FOUNDATION OF STOCHASTIC MODELING AND APPLICATIONS

A Thesis Presented to the Department of Pure and Applied Mathematics, African University of Science and Technology

In Partial Fulfilment of the Requirements for the Degree of

Master of Science

by

Sani Rahama Abdullahi

Abuja, Nigeria

June, 2019.

Foundation of Stochastic Modeling and Applications

Certification

This is to certify that the thesis titled "**Foundation of Stochastic Modeling and Applications**" submitted to the school of postgraduate studies, African University of Science and Technology (AUST), Abuja, Nigeria for the award of the Master's degree is a record of original research carried out by Sani Rahama Abdullahi in the department of Pure and Applied Mathematics.

Approval

FOUNDATION OF STOCHASTIC MODELLING AND APPLICATIONS

By

Sani Rahama Abdullahi

A THESIS PRESENTED TO THE DEPARTMENT OF PURE AND APPLIED MATHEMATICS

RECOMMENDED: ==================

Supervisor, Prof. Gane Samb Lo

==================

Head, Department of Pure and Applied Mathematics

APPROVED: ==================

Chief Academic Officer, Prof. C. E. Chidume

=================

Date

Abstract

This thesis presents an overview on the theory of stopping times, martingales and Brownian motion which are the foundations of stochastic modeling. We started with a detailed study of discrete stopping times and their properties. Next, we reviewed the theory of martingales and saw an application to solving the problem of "extinction of populations". After that, we studied stopping times in the continuous case and finally, we treated extensively the concepts of Brownian motion and the Wienner integral.

Key Words. *Stochastic Processes, Stopping times, Martingales, Galton-Watson branching process, Brownian motion.*

Dedication

Dedicated to my parents, Mr & Mrs Sani Abdullahi, my siblings Abba, Hajia and Maryam, my husband Mr Buhari Salisu and my daughter Aisha Buhari.

Acknowledgements

I give thanks and praise unto Allah thew Creator and Cherisher of the universe for guiding meand seeing me through all my undertakings in the African University of Science and Technology, Abuja, may Allah's blessings and peace be upon His Prophet Muhammad(S.A.W)
I would like to express my deep appreciation and gratitude to my supervisor, Professor Gane Samb Lo, who has the poise and the substance of a genius. During the course of this work, he continually and convincingly inculcated in me the spirit of adventure in research and an excitement in studying mathematics. Without his guidance and persistent help, this thesis would not have been possible.

I am extremely thankful to Professor Charles E. Chidume, Vice-President (Academics) and Director of mathematics Institute, African University of Science and Technology (AUST), Abuja for providing me an opportunity and the time necessary to do this thesis. Also, I thank Miss Amaka Udigwe, Administrative assistant to the Vice-President (Academics), AUST, Abuja for her consistent availability and readiness to assist me at all times. To the entire faculty members of the Mathematics Institute, AUST, I owe you a big thanks.

I am equally grateful to my parents,siblings and husband for their love, encouragements and prayers during the course of this work.

I would like to express my special gratitude to African University of Science and Technology (AUST), Abuja and the African Development Bank(ADB) for their financial support. It was their scholarships that supported my study throughout my master programme.

Finally, my thanks and appreciations also go to my colleague and friend, Miss Hassan Zulaihat and Miss Doumbia Fatima for their understanding and co-operation during the course of this work.

Contents

CHAPTER 1

General Introduction

1. The context

The present dissertation should be placed in the project to build within the African University of Sciences and Technologies a research team in Stochastics and Statistics.

For a significant number of years, the course Measure Theory and Integration (MTI) is taught. In the two precedent Master classes, the course (MTI) has been extensively developed. The time allocated to this course allows now to cover the contents of the main reference of the course which is the exposition of Lo (2018).

That content exposed in seven hundred pages is intended to allow the reader to train himself on the knowledge broken into exercises.

This full course of (MTI) should be the basis of two teams of research in AUST:
(A) a team of research in Abstract integration and in Set-valued Integrations.

(B) a team on Stochastics and applications in Finance, Biology, Genetics, Population, etc.

The basis in Probability theory which is beneath (B) will lead to a branch of research in :

(C) Statistical Methods and Applied Statistics.

In setting up the described process, in its Probability theory component, the first step consisted in the development of the course of *Foundation of Probability Theory (MFPT)* (Lo (2018)).

This book was exposed in 2019 as a PhD course in AUST.

The aim of this dissertation is to gather the mathematical tools for stochastic modeling, or at least to gather a great deal of them in a consistent text based on the books of (MTI) and (MFPT).

So, the dissertation will open the doors of first thesis in Stochastics in AUST or will serve the future candidates for theses in Stochastics In AUST.

2. Stochastic Modeling

In real, many phenomena are described by sequence of random variables or family of random variables. Those described by a sequence require discrete stochastic modeling while those described by an arbitrary family requires continuous stochastic modeling. For example :

(a) In gambling, the surplus of a gambler at a discrete time n is a random variable X_n. Here one may be interested in the possibility of the gambler losing all of his money and to get ruined.

(b) Let us assume that some population begins with a patriarch which reproduces a random number offspring at time $n=1$. At any time $n+1$, each of the offspring reproduced at time n gives a random number of offspring. So the total number of new members at time n is a random number X_n. A natural question is : is there any possibility that the population comes to extinction, that is no offspring are made at some time N. We might also want to have an estimation of the number of offspring for large values of n, whether X_n becomes stable or increases to infinity (case of China in the past) or decreases to zero (actual situation in some European countries).

In these two cases, we face discrete stochastic modeling.

(c) Let us suppose that an insurance company has a surplus S_t at time t. It continues collecting premiums from clients with P_t the total of premium collected at time t, the return of its investments of the premium with C_t the total investments returns at time t and paying the claims to clients with L_t the total amount of losses payed to clients. The surplus of the company at time t is

$$S_t = u + P_t + C_t - L_t,$$

where u is the initial surplus at tome $t = 0$ or capital. The worse event the company wants to avoid is the ruin situation at time a t_0, which is the first time where $S_t \leq 0$.

Dealing with Situation (c) is done through continuous time stochastic modeling.

In this dissertation, we will provide interesting parts of the theory beneath such stochastic modeling.

3. Scope of the dissertation

We divide the dissertation into three parts.

$\bowtie$ The first part deals with discrete stochastic modeling. We will introduce two very important notions, that is, the notion of stopping times and theory of martingales.

As a first example, we study the extinction question of a sequence of a population, as described in Situation (b) above in specific conditions.

⋈ The second part deals with continuous stochastic modeling. Here again, We will introduce to continuous versions for stopping times and most importantly, we are going to complete this section with an introduction to Brownian Motion and present a thorough study of it.

⋈ The third part is an opening to Stochastic Integration and Stochastic Differential equations.

Generally, the contents I summarized here can be found in the most important books of the discipline. However, I particularly used Loève (1997), Chung (1974), Neveu (1965) and Lo (2018) for the fundamental modern probability theory, Neveu (1975) for discrete martingale, Billingsley (1995), Taylor and Karlin (1987) for the introduction to stochastic processes and Kuo (2000) for the stochastic calculus. Gathering all this materials and using them in a coherent way was possible in the frame of the series on probability and statistics in which Professor Lo introduces to the most inner secret of those disciplines in a series of books (Lo (2018), Lo (2018), Lo (2019), etc.) I am grateful to be able to benefit from that frame that helped me to reach so many things in a few months.

I am aware that reading and mastering the the key elements of Stochastics and trying to realize the described content is a very difficult and heavy challenge. But with the help of the leaders of AUST, especially the HOD of Pure and Applied Mathematics, with the full supervision of professor Gane Samb Lo, we humbly think that we had a firm introduction to stochastic modeling and we are ready to go further to research activities.

Part 1

Discrete Stochastic Modeling

CHAPTER 2

Stopping times

1. Introduction

A stopping time intuitively is a stopping rule. It is used in deciding whether to continue or stop a process on the basis of the present position and past events. It also plays an important role in stochastic calculus as it adds more elegance to the theory of martingales. As an example, let us use the situation of a gambler. Suppose a gambler decides that he is going to stop gambling whenever he has no money left. If X_n denotes the amount of money he has left at the nth round and V denotes the time that he stops gambling, then V is a stopping time since the time he stops depends only on how much he has at the present round and not on future happenings.

Before we see what stopping times are and their propertuies, let us begin by giving a brief explanation on stochastic processes.

2. Stochastic Processes

Definition 2.1. *A stochastic process is defined by the triplet* $(\Omega, \mathcal{A}, \mathbb{P}), (X_t)_{t\in T}, (E, \mathcal{B})$ *where*

(a) $(\Omega, \mathcal{A}, \mathbb{P})$ *is a probability space.*

(b) The time space T is arbitrary.

(c) The state space $(E, \mathcal{B}(E))$ which is a measurable space.

(d) paths of the stochastic process for any fixed $\omega \in \Omega$ defined as the mapping

$$\begin{aligned} T &\to E \\ t &\mapsto X_t(\omega) = X(t, \omega) \end{aligned}$$

such that for any $t \in T$, the mapping

$$\begin{aligned} X_t : (\Omega, \mathcal{A}) &\to (E, \mathcal{B}) \\ \omega &\mapsto X_t(\omega) \end{aligned}$$

is measurable.

Let T be an ordered set. Then we have the following definition.

Definition 2.2 (Filtration). *A filtration is a non-decreasing family $(\mathcal{F}_t)_{t \in T}$ of σ-algebras. i.e., for all $(s, t) \in T^2$, $s \leq t$ implies*

$$\mathcal{F}_s \subseteq \mathcal{F}_t$$

Definition 2.3. *A stochastic process $(X_t)_{t \in T}$ is said to be adapted to a filtration $(\mathcal{F}_t)_{t \in T}$ (or simply $(\mathcal{F}_t)_{t \in T}$-adapted if for every $t \in T$, X_t is $\mathcal{F}_t$ measurable.*

Given a stochastic process $(X_t)_{t \in T}$, where T is ordered, there is a natural filtration $(\mathcal{F}_t)_{t \in T}$ to which $(X_t)_{t \in T}$ is adapted. This filtration is defined for every $t \in T$ by

$$\mathcal{F}_t = \sigma\{X_s^{-1}(B), B \in \mathcal{B}, s \leq t\}.$$

Now, given an ordered set T, we may add a fourth element to get a broader definition.

Definition 2.4. *A stochastic process is defined by the* 4*-tuple* $(\Omega,\mathcal{A},\mathbb{P}),(X_t)_{t\in T},(E,\mathcal{B}),(\mathcal{F}_t)_{t\in T}$ *where*
(a) $(\Omega,\mathcal{A},\mathbb{P})$ *is a probability space.*
(b) The time space T *is arbitrary.*
(c) The state space $(E,\mathcal{B}(E))$ *which is a measurable space.*
(d) paths of the stochastic process for any fixed $\omega\in\Omega$ *defined as the mapping*

$$\begin{aligned} T &\rightarrow E \\ t &\mapsto X_t(\omega)=X(t,\omega) \end{aligned}$$

such that for any $t\in T$*, the mapping*

$$\begin{aligned} X_t:(\Omega,\mathcal{A}) &\rightarrow (E,\mathcal{B}) \\ \omega &\mapsto X_t(\omega) \end{aligned}$$

is measurable.
(e) $(X_t)_{t\in T}$ *is* $(\mathcal{F}_t)_{t\in T}$*-adapted*

Now, suppose we have a family $\mathcal{P}$ of marginal probability laws. i.e.,

$$\mathcal{P}=\{\mathbb{P}_{(t_1,\cdots,t_k)}:(t_1,\cdots,t_k) \text{ is an ordered finte subset of } T\}.$$

Then this family is said to be coherent if
(i) for every finite ordered subset of T, $(t_1,\cdots,t_k,t_{k+1})$, for any $(B_1,\cdots,B_k)\in\mathcal{B}^k$, we have that

$$\mathbb{P}_{(t_1,\cdots,t_k)}(B_1\times\cdots\times B_k)=\mathbb{P}_{(t_1,\cdots,t_k,t_{k+1})}(B_1\times\cdots\times B_k\times E).$$

(ii) for every finite ordered subset $(t_1, \cdots, t_k,)$ of T and for any permutation σ of $\{1, 2, \cdots, k\}$,

$$\mathbb{P}_{(t_{\sigma(1)}, t_{\sigma(2)}, \cdots, t_{\sigma(k)})}(B_{\sigma(1)} \times \cdots \times B_{\sigma(k)}) = \mathbb{P}_{(t_1, \cdots, t_k)}(B_1 \times \cdots \times B_k)$$

As a consequence of the Kolmogorov's extension theorem, we have the following theorem.

Theorem 2.5. Suppose $(E, \mathcal{B}(E))$ is a polish space(i.e., a complete and separable metric space). Given that $T \neq \emptyset$ and a coherent family $\mathcal{P}$ of probability measures. Then, there exists a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a stochastic process $(X_t)_{t \in T}$ such that for any ordered finite subset $(t_1, \cdots, t_k)$ of T, we have that

$$\mathbb{P}_{(X_{t_1}, \cdots, X_{t_k})} = \mathbb{P}_{(t_1, \cdots, t_k)}.$$

3. Basic Definitions

Let T be a subset of

$$\overline{\mathbb{R}}_+ = \{x \in \mathbb{R}, \ \ x \geq 0\} \cup \{+\infty\}.$$

General Definition of a stopping time. Given a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in T}$, a mapping

$$V : \ \ (\Omega, \mathcal{A}, \mathbb{P}) \to T$$

is stopping time with respect to the filtration $\mathcal{F}$ if and only if we have

$$\forall (t \in T), \ \ (V \leq t) \in \mathcal{F}_t. \Diamond \tag{3.1}$$

If T takes a countable number of values. i.e. $T \subseteq \mathbb{N} \cup \{+\infty\}$, then we have the following equivalent definition.

Definition of a discrete stopping time. Given a subset $\mathcal{N}$ of $\mathbb{N} \cup \{+\infty\}$, given a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathcal{N}}$, a mapping

$$V: \ (\Omega, \mathcal{A}, \mathbb{P}) \to \mathcal{N}$$

is stopping time with respect to the filtration $\mathcal{F}$ if and only if we have

$$\forall (n \in \mathcal{N}), \ (V = n) \in \mathcal{F}_n . \Diamond \tag{3.2}$$

Consistency. Let us show that when T is discrete, the two definitions are equivalent. Indeed, if Formula (3.1) holds, we have

$$\forall n \in T, (V \leq n) = \bigcup_{0 \leq p \leq n, \ \ p \in T} (V = p) \in \mathcal{F}_n.$$

Now, assume Formula (3.2) holds. Let $n \in T$. Either n has no predecessor in T and we have

$$(V = n) = (V \leq n) \in \mathcal{F}_n,$$

or n has a predecessor p in T and we have

$$(V = n) = (p < V \leq n) = (V \leq n) \setminus (V \leq p) \in \mathcal{F}_n.$$

Let us now see two easy examples of stopping times.

Two Examples of Stopping Times.

Example 1. Let T be the first time X_n is below a, that is :

$$T = \begin{cases} +\infty & if \;\; \forall (n \geq 1), X_n > a \\ \inf\{n \geq 1, X_n \leq a\} & \text{otherwise} \end{cases}$$

Then T is a stopping time.
Indeed, for any $n \geq 1$, we have

$$(T = n) = (X_1 > a, ..., X_{n-1} > a, X_n \leq a) \in \mathcal{F}_n$$

So,

$$(T \leq n) = \bigcup_{p=1}^{n} (T = p) \in \mathcal{F}_n.$$

Example 2. Let S be the first time, after T, X_n exceeds b, that is :

$$S = \begin{cases} +\infty & if \;\; \forall (n > T), X_n < b \\ \inf\{n > T, X_n \geq b\} & \text{otherwise} \end{cases}.$$

Then, S is also a stopping time.
Here also, we have for any $n \geq 1$, T can take on the values $1, ..., n-1$ on $(S = n)$. So we have

$$(S = n) = \sum_{m=1}^{n-1} (S = n) \cap (T = m)$$

$$= \sum_{m=1}^{n-1} (X_{m+1} < b, ..., X_{n-1} < b, X_n \geq b) \cap (T = m) \in \mathcal{F}_n,$$

Therefore,

$$\forall (n \geq 1), (S \leq n) \in \mathcal{F}_n.$$

In the next section, we are going to study the σ-algebra associated with a stopping time T.

4. σ-algebra generated by a stopping time

We begin by this proposition.

Proposition 2.6. *Let V be a stopping time with respect to a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in T}$ with values in T. Then the following collection of subset of Ω :*

$$\mathcal{A}_V = \{A \subset \Omega, \forall (t \in T),\ A \cap (V \leq t) \in \mathcal{F}_t\},$$

is a σ-algebra which is included in $\mathcal{F}_\infty$. If T is discrete, then

$$\mathcal{A}_V = \{A \subset \Omega, \forall (n \in T), A \cap (V = n) \in \mathcal{F}_n\}.\Diamond$$

Before we give the proof, let us have the following definition
:

Definition. For any stopping time V, the σ-algebra $\mathcal{A}_V$ is called the σ-algebra associated with the stopping time V or simply a stopped σ-algebra.

Proof of Proposition 2.6. To show that $\mathcal{A}_V$ is a σ-algebra. We have to check three things:

(i) $\Omega \in \mathcal{A}_V$

(ii) For any $A \in \mathcal{A}_V$, $A^c \in \mathcal{A}_V$

(iii) For any sequence $(A_n)_{n\geq 0} \in \mathcal{A}_V$, $\bigcup_{n\geq 0} A_n \in \mathcal{A}_V$.

(i) We have that $\Omega \in \mathcal{A}_V$, since

$$\forall t \in T,\ \ \Omega \cap (V \leq t) = (V \leq t) \in \mathcal{F}_t.$$

(ii) Let $A \in \mathcal{A}_V$. Then, for all $t \in T$, $A \cap (V \leq t) \in \mathcal{F}_t$. So,

$$A^c \cap (V \leq t) = (V \leq t) \ \ \setminus \ \ A(V \leq t) \in \mathcal{F}_t.$$

and thus, $A^c \in \mathcal{A}_V$.

(iii) Let $\{A_n, n \geq 0\} \subset \mathcal{A}_V$. We show that

$$A = \bigcup_{n\geq 0} A_n \in \mathcal{A}_V.$$

To see this, we say that for all $t \in T$

$$A \cap (V \leq t) = \bigcup_{n\geq 1} A_n \cap (V \leq t) \in \mathcal{F}_t.$$

So, $A \in \mathcal{A}_V$. Therefore, $\mathcal{A}_V$ is a σ-algebra. To finish the proof, we show that $\mathcal{A}_V \subseteq \mathcal{F}_\infty$. Since

$$\Omega = (V \leq +\infty) + (V = +\infty) = \bigcup_{n<+\infty, n\in T} (V \leq n) + (V = +\infty). \tag{4.1}$$

We have that for all $A \in \mathcal{A}_V$,

$$A = A \cap (V = +\infty) + \bigcup_{n<+\infty, n \in T} A \cap (V \leq n).$$

We have two cases. Either $+\infty \in T$ and then $A \cap (V = +\infty) \in \mathcal{F}_\infty$ and next

$$A = A(V = +\infty) + \bigcup_{n<+\infty, n \in T} A(V \leq n) \in \mathcal{F}_\infty.$$

Or, $+\infty \notin T$ and hence $(V = +\infty) = \emptyset$ and next

$$A = \bigcup_{n \geq 1} A \cap (V \leq n) = \mathcal{F}_\infty.\Diamond$$

Now it is important to a simple rule on the measurability of mappings with respect to a discrete stopped σ-algebra. We have the following proposition.

Proposition 2.7. *Let $V : \Omega \to T \subset \overline{\mathbb{N}}$ be a discrete stopping time and $h : \Omega \to \overline{\mathcal{N}}$ be a real-valued mapping. Then h is $\mathcal{A}_V$-measurable if and only if*

$$\forall n \in T,\ h1_{(V=n)}\ is\ \mathcal{F}_n - measurable.$$

Remarkable elements of stopped σ-algebras.

Let us consider stopping times with the same time space T.

(1) If V is a stopping time, we have

$$\forall s \in \mathbb{R}_+, \ \ (V \leq s) \in \mathcal{A}_V.$$

To see this, let $(s,t) \in \mathbb{R}_+^2$, then

$$(V \leq s)(V \leq t) = (V \leq min(s,t)) \in \mathcal{F}_{min(s,t)} \subset \mathcal{F}_t. \ \square$$

(2) Let V and S be two discrete stopping times, the sets $(V < S)$, $(V \leq S)$ and $(V = S)$ are both in $\mathcal{A}_V$ and in $\mathcal{A}_S$. Indeed, we have for $n \in T$,

$$(V < S) \cap (V = n) = (S > n) \cap (V = n) \in \mathcal{F}_n$$

Then $(V < S) \in \mathcal{A}_V$. Also, for any $n \in T$,

$$(V = S) \cap (V = n) = (S = n) \cap (V = n) \in \mathcal{F}_n$$

Then $(V = S)$ is in $\mathcal{A}_V$. Similarly, we have $(S = V)$ is in $\mathcal{A}_S$. So,

$$(S = V) \in \mathcal{A}_V \cap \mathcal{A}_S.$$

Now,

$$(V \leq S) = (V < S) + (V = S) \in \mathcal{A}_V$$

and for $n \in T$,

$$(V \leq S) \cap (S = n) = (V \leq n) \cap (S = n) \in \mathcal{F}_n.$$

So, $(V \leq S) \in \mathcal{A}_S$. Therefore,

$$(V \leq S) \in \mathcal{A}_V \cap \mathcal{A}_S.$$

Finally, we have

$$(V < S) = (V \leq S) \backslash (V = S) \in \mathcal{A}_V \cap \mathcal{A}_S.$$

Conditional Mathematical Expectation with respect to the σ-algebra generated by a stopping time. Given a discrete stopping time V, taking its values in T, and a real-valued random variable X which has a constant sign or which is integrable, then we have

$$\mathbb{E}(X/\mathcal{F}_V) = \sum_{n \in T} \mathbb{E}(1_{(V=n)} X/\mathcal{F}_n). \tag{4.2}$$

5. Operations on Stopping Times

We have the following properties of stopping times.

(1) A constant time $V \equiv v_0 > 0$ is a stopping time.

(2) Let S be a stopping time and let t_0 be a non-negative number such that $S + t_0 \in T$. Then $S + t_0$ is a stopping time.

(3) Let S and V be two stopping times, then $min(S, V)$ and $max(S, V)$ are stopping times.

(4) Let S and V be two discrete stopping times. Then

$$V \leq S \Leftrightarrow \mathcal{A}_V \subset \mathcal{A}_S. \tag{5.1}$$

(5) Let $(V_n)_{\geq 1}$ be a sequence of stopping times. Then

$$\sup_{n\geq 1} V_n$$

is also a stopping time. $\Diamond$

PROOF. **Proof of (1).** Let $t \in T$. If $v_0 \leq t$ then $(v_0 \leq t) = \Omega \in \mathcal{F}_t$ and if $v_0 > t$ then $(v_0 \leq t) = \emptyset \in \mathcal{F}_t$. In both cases, $(v_0 \leq t)\mathcal{F}_t$. $\square$

Proof of (2). For all $t \in T$,

$$(S + t_0 \leq t) = (S \leq t - t_0) \in \mathcal{F}_{t-t_0} \subseteq \mathcal{F}_t. \ \square$$

Proof of (3). For all $t \in T$:

$$(\max(S, V) \leq t) = (S \leq t) \cap (V \leq t) \in \mathcal{F}_t$$

and

$$(\min(S, V) \leq t) = (S \leq t) \cup (V \leq t) \in \mathcal{F}_t.$$

Proof of (4) Assume $V \leq S$. Let $A \in \mathcal{A}_V$, then for any n in T,

$$A \cap (S = n) = A \cap (V = n) \in \mathcal{F}_n,$$

if n has no predecessor in T. So, $A \in \mathcal{A}_S$. If n has a predecessor in T, then

$$A \cap (S = n) = \sum_{m \in T,\ m \leq n} (A \cap (V = m))\,(S = n) \in \mathcal{F}_n.$$

This is because $(A \cap (V = m)) \in \mathcal{F}_n$ and $(S = n) \in \mathcal{F}_n$. Thus any element of $\mathcal{A}_V$ is in $\mathcal{A}_S$.

Proof of (5) We have that

$$(\sup_{n\geq 1} V_n \leq t) = \bigcap_{n\geq 1} (V_n \leq t) \in \mathcal{F}_t.$$

□

6. Mesurability of a stopped stochastic process

In this section, we are faced with the following question. Given a stochastic process $(X_t)_{t\in T}$ which is adapted to a filtration $\mathcal{F}_t)_{t\in T}$ and a stopping time V with respect to the same filtration, do we have that the mapping

$$\omega \hookrightarrow X_{V(\omega)}(\omega)$$

is measurable? This mapping is called the **stopped stochastic process at the stopping time V.**

Here, we give the answer for the discrete case. The continuous case is more complicated and will be treated in the secondd part of this work.

Suppose that V is discrete. Then $T \subseteq \overline{\mathbb{N}}$. For any n in T,

$$(X_V \in B) \cap (V = n) = (X_n \in B) \cap (V = n) = X_n^{-1}(B) \cap (V = n) \in \mathcal{F}_n.$$

Hence

$$X_V^{-1}(B) \in \mathcal{A}_V \in \mathcal{F}_\infty \subset \mathcal{A}.$$

CHAPTER 3

Martingales

1. Introduction

The theory of Martingales has had a profound effect on modern probability theory and in statistics. Branches of Probability theory such as stochastic calculus rest on martingale foundations. The theory is extremely applicable and powerful as it has a lot of amazing consequences. Thus, it is necessary for every user of probability to at least know the basics of Martingale theory.

The term martingale originates in gambling theory. In a series of gambles organized in a Casino at times $n \in 1, 2, ..$ in some night, if at each game, X_n represents the money he plays and deposit in the playing table at the game n, We have that as the game goes on, the player accumulates information and facts from the outcomes of the former gambles. At each gamble n, if $\mathcal{F}_n$ represents the whole information he has. Before playing, he has to estimate its conditional return which is

$$\mathbb{E}(X_{n+1}/\mathcal{F}_n)$$

conditionally on the amount of available information $\mathcal{F}_n$.

The game is fair, for Casino's owner and the player if we have

$$X_n = \mathbb{E}(X_{n+1}/\mathcal{F}_n).$$

The game is unfair and is favorable to the Casino's owner if

$$X_n \geq \mathbb{E}(X_{n+1}/\mathcal{F}_n).$$

The game is unfair and favorable to the player if

$$X_n \leq \mathbb{E}(X_{n+1}/\mathcal{F}_n).$$

In the situation where the Casino's owner is favored, we say that we are in a supermartingale, i.e. favourable to the Casino's owner. In the case the player has the favor, the owner faces a submartingale, unfavorable to to Casino's owner. If neither of the casino's owner or the player is favored, that fairness is described by the term martingale.

In this chapter, we are going to start giving a summary of the concept of conditional expectations which will be a central tool used in defining martingales. Next, we are going to provide a number of equivalent definitions and some examples of martingales. This is then followed by the establishment of various properties. We shall focus on discrete martingales (when $T = \mathbb{N}$). But many properties will be presented in the general case of continuous time (when $T = \mathbb{R}_+$) if the particular handling of continuous times does not require significantly different methods.

