quad \forall x \in \mathcal{C}
\end{aligned}$$

If $X > K$, then $X^p > K^{p-1}X$.

Therefore, for any given $\varepsilon > 0$, choose K to be such that $\frac{A}{K^{p-1}} \leq \varepsilon$.

∎

Theorem 3.2 Suppose that a class $\mathcal{C}$ of random variables are dominated by an integrable random variable Y:

$$|X(\omega)| \leq Y(\omega), \quad \forall \omega \in \Omega, \forall X \in \mathcal{C}, \mathbb{E}|Y| < \infty$$

then the class $\mathcal{C}$ is UI.

Proof. Note that since $|X(\omega)| \leq Y(\omega), \forall \omega$, then

$$\{\omega \mid |X(\omega) > K|\} \subset \{\omega \mid |Y(\omega)| > K\}$$

Therefore,

$$\int_{|X|>K} |X| \, d\mathbb{P} \leq \int_{|Y|>K} |X| \, d\mathbb{P} \leq \int_{|Y|>K} |Y| \, d\mathbb{P}$$

Since Y is integrable, by Corollary 2.5.2, for any given $\varepsilon > 0$, there exists $K \geq 0$ such that

$$\int_{|Y|>K} |Y| \, d\mathbb{P} < \varepsilon.$$

This implies that $\forall X \in \mathcal{C}$,

$$\int_{|X|>K} |X| \, \mathrm{d}\mathbb{P} < \varepsilon.$$

■

Theorem 3.3 Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, and $\{\mathcal{G}_\alpha\}_{\alpha \in \mathcal{A}}$ be a sequence of sub-σ-algebra of f. Denote the class

$$\mathcal{C} := \{\mathbb{E}[X \mid G_\alpha]\}_{\alpha \in \mathcal{A}}$$

Then the class $\mathcal{C}$ is UI.