2. Conditional Expectation

Construction of the conditional expectation

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let X be a quasi integrable real valued random variable(i.e., $\int X^+ d\mathbb{P} < +\infty$ or $\int X^- d\mathbb{P} < +\infty$). Let $\mathcal{B}$ be a σ-sub-algebra of $\mathcal{A}$. Define the mapping

$$\begin{aligned} \psi : \mathcal{B} &\to \overline{\mathbb{R}} \\ A &\mapsto \psi(A) = \int_A X d\mathbb{P} \end{aligned}$$

Then ψ is a σ-additive mapping which is continuous with respect to $\mathbb{P}$. So, we have by Radon Nikodym's theorem that there exists a random variable $Z : (\Omega, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that for any $B \in \mathcal{B}$,

$$\int_B X d\mathbb{P} = \psi(B) = \int_B Z d\mathbb{P}.$$

This mapping Z is what is called the conditional expectation of X given the σ-algebra $\mathcal{B}$. We may formally state the definition as follows.

Definition 3.1. *Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a quasi-integrable real valued random variable. Suppose $\mathcal{B}$ is a σ-sub-algebra of $\mathcal{A}$, then the conditional expectation of X given $\mathcal{B}$ denoted $E(X/\mathcal{B})$ is any mapping Z satisfying* *(i)* $Z : (\Omega, \mathcal{B}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *is a random variable.*

(ii) $\int_B X = \int_B Z$ *for any* $B \in \mathcal{B}$

REMARK 3.2. *It is immediate from the definition that for any quasi-integrable random variable* X, $\mathbb{E}(X) = \mathbb{E}(\mathbb{E}(X/\mathcal{B}))$.

Let us now see, some interesting and important properties of the conditional expectation.

Important Properties.

(i) If X is $\mathcal{B}$ measurable, then $\mathbb{E}(X/\mathcal{B}) = X$ *a.s.*

(ii) The mathematical conditional expectation is a non-negative linear and non-decreasing operator. Non-negativity here means

$$X \geq 0 \quad implies \quad \mathbb{E}(X/\mathcal{B}) \geq 0$$

(iii) $|\mathbb{E}(X/\mathcal{B})| \leq \mathbb{E}(|X|/\mathcal{B})$

(iv) Projection properties: If $\mathcal{B}_1$ and $\mathcal{B}_2$ are two σ-sub-algebras of $\mathcal{A}$ such that $\mathcal{B}_1 \subseteq \mathcal{B}_2 \subseteq \mathcal{A}$. Then, we have

$$\mathbb{E}(\mathbb{E}(X/\mathcal{B}_1)/\mathcal{B}_2) = \mathbb{E}(\mathbb{E}(X/\mathcal{B}_2)/\mathcal{B}_1) = \mathbb{E}(X/\mathcal{B}_1)$$

(v) A mapping X is said to be independent of $\mathcal{B}$ if for every $\mathcal{B}$-measurable mapping $Z : \Omega \to \mathbb{R}$ and for any measurable mapping $h : \mathbb{R} \to \mathbb{R}$, we have that

$$\mathbb{E}(Z \times h(X)) = \mathbb{E}(Z) \times \mathbb{E}(h(X)).$$

If X is independent of $\mathcal{B}$, then $\mathbb{E}(X/\mathcal{B}) = \mathbb{E}(X)$.

(vi) Monotone convergence Theorem for conditional expectation: Let $(X_n)_{n\geq 0} \subseteq L_1$. Assume $X_n \nearrow X$, then we have that

$$\lim_{n\to\infty} \mathbb{E}(X_n/\mathcal{B}) = \mathbb{E}(X/\mathcal{B})$$

(vii) Fatou Lebesgue Theorems for Conditional Expectation: Let $(X_n)_n$ be a quasi integrable real-valued random variable which is *a.s.* bounded below by an integrable random variable, then

$$\mathbb{E}(\liminf_{n\to\infty} X_n/\mathcal{B}) \leq \liminf_{n\to\infty} \mathbb{E}(X_n/\mathcal{B}).$$

If the sequence is *a.s.* bounded above by an integrable random variable, then

$$\mathbb{E}(\limsup_{n\to\infty} X_n/\mathcal{B}) \geq \limsup_{n\to\infty} \mathbb{E}(X_n/\mathcal{B}).$$

If the sequence is uniformly *a.s.* bounded by an integrable random variable Z and converges *a.s.* to X, then

$$\lim_{n\to\infty} \mathbb{E}(X_n/\mathcal{B}) = \mathbb{E}(X/\mathcal{B})$$

and

$$\mathbb{E}(|X|/\mathcal{B}) \leq \mathbb{E}(|Z|/\mathcal{B}).$$

THEOREM 3.3 (Jensen's Inequality for Mathematical Expectation). Let X be a random variable supported by an interval I on which a real valued convex function ϕ is defined. Suppose that X and $\phi(X)$ are integrable. Then for any σ-sub-algebra $\mathcal{B}$ of $\mathcal{A}$, we have

$$\phi(\mathbb{E}(X/\mathcal{B}) \leq \mathbb{E}(\phi(X)/\mathcal{B}).$$

Let us now begin this section by giving the definitions of Martingales.

3. Definitions and Basic Properties

Definition 3.4. *Let $(X_t)_{t\in T} \subseteq L^1(\Omega, \mathcal{A}, \mathbb{P})$ be $(\mathcal{F}_t)_{t\in T}$ adapted. Then:*

(a) $(X_t)_{t\in T}$ is a supermartingale if

$$\forall (s,t) \in T^2,\ s \le t,\ \mathbb{E}(X_t/\mathcal{F}_s) \le X_s\ a.s.$$

(b) $(X_t)_{t\in T}$ is a submartingale if

$$\forall (s,t) \in T^2,\ s \le t,\ \mathbb{E}(X_t/\mathcal{F}_s) \ge X_s\ a.s.$$

(c) $(X_t)_{t\in T}$ is a martingale if and only if it is both a supermartingale and a submartingale. i.e.,

$$\forall (s,t) \in T^2,\ s \le t,\ \mathbb{E}(X_t/\mathcal{F}_s) = X_s\ a.s.$$

Remark 3.5. *The above definition is equivalent to:*
$(X_t)_{t\in T}$ is a supermartingale if

$$\forall (s,t) \in T^2,\ s \le t, \forall B \in \mathcal{F}_s,\ \int_B X_t d\mathbb{P} \le \int_B X_s d\mathbb{P}.$$

$(X_t)_{t\in T}$ is a submartingale if

$$\forall (s,t) \in T^2,\ s \le t,\ \forall B \in \mathcal{F}_s,\ \int_B X_t d\mathbb{P} \ge \int_B X_s d\mathbb{P}.$$

$(X_t)_{t\in T}$ is a martingale if

$$\forall (s,t) \in T^2,\ s \leq t,\ \forall B \in \mathcal{F}_s,\ \int_B X_t d\mathbb{P} = \int_B X_s d\mathbb{P}.$$

Now, let $(X_n)_{n\geq 0} \subset L^1(\Omega, \mathcal{A}, \mathbb{P})$ be a sequence of random variables. Then, the following are immediate from the definition.

(a) $(X_n)_{n\geq 0}$ is a supermartingale if and only if

$$\forall n \geq 0,\ \ \mathbb{E}(X_{n+1}/\mathcal{F}_n) \leq X_n$$

(b) $(X_n)_{n\geq 0}$ is a submartingale if and only if

$$\forall n \geq 0,\ \ \mathbb{E}(X_{n+1}/\mathcal{F}_n) \geq X_n$$

(c) $(X_n)_{n\geq 0}$ is a martingale if and only if

$$\forall n \geq 0,\ \ \mathbb{E}(X_{n+1}/\mathcal{F}_n) = X_n$$

We also have the following facts:

For a martingale $(X_n)_{n\geq 0}$, the sequence of expectation $(\mathbb{E}(X_n))_{n\geq 1}$ is stationary.

For a supermartingale $(X_n)_{n\geq 0}$, the sequence of expectation $(\mathbb{E}(X_n))_{n\geq 1}$ is non-increasing .

For a submartingale $(X_n)_{n\geq 0}$, the sequence of expectations $(\mathbb{E}(X_n))_{n\geq 1}$ is non-decreasing.

Examples of Martingales.

(1) Sums of independent random variables. Let $(X_n)_{n\geq 1}$ be a sequence of independent random variables and let $(\mathcal{F}_n)_{n\geq 1}$ be the natural filtration. Define the partial sums $S_n = \sum_{i=1}^{n} X_i$, $n \geq 1$. Then each S_n is $\mathcal{F}_n$-measurable and X_{n+1} is independent of $\mathcal{F}_n$. So

$$\begin{aligned}
\forall n \geq 1,\ \mathbb{E}(S_{n+1}/\mathcal{F}_n) &= \mathbb{E}((S_n + X_{n+1})/\mathcal{F}_n) \\
&= \mathbb{E}((S_n + X_{n+1})/\mathcal{F}_n) \\
&= \mathbb{E}(S_n/\mathcal{F}_n) + \mathbb{E}(X_{n+1}/\mathcal{F}_n) \\
&= S_n + \mathbb{E}X_{n+1}
\end{aligned}$$

We conclude as follows. The sequence $(S_n)_{n\geq 1}$ is:
a supermartingale if $\mathbb{E}(X_n) \leq 0$ for every n.
a submartingale if $\mathbb{E}(X_n) \geq 0$ for every n.
a martingale if $\mathbb{E}(X_n) = 0$ for every n.

(2) Sequence of conditional expectation of the same random variable.
Suppose that X is integrable. Let us define

$$X_n = \mathbb{E}(X/\mathcal{F}_n), n \geq 0.$$

Then $(X_n)_{n\geq 0}$ is a martingale since, by the projection properties, we have for $n \leq m$

$$\mathbb{E}(X_m/\mathcal{F}_n) = \mathbb{E}(\mathbb{E}(X/\mathcal{F}_m)/\mathcal{F}_n) = \mathbb{E}(X/\mathcal{F}_n) = X_n.$$

Lemma 3.6. *Let $(X_n)_{n\geq 0} \subset \mathcal{L}_1$ and g be a convex function:*

1. *If $(X_n)_{n\geq 0}$ is a martingale, then $(g(X_n))_{n\geq 0}$ is a submartingale.*
2. *If $(X_n)_{n\geq 0}$ is a submartingale and in addition, g is non-decreasing , then $(g(X_n))_{n\geq 0}$ is a submartingale.*

Proof:

(1) Suppose $(X_n)_{n\geq 0}$ is a martingale and g is convex. Then, we have by the Jensen's inequality for conditional expectation that

$$\mathbb{E}\left(g(X_{n+1})/\mathcal{F}_n\right) \geq g\left(\mathbb{E}\left((X_{n+1})/\mathcal{F}_n\right)\right) = g(X_n) \quad \forall n \geq 1, \ .$$

(2) Suppose now that $(X_n)_{n\geq 0}$ is a submartingale and g is convex and non-decreasing. Then, by the non-decreasing property of g and again Jensen's inequality for conditional expectation,

$$\mathbb{E}\left(g(X_{n+1})/\mathcal{F}_n\right) \geq g\left(\mathbb{E}\left((X_{n+1})/\mathcal{F}_n\right)\right) \geq g(X_n) \quad \forall n \geq 1, \ .$$

Easy properties.

Let $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ be integrable sequences. Then,

(i) $(X_n)_{n\geq 0}$ is a supermartingale if and only if $(-X_n)_{n\geq 1}$ is a submartingale.

(ii) If $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are martingales, then for any two real numbers a and b, $aX_n + bY_n$ is a martingale.

(iii) If $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are both supermartingales(submartingales) then for $a, b \geq 0$, $aX_n + bY_n$ is a supermartingale(submartingale).

(iv) If $(X_n)_{n\geq 0}$ is a submartingale, then so is $(X_n^+)_{n\geq 0}$.

(v) If $(X_n)_{n\geq 0}$ is a non-negative submartingale then, for any $p \geq 1$ so is $(X_n^p)_{n\geq 0}$

(vi) If $(X_n)_{n\geq 0}$ is a martingale then $(X_n^+)_{n\geq 0}$ and $(X_n^-)_{n\geq 0}$ are submartingales. Therefore, $(|X_n|^p)_{n\geq 0}$ is a submartingale for any $p \geq 1$.

Switching principles. Let ν be an $(\mathcal{F}_n)_n$-stopping time.

(a) Suppose that $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are $(\mathcal{F}_n)_n$-adapted supermartingales and $X_\nu \geq Y_\nu$ on the set $(\nu < +\infty)$. Then

$$Z_n = X_n 1_{(\nu > n)} + Y_n 1_{(\nu \leq n))}$$

is also an $(\mathcal{F}_n)_n$-adapted supermartingale.

(b) Suppose that $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are $(\mathcal{F}_n)_n$-adapted submartingales and $X_\nu \geq Y_\nu$ on $(\nu < +\infty)$. Then

$$Z_n = X_n 1_{(\nu \leq n)} + Y_n 1_{(\nu > n)}$$

is an $(\mathcal{F}_n)_n$-adapted submartingale.

(c) Suppose that $(X_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ are $(\mathcal{F}_n)_n$-adapted martingales and $X_\nu = Y_\nu$ on $(\nu < +\infty)$. Then

$$Z_n = X_n 1_{(\nu \leq n)} + Y_n 1_{(\nu > n))}$$

is an $(\mathcal{F}_n)_n$-adapted martingale.

4. Maximal Inequalities

Let us begin by the maximal inequalities for submartingales.

Maximal inequality for non-negative submartingales.

Theorem 3.7. Let $(X_k)_{1\leq k\leq n}$ be a non-negative submartingale, $n \geq 1$. Then for any $\varepsilon > 0$, we have

$$\mathbb{P}(\max(X_1, X_2, ..., X_n) \geq \varepsilon) \leq \varepsilon^{-1} \int_{(\max(X_1, X_2, ..., X_n) \geq \varepsilon)} X_n \, d\mathbb{P}$$

and

$$\mathbb{P}(\max(X_1, X_2, ..., X_n) \geq \varepsilon) \leq \varepsilon^{-1} \mathbb{E}(X_n).$$

Proof:

Let $C = (\max(X_1 + \cdots + X_n) \geq \varepsilon)$. Then

$$C = \bigcup_{1 \leq k \geq n} (X_k \geq \varepsilon).$$

We may then transform this union into a sum of sets by setting

$$C_1 = (X_1 \geq \varepsilon), \quad C_2 = (X_1 < \varepsilon, X_2 \geq \varepsilon), \quad C_k = (X_1 < \varepsilon, ..., X_{k-1} < \varepsilon, X_k \geq \varepsilon), \quad k \geq 3.$$

So that

$$C = \sum_{1 \leq k \leq n} C_k.$$

Also, for any $k \in \{1, 2, \cdots, n\}$ we have that $C_k \in \sigma(X_1, ..., X_m)$ for any $m \leq k$.

Now,

$$\begin{aligned}
\mathbb{E}(X_n) &= \int X_n \ \ d\mathbb{P} \geq \int_C X_n \ \ d\mathbb{P} \\
&= \sum_{1 \leq k \leq n} \int_{C_k} X_n \ \ d\mathbb{P} \\
&\geq \sum_{1 \leq k \leq n} \int_{C_k} X_k \ \ d\mathbb{P} \\
&\geq \sum_{1 \leq k \leq n} \varepsilon \mathbb{P}(C_k) \\
&= \varepsilon \sum_{1 \leq k \leq n} \mathbb{P}(C_k) = \varepsilon \mathbb{P}(C)
\end{aligned}$$

Therefore,

$$\mathbb{P}(\max(X_1 + \cdots + X_n) \geq \varepsilon) \leq \varepsilon^{-1} \int_{(\max(X_1+\cdots+X_n)\geq\varepsilon)} X_n \ d\mathbb{P}. \tag{4.1}$$

Since $(X_k)_{1\leq k\leq n}$ is non-negative and $C \subseteq \Omega$, (4.2) gives us the second part of the theorem. i.e.,

$$\mathbb{P}(\max(X_1 + \cdots + X_n) \geq \varepsilon) \leq \varepsilon^{-1}\mathbb{E}(X_n).$$

Maximal inequality for non-negative supermartingales.

We have :

Theorem 3.8. *Let $(X_n)_{n\geq 0}$ be a non-negative supermartingale, $n \geq 1$. Then for any $a > 0$, we have*

$$\mathbb{P}((\sup_{n\geq 0} X_n \leq a)/\mathcal{F}_0) \leq min(X_0/a, 1).$$

Proof of Theorem 3.8.

If $\mathcal{B}$ is a σ-sub-algebra of $\mathcal{A}$ and for any $A \in \mathcal{A}$ we denote $\mathbb{E}(1_A/\mathcal{B})$ by $\mathbb{P}(A/\mathcal{B})$, then we have that for any $B \in \mathcal{B}$,

$$\int_B \mathbb{P}(A/\mathcal{B})d\mathbb{P} = \int_B 1_A d\mathbb{P} = \mathbb{P}(A \cap B).$$

So,

$$\forall A \in \mathcal{A}, \ \ \forall B \in \mathcal{B}, \ \ \mathbb{P}(A \cap B) = \int_B \mathbb{P}(A/\mathcal{B})d\mathbb{P}.$$

Now, let $a > 0$ and define

$$\nu_a = n1_{(n=min\{p\geq 0, X_p>a\})} + (+\infty)1_{(\sup\{X_p, p\geq 0\}\leq a)}$$

Then ν_a is a stopping time and

$$(\nu_a < +\infty) = \bigcup_{n\geq 0}(\nu_a \leq n) = (\sup_{n\geq 0} X_n > a)$$

To see that ν_a is a stopping time, let $n \geq 0$
if $n = 0$, then

$$(\nu_a = 0) = (X_0 > 0) \in \mathcal{F}_0$$

and if $1 \leq n < \infty$, then

$$(\nu_a = n) = (X_n > a) \cap \cup_{0\leq k\leq n-1}(X_k \leq a) \in \mathcal{F}_n.$$

if $n = \infty$, then

$$(\nu_a = +\infty) = \bigcup_{n\geq 0}(X_n \leq a) \in \bigcup_{n\geq 0}\mathcal{F}_n \subseteq \mathcal{F}_\infty.$$

Now, on $(\nu_a < \infty)$, we have that $\min\{p \geq 0, X_p > a\} = \nu_a$. So, $X_{\nu_a} \geq a$. Thus, we may define

$$Y_n = X_n 1_{(\nu_a > n)} + a1_{(\nu_a \leq n)}, \quad n \geq 0. \tag{4.2}$$

Then by the switching principle, $(Y_n)_n \geq 0$ is a supermartingale.

Now, we can deduce the following from 4.2.
if $n = 0$ we have two cases
(i) $Y_0 = X_0$: We have in this case that $(\nu_a > 0) = \Omega$. So,

$$\forall\omega \in \Omega, \quad \min\{p \geq 0, X_p > a\} > 0.$$

Thus, $X_0(\omega) \leq a \quad \forall\omega \in \Omega.$

Therefore, $Y_0 = \min(X_0, a)$

(ii) $Y_0 = a$: In this case, $(\nu_a \leq 0) = \Omega$. So,

$$\forall\omega \in \Omega, 0 \leq \nu_a(\omega) \leq 0.$$

i.e., $\nu_a = 0$. This means

$$\forall\omega \in \Omega, \quad \min\{p \geq 0, X_p > a\} = 0.$$

So, $X_0 > a$ and then $Y_0 = min(X_0, a)$.

So, in any case $Y_0 = min(X_0, a)$ and for any $n \geq 0$, $Y_n \geq a1_{(\nu_a \leq n)}$. Therefore, by taking conditional expectations with respect to $\mathcal{F}_0$ we have that for all $n \geq 0$,

$$\mathbb{E}(a1_{(\nu_a \leq n)}/\mathcal{F}_0) \leq \mathbb{E}(Y_n/\mathcal{F}_0) \leq Y_0 \leq min(X_0, a), \quad a.s.$$

So,

$$\mathbb{P}((\nu_a \leq n)/\mathcal{F}_0) \leq min(X_0/a, 1), \quad a.s.$$

Since $1_{(\nu_a \leq n)} \nearrow 1_{(\nu_a < \infty)}$, we have by the monotone convergence theorem for conditional expectations that

$$\mathbb{P}(\nu_a \leq n)/\mathcal{F}_0) = \mathbb{E}(1_{(\nu_a \leq n)}/\mathcal{F}_0) \nearrow \mathbb{E}(1_{(\nu_a < \infty)}/\mathcal{F}_0) = \mathbb{P}(\nu_a < +\infty)/\mathcal{F}_0).$$

So,

$$\mathbb{P}(\nu_a < +\infty)/\mathcal{F}_0) \leq min(X_0/a, 1), \ \ a.s.$$

which is equivalent to

$$\mathbb{P}((\sup_{n\geq 0} X_n > a)/\mathcal{F}_0) \leq \min(X_0/a, 1), \ \ a.s.$$

Now we have an important consequence of this theorem

Important consequence. We have by integrating both sides of the maximal inequality over the set $(X_0 < +\infty)$ that

$$\mathbb{P}(\sup_{n\geq 0} X_n > a, \ \ X_0 < +\infty) \leq \int_{(X_0<+\infty)} \min(X_0/a, 1) \ \ d\mathbb{P}.$$

and finally, by letting $a \to +\infty$, we get

$$\mathbb{P}(\sup_{n\geq 0} X_n = +\infty, \ \ X_0 < +\infty) = 0. \ \blacksquare$$

5. Almost sure convergence of Super or Sub-Martingale and Krickeberg Decomposition

Let us begin by Supermartingales.

Almost-sure convergence of non-negative supermartingales.

We have the following theorem

Theorem 3.9. Let $(X_n)_{n\geq 0}$ be a non-negative Supermartingale. Then it converges *a.s.*, say to X_∞ and this limit satisfies

$$\forall m \geq 1, \ \mathbb{E}(X_\infty/\mathcal{F}_m) \leq X_m.$$

(B) - Krickeberg Decomposition for Submartingales.

Theorem 3.10. Let $(X_n)_{n\geq 0}$ be a submartingale such that

$$\sup_{n\geq 0} \mathbb{E}(X_n^+) < +\infty.$$

Then it is the difference of a non-negative martingale $(M_n)_{n\geq 0}$ and a non-negative super-martingale $(Y_n)_{n\geq 0}$, that is

$$\forall n \geq 0, X_n = M_n - Y_n,$$

and for all $n \geq 0$, $M_n \geq X_n^+$.

Proof. Let $(X_n)_{n\geq 0}$ be a submartingale. Then $(X_n^+)_{n\geq 0}$ is also a submartingale and so is integrable. Define

$$M_n = \lim_{p\to+\infty} \mathbb{E}(X_p^+/\mathcal{F}_n).$$

Then $(M_n)_n \geq 0$ is well defined since for $n \geq 0$ fixed, we have that for $p \geq n$,

$$\mathbb{E}(X_{p+1}^+/\mathcal{F}_n) = \mathbb{E}\biggl(\mathbb{E}(X_{p+1}^+/\mathcal{F}_p)/\mathcal{F}_n\biggr),$$

and $(X_n^+)_{n\geq 0}$ is a submartingale gives

$$\mathbb{E}(X_{p+1}^+/\mathcal{F}_n) \geq \mathbb{E}(X_p^+/\mathcal{F}_n).$$

So, the sequence $\bigl(\mathbb{E}(X_p^+/\mathcal{F}_n)\bigr)_{(p\geq n)}$ in non-decreasing and thus, has a limit.

Since $\left(\mathbb{E}(X_p^+/\mathcal{F}_n)\right)_{(p\geq n)}$ is non--decreasing, we have that

$$M_n = \sup_{p\geq n} \mathbb{E}(X_p^+/\mathcal{F}_n).$$

So,

$$X_n \leq X_n^+ = \mathbb{E}(X_n^+/\mathcal{F}_n) \leq M_n$$

Now, the M_n's are integrable since for $n \geq 0$, by the monotone convergence theorem

$$\int M_n d\mathbb{P} = \lim_{p\to+\infty} \int \mathbb{E}(X_p^+/\mathcal{F}_n) d\mathbb{P} \leq \sup_{n\geq 0} \mathbb{E}(X_n^+) < +\infty.$$

Also, $\mathbb{E}(X_p^+/\mathcal{F}_n)$ is $\mathcal{F}_n$ measurable. So, M_n as a limit of $\mathcal{F}_n$ measurable functions is also $\mathcal{F}_n$ measurable. Therefore, $(M_n)_n$ is $(\mathcal{F}_n)_n$ adapted. Also, $(M_n)_n$ is a martingale since, by the monotone convergence theorem for conditional expectations and by the projection property,

$$\begin{aligned}
\mathbb{E}(M_{n+1}/\mathcal{F}_n) &= \mathbb{E}\left(\lim_{p\to+\infty} \mathbb{E}(X_p^+/\mathcal{F}_{n+1})/\mathcal{F}_n\right) \\
&= \lim_{p\to+\infty} \mathbb{E}\left(\mathbb{E}(X_p^+/\mathcal{F}_{n+1})/\mathcal{F}_n\right) \\
&= \lim_{p\to+\infty} \mathbb{E}(X_p^+/\mathcal{F}_n) \\
&= M_n.
\end{aligned}$$

Now, define

$$Y_n = M_n - X_n.$$

Then, each Y_n is integrable as a sum of integrable functions. Also, $(Y_n)_n$ is the sum of two supermartingales which is also a supermartingale. Since, $M_n \geq X_n$ for every n, we have that Y_n is non-negative. And therefore,

$$X_n = Y_n + M_n$$

is the desired decomposition.

(C) - Convergence of submartingales.

Theorem 3.11. Let $(X_n)_{n\geq 0}$ an integrable submartingale such that

$$\sup_{n\geq 0} \mathbb{E}(X_n^+) < +\infty. \quad (LC1)$$

Then $(X_n)_{n\geq 0}$ *a.s* converges to an integrable random variable X_∞.

If, moreover, $(X_n)_{n\geq 0}$ is a martingale, Condition (LC1) is equivalent to

$$\sup_{n\geq 0} \mathbb{E}(|X_n|) < +\infty. \quad (LC2)$$

Proof of Theorem (3.11). Since X_n satisfies the assumptions of the Krickeberg decomposition theorem, we have that there exists two non negative supermartingales $(M_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ such that $X_n = M_n - Y_n$, $n \geq 0$. The supermartingales $(M_n)_{n\geq 0}$ and $(Y_n)_{n\geq 0}$ converge *a.s.* to some limits, say M_∞ and Y_∞ respectively. Since $\mathbb{E}(M_\infty/\mathcal{F}_n) \leq M_n$ for every n, we get that

$$\mathbb{E}(M_\infty) = \mathbb{E}(\mathbb{E}(M_\infty/\mathcal{F}_n)) \leq \mathbb{E}(M_n) < \infty.$$

Similarly, we have that $\mathbb{E}(Y_\infty) < \infty$. So, M_∞ and Y_∞ are finite *a.s.* Hence we have $X_n \to X_\infty = M_\infty - y_\infty$.

Now, assume $(X_n)_{n\geq 0}$ is a martingale. Then, for all $n \geq 0$, $\mathbb{E}X_n = \mathbb{E}X_0$ and since $X_n^- = X_n^+ - X_n$

$$\mathbb{E}|X_n| = 2\mathbb{E}X_n^+ - \mathbb{E}X_0.$$

So, $\sup_{n\geq 0} \mathbb{E}(X_n^+) < +\infty$ if and only if $\sup_{n\geq 0} \mathbb{E}(|X_n|) < +\infty$.

6. L^1 convergence and Regular Martingales

.

We have the following theorem.

Theorem 3.12. Let $(X_n)_n$ be a sequence of random variables. Assume $X_n \to X_\infty$ in probability. If in addition, either:
(E1) The sequence $(|X_n|^p)_{n\geq 0}$ is $\mathbb{P}$-uniformly and absolutely integrable or
(E2) The sequence $(|X_n - X|^p)_{n\geq 0}$ is $\mathbb{P}$-uniformly and absolutely integrable or
(E3) The sequence $(|X_n|^p)_{n\geq 0}$ is $\mathbb{P}$-uniformly and continuously integrable or
(E4) $\|X_n\|_p \to \|X\|_p < +\infty$ holds.
Then, $(X_n)_{n\geq 0}$ converges to X_∞ in L^p

Since a submartingale converges *a.s.* (and thus in probability) under condition (LC1) of Theorem 3.11, one of the conditions above is enough to have L^1 convergence. For martingales, we have

Theorem 3.13. Let $(X_n)_{n\geq 0}$ be a martingale. The following conditions are equivalent :

(1) $(X_n)_{n\geq 0}$ converges in L^1.

(2) Condition (LC1) of Theorem 3.11 holds and there exists an integrable random variable such that

$$\forall n \geq 1, X_n = \mathbb{E}(X/\mathcal{F}_n),\ \ X \in L^1. \tag{6.1}$$

(3) There exists an integrable random variable such that Formula (6.1) holds.

(4) $(|X_n|)_{n\geq 0}$ is a uniformly and continuously integrable sequence.

Before we give the proof, let us establish two interesting properties.

Lemma 3.14. *Let X be an integrable real-valued random variable defined on $(\Omega, \mathcal{A}, \mathbb{P})$ and Γ be the class of all σ-sub-algebra of $\mathcal{A}$. Then the family $\{\mathbb{E}(X/\mathcal{B}),\ \mathcal{B} \in \Gamma\}$ is uniformly continuously integrable, that is*

$$\sup_{\mathcal{B}\in\Gamma} \int_{(|\mathbb{E}(X/\mathcal{B})|>c)} |\mathbb{E}(X/\mathcal{B})|\ d\mathcal{P} \rightarrow 0 \ as\ c \nearrow +\infty.$$

Proof of Lemma 3.14. Let $\mathcal{B} \in \Gamma$, we have that $0 \leq |\mathbb{E}(X/\mathcal{B})| \leq \mathbb{E}(|X|/\mathcal{B})$. So, for any $c, d > 0$,

$$\begin{aligned}
\int_{(|\mathbb{E}(X/\mathcal{B})|>c)} |\mathbb{E}(X/\mathcal{B})| \ d\mathcal{P} &\leq \int_{(\mathbb{E}(|X|/\mathcal{B})>c)} |\mathbb{E}(|X|/\mathcal{B})| \ d\mathcal{P} \\
&= \int_{\mathbb{E}(|X|/\mathcal{B})>c} |X| \ d\mathcal{P} \\
&= \int_{(\mathbb{E}(|X|/\mathcal{B})>c)\cap(|X|\leq d)} |X| \ d\mathcal{P} + \int_{(\mathbb{E}(|X|/\mathcal{B})>c)\cap(|X|>d)} |X| \ d\mathcal{P} \\
&\leq d\mathbb{P}(|X|/\mathcal{B}) > c) + \int_{(|X|>d)} |X| \ d\mathcal{P} \\
&\leq dc^{-1}\mathbb{E}(\mathbb{E}(|X|/\mathcal{B})) + \int_{(|X|>d)} |X| \ d\mathcal{P} \\
&= dc^{-1}\mathbb{E}(|X|) + \int_{(|X|>d)} |X| \ d\mathcal{P}
\end{aligned}$$

letting $c \nearrow +\infty$, we get that

$$\lim_{c\to+\infty} \int_{(|\mathbb{E}(X/\mathcal{B})|>c)} |\mathbb{E}(X/\mathcal{B})| \ d\mathcal{P} \leq \int_{(|X|>d)} |X| \ d\mathcal{P}$$

Finally, letting $d \nearrow +\infty$, we get

$$\int_{(|\mathbb{E}(X/\mathcal{B})|>c)} |\mathbb{E}(X/\mathcal{B})| \ d\mathcal{P} \to 0 \ \ as \ \ c \nearrow +\infty.$$

Since $\mathcal{B} \in \Gamma$ was arbitrary, we conclude that

$$\sup_{\mathcal{B}\in\Gamma} \int_{(|\mathbb{E}(X/\mathcal{B})|>c)} |\mathbb{E}(X/\mathcal{B})| \ d\mathcal{P} \to 0 \ \ as \ \ c \nearrow +\infty.$$

which is the desired result.

Lemma 3.15. *For any σ-sub-algebra $\mathcal{B}$ of $\mathcal{A}$, the linear operator $X \mapsto \mathbb{E}(X/\mathcal{B})$ is continuous on L^1.*

Proof of Lemma 3.15. Let us suppose that X_n converges to X in L^1.Then, we have that

$$|\mathbb{E}(X_n/\mathcal{B}) - \mathbb{E}(X/\mathcal{B})| \leq \mathbb{E}(|X_n - X|/\mathcal{B}).$$

So,

$$\int |\mathbb{E}(X_n/\mathcal{B}) - \mathbb{E}(X/\mathcal{B})| \ \ d\mathbb{P} \leq \int |X_n - X| \ \ d\mathbb{P} \to 0$$

Therefore, $\mathbb{E}(X_n/\mathcal{B}) \to \mathbb{E}(X/\mathcal{B})$ in L^1. Hence, the conditional expectation operator is continuous on L^1. $\square$

Now we may give the proof of the Theorem.

Proof of Theorem 3.13.

Proof of $(1) \longrightarrow (2)$. Assume $X_n \to X$ in L^1. Then,

$$\left| \int |X_n| - \int |X| \right| \leq \int ||X_n| - |X|| \leq \int |X_n - X|$$

This gives us $\mathbb{E}(|X_n|) \to \mathbb{E}(|X|)$. So, the sequence $(\mathbb{E}|X_n|)_{n\geq 0}$ is bounded. Since $(X_n)_{n\geq 0}$ is a martingale, we have that it has an a.s. limit , say X_∞. So, $X_\infty = X$ as two limits in probability of the same sequence. Thus X_n converges to X_∞ in L^1.

Now, by lemma 3.15, we have that the operator $\mathbb{E}(\circ/\mathcal{F}_n)$ is continuous. So, $\mathbb{E}(X_p/\mathcal{F}_n)$ converges to $\mathbb{E}(X_\infty/\mathcal{F}_n)$ in L^1. But, $\mathbb{E}(X_p/\mathcal{F}_n) = X_n$ for $p \geq n$. So, $\mathbb{E}(X_p/\mathcal{F}_n)$ also goes to X_n a.s. So,

$$\forall n \geq 0, \ \ X_n = \mathbb{E}(X_\infty/\mathcal{F}_n).\square$$

Proof of $(2) \longrightarrow (3)$. Immediate.

Proof of $(3) \longrightarrow (4)$. Here, we have to show that the family $\{X_n, n \geq 0\}$ is uniformly continuously integrable. But since $X_n = \mathbb{E}(X/\mathcal{F}_n)$, we now have to prove that the family $\{\mathbb{E}(X/\mathcal{F}_n),\ n \geq 0\}$ is uniformly continuously integrable. But this follows immediately from Lemma 3.14 above.

Proof of $(4) \longrightarrow (1)$. Assume $(X_n)_n$ is uniformly continuously integrable. Then,

$$\sup_{n\geq 0} \int_{(|X_n|>c)} |X_n| \ d\mathbb{P} \to 0 \ \ as \ \ c \ \ \to \infty$$

So, there exists $c_0 > 0$ s.t.,

$$\sup_{n\geq 0} \int_{(|X_n|>c_0)} |X_n| \ d\mathbb{P} \leq 1$$

But for $n \geq 0$,

$$\mathbb{E}|X_n| = \mathbb{E}(|X_n| 1_{(|X_n|\leq c_0)}) + \mathbb{E}(|X_n| 1_{(|X_n|>c_0)}) \leq c_0 + \sup_{n\geq 0} \int_{(|X_n|>c_0)} |X_n| \leq c_0 + 1 \ \ d\mathbb{P}.$$

So the sequence $\mathbb{E}|X_n|$ is bounded by $1+c_0$. This gives us that $(X_n)_{n\geq 0}$ converges to some X_∞ *a.s.* and thus in probability. Therefore, X_n converges to X_∞ in L^1. The proof is now complete. ∎

Definition. The sequence $(X_n)_{n\geq 0}$ is called a regular martingale if any of the conditions in theorem 3.13 holds.

Here are some interesting facts about regular martingales.

Proposition 3.16. *Let $(X_n)_{n\geq 0}$ be a regular martingale with a.s limit and limit in L^1, Z. If V is an associated stopping time, then*

$$X_V = \mathbb{E}(Z/\mathcal{F}_V).$$

Moreover, if V_1 and V_2 are two stopping times such that $V_1 \leq V_2$, then

$$\mathbb{E}(X_{V_2}/\mathcal{F}_{V_1}) = X_{V_1}.$$

Proof. Let $(X_n)_{n\geq 0}$ be a regular martingale. We have to justify the existence of X_V when $V = +\infty$. Since the *a.s.* limit of the $(X_n)_{n\geq 0}$ is Z, we may set $X_{+\infty} = Z$ on $(V = +\infty)$. Thus, we have

$$X_V = \sum_{V=n} X_n 1_{(V=n)} + Z 1_{(V=+\infty)}, \tag{6.2}$$

which is $\mathcal{F}_\infty$ measurable. Indeed, $\sum_{V=n} X_n 1_{(V=n)}$ is clearly measurable as a countable sum of measurable functions and $(V = +\infty) = \cup_{n\geq 0}(V > n) \in \mathcal{F}_\infty$, so, $Z1_{(V=+\infty)}$ is also measurable.

Now, from Formula 6.2, we have

$$\begin{aligned} X_V &= \sum_{V=n} X_n 1_{(V=n)} + Z1_{(V=+\infty)} \\ &= \sum_{V=n} \mathbb{E}(Z/\mathcal{F}_n)1_{(V=n)} + \mathbb{E}(Z/\mathcal{F}_\infty)1_{(V=+\infty)} \quad (L2) \\ &= \sum_{V=n} \mathbb{E}(1_{(V=n)}Z/\mathcal{F}_n) + \mathbb{E}(1_{(V=+\infty)}Z/\mathcal{F}_\infty) \quad (L3) \end{aligned}$$

where, in Line (L3), we used the regularity of the martingale for $V < +\infty$ and the fact that Z is $\mathcal{F}_\infty$-measurable. From there, using Property of discrete stoppiing times ((4.2),page 19), we get

$$X_V = \mathcal{E}(Z/\mathcal{F}_V).$$

So Formula is proved. Now, using that formula and the inclusion $\mathcal{F}_{V_1} \subset \mathcal{F}_{V_2}$, we have

$$\mathbb{E}(X_{V_2}/\mathcal{F}_{V_1}) = \mathbb{E}\biggl(\mathbb{E}(Z/\mathcal{F}_{V_2})/\mathcal{F}_{V_1}\biggr) = \mathbb{E}(Z/\mathcal{F}_{V_1}) = X_{V_1}. \ \square$$

Let us conclude this section by this property of regular martingales.

Proposition 3.17. *Let* $(X_n)_{n\geq 0}$ *be a* p*-integrable martingale,* $p > 1$, $1 - 1/p = 1/q$ *such that*

$$\sup_{n\geq 0} \|X_n\|_p < +\infty. \ (LC4)$$

Then the random variable $\sup_{n\geq 0}|X_n|$ *belongs to* L^p *and satisfies*

$$\|\sup_{n\geq 0}|X_n|\|_p \leq q \sup_{n\geq 0}\|X_n\|_p.$$

7. Doob's Decomposition for a submartingale

The Doob's Decomposition is one of the most important results and tools in Martingale studies.

Theorem 3.18. Any submartingale $(X_n)_{n\geq 0}$ is a sum of a martingale $(M_n)_{n\geq 0}$ and a non-decreasing sequence of random variables $(A_n)_{n\geq 0}$ which satisfies

$$\forall n \geq 0, A_n \textit{ and } A_{n+1} \textit{ are } \mathcal{F}_n \textit{ adapted}$$

that is

$$\forall n \geq 1,\ X_n = M_n + A_n \quad .$$

Proof of Theorem 3.18. Let $(X_n)_{n\geq 0}$ be a submartingale. Define

$$M_0 = X_0 \quad \textit{and} \quad M_n = M_{n-1} + \Big(X_n - \mathbb{E}(X_n/\mathcal{F}_{n-1})\Big), \ \ n \geq 1.$$

Define also

$$A_0 = 0 \quad \textit{and} \quad A_n = A_{n-1} + \Big(\mathbb{E}(X_n/\mathcal{F}_{n-1}) - X_{n-1}\Big), \ \ n \geq 1.$$

Then we claim that $(M_n)_{n\geq 0}$ and $(A_n)_{n\geq 0}$ are:

(i) both $(\mathcal{F}_n)_n-$adapted.

(ii) the desired composition.

(i) Let us start with $(M_n)_{n\geq 0}$. We proceed by induction. It is immediate that M_0 is $\mathcal{F}_0$ measurable. If we suppose that M_{n-1} is $\mathcal{F}_{n-1}$-measurable, It follows from the definition that M_n is also $\mathcal{F}_n$-measurable.

In the same manner, we may show that $(A_n)_{n\geq 0}$ is $(\mathcal{F}_n)_n$-adapted and $(\mathcal{F}_{n-1})_n$-adapted.

Since $(X_n)_n$ is a submartingale, we have that $\mathbb{E}(X_n/\mathcal{F}_{n-1})-X_{n-1} \geq 0$. So, for every $n \geq 1$, $A_n \geq A_{n-1}$.

Now, we have $n \geq 1$,

$$\mathbb{E}(M_n - M_{n-1}/\mathcal{F}_{n-1}) = \mathbb{E}(X_n/\mathcal{F}_{n-1}) - \mathbb{E}(X_n/\mathcal{F}_{n-1}) = 0.$$

So, $\mathbb{E}(M_n/\mathcal{F}_{n-1}) - M_{n-1} = 0$. Therefore, $(M_n)_n$ is a martingale.

Finally, we have to check that $X_n = M_n + A_n$ for all $n \geq 1$. Here also, we use an induction reasoning. First, remark that $X_0 = M_0 + A_0$. Next, we suppose that $X_{n-1} = M_{n-1} + A_{n-1}$, for a fixed $n \geq 1$. Then,

$$M_n + A_n = M_{n-1} + A_{n-1} - X_{n-1} + X_n = X_n.$$

The proof is finished. ■

CHAPTER 4

Watson-Galton Stochastic process : Extinction of populations

1. Introduction

Let us consider some organism which reproduces itself by a random and non-negative integer W of offspring with probability law $\mathbb{P}_W$ given by its *pdf* with respect to the counting measure . Let X_n denote the number of organisms we have at generation n. Then

$$X_0 = 1 \ \ \textit{and} \ \ X_1 = W$$

. Now, suppose that each of these X_1 organisms produces $W_{1,j}$ organisms, then

$$X_2 = \sum_{1 \leq j \leq X_1} W_{1,j}.$$

So, at any generation $n \geq 1$, the size of population is

$$X_{n+1} = \sum_{1 \leq j \leq X_n} W_{n,j}. \quad (REC)$$

The following hypotheses are assumed:

i The lifespans of the member of one generation is the same.

ii All of the $W_{i,j}$'s follow the probability law of $\mathbb{P}_W$.

iii The random variables $W_{i,j}$ are independent.

iv W has a finite mean μ and variance σ^2

Equation (*REC*) represents the general model of *Galton-Watson Branching Process*.

The main questions are : estimating the ultimate size of the population and determining the extinction probability of the population. Here, we address the first questions. Let us begin to study this branching process by a martingale approach.

2. Martingale Approach

Suppose that W has a finite expectation $\mu > 0$ and a finite second moment $r = \mathbb{E}(W^2)$ and the variance $\sigma^2 = r - \mu^2$. We are going to work with respect to the natural filtration $(\mathcal{F}_n)_{n\geq 0}$ associated with $(X_n)_{n\geq 0}$. For each $n \geq 0$, If we denote by μ_n, r_n and σ_n^2, the expectation, second moment and the variance of X_n respectively, we have, using conditional expectations that

$$\mathbb{E}(X_{n+1}/X_n = z) = \mathbb{E}\left(\sum_{1\leq j\leq X_n} W_{n,j}/X_n = z\right) = \mathbb{E}\left(\sum_{1\leq j\leq z} W_{n,j}/X_n = z\right) = z\mu.$$

Also,

$$X_{n+1}^2 = \sum_{1\leq j\leq X_n} W_{n,j}^2 + \sum_{1\leq j\neq h\leq X_n} W_{n,j}W_{n,h}$$

So, given $X_n = z$, we have

$$
\begin{aligned}
\mathbb{E}(X_{n+1}^2) &= \sum_{1\le j\le z} \mathbb{E}(W_{n,j}^2) + \sum_{1\le j\ne h\le z} \mathbb{E}(W_{n,j})\mathbb{E}(W_{n,h}) \\
&= \sum_{1\le j\le z} r + \sum_{1\le j\ne h\le z} \mu^2 \\
&= zr + z(z-1)\mu^2
\end{aligned}
$$

Therefore,

$$\mathbb{E}(X_{n+1}/X_n) = \mu X_n \ \ and \ \ \mathbb{E}(X_{n+1}^2/X_n) = rX_n + \mu^2(X_n^2 - X_n).$$

Taking expectations, we get that

$$\mu_{n+1} = \mu\mu_n \ \ and \ \ r_{n+1} = r\mu_n + \mu^2(r_n - \mu_n), \ \ n \ge 1.$$

So,

$$
\begin{aligned}
r_{n+1} &= (\sigma^2+\mu^2)\mu_n + \mu^2 r_n - \mu^2\mu_n \\
&= (\sigma^2+\mu^2)\mu_n + \mu^2(\sigma_n^2+\mu_n^2) - \mu^2\mu_n \\
&= \sigma^2\mu_n + \mu^2\mu_n + \mu^2\sigma_n^2 + \mu^2\mu_n^2) - \mu^2\mu_n \\
&= \sigma^2\mu_n + \mu^2\sigma_n^2 + \mu^2\mu_n^2,
\end{aligned}
$$

This implies that $\sigma_{n+1}^2 = r_{n+1} - (\mu\mu_n)^2 = \sigma^2\mu_n + \mu^2\sigma_n^2$. So, we have the two recurrence formulas

$$\mu_{n+1} = \mu\mu_n \ \ and \ \ \sigma_{n+1}^2 = \mu^2\sigma_n^2 + \sigma^2\mu_n, \ \ n \ge 0.$$

By considering the fact that $\mu_1 = \mu$ and $\sigma_0^2 = 0$ (since $X_0 = 1$), we get by inductively substituting into the recurrence relations the general expressions of μ_n and σ_n^2 which are:

$$\forall n \geq 1, \ \ \mu_n = \mu^n$$

and

$$\sigma_0^2 = 0, \ \ \forall n \geq 1, \ \ \sigma_n^2 = \sigma^2 \mu^{n-1} \sum_{k=0}^{n} \mu^k, \ \ n \geq 1.$$

Now, set

$$M_n = \frac{X_n}{\mu_n} = \frac{X_n}{\mu^n}, \ \ n \geq 0.$$

Then for all $n \geq 0$,

$$\begin{aligned}
\mathbb{E}(M_{n+1}/M_n) &= \mathbb{E}\left(\frac{X_{n+1}}{\mu^{n+1}}/M_n\right) \\
&= \frac{1}{\mu^{n+1}} \mathbb{E}\left(X_{n+1}/M_n\right) \\
&= \frac{1}{\mu^{n+1}} \mathbb{E}\left(X_{n+1}/X_n\right) \\
&= \mu \frac{X_n}{\mu^{n+1}} \\
&= \frac{X_n}{\mu^n} = M_n.
\end{aligned}$$

Therefore, $(M_n)_{n\geq 0}$ is a martingale. We also have that

$$\mathbb{E}(M_n) = \mathbb{E}\left(\frac{X_n}{\mu^n}\right) = 1$$

for all $n \geq 1$. Hence $(M_n)_{n\geq 0}$ has an *a.s.* limit M_∞. Moreover, using $\mathbb{E}(X_n^2) = \sigma_n^2 + \mu^{2n}$, we get that

$$\begin{aligned}
\mathbb{E}(M_n^2) &= \mathbb{E}\left(\frac{X_n^2}{\mu^{2n}}\right) \\
&= \frac{1}{\mu^{2n}}(\sigma_n^2 + \mu^{2n}) \\
&= \frac{\sigma^2}{\mu^{n+1}}\left(\sum_{k=0}^{n} \mu^k\right) + 1.
\end{aligned}$$

Let us discuss two clear cases.

Case $\mu < 1$. Here $\sigma_n^2 = \sigma^2\mu^{n-1}\sum_{k=0}^{n}\mu^k \to 0$ as $n \to +\infty$ and for any $\varepsilon > 0$,

$$\mathbb{P}(X_n > \varepsilon) \leq \mathbb{P}(|X_n - \mu_n| > \varepsilon/2) + \mathbb{P}(\mu_n > \varepsilon/2).$$

Since by Tchebychev's inequality

$$\mathbb{P}(|X_n - \mu_n| > \varepsilon/2) \leq \varepsilon^{-2}\sigma_n^2 \to 0$$

and $\mu_n \to 0$ implies $\mathbb{P}(\mu_n > \varepsilon/2)$ converges to zero. We have that $X_n \to_{\mathbb{P}} 0$. This then implies that

$$X_n \textit{ converges a.s. to } 0 \textit{ as } n \to +\infty.$$

So for $\mu < 1$, the population surely extincts at infinity.

Case $\mu > 1$. We have that for $\mu > 1$, the sequence $\left(\frac{\sigma^2}{\mu^{2n}(\mu-1)}(\mu^n - 1)\right)_n$ converges. So, it is bounded. Therefore,

$$\sup_{n\geq 0} \mathbb{E}(M_n^2) = \sup_{n\geq 0} \left(\frac{\sigma^2}{\mu^{2n}(\mu-1)}(\mu^n - 1) \right) < +\infty$$

This implies that the sequence $(|M_n|)_{n\geq 0}$ is uniformly continuously integrable. So, we get that $M_n \to_{L^1} M_\infty$. By the continuity of the L^1-norm, we have

$$\mathbb{E}M_\infty^2 = \lim_{n\to+\infty} \mathbb{E}M_n^2 = 1 + \frac{\sigma^2}{\mu(\mu-1)} > 0.$$

So M_∞^1 is not *a.s.* zero and

$$X_n \sim \mu^n M_\infty asn \to +\infty.$$

The first conclusion is that, X_n has the order of μ^n on a non-null set.

The case $\mu = 1$ is more complicated and concerns the coefficients p_0, p_1 and $1-p_0-p_1$. Now, let us see the extinction probability approach.

3. Extinction Probability Approach

Let us assume the hypothesis we already set. Define the event

$$E_n = (The\ \ population\ \ is\ \ extincted\ \ at\ \ generation\ \ n) = (X_n = 0), n \geq 1.$$

and extinction probabilities

$$u_n = \mathbb{P}(E_n), \quad n \geq 1.$$

It is clear that $u_0 = 0$ since $X_0 = 1$ and $u_1 = \mathbb{P}(W_1 = 0) = a_0$. We may also define, if it exists, the time of extinction T as

$$min\{n \geq 0, X_n = 0\}$$

and we have for all $n \geq 1$

$$E_n \subset (T \leq n) \;\; and \;\; (T < +\infty) = \bigcup_{n \geq 1} E_n.$$

If the population is extinct at generation n, then it is extinct at generation $n+1$. So, we have that $E_n \subseteq E_{n+1}$ for every n. Thus, $(E_n)_{n\geq 1}$ is a non-decreasing sequence. The extinction probability at a finite time $u = \mathbb{P}(T < +\infty)$ is the limit of the sequence $(\mathbb{P}(E_n))_{n\geq 1}$, which is also non-decreasing. Therefore,

$$u = \mathbb{P}(T < +\infty) = \lim_{n \to +\infty} u_n. \tag{3.1}$$

To begin, we remark that, for n fixed, the probability u_n depends only on the probability of the branching process $(X_n)_{n\geq 0}$, which in turns, depends the initial condition $X_0 = 1$ and the common probability law of the $W_{i,n}$'s.

Now, given that is the patriarch gives birth to k offsprings. We have that each of the k offsprings begins a new replication of the same branching process. The population of the patriarch family extincts at the generation n (that is, the event E_n occurs) if and only if the k families of its offsprings extinct at the same time, but after $(n-1)$ generations(that is the extinctions event $E_{1,n-1}$, ..., $E_{k,n-1}$ occur. Since the processes beginning from those k offsprings are independent and have the same probability law, we have that given $W_1 = k$,

$$E_n = \bigcap_{1 \leq j \leq k} E_{j,n-1}$$

which entails

$$\mathbb{P}(E_n/(W_1 = k)) = \prod_{1 \leq j \leq k} \mathbb{P}(E_{j,n-1}) = u_{n-1}^k.$$

So,

$$\mathbb{P}(E_n) = \sum_{k \geq 0} \mathbb{P}(E_n/(W_1 = k))P(W_1 = k) = \sum_{k \geq 0} \mathbb{P}(E_n/(W_1 = k))p_k.$$

This gives

$$u_n = \sum_{k \geq 0} p_k u_{n-1}^k = \mathbb{E}(u_{n-1}^W), \quad n \geq 2.$$

which is the *pgf* of the random variable W which is defined on $[0, 1]$. i.e.,

$$[0,1] \ni s \mapsto \Phi_W(s) = \mathbb{E}(s^W).$$

So we have

$$\forall n \geq 2, \ \ u_n = \Phi_W(u_{n-1}).$$

But the function Φ_W is continuous, differentiable twice and we have

$$\forall s \in]0,1], \ \ \Phi_W'(s) = \mathbb{E}(W s^{W-1}) = \sum_{k\geq 1} k p_k s^{k-1} > 0,$$

$$\forall s \in]0,1], \ \ \Phi_W''(s) = \mathbb{E}(W(W-1)s^{W-2}) = \sum_{k\geq 2} k(k-1) p_k s^{k-2}$$

and and when σ^2 is finite

$$\Phi_W(0) = p_0, \ \ \Phi_W(1) = 1, \ \ \Phi_W'(1) = \mu \ \ and \ \ \Phi_W''(1) = \mathbb{E}(W(W-1)) = \sigma^2 + \mu^2 - \mu.$$

Finally, letting $n \to \infty$ in , we see that the extinction probability u satisfies

$$u = \Phi_W(u), \ \ 0 \leq u \leq 1.$$

In other words, u is the smallest fixed point of Φ_W whenever it exists. We may discuss the existence of fixed points of Φ_W.

Now, if $p_0 > 0$. Then

$$\Phi_W(0) = p_0 > 0,$$

and for all $s \in [0,1]$

$$\Phi_W'(s) = \sum_{k\geq 1} kp_k s^{k-1} \leq \sum_{k\geq 1} kp_k = \mathbb{E}(W) = \mu$$

and

$$\Phi_W''(s) = \mathbb{E}(W(W-1)) = \sum_{k\geq 2} k(k-1)p_k s^{k-2} \leq \sum_{k\geq 2} k(k-1)p_k = \sigma^2 + \mu^2 - \mu.$$

We may see that $\Phi_W''(s) > 0$ for all $s \in]a,b]$ if one of the p_k, $k \geq 2$ is not zero, and is zero if

$$\forall k \geq 2, \ \ p_k = 0 \ \ and \ \ p_0 + p_1 = 1.$$

Now, the observations given above allow to solve completely the fixed point problem. We have the following cases:

Case $p_0 = 0$. Then $\mathbb{P}(W_{ij} = 0) = \mathbb{P}(W = 0) = 0$. So,the population will never die out since each member will leave after him one offspring at least. Then $u = 0$

Case $p_0 = 1$. Then, $\mathbb{P}(W = 0) = 0$. So, the patriarch will never have offspring and $u = 1$.

Case $0 < p_0 < 1$ and $p_0 + p_1 = 1$. Here, we have for all $s \in [0,1]$,

$$\Phi_W(s) = p_0 + s(1 - p_0)$$

Solving the equation $p_0+s(1-p_0) = s$ on $[0,1]$ with $p_0 > 0$ give the unique solution $u = 1$. So, the population also will die out at a finite time.

Case $0 < p_0 < 1$ and $p_0 + p_1 < 1$.
Claim: u is the smallest fixed point of Φ_W.
Proof: Since $u_n = \Phi_W(u_{n-1})$, we have by induction that $u_n = \Phi_W^n(u_0) = \Phi_W^n(0)$. Where Φ_W^n here means the composition of Φ_W n times. Now let a be another fixed point of Φ_W , then by taking into account that Φ_W^n is non-decreasing, we have that

$$a = \Phi_W(a) = \Phi_W^n(a) \geq \Phi_W^n(0) \quad \forall n \geq 1.$$

So, $a \geq \lim_{n\to\infty} \Phi_W^n(0) = \lim_{n\to\infty} u_n = u$. So, u is the smallest fixed point of Φ_W.
Now, we have that for all $s \in [0,1]$, $\Phi_W''(s) > 0$ and thus, the function Φ_W is convex on $[0,1]$. Now, we may study sub-cases.

We have the following sub-cases.

Sub-Case Case $0 < p_0 < 1$ and $p_0 + p_1 < 1$, $\mu < 1$. By Formula (F3) above, the mapping Φ_W is a contraction mapping and thus, has a unique fixed point . Since $\Phi_W(1) = 1$, we get that $u = 1$. So the population will die out surely.

This result may be checked graphically in the left figure in Table 1, where the non-decreasing function Φ_W is represented on $[0,1]$, with $u_0 = 0$ and $u_1 = \phi(u_0) = \Phi_W(u_0) = \Phi_W(0) = p_0 > 0$. The curve intersects with the line $y=x$ only at $y=x=1$.

Sub-Case Case $0 < p_0 < 1$ and $p_0 + p_1 < 1$, $\mu > 1$.

Define the map

$$[0,1] \ni s \to g(s) = \Phi_W(s) - s.$$

Then we have that

$$g(1) = 0 \quad andg(0) = p_0 > 0$$

Now,

$$g\prime(1) = \Phi_W\prime(1) - 1 = \mu - 1 > 0.$$

So, we have by continuity of g that there is a positive number δ such that $\forall s \in [1-\delta, 1]$, $g\prime(s) > 0$. So g is strictly increasing on $[1-\delta, 1]$. Therefore, we have, $\forall s \in [1-\delta, 1]$,

$$g(s) < g(1) = 0.$$

But $g(0) > 0$. So, by the intermediate value theorem, there exists $s \in]0,1[$ such that $g(s) = 0$. Therefore, we get that Φ_W has a fixed point in $]0,1[$.

To see this graphically, we use the right figure in Table 1 to see that we have two solutions, one of them being $u=1$. But the solution is the limit of the successive approximation method : $u_0 = 0$ (the probability that the population dies out at time zero is zero since atleast the partriach exists at time zero. $u_1 = \Phi_W(u_0)$, $u_2 = \Phi_W(u_1)$, $\cdots$, $u_n = \Phi_W(u_{n-1})$, $n \geq 1$. So the non-decreasing sequence $(u_n)_{n\geq 0}$ converges to the first fixed point of Φ_W.

So the probability u that the population dies out lies in $]0,1[$ and, by the figure, decreases to *zero* as μ increases to infinity.

Sub-Case Case $0 < p_0 < 1$ and $p_0 + p_1 < 1$, $\mu = 1$.
In this case, we again consider g as defined in the previous case. We have that

$$g\prime(s) = \Phi_W\prime(s) - 1 < \Phi_W\prime(1) - 1 = 0, \ \ \forall s \in [0,1[.$$

So, g is strictly decreasing on $[0,1[$. Therefore,

$$\Phi_W(s) - s = g(s) > g(1) = 0, \ \ \forall s \in [0,1[.$$

Hence, Φ_W has no fixed point in $[0,1[$ and we conclude that $u = 1$.
In this case too, we use the left figure in Table 1, with the tangent being the bisector $y = x$. Hence the approximation method, as shown in the previous sub-case, converges at $u = 1$.

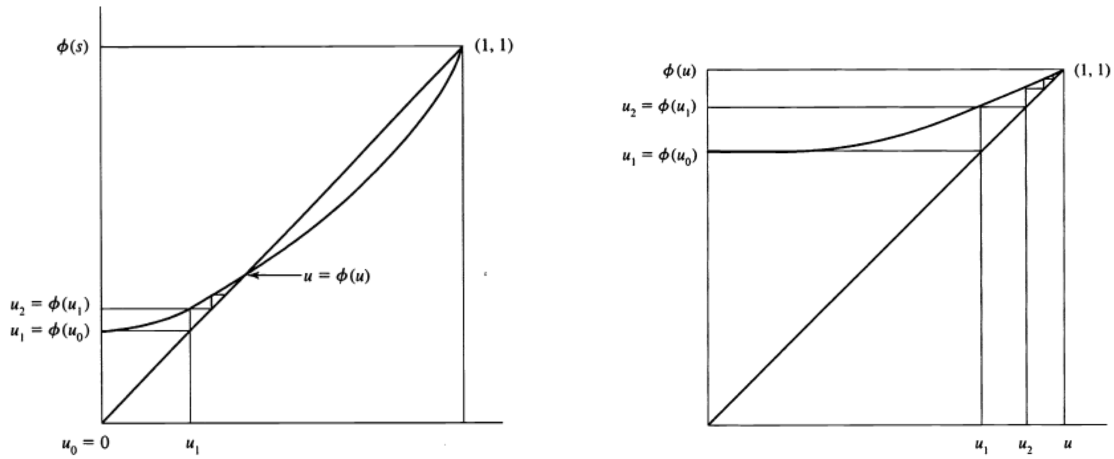

TABLE 1. Figures

Part 2

Continuous Stochastic Modeling

CHAPTER 5

Stopping Time and Measurable Stochastic Processes

An introduction to stopping times has been done in chapter 2. Here we are going to see some properties of stochastic processes and measurability of stopped stochastic process in the continuous case.

1. Stopped Stochastic processes in the continuous case

In this section, we take $T = \mathbb{R}_+$.

Regularity of Paths of Stochastic Processes

We endow T with $\mathcal{B}(T)$, the induced borel σ-algebra of $\mathbb{R}$ on it. We have two measurability types for stochastic processes. The global measurability and the progressive measurability.

Global Measurability. The stochastic process $(X_t)_{t \in T}$ is said to be globally measurable if when considered as a function of (t, ω), it is measurable. i.e., if

$$\begin{aligned} X : (T \times \Omega, \mathcal{B}(T) \otimes \mathcal{A}) &\rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ (t, \omega) &\mapsto X(t, \omega) \end{aligned}$$

is measurable.

Progressive Measurability. For any $u \in T$, we define $T_u = \{t \in T, t \leq u\}$ and endow T_u with $\mathcal{B}(T_u)$, the induced Borel σ-algebra on T_u. Consider the restricted stochastic preocess

$$(T_u \times \Omega) \ni (t, \omega) \to X_u(t, \omega) = X(t, \omega).$$

Now, given a filtration $(\mathcal{F}_t)_{t \in T}$ associated to the stochastic process $(X_t)_{t \in T}$, we say that $(X_t)_{t \in T}$ is progressively measurable with respect to the filtration $(\mathcal{F}_t)_{t \in T}$ if and only if for any $u \in T$, X_u is $\mathcal{B}(T_u) \otimes \mathcal{F}_u$ measurable. i.e., for any $B \in \mathcal{B}(\mathbb{R})$,

$$X^{-1}(B) = \{(t, \omega) \in T_u \times \Omega, X(t, \omega) \in B\} \in \mathcal{B}(T_u) \otimes \mathcal{F}_u.$$

Let us now see what continuity paths of stochastic processes mean.

Continuity paths Consider the path of the stochastic process for ω

$$\begin{array}{rcl} (T, \mathcal{B}(T)) & \to & (\mathbb{R}, \mathcal{B}(\mathbb{R})) \\ t & \mapsto & X(t, \omega) \end{array}$$

The stochastic process is said to have almost-everywhere continuous paths if

$$\mathbb{P}\left(\{\omega \in \Omega : \exists t_0, X(\circ, \omega) \;\; is \;\; not \;\; continuous \;\; at \;\; t_0\}\right) = 0$$

In a similar manner, almost-everywhere right continuous, left continuous and differentiable paths can be defined.

We are now going to see that the progressive measurability is more restrictive than the global measurability. and is extensively used in Stochastic Calculus.

Proposition 5.1. *A progressively measurable stochastic process is measurable.*

Proof. Let $(X_t)_{t\in T}$ be progressively measurable with respect to $(\mathcal{F}_t)_{t\in T}$. Since

$$[0,+\infty[\times\Omega = \bigcup_{n\geq 1}[0,n]\times\Omega.$$

Then, we have that for any $B\in\mathcal{B}(\mathbb{R})$,

$$\begin{aligned}
(X\in B) &= \bigcup_{n\geq 1}(X\in B)\cap([0,n]\times\Omega)\\
&\bigcup_{n\geq 1}\{(t,\omega)\in[0,+\infty[\times\Omega, X(t,\omega)\in B\}\cap([0,n]\times\Omega)\\
&= \bigcup_{n\geq 1}\{(t,\omega)\in[0,n]\times\Omega, X(t,\omega)\in B\} = \bigcup_{n\geq 1}\{(t,\omega)\in[0,n]\times\Omega, X_{[0,n]\times\Omega}(t,\omega)\in B\}\\
&= \bigcup_{n\geq 1}(X_{[0,n]\times\Omega}\in B).
\end{aligned}$$

Since $(X_t)_{t\in T}$ is progressively measurable, we have that for each n, $(X_{[0,n]\times\Omega}\in B)\in\mathcal{B}_n\otimes\mathcal{F}_n$. So, $\bigcup_{n\geq 1}(X_{[0,n]\times\Omega}\in B)\in\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{A}$. Therefore, $(X_t)_{t\in T}$ is globally measurable. $\square$

An important and interesting progressively measurable mapping is the right-continuous or left-continuous stochastic process. We have the following proposition

Proposition 5.2. *An almost sure right-continuous or left-continuous stochastic process* $\{X_t,\ t\geq 0\}\subset\mathbb{R}$ *is progressively measurable.*

Proof. Let $\{X_t,\ t \geq 0\}$ be an almost sure left-continuous stochastic process. Let us consider the decomposition

$$\mathbb{R}_+ = \sum_{k=1}^{\infty} [(k-1)/n, k/n[, \quad n \geq 1,$$

of $\mathbb{R}_+$. Define, for any $n \geq 1$, the discrete measurable mapping such that for any $(t,\omega) \in \mathbb{R}_+ \times \Omega$,

$$X^{(n)}(t,\omega) = X(\frac{k-1}{n}, \omega) \ \ if \ \ t \in [(k-1)/n, k/n],$$

and 0 otherwise. that is,

$$X_t^{(n)}(\omega) = \sum_{k=1}^{\infty} X(\frac{k-1}{n}, \omega) 1_{[(k-1)/n, k/n[}(t).$$

For any $t \in \mathbb{R}_+$, for any $n \geq 1$, there exists a unique $k = k_n(t)$, such that $t \in [(k-1)/n, k/n]$. By setting $x_n(t) = (k_n(t)-1)/n$, we have that

$$X_t^{(n)}(\omega) = X(x_n(t), \omega) \ \ and \ \ x_n(t) \leq t < x_n(t) + 1/n.$$

So, $x_n(t) \to t$ from the left. By the left-continuity of $X(\circ, \omega)$, we have that $X_t^{(n)}(\omega) \to X_t(\omega) a.s.$

Now, fix $n \geq 1$ and let $B \in \mathcal{B}(\mathbb{R})$. Then for any $u > 0$,

$$\begin{aligned}
(X^{(n)}_{|[0,u]\times\Omega} \in B) &= \{(t,\omega) \in [0,u] \times \Omega, X^{(n)}_{|[0,u]\times\Omega}(t,\omega) \in B\} \\
&= \{(t,\omega) \in [0,u] \times \Omega, X^{(n)}(t,\omega) \in B\} \\
&= \{(t,\omega) \in [0,u] \times \Omega, X(x_n(t),\omega) \in B\} \\
&= \bigcup_{\{k: \exists t, k = k_n(t)\}} \{(t,\omega) \in [0,u] \times \Omega, k_n(t) = k, \ \ X(\frac{k-1}{n}, \omega) \in B\}. \quad (L3)
\end{aligned}$$

Let $k_u \geq 1$ such that

$$\frac{k_0-1}{n} \leq u < \frac{k_0}{n}.$$

Then for every $t \in [0,u]$, $k_n(t) \leq k_u$. So, we have that

$$\begin{aligned}
(X^{(n)}_{|[0,u]\times\Omega} \in B) &= \bigcup_{\{k:\exists t, k=k_n(t)\}} \{(t,\omega) \in [0,u]\times\Omega, k_n(t)=k, \quad X(\frac{k-1}{n},\omega) \in B\} \\
&= \bigcup_{k \leq k_0} \{(t,\omega) \in [(k-1)/n, k/n[\times\Omega, \quad X_{(k-1)/n}(\omega) \in B\} \\
& \quad \bigcup_{k \leq k_0} [(k-1)/n, k/n[\times(X_{(k-1)/n})^{-1}(B).
\end{aligned}$$

Since each $[(k-1)/n, k/n[\times(X_{(k-1)/n})^{-1}(B)$ is in $\mathcal{B}_{k/n}\otimes\mathcal{F}_{(k-1)/n} \subset \mathcal{B}_u \otimes \mathcal{F}_u$ for $k \leq k_0$. We have that for any $n \geq 1$,

$$(X^{(n)}_{|[0,u]\times\Omega} \in B) \in \mathcal{B}_u \otimes \mathcal{F}_u.$$

So, $X_{[0,u]\times\Omega} = \lim_{n\to\infty} X^{(n)}_{|[0,u]\times\Omega}$ is $\mathcal{B}_u \otimes \mathcal{F}_u$ measurable. Hence, $(X_t)_{t\in T}$ is progressively measurable.

Let us now conclude this chapter with this result about the stopped stochastic process.

Theorem 5.3. Let $(X_t)_{t\in T}$, $T = \mathbb{R}_+$, be a progressively measurable stochastic process which is adapted to the filtration $\mathcal{F} = (\mathcal{F})_{t\in T}$. Let V be stopping time with respect to the filtration $\mathcal{F}$. Then the stopped stochastic process X_V is $\mathcal{A}_V$-measurable, where $\mathcal{A}_V$ is the σ-algebra generated by V.

Proof of Theorem . We have to prove that for any Borel set B of $\mathbb{R}$, that $(X_V \in B)$ belongs to $\mathcal{A}_V$, which means that

$$\forall(u \geq 0), (X_V \in B) \cap (V \leq u) \in \mathcal{F}_u.$$

Now, let $u \geq 0$, then we have that

$$(X_V \in B) \cap (V \leq u) = (X_{V_{|(V \leq u)}} \in B).$$

But $X_{V|(V \leq u)} = X_{|[0,u] \times \Omega} \circ \ell$, where

$$\begin{aligned} \ell : ((V \leq u), \mathcal{F}_u^*) &\rightarrow \quad [0, u] \times \Omega \\ \omega &\rightarrow \ell(\omega) = (V(\omega), \omega) \end{aligned}$$

and $\mathcal{F}_u^*$ is the σ-algebra induced by $\mathcal{F}_u$ on $(V \leq u)$. Since ℓ is measurable (because its coordinates are) and $X_{|[0,u] \times \Omega}$ is $\mathcal{B}(T_u) \otimes \mathcal{F}_u$-measurable by the progressive measurability of $(X_t)_{t \in T}$, we conclude that $X_{V|(V \leq u)}$is measurable. Hence, $(X_V \in B) \cap (V \leq u) \in \mathcal{F}_u$ $\blacksquare$.

CHAPTER 6

Introduction to the Brownian Motion

1. Kolmogorov Construction of the Brownian Motion

Family of finite distribution probability laws. Let $k \geq 1$ and let $0 = t_0 < t_1 < ... < t_k$ be k real numbers. Let $Y = (Y_{t_1}, Y_{t_2}, ..., Y_{t_k})^t$ be a random vector with k independent and centered random variables. Assume that for $i \in \{1, 2 \ldots, k\}$, $Y_{t_i} \sim \mathcal{N}(0, t_i - t_{i-1})$. Consider the transformation

$$X = (X_{t_1}, X_{t_2}, ..., X_{t_k})^t = (Y_{t_1}, Y_{t_1} + Y_{t_2}, ..., Y_{t_1} + Y_{t_2} + ... + Y_{t_k})^t.$$

Then,

$$Y = (X_{t_1}, X_{t_2} - X_{t_1}, ..., X_{t_k} - X_{t_{k-1}})^t.$$

Now, for any $(y_1, y_2, \ldots, y_k)^t \in \mathbb{R}^k$, we have that

$$\begin{aligned} f_Y(y) &= \prod_{i=1}^{k} f_{Y_{t_i}}(y_i) \\ &= \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp(-\frac{1}{2}\frac{y_i^2}{t_i - t_{i-1}}). \end{aligned}$$

and the Jacobian determinant of the transform

$$|J(y)| = 1$$

So,

$$\begin{aligned} f_X(x) &= f_Y(y)|J(y)| \\ &= f_Y(x_1, x_2 - x_1, \ldots, x_k - x_{k-1}) \\ &= \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp(-\frac{1}{2}\frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}) \end{aligned}$$

Now, for any $0 = t_0 < t_1 < ... < t_k$ and$(x_1, x_2, \ldots, x_k) \in \mathbb{R}^k$, set

$$f_{(t_1,t_2,\ldots,t_k)}(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp(-\frac{1}{2}\frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}$$

then the family $\{f_{(t_1,t_2,\ldots,t_k)}, t_1 < ... < t_k, k \geq 1\}$ of *pdf*'s is coherent.

Indeed, it is immediate from the construction that each $f_{(t_1,t_2,\ldots,t_k)}$ is a pdf. So, we only need to show that for $k \geq 1$, the *pdf* $f_{(t_1,\ldots,t_k)}$ is the marginal *pdf* of $f_{(t_1,\ldots,t_{k+1})}$. But by Fubini's Theorem, we

have that

$$\begin{aligned}
\int_{\mathbb{R}} f_{(t_1,\ldots,t_{k+1})}(x_1,\ldots,x_{k+1}) \ \ dx_{k+1} &= \int_{\mathbb{R}} \prod_{i=1}^{k+1} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(-\frac{1}{2}\frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}\right) \ \ dx_{k+1} \\
&= \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(-\frac{1}{2}\frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}\right) \\
&\times \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi(t_k - t_{k-1})}} \exp\left(-\frac{1}{2}\frac{(x_{k+1} - x_k)^2}{t_k - t_{k-1}}\right) \ \ dx_{k+1} \\
&= \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left(-\frac{1}{2}\frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}\right) \\
&= f_{(t_1,\ldots,t_k)}
\end{aligned}$$

Thus, the family is coherent. As a consequence of the Kolmogorov's extension theorem, we have the following theorem

Theorem 6.1. There exists a stochastic process $(\Omega, \mathcal{A}, \mathbb{P}), (B_t)_{t\geq 0}, (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ whose finite-distribution margins are characterized by their *pdf*'s : for any $x = (x_1, \ldots, x_k)^t \in \mathbb{R}^d$

$$f_{(B(t_1),\ldots,B(t_k))}(x_1,\ldots,x_k) = \prod_{i=1}^{k} \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp(-\frac{1}{2}\frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}),$$

The stochastic process defined in theorem 6.1 is called the Brownian motion.

Remark 6.2. *The following properties are immediate from the construction*

(1) *The Brownian motion is a centred Gaussian process*

(2) *For any* $t \in \mathbb{R}$, $B_t \sim \mathcal{N}(0, t)$

(3) *The Brownian motion has independent increments i.e. For* $0 < s_1 < s_2 < s_3$, $B_{s_2} - B_{s_1}$ *and* $B_{s_3} - B_{s_2}$ *are independent.*

(4) *The Brownian motion has strong stationary increments : for* $0 \leq s < t$, $B_t - B_s \sim B(t-s) \sim \mathcal{N}(0, t-s)$..

(5) *For all* $(s,t) \in \mathbb{R}_+^2$, $\Gamma(s,t) = Cov(B_t, B_s) = \min(s,t)$.

To see (5), consider $0 \leq s \leq t$. *We have that*

$$\begin{aligned}
\mathbb{C}ov(B_t, B_s) &= \mathbb{E}(B(s)B(t)) - \mathbb{E}(B(s))E(B(t)) \\
&= \mathbb{E}(B(s)B(t)) \\
&= \mathbb{E}(B(s)(B(s) + (B(t) - B(s))) \\
&= \mathbb{E}(B(s)^2) + \mathbb{E}B(s)(B(t) - B(s)) \\
&= s + \mathbb{E}(B(s))\mathbb{E}((B(t) - B(s))) \\
&= s
\end{aligned}$$

Therefore,

$$\mathbb{C}ov(B_t, B_s) = s = min(s,t). \square$$

2. Characterizations and Tranformations of the Brownian Motion

Proposition 6.3. *The following characterizations hold.*

(1) Any centered Gaussian Process with time space $\mathbb{R}_+$, *with state space* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, *and of variance-covariance function*

$$\Gamma(s,t) = min(s,t),\ (s,t) \in \mathbb{R}_+^2,$$

has the same probability law as the Brownian Motion.

(2) Any Stochastic Process $(B_t)_{t\geq 0}$ with time space $\mathbb{R}_+$, state space $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ satisfying the following properties :

(i) Each margin B_t is a centered Gaussian random variable with variance t.

(ii) B has independent increments : for any ordered and finite subset $0 = t_0 < t_1 < ... < t_k$, $k \geq 1$, of $\mathbb{R}_+$, the increments $(B_{t_1} - B_{t_0},\ B_{t_2} - B_{t_1},\ ...,\ B_{t_k} - B_{t_{k-1}})$ are independent.

has the same probability law as the Brownian Motion.

Proof of (2). Let $(B(t))_{t\geq 0}$ be a stochastic process satisfying the two conditions. We need to show that $(B(t))_{t\geq 0}$ is a centred Gaussian process satisfying

$$\forall (s,t) \in \mathbb{R}_+^2, Cov(B_t, B_s) = \min(s,t).$$

Since for any ordered and finite subset $0 = t_0 < t_1 < ... < t_k$, $k \geq 1$, of $\mathbb{R}_+$, the vector $(B(t_1), ..., B(t_k))$ is a linear transform of the vector of increments $(B_{t_1} - B_{t_0}, B_{t_2} - B_{t_1}, ..., B_{t_k} - B_{t_{k-1}})$ which is Gaussian as a vector of independent Gaussian components, we have that$(B(t_1), ..., B(t_k)$ is also Gaussian. This and the fact that each B_t is centred gives us that $(B(t))_{t\geq 0}$ is a centred Gaussian process. Also, with condition (ii) at hand, we may use the same technique as used in computing the the variance covariance function to get that

$$\forall (s,t) \in \mathbb{R}_+^2, Cov(B_t, B_s) = \min(s,t),$$

So we have by (1) that $(B(t))_{t\geq 0}$ is a Brownian motion.

3. Tranformations

The motion movement can be transformed while keeping the same law.

A - Time re-scaling. Given a Brownian motion $(B(t))_{t\geq 0}$, we have the following interesting re-scaling property.

Fix any $c > 0$ and define for all $t \geq 0$

$$B^{(1)}(t) := c^{-1/2} B(ct).$$

Then $(B^{(1)}(t))_{t\geq 0}$ also follows a Brownian motion. Indeed, for any

$t \geq 0$, we have

$$B^{(1)}(t) \sim \mathcal{N}(0,t)$$

and for any $s < t$,

$$\begin{aligned}
\mathbb{E}(B^{(1)}(s)(B^{(1)}(t) - B^{(1)}(s))) &= \mathbb{E}(c^{-1/2}B(cs)(c^{-1/2}B(ct) - c^{-1/2}B(cs))) \\
&= c^{-1}\mathbb{E}(B(cs)(B(ct) - B(cs)) \\
&= 0 \\
&= \mathbb{E}(B^{(1)}(s))\mathbb{E}((B^{(1)}(t) - B^{(1)}(s))).
\end{aligned}$$

Thus, $(B^{(1)}(t))_{t\geq 0}$ is a Brownian motion.

REMARK 6.4. *$B^{(1)}$ is continuous if B is.*

If we denote the operator which transforms B into $B^{(1)}$ by

$$T_c(B) = B^{(1)},$$

then we have that

$$B^{(1)} = T_c(B) \Leftrightarrow B = T_{1/c}(B^{(1)}).$$

B - Increment of translation For $t \geq 0$ fixed, define $\forall s \geq 0$,

$$B^{(2)}(s) = B(s+t) - B(t).$$

Then $(B^{(2)}(t))_t$ is also a Brownian motion. Indeed, stationarity of the Brownian motion gives us

$$B(s+t) - B(t) \sim \mathcal{N}(0, s)$$

and for $0 \leq s_1 < s_2 < s_3$, we have that

$$B^{(2)}(s_2) - B^{(2)}(s_1) = B(s_2+t) - B(s_1+t)$$

and

$$B^{(2)}(s_3) - B^{(2)}(s_2) = B(s_3+t) - B(s_2+t)$$

which are independent by the independence of increments of the Brownian motion.
Here again, we can define the operator

$$T_{(t)}(B) = B^{(2)}.$$

We also can see that

$$B^{(2)} = T_{(t)}(B) \Leftrightarrow B = T_{(-t)}(B^{(1)}). \tag{3.1}$$

The composition of the two operators T_c and $T_{(t)}$ is commutative, that is,

$$T_c \circ T_{(t)} = T_{(t)} \circ T_c.$$

4. Standard Brownian Motion

In this section, we are going to construct an *a.s* continuous version of the Brownian motion which is known as the standard Brownian motion. Almost sure continuity of the Brownian motion is important as it plays an important role in Stochastic integration. So, we will need to see how the continuous Brownian motion is constructed.

Let us start by constructing a null set where outside the null set, the Brownian motion has uniformly continuous paths on Dyadic numbers.

Step 1 : Construction of a null set where on its complement, B has locally uniform paths on Dyadic numbers.

Consider the class of all dyadic numbers

$$D = \{\frac{k}{2^n} : k \in \mathbb{Z}, n \in \mathbb{N}\}.$$

Then D is dense in $\mathbb{R}$.

Indeed, for any $a, b \in \mathbb{R}$ such that $a < b$, we have that $(b-a)2^n \to \infty$ as $n \to \infty$. So, there is some natural number N such that

$$(b-a)2^N > 1$$

i.e., the length of the interval $(a2^N, b2^N)$ is greater than 1. So, there is an integer K such that

$$a2^N < K < b2^N$$

i.e.,

$$a < \frac{K}{2^N} < b.$$

Therefore, D is dense in $\mathbb{R}$.

Now, define

$$I_{k,n} = [\frac{k}{2^n}, \frac{k+2}{2^n}], \quad (k,n) \in \mathbb{Z} \times \mathbb{N},$$

$$M_{k,n} = \sup_{r \in I_{k,n} \cap D} |B(r) - B\left(\frac{k}{2^n}\right)|$$

and

$$M_n = \sup_{k \leq (n+1)2^n} M_{k,n}.$$

Claim $\mathbb{P}(M_n > n^{-1}, \ i.o.) = 0$.

Proof. Let $k \in \mathbb{Z}$ and $n \in \mathbb{N}$. Since $D \cap I_{k,n} \subset D$ which is countable, we have that there exists a sequence $(r_p)_{p \geq 1}$ such that $D \cap I_{k,n} = (r_p)_{p \geq 1}$. Also, we have that

$$(4.1) \qquad M_{k,n,p} = \sup_{1 \leq i \leq p} \left|B\left(r_i\right) - B\left(\frac{k}{2^n}\right)\right| \nearrow M_{k,n} \ \ as \ \ p \nearrow +\infty.$$

Now, fix $p \in \mathbb{N}$ and set $r_0 = \frac{k}{2^n}$. Define

$$X_i = B\left(r_i\right) - B\left(r_{i-1}\right).$$

Then by the independence of increments of the Brownian motion, the X_i's are independent. Also, We have that

$$X_1 + ... + X_i = B\left(r_i\right) - B\left(\frac{k}{2^n}\right).$$

So, by Etemadi's Formula, we have that for any $\alpha > 0$,

$$\mathbb{P}\left(\max_{1 \leq i \leq p} |X_1 + \cdots + X_i| \geq 3\alpha\right) \leq 3 \max_{1 \leq i \leq p} \mathbb{P}\left(|X_1 + \cdots + X_i| \geq \alpha\right)$$

i.e., for all $\alpha > 0$,

$$\mathbb{P}(M_{k,n,p} \geq 3\alpha) \leq 3 \max_{1 \leq i \leq p} \mathbb{P}(|X_1 + \cdots + X_i| \geq \alpha).$$

Now,

$$
\begin{aligned}
\mathbb{P}(M_{k,n,p} \geq 3\alpha) &\leq 3 \max_{1 \leq i \leq p} \mathbb{P}\left(|B\left(r_i\right) - B\left(\frac{k}{2^n}\right)| \geq \alpha\right) \\
&= 3 \max_{1 \leq i \leq p} \mathbb{P}\left(|N(0,1)| \geq \alpha (r_i - \frac{k}{2^n})^{-\frac{1}{2}}\right) \\
&= 3 \max_{1 \leq i \leq p} \mathbb{P}\left(|N(0,1)|^4 \geq \alpha^4 (r_i - \frac{k}{2^n})^{-2}\right) \\
&\leq 9 \max_{1 \leq i \leq p} \alpha^{-4} (r_i - \frac{k}{2^n})^2 \\
&\leq 9\alpha^{-4} 2^{-2n}
\end{aligned}
$$

Since the sequence $\{\mathbb{P}(M_{k,n,p} \geq 3\alpha)\}$ is nondecreasing and thus has a limit, we get that

$$\lim_{p \to \infty} \mathbb{P}(M_{k,n,p} \geq 3\alpha) \leq 9\alpha^{-4} 2^{-2n}.$$

Continuity of probability then gives

$$\mathbb{P}(M_{k,n} \geq 3\alpha) \leq 9\alpha^{-4} 2^{-2n}.$$

Finally, we have

$$\mathbb{P}(M_n \geq 3\alpha) \leq \sum_{k=1}^{(n+1)2^n} \mathbb{P}(M_{k,n,p} \geq 3\alpha) \leq 9(n+1)\alpha^{-4} 2^{-n}.$$

Set $\alpha = n^{-1}/3$, then the last inequality becomes

$$\mathbb{P}(M_n \geq n^{-1}) \leq 3^{-2}(n+1) n^4 2^{-n}.$$

This gives us that the series $\sum_{n\geq 1} \mathbb{P}(M_{k,n} \geq n^{-1})$ is convergent. Hence, by the Borel- Cantelli lemma, we conclude that $\mathbb{P}(M_n > n^{-1},\ i.o.) = 0$ $\square$

Next, we show that the Brownian motion has uniformly continuous paths on dyadic numbers outside $(M_n > n^{-1},\ i.o.)$.

Set $N = \{M_n > n^{-1},\ i.o.\}$ and fix $\omega \in N^c$, $t \in \mathbb{R}_+$. We show that B is uniformly continuous on $D \cap [0,t]$.
Let $\varepsilon > 0$. $\omega \in N^c$ implies $M_n(\omega) \geq n^{-1}$ finitely often. So, there is a natural number $n_0 = n(\omega, \varepsilon)$ such that for n larger than n_0, we have that $M_n(\omega) \leq n^{-1}$. We may choose n_1 such that

$$n_1 > n_0, n_1 > t \quad and \quad 2n_1^{-1} \leq \varepsilon.$$

So, the interval $[0,t]$ could be partitioned as

$$\{0\} + \sum_{0 \leq k \leq K(t)-1}]\frac{k}{2^{n_1}}, \frac{k+1}{2^{n_1}}] +]\frac{K(t)}{2^{n_1}}, t], \quad for \quad \frac{K(t)}{n} \leq t < \frac{k+1}{n}.$$

Let $r_1, r_2 \in [0,t] \cap D$ such that $|r_1 - r_2| < 2^n$, then either r_1 and r_2 are in one interval of the decomposition or in two adjacent intervals.

which implies

$$k \leq (n+1)2^n.$$

It follows that

$$\begin{aligned}
|B(r_1,\omega)-B(r_2,\omega)| &= \left|B(r_1)-B\left(\frac{k}{2_{n_1}}\right)\right|+\left|B(r_2)-B\left(\frac{k}{2_{n_1}}\right)\right| \\
&\leq 2M_{k,n_1}(\omega) \\
&\leq 2M_{n_1}(\omega) \\
&\leq 2n_1^{-1} \\
&\leq \varepsilon.
\end{aligned}$$

Therefore, $D \ni r \to B(\omega, r)$ is uniformly continuous on each $[0,t]$, $t>0$, outside the null set N.

Finally, we are going to use the theory of Weak Convergence to conclude.

Let $\omega \in N^c$. Let $t \geq 0$. Since $B(\omega, \circ)$ is uniformly continuous on $D \cap [0,t]$, we have that there exists a unique continuous extension of $B(\omega,\circ)$ on $\overline{D\cap[0,t]}=[0,t]$. Let us denote this extension by $\tilde{B}$. Then we have that $\tilde{B}(w,\circ)$ is continuous on $\mathbb{R}_+$ Furthermore, $\tilde{B}(w,\circ)$ id defined by

$$\tilde{B}(w,t)=\lim_{D\ni r\searrow t} B(w,r)$$

Now for any $0=t_0<t_1<...<t_k$, $k\geq 1$, we have that

$$(B(r_1),B(r_2),...,B(r_k)) \to (\tilde{B}(t_1),\tilde{B}(t_2),...,\tilde{B}(t_k)) \ \ as \ \ r_1 \searrow t_1, r_1 \searrow t_1, ..., r_k \searrow t_k,$$

outside the null-set N. So, by the comparison theorem, we get that

$$(B(r_1), B(r_2), ..., B(r_k)) \rightsquigarrow (\tilde{B}(t_1), \tilde{B}(t_2), ..., \tilde{B}(t_k)) \ \ as \ \ r_1 \searrow t_1, r_1 \searrow t_1, ..., r_k \searrow t_k,$$

Next, we apply the Portmanteau Theorem to get that the *cdf's* converge. i.e., for any continuity point $x = (x_1, ..., x_k)$ of the *cdf* $F_{\tilde{B}(t_1),\tilde{B}(t_2),...,\tilde{B}(t_k)}$, we have

$$F_{\tilde{B}(t_1),\tilde{B}(t_2),...,\tilde{B}(t_k)}(x_1, x_2, ..., x_k) = \lim_{r_1 \searrow t_1,...,r_k \searrow t_k} F_{B(r_1),B(r_2),...,B(r_k)}(x_1, x_2, ..., x_k)$$

But by Scheffé's Theorem, we have

$$F_{B(r_1),B(r_2),...,B(r_k)}(x_1, x_2, ..., x_k) = \int_{]-\infty,x]} f_{(r_1,...,r_k)}(s_1, ..., s_k) ds_1...ds_k$$

converges also, for all $x \in \mathbb{R}^k$,

$$\int_{]-\infty,x]} f_{(t_1,...,t_k)}(s_1, ..., s_k) ds_1...ds_k = F_{B(t_1),B(t_2),...,B(t_k)}(x_1, x_2, ..., x_k).$$

Therefore, we conclude that on the continuity points of $F_{\tilde{B}(t_1),\tilde{B}(t_2),...,\tilde{B}(t_k)}$,

$$F_{B(t_1),B(t_2),...,B(t_k)} = F_{\tilde{B}(t_1),\tilde{B}(t_2),...,\tilde{B}(t_k)},$$

Hence, B and $\tilde{B}$ are equal in distribution.

Conclusion. The stochastic process $\tilde{B}$ is a Brownian motion with *a.s.* continuous paths.

Characterization of standard Brownian motion. In the former characterization, we needed to have Gaussian margins at least. But if we suppose that B is continuous, we have a more powerful following result.

Proposition 6.5. *Let $\{B(t),\ t \geq 0\}$ be a continuous stochastic process with independent increments such that $\mathbb{E}B(t) = 0$ and $\mathbb{E}B(t)^2 = t$. Then B is a standard Brownian motion.*

Proof of Proposition 6.5. Let B be a stochastic process satisfying the assumptions. Then we have that

$$\mathbb{E}B(0)^2 = 0.$$

So, $B_0 = 0$ a.s. Now, let $0 \leq s \leq t$, then

$$\mathbb{E}(B(t) - B(s))^2 = \mathbb{E}(B(t)^2 + B(s)^2 + 2B(s)B(t)).$$

But,

$$\mathbb{E}(B(t)B(s)) = \mathbb{E}(B(s)^2 + (B(s) - B(0))(B(t) - B(s))) = \mathbb{E}(B(s)^2) = s$$

So,

$$\mathbb{E}(B(t) - B(s))^2 = t + s - 2s = t - s.$$

Therefore, $B(t) - B(s)$ has mean 0 and variance $t - s$. But, for us to conclude that B is a Brownian motion, we have to show that

$B(t)-B(s)\sim\mathcal{N}(0,t-s)$. Let us prove it for $s=0$, $0\le t<1$.

Define for $u\in\mathbb{R}$, $0\le t<1$, g by

$$g(t,u)=\mathbb{E}\exp(iuB(t)).$$

We show that g is the solution of the i.v.p.

$$\frac{\partial g(t,u)}{\partial t}=-\frac{u^2}{2}g(t,u),\qquad g(0,u)=1.$$

Fix $u\in\mathbb{R}$ and set $\Delta_{s,t}=B(t)-B(s)$, $0\le s\le t$. We recall that

$$e^{ix}=1+ix-x^2/2+c(x),$$

where $|c(x)|\le|x|^3$. We choose an h small enough so that $t+h<1$, then

$$\begin{aligned}
g(t+h,u)-g(t,u) &= \mathbb{E}\Big(\exp(iuB(t+h))-\exp(iuB(t))\Big)\\
&= \mathbb{E}\Big(\exp(iuB(t))\Big(\exp(iu\Delta_{t,t+h})-1\Big)\Big)\\
&= \mathbb{E}\Big(\exp(iuB(t))\Big(iu\Delta_{t,t+h}-\frac{u^2}{2}\Delta^2_{t,t+h}+c(\Delta_{t,t+h})\Big)\Big)\\
&= \mathbb{E}\Big(iu\Delta_{t,t+h}\exp(iuB(t))\Big)-\mathbb{E}\Big(\frac{u^2}{2}\Delta^2_{t,t+h}\exp(iuB(t))\Big)\\
&+ \mathbb{E}\Big(c(\Delta_{t,t+h})\exp(iuB(t))\Big)\\
&= -\frac{u^2}{2}hg(t,u)+\mathbb{E}\Big(\exp(iuB(t))c(\Delta_{t,t+h})\Big)\\
&= -\frac{u^2}{2}hg(t,u)+\mathbb{E}(\exp(iuB(t)))\mathbb{E}(c(\Delta_{t,t+h}))
\end{aligned}$$

Now,

$$|\mathbb{E}(\exp(iuB(t)))\mathbb{E}(c(\Delta_{t,t+h}))| \leq \mathbb{E}\,|c(\Delta_{t,t+h})| \leq \mathbb{E}\,|\Delta_{t,t+h}|^3.$$

We claim that $\frac{1}{h}\mathbb{E}\,|\Delta_{t,t+h}|^3 \to 0$ as $h \to 0$.

Indeed, for $n \geq 1$, fixed, we have

$$\Delta_{t,t+h} = B(t+h) - B(t) = \sum_{1\leq j\leq n} B\left(t + \frac{hj}{n}\right) - B\left(t + \frac{h(j-1)}{n}\right).$$

For $j \in \{1, 2, \ldots, n\}$ define

$$X_{j,n} = B\left(t + \frac{j}{n}h\right) - B\left(t + \frac{j-1}{n}h\right),$$

and

$$S_{j,n} = X_{1,n} + \cdots + X_{j,n} = B\left(t + \frac{j}{n}h\right) - B(t), \quad j \in \{0, \cdots, n\}$$

Then the following are immediate,

$$S_{n,n} = \Delta_{t,t+h} \ \textit{and} \ (S_{n,n} - S_{j,n}) = B(t+h) - B\left(t + \frac{j}{n}h\right), \quad j \in \{0, \cdots, n\}$$

Set $S_{0,n} = 0$ and define $M_n^{\perp}$ by

$$M_n^{\perp} = \max_{0\leq j\leq n} \min(|S_{j,n}|, \ \ |S_{n,n} - S_{j,n}|).$$

Then we have that

$$|\Delta_{t,t+h}| = |S_{n,n}| \leq M_n = \max_{0\leq j\leq n} |S_{j,n}| \leq 3M_n^{\perp} + \max_{1\leq j\leq n} |X_{i,n}|. \tag{4.2}$$

We also have that for $\gamma = 2$,

$$\begin{aligned}
&\mathbb{E}(|S_{j,n} - S_{i,n}|^{\gamma}|S_{k,n} - S_{j,n}|^{\gamma}) \\
&= \mathbb{E}\left(\left|B\left(t+\frac{j}{n}h\right) - B\left(t+\frac{i}{n}h\right)\right|^2 \left|B\left(t+\frac{k}{n}h\right) - B\left(t+\frac{j}{n}h\right)\right|^2\right) \\
&= \frac{h^2}{n}(j-i)(k-j) = \left(\sum_{i<\ell\leq j} u_{\ell,n}\right)^{\alpha} \left(\sum_{j<\ell\leq k} u_{\ell,n}\right)^{\alpha},
\end{aligned}$$

where $\alpha = 1$ and $u_{\ell,n} = h/n$ for all $\ell \in \{1, \cdots, n\}$ and $u = u_{1,n} + \cdots + u_{n,n} = h$. So, there exists K independent of n such that for all $z > 0$,

$$\mathbb{P}\left(M_n^{\perp} \geq z\right) \leq \frac{K}{z^{2\gamma}}(u)^{2\alpha}.$$

This and Formula (4.2), gives

$$\mathbb{P}(|\Delta_{t,t+h}| \geq x) \leq \mathbb{P}\left(M_n^{\perp} \geq \frac{x}{4}\right) + \mathbb{P}\left(\max_{1\leq j\leq n} |X_j| \geq \frac{x}{4}\right), \forall x > 0.$$

which gives, for $x > 0$,

$$\mathbb{P}(|\Delta_{t,t+h}| \geq x) \leq \frac{16Kh^2}{x^{2\gamma}} + \mathbb{P}\left(\max_{1\leq j\leq n} |B\left(t+\frac{jh}{n}\right) - \left(B\left(t+\frac{(j-1)h}{n}\right)|\right) \geq \frac{x}{4}\right).$$

But, we have that the continuity modulus of B of radius $1/n$, $\delta(B, 1/n) \to 0$ a.s and thus in probability as $n \to \infty$. So,

$$\mathbb{P}(|\Delta_{t,t+h}| \geq x) \leq \frac{16Kh^2}{x^{2\gamma}} + \mathbb{P}\left(\max_{1\leq j\leq n} |\delta(B, 1/n)| \geq \frac{x}{4}\right)$$

implies

$$\mathbb{P}(|\Delta_{t,t+h}| \geq x) \leq \frac{16Kh^2}{x^4}.$$

Finally, setting $K_1 = 48K$, we have that for any $a > 0$

$$\begin{aligned}
\mathbb{E}|\Delta_{t,t+h}|^3 &= \int_0^{+\infty} \mathbb{P}(|\Delta_{t,t+h}|^3 > x) \ dx \\
&= \int_0^{a} \mathbb{P}(|\Delta_{t,t+h}| > x^{1/3}) \ dx + \int_a^{+\infty} \mathbb{P}(|\Delta_{t,t+h}| > x^{1/3}) \ dx \\
&\leq a + 16Kh^2 \int_a^{+\infty} x^{-4/3} \ dx \\
&= a + K_1 h^2 a^{-1/3}.
\end{aligned}$$

The derivative of the function $0 < a \to a + K_1 h^2 a^{-1/3}$ is $1 - \frac{K_1 h^2 a^{-4/3}}{3}$. Set $a_0 = (K_1/3)^{3/2} h^3 = K_2 h^3$, we have that the derivative is negative before a_0, positive after a_0 and zero at a_0. So, the function attains its minimum at a_0. Thus, it has a minimal value $K_3 h^{3/2}$ where $K_3 = K_2 + K_1 K_2^{-1/3}$. We now get that

$$\frac{1}{h}\mathbb{E}|\Delta_{t,t+h}|^3) \leq K_3 h^{1/2}$$

which gives us

$$\frac{1}{h}\mathbb{E}|\Delta_{t,t+h}|^3 \to 0 \quad as \ \ h \quad \to 0$$

Finally, we have that

$$\begin{aligned}
\frac{\partial g(t,u)}{\partial t} &= \lim_{h\to 0}\frac{g(t+h,u)-g(t,u)}{h}\\
&= -\frac{u^2}{2}g(t,u)+\lim_{h\to 0}\frac{1}{h}\mathbb{E}(\exp(iuB(t))c(\Delta_{t,t+h})\\
&= -\frac{u^2}{2}g(t,u)
\end{aligned}$$

This gives

$$\forall u\in\mathbb{R},\quad \frac{\partial g(t,u)}{\partial t}=-\frac{u^2}{2}g(t,u),$$

whose solution is, for u fixed,

$$\log g(t,u)=-\frac{u^2t}{2}+C(u),\qquad g(0,u)=1.$$

Thus, the constant $C(u)$ is zero and $B(t)\sim\mathcal{N}(0,t)$. The proof is over. ■

5. Elements of random Analysis using the standard Brownian motion

We are to see some important facts on the paths of the Brownian Motion $\{B(t,\omega),\ t\in\mathbb{R}_+\}$. In Analysis on real functions, at least we study the concepts of continuity, of differentiability, of total variation and of integrability of a function f and of the Riemann-Stieljes integration of a function f with respect to a df. From that analysis, we will see that that limitation of the total variation constitutes to the cause of the development of the important field of Îto's calculus which is now known as the stochastic calculus.

We already show the existence of a continuous version of a Brownian motion, called now standard Brownian motion. We will now show that the Brownian motion is nowhere differentiable.

A - Non-differentiability at right or at left of the standard Brownian motion.

Suppose that at some point $t \in \mathbb{R}_+$, $\omega \in \Omega$, that there exists a positive real number $0 < K(t,\omega)$ such that the path $B(t,\omega)$ satisfies

$$(H) \qquad -(K/2) \leq \liminf_{s \searrow t} \frac{B(s,\omega) - B(t,\omega)}{s-t} \leq \limsup_{s \searrow t} \frac{B(s,\omega) - B(t,\omega)}{s-t} \leq (K/2)$$

We note that differentiability at left implies condition (H) with equality. If we can show that condition (H) holds only on atmost a null set, we may conclude that $\{B(t,w), t \in \mathbb{R}\}$ is left differentiable on atmost a null set.

Now, fix t and suppose that (H) holds. Let $\liminf_{s \searrow t} \frac{B(s,\omega)-B(t,\omega)}{s-t} = A$. Then $A + K > A + \frac{K}{2} \geq 0$. So, there exists $\delta_1 > 0$ such that for for all $t \leq s \leq t + \delta_1$, we have

$$\frac{B(s,\omega) - B(t,\omega)}{s-t} > A - (A+K) = -K$$

Similarly, we get from the other part of (H) that there exists $\delta_2 > 0$ such that for for all $t \leq s \leq t+\delta_2$, we have

$$\frac{B(s,\omega)-B(t,\omega)}{s-t} < K$$

So, by combining the two inequalities, we get that there exists δ such that for for all $t \leq s \leq t+\delta$, we have

$$|B(s,\omega)-B(t,\omega)| \leq K|s-t|.$$

Now, pick $n \geq 1$ large enough to have

$$4 \times 2^{-n} < \delta, \quad 8K < n \quad and \quad n > t.$$

-

and we consider the partition of $\mathbb{R}_+$,

$$\{0\} + \sum_{k\geq 1} \left]\frac{k-1}{2^n}, \frac{k}{2^n}\right]$$

There exists a unique number $k_n \geq 1$ such that $t \in](k_n-1)2^{-n},\ k_n 2^{-n}]$, that is, $|k_n 2^{-n} - t| \leq 2^{-n}$. Next, for $i \in \{0,1,2,3\}$,

$$\begin{aligned} |(k+i)2^{-n}-t| &\leq |k2^{-n}-t| + i2^{-n} \\ &\leq 2^{-n} + i2^{-n} \\ &\leq 4 \times 2^{-n} < \delta. \end{aligned}$$

So,

$$|(k+i)2^{-n}-(k+i-1)2^{-n}| \leq |(k+i)i2^{-n}-t| + |(k+i-1)2^{-n}-t| \leq 8\times 2^{-n}, \quad 1 \leq i \leq 3$$

Therefore,

$$X_{kn} = \max_{1 \leq i \leq 3} |B\left(\frac{k+i}{2^n}\right) - B\left(\frac{k+i-1}{2^n}\right)| \leq 8K2^{-n} < n2^n$$

This gives

$$\max_{k \leq n2^n} X_{nk} \leq n2^{-n}.$$

Set

$$Y_n = \max_{k \leq n2^n} X_{nk} \quad and \quad A_n = (Y_n \leq n2^{-n}).$$

Then we have that $\omega \in A_n$. Thus, $D \subseteq A_n$.

Since the X_{nk}'s are independent, we have that

$$\begin{aligned}
a &= \mathbb{P}(X_{nk} \leq n2^{-n}) \\
&= \mathbb{P}(\cap_{1 \leq i \leq 3}(|B\left(\frac{k+i}{2^n}\right) - B\left(\frac{k+i-1}{2^n}\right)| \leq n2^{-n})) \\
&= \prod_{i=1}^{3} \mathbb{P}|B\left(\frac{k+i}{2^n}\right) - B\left(\frac{k+i-1}{2^n}\right)| \leq n2^{-n}) \\
&= \prod_{i=1}^{3} \mathbb{P}\left(|\mathcal{N}(0,1)| \leq n2^{-n}2^{n/2}\right) \\
&\leq \left(\int_0^{n2^{-n}2^{n/2}} \phi(x)dx\right)^3
\end{aligned}$$

where ϕ is the*pdf* of $\mathcal{N}(0,1)$-random variable which is less than or atmost equal to 1. Therefore,

$$a \leq n2^{-3n}2^{3n/2}.$$

Also,

$$\begin{aligned}
\mathbb{P}(Y_n \leq n2^n) &= \mathbb{P}\left(\max_{k \leq n2^n} X_{nk} \leq n2^{-n}\right)\\
&= n2^n \mathbb{P}\left(X_{nk} \leq n2^{-n}\right)\\
&\leq n2^n(n2^{-3n})2^{3n/2}\\
&= n^2 2^{-2n-3n/2}
\end{aligned}$$

So,

$$\mathbb{P}(A_n) \to 0 \quad as \quad n \to \infty$$

Therefore, $\mathbb{P}(D)$ is 0. Hence, D is negligible. ■

B - B is non-where of bounded Variation on any interval $[0,T]$.

Let us define as $\mathcal{P}(T)$ the set of all subdivisions c of $[0,T]$. Let us take one subdivision $t = (c_0, c_1, \cdots, c_k = T) \in \mathcal{P}(T)$, that is

$$[0,T] = \{t_0\} + \sum_{j=1}^{k}]c_{j-1}, c_j].$$

Let $F : [0,T] \to \mathbb{R}$ be a real-valued function. The variation of F over $[0,T]$ is defined by

$$V_T(F,c) = \sum_{j=1}^{k} |F(c_j) - F(c_{j-1})|.$$

Going back the Measure Theory and Integration, the Riemann-Stieltjes integral associated to F, which gives for a function $f : [0, T] \to \mathbb{R}$ the integral value

$$I = \int_0^T f(x) \; dF(x),$$

is not trivial in the sense that continuous functions f are integrable if F is of Total Variation. So we can be tempted to try to define and integral with respect to the Brownian motion path-wise, that to construct for (say an *a.s* continuous) random function $f(\omega, \circ) : [0, T] \to f(\omega, t) \in \mathbb{R}$ of the form

$$\omega \to \int_0^T f(\omega, t) \; dB(\omega, t).$$

To proceed, we need to ensure at least $B(\omega, \circ)$ has a bounded total variation over $[0, T]$ for a measurable and non-negligible set of ω.But this leads to a dead end since we have the following fact.

$$V_T(F) = \sup_{c \in \mathcal{P}(T)} V_T(F, c)$$

By definition, F has a bounded total variation over $[0, T]$ if and only if $V_T(F) < \infty$.

Proposition 6.6. *For any $T > 0$, The Brownian motion is non-where of bounded total variation over $[0, T]$.*

Before we prove this result, let us first look at the quadratic variation of the Brownian motion.

C - Quadratic variation of B over a bounded interval $[0,T]$.

Let $T>0$. The quadratic variation of B over the partition $c=(c_0,c_1,\cdots,c_k)\in\mathcal{P}(T)$ is defined as

$$V_{Q,T}(B)=\sum_{j=1}^{k}(B(c_j)-B(c_{j-1}))^2.$$

CLAIM 1. *$V_{Q,T}(B)$ converges to T in L^2.*

Proof of Claim 1. We have

$$\begin{aligned}\mathbb{E}V_{Q,T}(B) &= \sum_{j=1}^{k}\mathbb{E}\left[\frac{B(c_j)-B(c_{j-1})}{\sqrt{c_j-c_{j-1}}}\right]^2(c_j-c_{j-1})\\ &= \sum_{j=1}^{k}(c_j-c_{j-1})^2=T\end{aligned}$$

and

$$
\begin{aligned}
\mathbb{E}V_{Q,T}^2(B) &= \sum_{j=1}^{k} \mathbb{E}\left[\frac{B(c_j)-B(c_{j-1})}{\sqrt{c_j-c_{j-1}}}\right]^4 (c_j-c_{j-1})^2 \\
&+ \sum_{1\le i\ne j\le k} \mathbb{E}\left(\left[\frac{B(c_j)-B(c_{j-1})}{\sqrt{c_j-c_{j-1}}}\right]^2\left[\frac{B(c_i)-B(c_{i-1})}{\sqrt{c_i-c_{i-1}}}\right]^2\right)(c_i-c_{i-1})(c_j-c_{j-1}) \\
&= 3\sum_{i=1}^{k}(c_i-c_{i-1})^2 + \sum_{i\ne j}(c_j-c_{j-1})(c_i-c_{i-1}) \\
&= 2\sum_{i=1}^{k}(c_i-c_{i-1})^2 + \left[\sum_{j=1}^{k}(c_j-c_{j-1})\right]^2 \\
&= 2\sum_{i=1}^{k}(c_i-c_{i-1})^2 + T^2.
\end{aligned}
$$

So,

$$
\begin{aligned}
\mathbb{E}(V_{Q,T}-T)^2 &= \mathbb{E}(V_{Q,T}^2) - 2T\mathbb{E}V_{Q,T} + T^2 \\
&= \mathbb{E}(V_{Q,T}^2) - 2T^2 + T^2 \\
&= 2\sum_{j=1}^{k}(c_j-c_{j-1})^2 \\
&\le 2\sup_{1\le j\le k}(c_j-c_{j-1})\sum_{j=1}^{k}(c_j-c_{j-1}) \\
&= 2Tm(c),
\end{aligned}
$$

where $m(c) = \sup_{1\le j\le k}(c_j-c_{j-1})$ is the modulus of the partition. So, when $m(c)\to 0$, we have that

$$\mathbb{E}(V_{Q,T}-T)^2 \le 2Tm(c) \to 0. \ \square$$

Therefore, $\|V_{Q,T}-T\| \rightarrow 0$. Hence, $V_{Q,T}(B)$ converges to T in L^2.

Proof of Proposition6.6. Suppose for contradiction that B is of total bounded variation, then $V_T(B) < \infty$. Let us take the partition $c=(c_0, c_1, \cdots, c_k) \in \mathcal{P}(T)$ of modulus $1/k$. Then,

$$\begin{aligned} V_{Q,T}(B) &= \sum_{j=1}^{k} \ (B(c_j) - B(c_{j-1}))^2 \\ &\leq \sup_{1\leq j\leq k} |B(c_j) - B(c_{j-1})| \sum_{j=1}^{k} \ B(c_j) - B(c_{j-1}) \\ &\leq |\delta(B, 1/k)| V_T(B) \end{aligned}$$

So, as the modulus of the partition $1/k \rightarrow 0$, we have that $\delta(B, 1/k) \rightarrow 0$. This gives us $V_{Q,T}(B) \rightarrow 0$ a.s. This is a contradiction since $V_{Q,T}(B)$ goes to T in L_2, if it converges a.s., then the limits must coincide.

E - Another special representation of the continuous Brownian Motion,
We have the third probability law preserving transform for the Brownian Motion as follows :
Let $(B_t)_t$ be a standard Brownian motion. Then, the process $(B_t^{(3)})_t$ defined by

$$B^{(3)}(t) = tB(t^{-1}) \quad t > 0$$

and for $t = 0$, $B^{(3)}(0) = 0$ is also a standard Brownian motion.

5.1. Wienner Integral.

(a) Preparations.

in this section, we are going to see how the Wienner integral which is going to lead us to stochastic integration is constructed. As in the case of the usual Lebesgue integral, we are also going to start with the class of elementary functions here. Let $a < b$ be two real numbers. Consider $([a,b], \mathcal{B}([a,b])$. It is known that elementary functions on $[a,b]$ which we shall denote by $\mathcal{E}(a,b)$ are functions of the form

$$f = \sum_{1 \leq j \leq p} \alpha_j 1_{A_j}, \tag{5.1}$$

where $(A_j)_{1 \leq j \leq p} \subseteq \mathcal{B}([a,b])$ partitions $[a,b]$ and $(\alpha_j)_{1 \leq j \leq p} \subseteq \mathbb{R}$ and p is some natural number.

We define the subclass $\mathcal{E}_I(a,b)$ of $\mathcal{E}(a,b)$ to be elementary functions of the form

$$f = \sum_{0 \leq j \leq p-1} \alpha_j 1_{]a_j, a_{j+1}]},$$

where $a_0 = a$, $a_p = b$ and

$$[a,b] = \{a\} + {}_{0 \leq j \leq p-1}]a_j, a_{j+1}].$$

Before we begin the construction, let us see the following lemma.

LEMMA 6.7. *$\mathcal{E}_I(a,b)$ is dense in $L^2([a,b],\ \mathcal{B}([a,b]),\ \lambda)$ where λ is the Lebesgue measure on $\mathbb{R}$.*

Proof. Let $f \in: L^2([a,b],\ \mathcal{B}([a,b]),\ \lambda)$. Let $\varepsilon > 0$. We find $f^* \in \mathcal{E}_I(a,b)$ such that

$$\|f - f^*\| < \varepsilon.$$

Since f is measurable, we have that there exist a sequence $(f_n)_n$ of elementary functions satisfying $f_n \to f$ as $n \to \infty$. We also have that for each n, $|f_n|^2 \in \mathcal{E}(a,b)$ and $|f_n|^2 \to |f|^2$

So, there is a natural number N such that $\forall n \geq N$,

$$\begin{aligned} |f_n|^2 &\leq (|f_n - f| + |f|)^2 \\ &\leq 2(|f_n - f|^2 + |f|^2) \\ &< 2(1 + |f|^2) \end{aligned}$$

This means $|f_n|^2$ is bounded by an integrable function. By the Dominated Convergence theorem, we have that

$$\|f_n\|_2^2 \to \|f\|_2^2 \quad as \;\; n \to +\infty.$$

Since $f_n^2 \to f^2$ a.s. and thus in probability, by classical rules of convergence in L^2, we get

$$f_n \to f \;\; in \;\; L^2.$$

So, there exists $n_1 \geq 1$ such that

$$\|f - f_{n_1}\| < \varepsilon/2. \qquad (A1)$$

Let us suppose that

$$f_{n_1} = \sum_{1 \leq j \leq p} \alpha_j 1_{A_j},$$

with $\alpha_j > 0$ for each j and $\alpha_j = 0$ on $(\cup_{1 \leq j \leq p} A_j)^c$.

Since $\lambda(A_j) < +\infty$ for each j, we have that there exists $I_j \subset [a, b]$ a finite sum of intervals such that

$$\lambda(A_j \Delta I_j) \leq \frac{\varepsilon^2}{(4p|\alpha_j|^2)}.$$

Set

$$h = \sum_{1 \leq j \leq p} \alpha_j 1_{I_j},,$$

Then,

$$\begin{aligned} |f_{n_1} - h| &\leq \sum_{1 \leq j \leq p} |\alpha_j| |1_{A_j} - 1_{I_j}| \\ &\leq \sum_{1 \leq j \leq p} |\alpha_j| 1_{A_j \Delta I_j} \end{aligned}$$

So,

$$\begin{aligned} \|f_{n_1} - h\|^2 &= \int_{[a,b]} |f_{n_1} - h|^2 d\lambda \\ &\leq \sum_{1 \leq j \leq p} (\alpha_j)^2 \lambda(A_j \Delta I_j) \\ &\leq \varepsilon^2/4. \end{aligned}$$

i.e.,

$$\|f_{n_1} - h\| < \frac{\varepsilon}{2}$$

Therefore,

$$\|f - h\| \leq \|f - f_{n_1}\| + \|f_{n_1} - h\| < \varepsilon.$$

Finally by opening the intervals in the expression of h at left and by closing them at right and eventually the one-point intervals, h is transformed into $f^* \in \mathcal{E}_I(a,b)$ that is equal to f^* outside a countable set, that is $h = f^*$ λ-a.e. and hence

$$\|f - f^*\| < \varepsilon.$$

Hence, $\mathcal{E}_I(a,b)$ is dense in L^2 in the sense of the L^2-convergence. $\square$

(b) Wienner Integrals of elements of $\mathcal{E}_I$.

Definition 6.8. *Let $h \in \mathcal{E}_I(a,b)$. Assume h has the representation*

$$h = \sum_{0 \leq j \leq p-1} \alpha_j 1_{]a_j, a_{j+1}]}.$$

The Wienner integral of h is defined as

$$\int_a^b h(t)\ dB(t) = \sum_{0 \leq j \leq p-1} \alpha_j (B(a_{j+1}) - B(a_j)).$$

Properties of the Wienner integral on $\mathcal{E}_I$.

(i) The definition does not depend on the representation of h. As in the case of the Lebesgue integral, we may use the superpositions of the subdivisions to show the coherence of the definition.

(ii) The integral is linear on $\mathcal{E}_I$.

(iii) Most importantly, we have that for any $h \in \mathcal{E}_I$

$$\mathbb{E}\left(\int_a^b h(t) \ \ dB(t)\right)^2 = \int_a^b h(t)^2 d\lambda(x),$$

i.e.,

$$\| \int_a^b h(t) \ \ dB(t)\|_{L^2(\Omega,\mathcal{A},\mathbb{P})} = \|f\|_{L^2([a,b],\mathcal{B}([a,b]),\lambda)}.$$

We say that the Wienner integral is an isometry from $L^2([a,b],\mathcal{B}([a,b]),\lambda)$ to $L^2(\Omega,\mathcal{A},\mathbb{P})$.

Proof of Property (ii) and (iii).

(iii) Let $h \in \mathcal{E}_I$ such that h has the representation

$$h = \sum_{0 \leq j \leq p-1} \alpha_j 1_{]a_j,a_{j+1}]}.$$

the wienner integral of h is a centered Gaussian random variable as a linear combination of centred Gaussian random variables.

Let Z be the Wienner integral of h. Then,

$$
\begin{aligned}
\mathbb{E}Z^2 &= \sum_{0\leq j\leq p-1} \alpha_j^2 \mathbb{E}\left(\frac{B(a_{j+1})-B(a_j)}{\sqrt{a_{j+1}-a_j}}\right)^2 (a_{j+1}-a_j)\\
&+ \sum_{0\leq h\neq j\leq p-1} \alpha_j\alpha_h \mathbb{E}(B(a_{h+1})-B(a_h)) \times \mathbb{E}(B(a_{j+1})-B(a_j))\\
&= \sum_{0\leq j\leq p-1} \alpha_j^2 (a_{j+1}-a_j)\\
&= \int h^2 \ d\lambda
\end{aligned}
$$

Therefore,

$$
\| \int_a^b h(t) \ \ dB(t) \|_{L^2(\Omega,\mathcal{A},\mathbb{P})} = \|f\|_{L^2([a,b],\mathcal{B}([a,b]),\lambda)}.
$$

(c) Wienner Integral of a real-valued square integrable function on $[a,b]$.

Definition 6.9. *For any $f \in L^2([a,b],\mathcal{B}([a,b]),\lambda)$, then there exists a sequence $(f_n)_{n\geq 1} \subset\in \mathcal{E}_I$ such that*

$$
f_n \rightarrow_{L^2([a,b],\mathcal{B}([a,b]),\lambda)} f \ as \ n \rightarrow +\infty.
$$

The Wienner integral of f, denoted

$$
\int_{[a,b]} f \ dB(t)
$$

is given as the limit

$$
\int_{[a,b]} f_n \ dB(t) \rightarrow_{L^2(\Omega,\mathcal{A},\mathbb{P})} \int_{[a,b]} f \ dB(t). \ as \ n \rightarrow +\infty.
$$

The Wienner integral of $f \in L^2([a,b],\mathcal{B}([a,b]),\lambda)$ is a centered Gaussian Random variable with variance

$$\|f\|^2_{L^2([a,b],\mathcal{B}([a,b]),\lambda)}.$$

We now show that the Wienner integral is well defined. To show this, we need to show two things:

(i) The sequence $\int_a^b f_n(t)\ dB(t)$ converges in $L^2(\Omega,\mathcal{A},\mathbb{P})$

(ii) If there is another sequence $(g_n)_{n\geq 1} \subset \mathcal{E}_I$ such that $g_n \to f$ in L^2, then

$$\lim_{n\to+\infty}\int_a^b f_n(t)\ dB(t) = \lim_{n\to+\infty}\int_a^b g_n(t)\ dB(t)..$$

proof (i) To show that the sequence $\int_a^b f_n(t)\ dB(t)$ converges, it suffices to show that it is Cauchy. Let $n \geq 1$ and $r \geq 0$, we have

$$\begin{aligned}
\mathbb{E}\left(\int_{[a,b]} f_n\ \ dB(t) - \int_{[a,b]} f_{n+r}\ \ dB(t)\right)^2 &= \mathbb{E}\left(\int_{[a,b]}(f_n - f_{n+r})(t)\ \ dB(t)\right)^2 \\
&= \int_{[a,b]}(f_n - f_{n+r})^2(t)d\lambda(t) \qquad (L2) \\
&= \|f_n - f_{n+r}\|^2_{L^2([a,b])}.
\end{aligned}$$

Since $(f_n)_n$ is convergent in $L^2([a,b])$, we have that it is Cauchy. So, $\|f_n - f_{n+r}\| \to 0$ as $n \to \infty$. Therefore, $\|\int_{[a,b]} f_n\ \ dB(t) - \int_{[a,b]} f_{n+r}\ \ dB(t)\|^2 \to 0$. Hence, the sequence is Cauchy.

(ii)

$$\begin{aligned}
\|\int_{[a,b]} f_n \ dB(t) - \int_{[a,b]} g_n \ dB(t)\|^2 &= \mathbb{E}\left(\int_{[a,b]} f_n \ dB(t) - \int_{[a,b]} g_n \ dB(t)\right)^2 \\
&= \mathbb{E}\left(\int_{[a,b]} (f_n - g_n)(t) \ dB(t)\right)^2 \\
&= \int_{[a,b]} (f_n - g_n)^2(t) d\lambda(t) \\
&= \|f_n - g_n\|^2_{L^2([a,b])} \\
&\leq 2\left(\|f_n - f\|^2_{L^2([a,b])} + \|g_n - f\|^2_{L^2([a,b])}\right).
\end{aligned}$$

So, if Z_1 is the limit of $\int_{[a,b]} f_n \ dB(t)$ and Z_2 is the limit of $\int_{[a,b]} g_n \ dB(t)$, we have that

$$\|Z_1 - Z_2\| \leq \|Z_1 - \int_{[a,b]} f_n \ dB(t)\| + \|Z_2 - \int_{[a,b]} g_n \ dB(t)\| + \|\int_{[a,b]} f_n \ dB(t) - \int_{[a,b]} g_n \ dB(t)\| \to 0$$

Hence, $Z_1 = Z_2$ a.s.

From (i) and (ii), we conclude that the definition of the Wienner integral is coherent.
But, as a consequence of the convergence of $(f_n)_{n\geq 1}$ to f in $L^2([a,b], \mathcal{B}([a,b])$ we also have as a second fact :

$$\sigma_n^2 = \|f_n\|^2_{L^2([a,b],\mathcal{B}([a,b]),\lambda)} \to \|f\|^2_{L^2([a,b],\mathcal{B}([a,b]),\lambda)} = \sigma^2,$$

From (i), we conclude that there exists $Z \in L^2(\Omega, \mathcal{A}, \mathbb{P})$ such that

$$Z_n = \int_{[a,n]} f_n \ dB(t) \to_{L^2(\Omega,\mathcal{A},\mathbb{P})} Z.$$

From the second fact, by using the characteristic functions for the weak convergence (implied by the L^2-convergence), we conclude that the sequences centered Gaussian random variables with respective variances $(\sigma_n^2)_{n\geq 1}$ converges weakly to a centered Gaussian random variable with variance σ^2.

We are now going to see some interesting properties of the Wienner integral.

Proposition 6.10. *We have the following properties of the Wienner integral.*

(a) The Wienner integral is linear.

(b) The Wienner integral on $[a,b]$is an isometry, in the following sense : for any $f \in L^2([a,b],\mathcal{B}([a,b]),\lambda)$,

$$\| \int_a^b f(t)\ dB(t) \|_{L^2(\Omega,\mathcal{A},\mathbb{P})} = \|f\|_{L^2([a,b],\mathcal{B}([a,b]),\lambda)}, \lambda. \tag{5.2}$$

(c) The Wienner integral preserves the inner product : $(f,g) \in L^2([a,b],\mathcal{B}([a,b]),\lambda)^2$,

$$\mathbb{E}\biggl(\biggl(\int_a^b f(t)\ dB(t) \biggr) \biggl(\int_a^b g(t)\ dB(t) \biggr) \biggr) = \int_{[a,b]} f(x) g(x)\ d\lambda(x). \tag{5.3}$$

(d) If f is a continuous function of bounded variation, the Wienner integral can be computed as a Riemann-Stieltjes integral through the integration by parts formula

$$(WI) - \int_{[a,b]} f(t)\ dB(t) = \Bigl[f(t)B(t) \Bigr]_a^b - \biggl((RS) - \int_{[a,b]} B(\omega) df(t) \biggr)$$

where (WI) stands for the Wiener integral and (RS) for the Riemann-Stieltjes integral.

Proof.

Proof of Point (c). Let $(f,g) \in L^2([a,b]$, and let $W(f)$ denote the Wienner integral of f and that of g, $W(g)$. By applying the linearity property and then the isometry property of the Wienner integral, we get that

$$\begin{aligned}
\mathbb{E}(W(f)+W(g))^2 &= \mathbb{E}(W(f+g))^2 \\
&= \int_{[a,b]} (f+g)(t) \ \ d\lambda(t) \\
&= \|f\|^2_{L^2([a,b]} + \|2\|^2_{L^2([a,b]} + 2\int_{[a,b]} f(t)g(t)d\lambda(t).
\end{aligned}$$

But by linearity of expectation and again, isometry, we have that

$$\begin{aligned}
\mathbb{E}(W(f)+W(g))^2 &= \mathbb{E}W(f)^2 + \mathbb{E}W(g))^2 + 2\mathbb{E}W(f)W(g) \\
&= \|f\|^2_{L^2([a,b]} + \|g\|^2_{L^2([a,b]} + 2\mathbb{E}W(f)W(g).
\end{aligned}$$

Equating the two expressions of $\mathbb{E}(W(f)+W(g))^2$, we get our conclusion.

CHAPTER 7

Poisson Stochastic Processes

In this chapter, we are going to study the Poisson stochastic from three points of view. This stochastic process is mainly used to described occurrence times of some random events over a continuous time. Here are some general examples :

(a) Arrival times to a desk in some bank.

(b) Occurrence times of failure of machines in a company.

(c) Arrival times of tasks to the central unity of a computer.

(d) etc.

Let us introduce the stochastic process from several points of view.

1. Description by exponential inter-arrival

Let us consider a bank desk which opens at 08H00 for example. Assuming that at that time there is no clients and we denote $Z = 0$ at that initial time taken as $t = 0$. In this model we want simple, clients arrive one after another at random times [Later, models in which several clients might come at the same times will be studied],

$$0 < Z_1 < Z_2 < .... < Z_n < ...$$

so that the inter-arrival times

$$X_1 = Z_1 - Z_0, X_2 = Z_2 - Z_1, X_3 = Z_3 - Z_2, ..$$

are independent and follow an exponential of parameters $\lambda > 0$, denoted

$$X_1, X_2, ...iid \sim \mathcal{E}(\lambda). \tag{1.1}$$

Let us begin by giving the finite-distributions of the sequences $(X_n)_{n\geq 1}$ and $(Z_n)_{n\geq 1}$. This will allow to derive the characteristics of this important stochastic process.

1.1. Probability law of arrival times and that of the inter-arrival times.

We have for all $n \geq 1$

$$f_{(X_1,...,X_n)}(x_1, ..., x_n) = \lambda^n \exp(-\lambda \sum_{i=1}^{n} x_i) \ 1_{(x_1 \geq 0,...,x_n \geq 0)}. \tag{1.2}$$

From there, we can derive the *pdf* of $Z^{(n)} = {}^t(Z_1, ..., Z_n)$. Indeed we have :

Proposition 7.1. *The finite-distributions of arrival times* $Z_1 < Z_2 < .... < Z_n < ...$ *are given as follows. For all* $n \geq 1$*,*

(1.3) $$f_{(Z_1,...,Z_n)}(z_1, ..., z_n) = \lambda^n \exp(-\lambda z_n) 1_{\Gamma_n}(z)$$

with

$$\Gamma_n = \{z = (z_1, ..., z_n) \in \mathbb{R}^n; 0 < z_1 < z_2 < ... < z_n\}.$$

Each $i \geq 1, Z_i$ *follows a Gamma of parameters* (i, λ) *:* $Z_i \sim \gamma(i, \lambda)$*.*

Proof. Let $n \geq 1$, each of the two $Z^{(n)} = Z^t(Z_1, ..., Z_n)$ and $X^{(n)} =^t (X_1, ..., X_n)$ are linear transformations of the others:

$$\begin{cases} Z_1 = & X_1 \\ Z_2 = & X_1 + X_2 \\ \ \ ... & ... \\ Z_n = & X_1 + X_2 + ... + X_n \end{cases} \Longleftrightarrow \begin{cases} X_1 = & Z_1 - Z_2 \\ X_2 = & Z_2 - Z_1 \\ \ \ ... & ... \\ X_n = & Z_n - Z_{n-1} \end{cases}.$$

Let us denote the matrices in the linear transformations above by A and B and we write

$$Z^{(n)} = AX^{(n)} \Leftrightarrow X^{(n)} = BZ^{(n)}.$$

The support of $(X_1, ..., X_n)$ is $D_n = \mathbb{R}_+^n$ and that of $(Z_1, ..., Z_n)$ is

$$\Gamma_n = \{z = (z_1, ..., z_n) \in \mathbb{R}^n; 0 < z_1 < z_2 < ... < z_n\}.$$

The mapping $X^{(n)} = BZ^{(n)}$ is a diffeomorphism between D_n and Γ_n and the jacobian determinant satisfies $|\det(B)| = 1$, The change of variables (see Lo (2018), Chapter 3) entails

(1.4) $$f_{(Z_1,...,Z_n)}(z_1, ..., z_n) = f_{(X_1,...,X_n)}(x_1, ..., x_n) \ \ 1_{\Gamma_n}(z_1, ..., z_n)$$

with

(1.5) $$(x_1, ..., x_n) = B(z_1, ..., z_n) = (z_1 - z_0, z_2 - z_1, ..., z_n - z_{n-1}).$$

By plugging (1.5) in (1.2), (1.4) becomes

$$f_{(Z_1,...,Z_n)}(z_1, ..., z_n) = \lambda^n \exp(-\lambda z_n) \ \ 1_{\Gamma_n}(z_1, ..., z_n).$$

The first approach is finished. $\square$

Let us go the second approach.

2. Counting function

Let use the independent and λ-exponentially distributed inter-arrival times X_j, $j \geq 1$ corresponding to the arrival times $Z_0 = 0 < Z_2 < .... < Z_n < ....$ We define the counting stochastic processes $N(t)$ as the number of arrivals up to time $t \geq 1$:

$$\forall t \geq 0, \quad N_t = N(]0,t]) = \#\{i \geq 1, Z_i \leq t\}.$$

We easily can express events of the form $(N_t = k)$ in function the arrival times Z_i. Indeed, we have

$$\forall t \geq 0, \quad (N_t = k) = (Z_k \leq t < Z_{k+1})$$

We also can write

$$N_t = \sum_{k \geq 1} 1_{]0,t]}(Z_i).$$

The counting function itself is called the Poisson process of intensity $\lambda > 0$.

We are going to give the characteristic properties of the counting function. To do short, we propose the following formula [which will be used in the computations later] as an exercise (see solution in page **??**)

Exercise. For all $k \geq 1$, show that

$$\int_{(0<z_1<z_2<...<z_k\leq t)} dz_1...dz_k = \frac{t^k}{k!}.$$

Here are interesting properties of the counting process.

2.1. One-dimensional margins of the counting process.

For $t=0$, we have $N_0=0$. For $t>0$, we have for $k\geq 0$,

$$\begin{aligned}
\mathbb{P}(N_t &= k) = \mathbb{P}(Z_k \leq t < Z_{k+1}) \\
&= \mathbb{P}(Z_1 \leq Z_2 \leq ... \leq Z_k \leq t < Z_{k+1}).
\end{aligned}$$

From there, we use the *pdf* of $(Z_1, Z_2, ..., Z_k, Z_{k+1})^t$ given in Formula (1.3) and the Fubini's formula to get

$$\begin{aligned}
\mathbb{P}(N_t &= k) = \lambda^{k+1} \int_{z_1\leq z_2 \leq ... \leq z_k \leq t < z_{k+1}} \exp(-\lambda z_{k+1}) dz_1....dz_k dz_{k+1} \\
&= \lambda^{k+1} \int_{z_1\leq z_2 \leq ... \leq z_k \leq t} dz_1....dz_k \int_{\leq t < z_{k+1}} \exp(-\lambda z_{k+1}) dz_{k+1} \\
&= \lambda^{k+1} \int_{z_1\leq z_2 \leq ... \leq z_k \leq t} dz_1....dz_k \left[-\frac{\exp(-\lambda z_{k+1})}{\lambda} \right]_{z_{k+1}=t}^{z_{k+1}=+\infty} \\
&= \lambda^k \exp(-\lambda t) \times \int_{z_1\leq z_2 \leq ... \leq z_k \leq t} dz_1....dz_k.
\end{aligned}$$

By using the exercise above, we arrive at

$$\mathbb{P}(N_t = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!},$$

which proves that $N(t) \sim \mathcal{P}(\lambda t)$. $\square$

This means that each margin N_t follows a Poisson Law of parameter λt : $N_t \sim \mathcal{P}(\lambda t)$.

The name *Poisson* given to that stochastic process certainly derives from that fact.

2.2. finite-distribution laws.

Proposition 7.2. *For all $k \geq 1$, for all $t_0 = 0 < t_1 < \ldots < t_k$, for all $(n_1, \ldots, n_k) \in \mathbb{N}^k$, we have (with the convention that $n_0 = 0$),*

(2.1)

$$\mathbb{P}(N_{t_1} = n_1, \ldots, N_{t_k} = n_k) = \prod_{j=1}^{k} \frac{(\lambda(t_j - t_{j-1}))^{n_j - n_{j-1}}}{\left(n_j - n_{j-1}\right)!} \exp(-\lambda(t_j - t_{j-1})) 1_{(0 \leq n_1 \leq n_2 \leq \ldots \leq n_k)}.$$

or, in other words,

(2.2)

$$\mathbb{P}(N_{t_1} = n_1, \ldots, N_{t_k} = n_k) = \exp(-\lambda t_k) \prod_{j=1}^{k} \frac{(\lambda(t_j - t_{j-1}))^{n_j - n_{j-1}}}{\left(n_j - n_{j-1}\right)!} 1_{(0 \leq n_1 \leq n_2 \leq \ldots \leq n_k)}.$$

Before we give the proof, we should remark this important fact about the memory loss property.

Remark (R) Based on the independence of the *iidness* inter-arrival times, we surely have that given $N_t = k$ at time t, the $N^*(u) =$

$N(t+u)-k$, $u \geq 0$ in again a counting process based on the interval times $Z_0^* = 0$, $Z_i^* = Z_{i+k}^* - k$, $i \geq 1$, with inter-arrival-times $Z_{i+1}^* - Z_i^* = Z_{k+i+1}^* - Z_{k+i}^*$ *iid* $\sim \mathcal{P}(\lambda)$, independent to events based on $(Z_1, ..., Z_k)$. So, we have

$$N_u \stackrel{d}{=} N_{t+u} - N_t.$$

As well, $N_t = k$, $N_{t+u} - N_t$ does not depend on k. We can generalize that for any $k \geq 1$, for any $t_0 = 0 < t_1 < ... < t_k$, we have

$$(N_{t_1} = n_1, ..., N_{t_k} = n_k) = (N_{t_1} - N_{t_0} = n_1 - n_0, N_{t_2} - N_{t_1} = n_2 - n_1 ..., N_{t_k} - N_{t_{k-1}} = n_{k-1} - n_k),$$

From there, we may get the following facts :

(a) Each $N_{t_j} - N_{t_{j-1}}$, $1 \leq j \leq k$, has the same law as $N_{t_j - t_{j-1}}$, that is $\mathcal{P}(\lambda(t_j - t_{j-1}))$.

(b) The $N_{t_j} - N_{t_{j-1}}$ are independent.

Since N_t non-decreasing in $t \geq 1$, the event $(N_{t_1} = n_1, ..., N_{t_k} = n_k)$ is possible if only $n_0 = 0 \leq n_1 \leq n_2 \leq ... \leq n_k$, So we get

$$\mathbb{P}(N_{t_1} = n_1, ..., N_{t_k} = n_k) = \prod_{j=1}^{k} \frac{(\lambda(t_j - t_{j-1}))^{n_j - n_{j-1}}}{\binom{n_j - n_{j-1}}{}!} \exp(-\lambda(t_j - t_{j-1})) 1_{(0 \leq n_1 \leq n_2 \leq ... \leq n_k)},$$

which entails

$$\mathbb{P}(N_{t_1} = n_1, ..., N_{t_k} = n_k) = \exp(-\lambda t_k) \prod_{j=1}^{k} \frac{(\lambda(t_j - t_{j-1}))^{n_j - n_{j-1}}}{(n_j - n_{j-1})!} 1_{(0 \leq n_1 \leq n_2 \leq ... \leq n_k)}.$$

From these facts, let us characterize the counting stochastic process.

2.3. Characterization of the counting process.

We already proved this.

Proposition 7.3. *The counting stochastic process satisfies the following properties.*

(IC)[Initial condition] $N_0 = 0$ *a.s.*

(PM) [Poisson margins] The margins N_t, $t \geq 1$, *follow Poisson laws* $\mathcal{P}(\lambda t)$.

(SI) [Stationarity of increments] N_t *has strong stationary increments, that is, for any* $k \geq 2$, *for any* $t_0 = 0 < t_1 < ... < t_k$,

$$(N_{t_j} - N_{t_{j-1}},\ 1 \leq j \leq k-1)^t =_d (N_{t_j - t_{j-1}},\ 1 \leq j \leq k-1)^t.$$

(I2) [Independent increments] N_t *has independent increments : for any* $k \geq 2$, *for any* $t_0 = 0 < t_1 < ... < t_k$, *the random variables* $N_{t_1} - N_{t_0}$, $N_{t_2} - N_{t_1}$, $\cdots$, $N_{t_k} - N_{t_{k-1}}$ *are independent.*

As explained in **?**, *Chapter 2 [Definitions of types of stationary], with condition (I2) and the fact that* $N(0) = 0$, *Property (SI) is equivalent to*

$$\forall t \geq 0,\ u \geq 1,\ \ N(t+u) - N(t) =_d N(u).$$

(RC) [Paths right-continuity] The paths $t \hookrightarrow N_t(\omega)$ *are right continuous.*

Le point (RC) comes of the fact that for any $k \geq 0$ and for any $t \geq 0$,

$$(N_t = k) = (Z_k \leq t < Z_{k+1}).$$

entails that for any $\omega \in \Omega$ and for $s_0 > 0$ so small that $t + s_0 < Z_{k+1}(\omega)$, we still have for any $0 \leq s < s_0$

$$(Z_k(\omega) \leq t < t + s < Z_{k+1}(\omega)),$$

and $N_{\omega,t+s} = N(\omega, t)$ for $0 \leq s < s_0$. $\square$

In terms of Probability laws, we will have the Properties (IC), (PM), (SI) and (I2) do exactly characterize the probability law of $N(t)$, A stochastic process $N(t)$ with time $t \geq 0$ and sate space $\mathbb{N}$ will be called a Poisson process of intensity $\lambda > 0$ if Conditions (IC), (PM), (SI) and (I2) hold. If, on top of them, (RC) holds *a.s.*, we say that a **standard** Poisson process of intensity $\lambda > 0$.

Let us use the Kolmogorov Existence Theorem ***(KET)*** to give birth the Poisson process from another point of view. But before we proceed to that, let us make the following remarks.

(A) The finite-distributions in Formula (2.1) hold if Conditions (PM), (SI) and (I2) hold.

(B) If the finite-distribution of a stochastic process process (N_t) are given by Formula (2.1), we easily prove the assertions (IM), (SI), $(I2)$ by the very simple definition of the independence (seeLo (2018), Chapter 2).

Let use the Kolmogorov approach.

3. Approach of the Kolmogorov Existence Theorem

From a probabilistic point of view, the best approach begins with the Kolmogorov's theorem with the characterization of finite distributions. However, this approach does not always guarantee properties we want to have for the trajectories. The approach of Kolmogorov only guarantees the probability laws. For example, the right-continuity of the paths of $(N_t)_{t\geq 0}$ must be established beyond the Kolmogorov's theorem to build a right-continuous version. We already did this for the Brownian movement for which we created a continuous version.

3.1. Construction of the Counting of the Poisson Process by the *KET*.

Theorem 7.4. Consider the family of discrete finite-distributions determined by their discrete *pdf*'s : for $k \geq 1$, for $t_0 = 0 < t_1 < ... < t_k$, par

(3.1)
$$f_{(t_1,...,t_k)}(n_1,...,n_k) = \prod_{j=1}^{k} \frac{(\lambda(t_j - t_{j-1}))^{n_j - n_{j-1}}}{(n_j - n_{j-1})!} \exp(-\lambda(t_j - t_{j-1}))1_{(0\leq n_1 \leq n_2 \leq ... \leq n_k)}.$$

Then this family is consistent and there exist a stochastic process

$$(\Omega, A, \mathbb{P}, (N_t)_{t\geq 0}, \mathbb{R}_+, \mathbb{N})$$

of finite-distributions defined by Formula (3.1) and that Properties (IC), (PM), (SI) and (I2) holds.

Proof. Let us establish the consistency condition (CHSD2) (see page **??**, Chapter **??**). We have to prove that for any $k \geq 1$, for any $t_0 = 0 < t_1 < ... < t_k < t_{k+1}$

$$\int_{\mathbb{N}} f_{(t_1,...,t_k,t_{k+1})}(n_1,...,n_k,n_{k+1})d\nu(n_{k+1}) = f_{(t_1,...,t_k,t_{k+1})}(n_1,...,n_k),$$

that is

$$\sum_{n_{k+1}\geq 0} f_{(t_1,...,t_k,t_{k+1})}(n_1,...,n_k,n_{k+1}) = f_{(t_1,...,t_k,t_{k+1})}(n_1,...,n_k).$$

Let us treat the nontrivial case: $0 \leq n_1 \leq n_2 \leq ... \leq n_k$, otherwise the equation is simply $0 = 0$. Furthermore, the terms $f_{(t_1,...,t_k,t_{k+1})}(n_1,...,n_k,n_{k+1})$ are null for $n_{k+1} < n_k$. So, $\sum_{n_{k+1}\geq 0} f_{(t_1,...,t_k,t_{k+1})}(n_1,...,n_k,n_{k+1})$ is equal to :

$$\begin{aligned}
&= \sum_{n_{k+1}\geq n_k} f_{(t_1,...,t_k,t_{k+1})}(n_1,...,n_k,n_{k+1}) \\
&= \sum_{n_{k+1}\geq n_k} \prod_{j=1}^{k+1} \frac{(\lambda(t_j-t_{j-1}))^{n_j-n_{j-1}}}{(n_j-n_{j-1})!} \exp(-\lambda(t_j-t_{j-1})) \\
&= \prod_{j=1}^{k} \frac{(\lambda(t_j-t_{j-1}))^{n_j-n_{j-1}}}{(n_j-n_{j-1})!} \exp(-\lambda(t_j-t_{j-1})) \\
&\times \sum_{n_{k+1}\geq n_k} \frac{(\lambda(t_{k+1}-t_k))^{n_{k+1}-n_k}}{(n_j-n_{j-1})!} \exp(-\lambda(t_{k+1}-t_k)).
\end{aligned}$$

By the change of variables $s = n_{k+1} - n_k$, we have

$$\begin{aligned}
&\sum_{n_{k+1}\geq n_k} \frac{(\lambda(t_{k+1}-t_k))^{n_{k+1}-n_k}}{(n_j-n_{j-1})!} \exp(-\lambda(t_{k+1}-t_k)) \\
&= \exp(-\lambda(t_{k+1}-t_k)) \sum_{s\geq 0} \frac{(\lambda(t_{k+1}-t_k))^s}{s!} \\
&+ \exp(-\lambda(t_{k+1}-t_k)) \times \exp(\lambda(t_{k+1}-t_k)) \\
&= 1.
\end{aligned}$$

So, we have the desired result. To finish, we apply the *KET* and Remark (R) above. □

By Remark (R), (page 117), Conditions (IC), (PM), (SI) and (I2) together determine the finite-distributions of $N(t)$ as in Formula (3.1) which in turns allows a unique extension to a Poisson Counting process N^*, which is equal to N in law. So, we have

Theorem 7.5. The counting stochastic process of a Poisson process is entirely characterized by Properties (IC), (PM), (SI) and (I2).

3.2. Right-Continuous version.

Now, we have created a Poisson counting stochastic process from the finite-distributions in (3.1), can we transform it into an *a.s* right-continuous Poisson counting stochastic process? It is possible according to **?**.

4. More properties for the Standard Poisson Process

It is time to get more deeper properties of the standard Poisson process (from the three points view of *iid* exponential inter-arrival times, its characterization by properties (IC), (PM), (SI) and (I2) and its its characterization by the finite-distributions in (3.1)).

4.1. Conditional Laws.

Proposition 7.6. *Conditionally on the event* $(N_t = n)$*, the vector* $(Z_1, ..., Z_n)$ *of arrival time has the following pdf*

$$f_{(Z_1,...,Z_n)}(z|N_t = n) = \frac{n!}{t^n} 1_{(0<z_1<z_2<...<z_n \leq t)}. \tag{4.1}$$

Proof. For all Borel subset of $\mathbb{R}^n$, we have

$$\begin{aligned} \mathbb{P}((Z_1, ..., Z_n) \ \in \ A|N_t = n) &= \frac{\mathbb{P}(Z_1, ..., Z_n) \in A, N_t = n)}{P(N_t = n)} \\ &= \frac{\mathbb{P}(Z_1, ..., Z_n) \in A, Z_n \leq t < Z_{n+1})}{P(N_t = n)}. \end{aligned}$$

Let us use the unconditional *pdf* of $(Z_1, ..., Z_{n+1})$ and The Funibi's theorem to get

$$\begin{aligned} \mathbb{P}((Z_1, ..., Z_n) \ \in \ A|N_t = n) &= \frac{\lambda^{n+1}}{\mathbb{P}(N_t = n)} \\ &\times \int_{(z_1,...,z_n)\in A, 0\leq z_1 \leq ... \leq z_n \leq t < z_{n+1}} \exp(-\lambda z_{n+1}) dz_1....dz_n dz_{n+1} \\ &= \frac{\lambda^{n+1}}{\mathbb{P}(N_t = n)} \int_{(z_1,...,z_n)\in A, 0\leq z_1 \leq ... \leq z_n \leq t} dz_1....dz_n \left(\int_{t<z_{n+1}} \exp(-\lambda z_{n+1}) dz_{n+1} \right) \\ &= \frac{\lambda^{n+1}}{\mathbb{P}(N_t = n)} \int_{(z_1,...,z_n)\in A, 0\leq z_1 \leq ... \leq z_n \leq t} dz_1....dz_n \left(\frac{\exp(-\lambda t)}{\lambda} \right) \\ &= \frac{\lambda^{n+1}}{(\lambda t)^n \exp(-\lambda t)/n!} \int_{(z_1,...,z_n)\in A, 0\leq z_1 \leq ... \leq z_n \leq t} dz_1....dz_n \left(\frac{\exp(-\lambda t)}{\lambda} \right) \\ &= \int_A \frac{n!}{t^n} 1_{\Gamma_n}(z_1, ..., z_n) \ \ dz_1....dz_n. \end{aligned}$$

By definition of the *pdf* with respect to the Lebesgue measure, we have for all $n \geq 1$,

$$d\mathbb{P}_{(Z_1,\ldots,Z_n)}(\circ|N_t = \frac{n!}{t^n}1_{\Gamma_n}.$$

4.2. Law of arrival times when the order of arrival is lost. A very interesting property of Poisson process is the following. Let us suppose that clients arrive to some bank desk and each time one clients arrives, his name is put on a sheet and his arrival time in an ordered list and the arrival time on a card. Once n clients has already arrived up to time t, they sat in a waiting room without any order. The only way to have the different arrival times is to use the ordered list. Let us suppose that the list which determines the the correspondence between the cards and their owners is lost. The bank officer decides then to take a random order of the clients and to shuffle the cards and to give a card to each client in the current order that was set by the bank officer. In that situation, we will proof that the arrival times are independent and are uniformly distributed on $[0,\ t]$.

Let S_n be the set of permutations $\{1,2,...,n\}$ endowed with the σ-algebra $\mathcal{P}(\{1,2,...,n\}) = \mathcal{P}_n$ and the discrete $\mathbb{U}$ on $(S_n, \mathcal{P}_n)$

$$\mathbb{U}(\{\sigma\}) = \frac{1}{n!},\ \ \sigma \in S_n.$$

Since the ordered arrival times is lost, a random ordering is denoted by

$$(Z_{\sigma(1)}, ..., Z_{\sigma(n)})(\omega) = \sigma(Z_1, ..., Z_n)(\omega)$$

with two elementary events : $\omega \in \Omega$ and $\sigma \in S_n$. The study takes place in the probability space

$$(S_n \times \Omega, \ \ \mathcal{P}_n \otimes \mathcal{A}, \ \ \mathbb{L}),$$

where

$$\mathbb{L} = \mathbb{U} \otimes \mathbb{P}(\cdot | N_t = n).$$

We have :

PROPOSITION 7.7. *Let us suppose that conditionally on $(N_t = n)$ and that we have lost the order of arrival times. Then the vector $(Z_1, ..., Z_n)$ has has margins independent and uniformly distributed on $[0, t]$, that is, pour tout $B \in \mathcal{B}(\mathbb{R}^n)$,*

$$\mathbb{L}(\sigma(Z_1, ..., Z_n) \in B) = \frac{1}{t^n} \int_B \prod_{i=1}^{n} 1_{[0,t]}(u_i) du_1 ... du_n.$$

Proof. Before we go further, we recall that the hyperplans of $\mathbb{R}^k$ are null sets with respect to the Lebesgue measure. Let us denote by λ_n by the Lebesgue measure on $\mathbb{R}^n$. Then we have

$$\lambda_n(\{(z_1, ..., z_n), \ \ \exists (1 \leq i \neq j \leq n), \ \ z_i = z_j\}) = 0. \tag{4.2}$$

Now we have

$$\begin{aligned}
&\mathbb{L}(\sigma(Z_1,...,Z_n)\in B)=\sum_{\sigma_0\in S_n}\mathbb{L}(\sigma(Z_1,...,Z_n)\in B,\sigma=\sigma_0)\\
&=\sum_{\sigma_0\in S_n}\mathbb{U}\otimes\mathbb{P}(\cdot|N_t=n)(\sigma=\sigma_0)\times\mathbb{U}\otimes\mathbb{P}(\cdot|N_t=n)(\sigma(Z_1,...,Z_n)\in B|(\sigma=\sigma_0))\\
&=\sum_{\sigma_0\in S_n}\mathbb{U}\otimes\mathbb{P}(\cdot|N_t=n)(\sigma=\sigma_0)\times\mathbb{U}\otimes\mathbb{P}(\cdot|N_t=n)(\sigma_0(Z_1,...,Z_n)\in B|(\sigma=\sigma_0)).
\end{aligned}$$

But $\mathbb{U}\otimes\mathbb{P}(\cdot|N_t = n)(\sigma = \sigma_0)$ does not depend on σ. Hence we use the marginal probability on $\mathbb{U}$ and we have

$$\mathbb{U}\otimes\mathbb{P}(\cdot|N_t=n)(\sigma=\sigma_0)=\frac{1}{n!}.$$

Also, we have

$$\mathbb{U}\otimes\mathbb{P}(\cdot|N_t=n)(\sigma_0(Z_1,...,Z_n)\in B|(\sigma=\sigma_0))$$

depends on ω (since σ fixed as σ_0), we utilize that the the marginal probability $\mathbb{P}(\cdot|N_t = n)$ and we have

$$\mathbb{U}\otimes\mathbb{P}(\cdot|N_t=n)(\sigma_0(Z_1,...,Z_n)\in B)=\mathbb{P}(\cdot|N_t=n)((Z_1,...,Z_n)\in\sigma_0^{-1}(B)).$$

Moreover, we have

$$\begin{aligned}\mathbb{L}(\sigma(Z_1,...,Z_n) &\in B) = \frac{1}{n!}\sum_{\sigma_0\in S_n}\mathbb{P}(\cdot|N_t=n)((Z_1,...,Z_n)\in\sigma_0^{-1}(B))\\ &= \frac{1}{n!}\sum_{\sigma_0\in S_n}\mathbb{P}(\cdot|N_t=n)((Z_1,...,Z_n)\in\sigma_0^{-1}(B)),\end{aligned}$$

since σ_0^{-1} runs over S_n if σ_0 does. By applying (4.1), we have

$$\begin{aligned}\mathbb{L}(\sigma(Z_1,...,Z_n) &\in B) = \frac{1}{n!}\int\sum_{\sigma_0\in S_n}\mathbb{P}(\cdot|N_t=n)((Z_1,...,Z_n)\in\sigma_0(B))\\ &= \frac{1}{t^n}\sum_{\sigma_0\in S_n}\int_{\sigma_0(B)}1_{(0<z_1<z_2<...<z_n\leq t)}dz_1....dz_n.\end{aligned}$$

Hence

$$\begin{aligned}\mathbb{L}(\sigma(Z_1,...,Z_n) &\in B) = \frac{1}{t^n}\sum_{\sigma_0\in S_n}\int 1_{\sigma_0(B)}(z_1...,z_n)1_{(0<z_1<z_2<...<z_n\leq t)}dz_1....dz_n\\ &= \frac{1}{t^n}\sum_{\sigma_0\in S_n}\int 1_B(\sigma_0^{-1}(z_1...,z_n))1_{(0<z_1<z_2<...<z_n\leq t)}dz_1....dz_n.\end{aligned}$$

In each integral, we make the change of variable $\sigma_0^{-1}(z_1...,z_n)=(u_1,...u_n)$. σ_0 is diffeomorphism on $\mathbb{R}^n$, is linear and the Jacobian of the diffeomorphism the $|J|=1$. Thus,

$$\begin{aligned}\mathbb{L}(\sigma(Z_1,...,Z_n) &\in B) = \frac{1}{t^n}\sum_{\sigma_0\in S_n}\int_B 1_{(0<u_{\sigma_0(1)}<u_{\sigma_0(2)}<...<u_{\sigma_0(n)}\leq t)}du_1....du_n\\ &= \frac{1}{t^n}\int_B\left\{\sum_{\sigma_0\in S_n}1_{(0<u_{\sigma_0(1)}<u_{\sigma_0(2)}<...<u_{\sigma_0(n)}\leq t)}\right\}du_1....du_n.\end{aligned}$$

But $]0,t]^n$ in the sum of

$$C_t = \{(u,...,u_n) \in]0,t]^n, \forall (1 \leq i \neq j \leq n), \quad u_i \neq u_j\}$$

and of sets included in hyperplans. We have

$$\begin{aligned} C_t &= \sum_{\sigma_0 \in S_n} \{(u_1,...,u_n) \in [0,t]^n, \quad u_{\sigma_0(1)} < u_{\sigma_0(2)} < ... < u_{\sigma_0(n)}\} \\ &= \sum_{\sigma_0 \in S_n} \{(u_1,...,u_n) \in \mathbb{R}^n, \quad 0 \leq u_{\sigma_0(1)} < u_{\sigma_0(2)} < ... < u_{\sigma_0(n)} \leq t\} \end{aligned}$$

Hence,

$$1_{C_t} = \left\{ \sum_{\sigma_0 \in S_n} 1_{(0<u_{\sigma_0(1)}<u_{\sigma_0(2)}<...<u_{\sigma_0(n)}\leq t)} \right\}$$

and thus

$$\mathbb{L}(\sigma(Z_1,...,Z_n) \in B) = \frac{1}{t^n} \int_B 1_{C_t}(u,...,u_n) \ \ du_1....du_n.$$

Finally, by (4.2), the set

$$[0,t]^n \backslash C_t = \{(z_1,...,z_n) \in [0,t]^n, \quad \exists (1 \leq i \neq j \leq n), \quad z_i = z_j\}$$

is null and that leads to

$$\begin{aligned} \mathbb{L}(\sigma(Z_1,...,Z_n) &\in B) = \frac{1}{t^n} \int_B 1_{[0,t]^n}(u,...,u_n) \ \ du_1....du_n \\ &= \int_B \left\{ \prod_{i=1}^{n} \left(\frac{1}{t} 1_{[0,t]}(u_i) \right) \right\} du_1...du_n. \end{aligned}$$

Here, we identify the *pdf* of a vector of dimension n with *iid* components uniformly distributed on $[0, t]$. $\square$

4.3. Superposition of Poisson Processes.

Proposition 7.8. *Let us assume that we have two independent Poisson processes with respective intensities λ_1 and λ_2 and with respective counting functions $N_t^{(1)}$ and $N_t^{(2)}$, taking place at the same time. By superposing the arrival times of the two processes, nous obtenons a Poisson process of intensity $\lambda = \lambda_1 + \lambda_2$.*

Proof. By superposing the arrival times of the two processes, the total number of arrival times N_t is on the form of $N_t^{(1)}$ and $N_t^{(2)}$:

$$N_t = N_t^{(1)} + N_t^{(2)}.$$

Now, each $N_t^{(i)}$, $i \in \{1, 2\}$, only depends on the arrivals times. Since the two stochastic processes $N_t^{(i)}$, $i \in \{1, 2\}$'s are independent.
So the law of their sum is the convolution product of each element od the sum. So,

$$N_t^{(1)} + N_t^{(2)} \sim \mathcal{P}((\lambda_1 + \lambda_2)t).$$

for all $t \geq 0$. Form there we easily get (IC), (PM), (SI) and (I2) listed in Subsection 2.3 (page 119). Finally, N_t is a counting process of intensity Poisson $\lambda = \lambda_1 + \lambda_2$, which right-continuous if both processes $N_t^{(i)}$, $i \in \{1, 2\}$, are. $\square$

4.4. Decomposition of Stochastic processes.

Proposition 7.9. *We consider a Poisson stochastic of arrival times of intensity $\lambda > 0$. Let us associate to the stochastic processes a Bernoulli trial of probability $p \in]0,1[$, which is independent of the arrival times. To each arrival, we perform the Bernoulli experience independently from the former performances. If we have a success, the client is directed to a desk A, and to a desk B otherwise. Let $N_t^{(1)}$ the counting functions at the desk A and by $N_t^{(2)}$ that of the desk B. We have :*

(a) $N_t^{(1)}$ *and* $N_t^{(2)}$ *are independent.*

(b) $N_t = N_t^{(1)} + N_t^{(2)}$.

(c) $N_t^{(1)}$ *is a Poisson stochastic process of intensity* $p\lambda$ *and* $N_t^{(2)}$ *Poisson stochastic process of intensity* $(1-p)\lambda$.

Proof. Let us denote the event $(N_t^{(1)} = k | N_t = n)$: after n arrival times, we have k successes in the Bernoulli trial. Hence, if $0 \leq k \leq n$, we have

$$\mathbb{P}(N_t^{(1)} = k | N_t = n) = \binom{n}{k} \; p^k (1-p)^{n-k}.$$

Hence

$$\begin{aligned}
\mathbb{P}(N_t^{(1)} &= k) = \sum_{n=0}^{\infty} \mathbb{P}(N_t = n) \times \mathbb{P}(N_t^{(1)} = k | N_t = n) \\
&= \sum_{n \geq k}^{\infty} \frac{n!}{p!(n-p)!} p^k (1-p)^{n-k} \times \frac{(\lambda t)^n}{n!} \exp(-\lambda t) \\
&= \frac{(\lambda t p)^k}{k!} \exp(-\lambda t) \sum_{n \geq k}^{\infty} \frac{(\lambda (1-p) t)^{n-k}}{(n-k)!} \\
&= \frac{(\lambda p t)^k}{k!} \exp(-\lambda t) \exp(\lambda (1-p) t) \\
&= \frac{(\lambda p t)^k}{k!} \exp(-\lambda p t),
\end{aligned}$$

leading to

$$N_t^{(1)} \sim \mathcal{P}(\lambda p t).$$

At the same time, we have

$$N_t^{(2)} \sim \mathcal{P}(\lambda (1-p) t).$$

Let us show that the two margins are independent. We notice that $(N_t^{(1)} = k, N_t^{(2)} = \ell)$ is included in $(N_t = k + \ell)$. We get

$$\begin{aligned}
\mathbb{P}(N_t^{(1)} &= k, N_t^{(2)} = \ell)) = P(N_t^{(1)} = k, N_t^{(2)} = n-k, N_t = k+\ell) \\
&= \mathbb{P}(N_t = k+\ell)P(N_t^{(1)} = k, N_t^{(2)} = n-k | N_t = k+\ell) \\
&= \frac{(\lambda t)^{k+\ell}}{(k+\ell)!} \times \binom{k+\ell}{k} p^k (1-p)^\ell = \frac{(\lambda p t)^k}{k!} \times \frac{(\lambda(1-p)t)^\ell}{\ell!} \\
&= \mathbb{P}(N_t^{(1)} = k) \times \mathbb{P}(N_t^{(2)} = \ell). \ \square
\end{aligned}$$

4.5. The Bus paradox.

Let us suppose that the Poisson arrival times are those of buses at a given station S_0 and that $\lambda > 0$ is the intensity of the stochastic process. Let us take two passengers arriving between two arrival times of buses, in a random inter-arrival time interval. The bus paradox says that : what ever be the arrival times of these two passengers in a random inter-arrival time interval, the mathematical expectations of their waiting times before the next bus are the same and are equal to $1/\lambda$.

Indeed, if a is the arrival time of the passenger between Z_ν and $Z_{\nu+1}$, where ν is the random variable defined by

$$\nu = \max\{k \geq 0, Z_k \leq a\}.$$

The next bus arrives at the time $Z_{\nu+1}$ and the waiting time

$$U = Z_{\nu+1} - a.$$

Let us compute the expectation of U. We may use a drawing to better understand the event $Z_\nu \leq a < Z_{\nu+1}$. Let us see the relation between

$$V = a - Z_{\nu+1} \ \ and \ \ U = Z_{\nu+1} - a.$$

We notice that for $u \geq 0$, $(U > u) = (Z_{v+1} > a+u)$. This event exactly means that there is no arrivals between a and $a+u$, Hence,

$$(U > u) = (N_{a+u} - N_a = 0)$$

As well, for all $v \geq 0$, we have $(V > u) = (Z_v < a - v)$. If $a - v \leq 0$, the event $(V > u) = (Z_v < a - v)$ is impossible since $Z_v > 0$ $a.s.$ and if $a - v > 0$, that is $Z_\nu \leq a - v \leq a$, it is exactly $N_a - N_{a-v} = 0$. Hence for $0 < v \leq a$, we get

$$(V > v) = (N_a - N_{a-v} = 0).$$

Hence,

$$\begin{aligned}\mathbb{P}(U &> u, V > v) = \mathbb{P}(N_{a+u} - N_a = 0, N_a - N_{a-v} = 0)\\ &= \mathbb{P}(N_{a+u} - N_a = 0) \times \mathbb{P}(N_a - N_{a-v} = 0)\\ &= e^{-\lambda u} \times e^{-\lambda v}\\ &= \mathbb{P}(U > u)\mathbb{P}(V > v).\end{aligned}$$

For $v \geq a$,

$$\begin{aligned}\mathbb{P}(U &> u, V > v) = \mathbb{P}(N_{a+u} - N_a = 0, N_a - N_{a-v} = 0)\\ &= \mathbb{P}(N_{a+u} - N_a = 0, \emptyset)\\ &= 0\\ &= \mathbb{P}(U > u)\mathbb{P}(V > v), \quad since \;\; (, \quad since \;\; (\mathbb{P}(V > v) = 0).\end{aligned}$$

with

$$\mathbb{P}(V > v) = \begin{cases} 0 & si \quad v \geq a \\ e^{-\lambda v} & si \quad 0 < v < a \end{cases}.$$

and for all $u > 0$,

$$\mathbb{P}(U > u) = e^{-\lambda u}.$$

So, for all $u > 0$, $v > 0$,

$$\mathbb{P}(U > u, V > v) = \mathbb{P}(U > u)\mathbb{P}(V > v).$$

It follows that U and V are independent with the concerned laws, We deduce that

$$\begin{aligned}\mathbb{E}(U) &= \int_0^{+\infty} \mathbb{P}(U > u)du = \int_0^{+\infty} e^{-\lambda v}dv \\ &= \lambda^{-1}.\end{aligned}$$

Hence the expectation of waiting times for two passengers arriving at the station in a random inter-arrival time interval, are the same.

Concerning V and the inter-arrival time $Z_{v+1} - Z_v$, it follows that

$$\begin{aligned}\mathbb{E}(V) &= \int_0^{+\infty} \mathbb{P}(V > v)dv = \int_0^{a} e^{-\lambda v}dv \\ &= \lambda^{-1}(1 - e^{-\lambda a}) \to \lambda^{-1},\end{aligned}$$

as $a \to +\infty$ and

$$\begin{aligned}\mathbb{E}(Z_{v+1} - Z_v) &= \mathbb{E}(Z_{v+1} - a + a - Z_v) \\ &= \mathbb{E}(U) + \mathbb{E}(V) \to \lambda^{-1}\end{aligned}$$

as $a \to +\infty$.

5. Kolmogorov equations

One of the most important and useful characterization of counting functions of Poisson processes resides in the expansions of the function

$$0 \leq t \rightarrow p_k(t) = \mathbb{P}(N_t = k), \quad t \geq 0, \quad k \geq 0$$

in neighborhoods of zero, in particular differential equations satisfied by those expansions. The following proposition gives interesting properties of $p(t)$, $t \geq 0$.

Proposition 7.10. *For any counting function of a Poisson stochastic processes* $\{N(t),\ t \geq 0\}$ *and by denoting*

$$p_k(t) = \mathbb{P}(N_t = k),\ t \geq 0,\ k \geq 0$$

we have, as $h \downarrow 0$,

(D1) $p_0(h) = 1 - \lambda h + o(h)$.

(D2) $p_1(h) = \lambda h + o(h)$.

(D3) $p_k(h) = o(h)$, *uniformly for* $k \geq 2$, *meaning*

$$\sup_{k \geq 2} p_k(h) = o(h).$$

Comments. Property (D1) means that the probability of having more than one arrival in a small interval is λ-proportional to the length of that interval as it goes to zero :

$$\mathbb{P}(N_h > 1) = 1 - P(N_h = 0) = h(\lambda + o(1)),$$

that is : the more the interval is small, the less the probability of finding more than one arrival in the interval, with a constant ratio λ.

Property (D2) says that, for all $t \geq 0$, for all $k \geq 0$,

$$\frac{\mathbb{P}(N_{t+h} = k+1) - \mathbb{P}(N_t = k)}{h} = \frac{\mathbb{P}(N_h = 1)}{h} \to \lambda,$$

meaning that the probability to have exactly one arrival in an interval of length $h > 0$ is proportional to h with ratio λ, meaning also that the probability of having one arrival at each instant is λ which become an instant probability (exactly as a *pdf*). This explains the name of intensity of λ.

Finally, (D3) means that the probability of having more than two arrivals in a small interval is infinitely small (of order 1) with respect to the length of that interval.

Proof. The formulas we want to prove derive from expansions of the exponential function at zero. Let us prove each property.

(D1) For $k = 0$,

$$p_0(h) = \frac{(\lambda h)^0}{0!} \exp(-\lambda h) = 1 - \lambda h + o(h).$$

(D2) For $k = 1$,

$$\begin{aligned} p_1(h) &= \frac{(\lambda h)^1}{1!} \exp(-\lambda h) = \lambda h(1 - \lambda h + o(h)) = \lambda h - \lambda^2 h^2 + o(\lambda h^2) \\ &= \lambda h + o(h). \end{aligned}$$

(D3) We have

$$\mathbb{P}(N_h \geq 2) = 1 - \mathbb{P}(N_h = 0) - \mathbb{P}(N_h = 1).$$

By using the two previous properties, we get

$$\begin{aligned} \mathbb{P}(N_h &\geq 2) = 1 - (1 - \lambda h + o(h)) - (\lambda h + o(h)) \\ &= o(h). \end{aligned}$$

By decreasingness of the events $(N_h \geq k)$ in k, it follows that

$$\forall k \geq 2, \ \ (N_h \geq k) \subset (N_h \geq 2)$$

and hence

$$\sup_{k\geq 2} \mathbb{P}(N_h \geq k) \leq \mathbb{P}(N_h \geq 2) = o(h).\square$$

Now, we are going to do the reverse way.

Proposition 7.11. *A stochastic process $(N_t)_{t\geq 0}$ with non-negative integer values and satisfying Properties (D1), (D2), (D3),*

(IC) $N_0 = 0$ a.s.,

(SI) $(N_t)_{t\geq 0}$ has stationarity increments

and

(I2) $(N_t)_{t\geq 0}$ has independent increments.

Then $(N_t)_{t\geq 0}$ is a counting function of a stochastic process, that is we have, on top of (IC), (SI) and (I2),
(PM) for all $t \geq 0$, $N_t \sim \mathcal{P}(\lambda t)$.

Proof. Let as assume that the properties (D1), (D2),(D3), (IC), (SI) and (I2) hold. We fix $t \geq 0$ and $k \geq 0$. So on only have to prove (PM) to complete the proof. We are going to determine the marginal laws *through* the Kolmogorov equations. We have :

$$\begin{aligned}
p_0(t+h) &= \mathbb{P}(N_{t+h}=0) \\
&= \sum_{m\geq 0} \mathbb{P}(N_{t+h}=0, N_h=m) \\
&= \mathbb{P}(N_{t+h}=0, N_h=0) = \mathbb{P}(N_{t+h}-N_h=0, N_h=0) \\
&= \mathbb{P}(N_{t+h}-N_h=0) \times \mathbb{P}(N_t=0) \\
&= p_0(t)p_0(h) = p_0(t) - \lambda h p_0(t) + o(h),
\end{aligned}$$

It follows by (D1) that

$$\frac{p_0(t+h)-p_0(t)}{h} = -\lambda p_0(t) + o(1) \rightarrow p'(t) = -\lambda p_0(t).$$

Hence for $k \geq 1$,

$$\begin{aligned}
p_k(t+h) &= \mathbb{P}(N_{t+h}=k) \\
&= \sum_{m\geq 0} \mathbb{P}(N_{t+h}=k, N_h=m) \\
&= \sum_{m\geq 0} \mathbb{P}(N_{t+h}-N_h=k-m, N_h=m) \\
&= \mathbb{P}(N_{t+h}-N_h=k, N_h=0) + \mathbb{P}(N_{t+h}-N_h=k-1, N_h=1) + A, \qquad (5.1)
\end{aligned}$$

with

$$A = \sum_{m\geq 2} \mathbb{P}(N_{t+h}-N_h=k-m, N_h=m).$$

In the two last lines, we have decomposed the expression according to the different values of m : $m=0$, $m=1$ and $m \geq 2$. It follows that

$$
\begin{aligned}
(5.2) \qquad A &= \sum_{m\geq 2} \mathbb{P}(N_{t+h}-N_h = k-m, N_h = m) \\
&= \sum_{m\geq 2} \mathbb{P}(N_{t+h}-N_h = k-m,) \times \mathbb{P}(N_h = m) \\
&= \sum_{m\geq 2} \mathbb{P}(N_{t+h}-N_h = k-m \times \mathbb{P}(N_h = m).
\end{aligned}
$$

Since $\mathbb{P}(N_h = m) = p_m(h) = o(h)$ uniformly in m, it comes that

$$
\begin{aligned}
(5.3) \qquad A &= o(h) \sum_{m\geq 2} \mathbb{P}(N_{t+h}-N_h = k-m) \\
&= o(h)\mathbb{P}(N_{t+h}-N_h \geq k-2) = o(h).
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
(5.4) \qquad \mathbb{P}(N_{t+h}-N_h &= k, N_h = 0) = \mathbb{P}(N_{t+h}-N_h = k) \times \mathbb{P}(N_h = 0) \\
&= p_k(t)p_0(h) = p_k(t) - \lambda h p_k(t) + o(h)
\end{aligned}
$$

and, finally,

$$
\begin{aligned}
(5.5) \qquad \mathbb{P}(N_{t+h}-N_h &= k-1, N_h = 1) \\
&= \mathbb{P}(N_{t+h}-N_h = k-1) \times \mathbb{P}(N_h = 1) \\
&= p_{k-1}(t)p_1(h) = \lambda h p_{k-1}(t) + o(h).
\end{aligned}
$$

By putting Formulas (5.1), (5.3), (5.4) and (5.5) together, we get

$$\frac{p_k(t+h) - p_k(t)}{h} = -\lambda(p_k(t) + p_{k-1}(t)) + o(1).$$

Thus, $p_k(t)$ is differentiable and

$$p_k'(t) = \lambda(p_{k-1}(t) - p_k(t)).$$

So we have obtained the following equations, called Kolmogorov Equations,

$$\forall t \geq 0, \forall k \geq 0, \ \ \frac{1}{\lambda} p_k'(t) = p_{k-1}(t) - p_k(t)),$$

with the convention that $p_{k-1}(t) \equiv 0$.

In the second part of this proof, we are going to solve the Kolmogorov Equations by determining the moment function of each N_t, By definition,

$$f(t,z) = \mathbb{E}(z^{N_t}) = \sum_{k \geq 0} z^k \mathbb{P}(N_t = k) = \sum_{k \geq 0} z^k p_k(t), z \geq 0.$$

This function is uniformly summable on the domain $|z| \leq 1$. It is differentiable on it term by term. By the Kolmogorov equations, we have

$$\begin{aligned}
f'(t,z) &= \sum_{k\geq 0} z^k p_k'(t) \\
&= \lambda \sum_{k\geq 0} z^k (p_{k-1}(t) - p_k(t)) \\
&= \lambda \left\{ \sum_{k\geq 0} z^k p_{k-1}(t) - \sum_{k\geq 0} z^k p_k(t)) \right\} \\
&= \lambda \left\{ \sum_{k\geq 1} z^k p_{k-1}(t) - \sum_{k\geq 0} z^k p_k(t)) \right\} \\
&= \lambda \left\{ z \sum_{k\geq 0} z^{k-1} p_{k-1}(t) - \sum_{k\geq 0} z^k p_k(t)) \right\} \\
&= \lambda(z-1) f(t,z).
\end{aligned}$$

It follows

$$\frac{f'(t,z)}{f(t,z)} = \lambda(z-1).$$

The general solution of that equation is

$$f(t,z) = K(t)\exp(\lambda t(z-1)), \tag{5.6}$$

where $K(t)$ is the integration constant. For $t \geq 0$, for $z = 1$,

$$f(t,1) = \mathbb{E}(1^{N_t}) = \sum_{k\geq 0} p_k(t) = 1$$

and Equation (5.6) gives $z = 1$

$$K(t) = 1.$$

Hence

$$f(t, z) = \exp(\lambda(z - 1)), z \geq 0 \tag{5.7}$$

and since this is the moment function of a Poisson law of parameter λt, we get the marginal law of N_t. We conclude by using Proposition 7.3. ■

Part 3

Stochastic Integration

CHAPTER 8

Itô Integration or Stochastic Calculus

This chapter and the one that follows have a special place in this book since the functions to integrate have two arguments, one of them being random. In clear, we deal with functions of the form

$$\mathbb{R} \ni t \to f(t,\omega) \in \mathbb{R}, \tag{0.1}$$

where ω is element of a probability space $(\Omega, \mathcal{A}, \mathbb{R})$. For each fixed ω, Formula 0.1 constitutes a path of f, which is defined as a stochastic process of space time $\emptyset \neq T \subset \mathbb{R}$ if for all $t \in T$,

$$\omega \to f(t,\omega),$$

is measurable as a mapping from $(\Omega, \mathcal{A})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Actually, this chapter is destined to readers who have gone through the book of **?** of the present series. We begin by recalling basic notions about the paths of stochastic processes.

1. Regularity of paths of stochastic processes

(1) Measurable stochastic processes. We endow T with $\mathcal{B}(T)$, the induced Borel σ-algebra of $\mathbb{R}$ on it.

(1a) Global measurability. The stochastic process, considered as a function of (t, ω) is measurable when

$$\begin{array}{rlcl} X: & (T \times \Omega, \mathcal{B}(T) \otimes \mathcal{A}) & \longmapsto & (\mathbb{R}, \mathcal{B}(T)) \\ & (t, \omega) & \longrightarrow & X(t, \omega) \end{array}$$

is measurable. By using the characterization of the measurability on the product space $(T\times\Omega, \mathcal{B}(T)\otimes\mathcal{A})$, the stochastic process is measurable if and only if

$$\forall A \in \mathcal{B}(\mathbb{R}), X^{-1}(A) = \{(t, \omega) \in T \times \Omega, \ \ X(t, \omega) \in A\} \in \mathcal{B}(T) \otimes \mathcal{A}.$$

(1b) Progressive Measurability. For any $u \in T$, we define $T_u = \{u \in T, \ t \leq u\}$ endowed with $\mathcal{B}(T_u)$, the induced Borel σ-algebra on T_u. Finally, let us consider the restricted stochastic process

$$(T_u \times \Omega) \ni (t, \omega) \to X_u(t, \omega) = X(t, \omega).$$

Now, given an associated filtration $\mathcal{F}$ with the stochastic process, we say that the stochastic process X is progressively measurable with respect to the filtration $\mathcal{F}$ if and only if, for any $u \in T$, X_u is $(\mathbb{B}(T_u)\otimes\mathcal{F}_u)$-measurable, that is for any $B \in \mathcal{B}(\mathbb{R})$

$$X^{-1}(B) = \{(t, \omega) \in T_u \times \Omega, \ \ X(t, \omega) \in B\} \in \mathcal{B}(T_u) \otimes \mathcal{F}_u.$$

The notion of progressive measurability is important in Stochastic Calculus as we will see it.

(2) Continuity paths. Consider the path of the stochastic process for ω

$$\begin{array}{lcl} (T, \mathcal{B}(T) & \longmapsto & (\mathbb{R}, \mathcal{B}(T)) \\ t & \longrightarrow & X(t, \omega) \end{array}$$

Smoothness properties of such paths are important elements of Random Analysis. Here are some examples :

Continuous Paths. The stochastic process is said to have almost-everywhere continuous paths if

$$\mathbb{P}(\{\omega, \exists t_{0,} t \longrightarrow X(t, \omega) \text{ not continuous at } t_0\}) = 0,$$

whenever the set $\{\omega, \exists t_{0,} t \longrightarrow X(t, \omega) \text{ not continuous at } t_0\}$ is measurable. Such sets are handled in the frame of separable stochastic processes in Random Analysis. Otherwise, one should use the outer probability.

$$\mathbb{P}^*(\{\omega, \exists t_0, t \longrightarrow X(t, \omega) \text{ not continuous at } t_0\}) = 0,$$

Likewise, we define stochastic process processes having *a.s.* right-continuous paths if

$$\mathbb{P}^*(\{\omega, \exists t_0,\ t \longrightarrow X(t,\omega) \text{ not right-continuous at } t_0\}) = 0,$$

or having *a.s.* differentiable paths if

$$\mathbb{P}^*(\{\omega, \exists t_0,\ t \longrightarrow X(t,\omega) \text{ not differentiable at } t_0\}) = 0,$$

etc. All the real analysis can be and is studied in the random frame. This branch of Analysis is called Random analysis.

(3)- (RC) or (RL) and Stopped stochastic processes.

We have to use these important results (proved in **?**).

Proposition 8.1. *An almost sure right-continuous (RC) or left-continuous (LL) stochastic process* $\{t_t,\ t \geq 0\} \subset \mathbb{R}$ *is progressively measurable.*

Theorem 8.2. Let $(f_t)_{t\in T}$, $T = \mathbb{R}_+$, be a progressively measurable stochastic process which is associated with the filtration $\mathcal{F} = (\mathcal{F}_\sqcup)_{t\in T}$. Let V be stopping time with respect to the filtration $\mathcal{F}$. Then the mapping f_V is $\mathcal{A}_V$-measurable, where $\mathcal{A}_V$ is the σ-algebra generated by V.

2. Definition and justification of the Itô Stochastic integrals

Let us fix a Brownian motion $\{B(t),\ t \geq 0\}$ on $(\Omega, \mathcal{A}, \mathbb{R})$, we explained in **?** why a theory of path-wise Riemann-Stieltjes integration (say from a to $b > a$) of the form

$$\int_a^b f(\omega, t)\ dB(\omega, t)$$

is not meaningful. Because the Brownian motion is non-where of finite totally bounded variation. The integral we are introducing has been created by **?**.

Not all stochastic processes have Itô's integrals. Rather, we only focus on square integrable and progressively measurable mappings .

Le us fix two reals numbers a and b such that $a \geq b$. We suppose we have a filtration $(\mathcal{F}_t)_{t \in [a,b]}$ such that
(a) For all $(t, s) \in [a, b]$ with $a < b$, $B(t) - B(s)$ is independent of $\mathcal{F}_t$.

Let us denote by $\mathcal{C}_0([a, b])$ the class of stochastic processes $\{f(t, \omega),\ \omega \in \Omega,\ t \in [a, b]\}$ such that :

(b) $\{f(t, \omega),\ \omega \in \Omega,\ t \in [a, b]\}$ is $\mathcal{F}$-adapted,

and

(c)

$$\int_{\Omega} \mathbb{E}\left(|f(\circ,t)|^2\right) \ \ dt < +\infty. \qquad\qquad (d).$$

Now, let us denote par $\mathcal{C}_{pm}([a,b]$ the class of simple functions of the form below.

$$h \in \mathcal{E}_{pm}([a,b]), \ \ h = \sum_{i=1}^{p} Z_{i-1} 1_{]t_{i-1},t_i]}, \tag{2.1}$$

where the real-valued random variables $(Z_i)_{0\leq i \leq p-1}$, defined on $(\Omega, \mathcal{A}, \mathbb{P})$, are such that each Z_i is $\mathcal{F}_{t_{i-1}}$ measurable and $a = t_0 < t_1 < \cdots < t_p = b)$ is a partition of $[a,b]$.

We have the following density lemma.

Lemma 8.3. *For any $f \in \mathcal{C}_0([a,b]$, there exists a sequence $(f_n)_{n\leq 1} \subset \mathcal{C}_{pm}([a,b]$ such that, as $n \to 0$,*

$$\|f - f_n\|_{L^2([a,b]\times\Omega, \mathcal{B}([a,b])\otimes\mathcal{A}, \mathbb{P}\otimes\lambda)} \to 0$$

Proof. It is given in the Appendix (page 169).

For short, we write the L^2 spaces as follows.

$$\begin{aligned} L^2([a,b]\times\Omega) &= L^2([a,b]\times\Omega, \mathcal{B}([a,b])\otimes\mathcal{A}, \mathbb{P}\otimes\lambda) \\ L^2([a,b]) &= L^2([a,b], \mathcal{B}([a,b]), \lambda) \\ L^2(\Omega]) &= L^2(\Omega, \mathcal{A}, \mathbb{P}). \end{aligned}$$

We have the important property that both $\mathcal{C}_{pm}([a,b])$ and $\mathcal{C}_{pm}([a,b])$.

2.1. Definition - Proposition. We define the Itô integral by two steps using properties to be proved.

(a) For $h \in \mathcal{E}_{pm}([a,b])$ of the form

$$h = \sum_{i=1}^{p} Z_{i-1} 1_{]t_{i-1},t_i]}, \quad Z_{i-1} \;\; \mathcal{F}_{t_{i-1}} - measurable, a = t_0 < t_1 < \cdots < t_p = b,$$

we define

$$I(h) = \sum_{i=1}^{p} Z_{i-1}(B(t_i) - B(t_{i-1}))$$

and $I(h)$ is a centered square integrable random variable of variance $\|h\|^2_{L^2([a,b]\times\Omega)}$ that does not depend on a specific representation of h.

(b) For $f \in \mathcal{C}_0([a,b])$, there exists a sequence $(f_n)_{n\leq 1} \subset \mathcal{C}_{pm}([a,b]$ such that, as $n \to 0$,

$$\|f - f_n\|_{L^2([a,b]\times\Omega,\mathcal{B}([a,b])\otimes\mathcal{A},\mathbb{P}\otimes\lambda)} \to 0$$

and we define $I(f)$ by the limit

$$I(f_n) \to_{L^2(\Omega])} I(f)$$

and $I(f)$ is a centered square integrable random variable of variance $\|f\|^2_{L^2([a,b]\times\Omega)}$ that does not depend on a specific approximating sequence $(f_n)_{n\leq 1}$. $\Diamond$

2.2. Proofs of the claims in Subsection 2.1. .

Let proceed by steps.

1. $\mathcal{C}_{pm}$ is a linear space. We use the superposition method widely used in Lo (2018) and so we omit the details. Let h and k be two elements of $\mathcal{C}_{pm}$ written as follows with the required conditions

$$h = \sum_{i=1}^{p} Z_{i-1} 1_{]t_{i-1},t_i]}, \ \ and \ \ k = \sum_{i=1}^{q} U_{j-1} 1_{]t_{j-1},s_j]}.$$

By denoting $A_i =]t_{i-1}, t_i]$, $B_j =]s_{j-1}, t_j]$ and $H = \{(i,j) \in [1,p]\times[1,q], \ \ A_iB_j \neq \emptyset\}$, we have

$$h = \sum_{(i,j)\in H} Z_{i-1} 1_{A_iB_j}, \ \ and h = \sum_{(i,j)\in H} U_{j-1} 1_{A_iB_j}. \tag{2.2}$$

But for $(i,j) \in H$, each A_iB_j is an interval of the form $A_iB_j =]u_{\ell-1}, u_\ell]$ with $u_{\ell-1} = max(t_{j-1}, s_{j-1})$ and $u_\ell = max(t_j, s_j)$. So Z_{i-1} and

U_{i-1} are both $\mathcal{F}_{u_{\ell-1}}$ measurable.

Now for $(\alpha, \beta) \in \mathbb{R}^2$,

$$\alpha h + \beta k = \sum_{(i,j)\in U} (\alpha Z_{i-1} + \beta U_{j-1}) 1_{A_i B_j}, \tag{2.3}$$

where $(\alpha Z_{i-1} + \beta U_{j-1})$ is $\mathcal{F}_u$ measurable with $A_i B_j =]u, v]$. So $\alpha h + \beta k \in \mathcal{C}_{pm}$. $\square$

2. Independence of $I(H)$ on a specific representations. Let us suppose that $h \in \mathcal{C}_{pm}$ has two representation of the form

$$h = \sum_{i=1}^{p} Z_{i-1} 1_{]t_{i-1}, t_i]} = \sum_{i=1}^{q} U_{j-1} 1_{]t_{j-1}, s_j]}.$$

We use the superposition above to get the follow presentations based on the same partition

$$h = \sum_{(i,j)\in H} Z_{i-1} 1_{A_i B_j} = \sum_{(i,j)\in H} U_{i-1} 1_{A_i B_j} = \sum_{(i,j)\in H} V_{ij} 1_{A_i B_j},$$

where for $(i,j) \in H$, $h(t, \omega) = Z_{i-1} = U_{i-1}$ for $t \in A_i B_j$ and then

$$V_{ij} = Z_{i-1} 1_{A_i B_j} = V_{i-1} 1_{A_i B_j}.$$

For empty sets $A_i B_j$, we put $V_{ij} = 0$. Moreover the sets $A_i B_j$, for $= (i,j) \in H$, are intervals $]u_{ij}, v_{ij}]$, $u_{ij} <_{ij}$, forming a partition of $]a, b]$ and are ordered from left to right. For empty sets $A_i B_j$, we set $]u_{ij}, v_{ij}]$ with an arbitrary number $u_{ij} = v_{ij}$. It is simple to see that for $i \in [1, p]$ and $j \in [1, q]$,

$$]t_{i-1}, t_i] =]t_{i-1}, t_i] \cap \sum_{(i,j)\in H}]u_{ij}, v_{ij}] = \sum_{z\in[1,q]:(i,z)\in H}]u_{ij}, v_{ij}] \tag{2.4}$$

$$]s_{j-1}, t_j] =]t_{j-1}, t_j] \cap \sum_{(i,j)\in H}]u_{ij}, v_{ij}] = \sum_{z\in[1,p]:(z,j)\in H}]u_{ij}, v_{ij}] \tag{2.5}$$

We have

$$\sum_{(i,j)\in H} V_{ij}(B(v_{ij}) - B(v_{ij}))$$

$$= \sum_{(i,j)\in H} Z_{i-1}(B(v_{ij}) - B(v_{ij})) \tag{2.6}$$

$$= \sum_{i=1}^{p}\sum_{j=1}^{q} Z_{i-1}(B(v_{ij}) - B(v_{ij})) \tag{2.7}$$

$$= \sum_{i=1}^{p} Z_{i-1}\biggl(\sum_{j=1}^{p} Z_{i-1}(B(v_{ij}) - B(v_{ij})\biggr) \tag{2.8}$$

$$= \sum_{i=1}^{p} Z_{i-1} \sum_{j\in[1,q]:(i,j)\in H} (B(v_{ij}) - B(v_{ij}) \tag{2.9}$$

$$= \sum_{i=1}^{p} Z_{i-1}B(t_i) - B(t_{i-1}). \tag{2.10}$$

Let us make some comments. In Line 2.6, we replaced V_{ij} by Z_{i-1} since for A_iB_j (which does not depend on $omega$), we have $V_{ij} = Z_{i-1} = U_{i-1}$ on Ω. In Line 2.7, we extended the summation on all $(i,j) \in [1,p]\times[1,q]$ since the contribution of the (i,j) not in H is zero. In Line 2.8, we used the Fubini rule for finite sums with independent summations indices. In Line 2.9, we dropped the null terms corresponding to j such that $(i,j) \notin H$. In Line 2.10, we used 2.4.

Now, by the symmetry of the roles of the two representations, we easily get that

$$\sum_{(i,j)\in H} V_{ij}(B(v_{ij}) - B(v_{ij})) = \sum_{j=1}^{q} Z_{j-1} B(s_j) - B(s_{j-1}). \tag{2.11}$$

The proof is over. $\square$

3. $I(\circ)$ is linear on $\mathcal{C}_{pm}$. Let us consider h and k as defined in Formula 2.2, α and β two real numbers. By using 2.3, and the notation above about the A_iB_j's, we have

$$\begin{aligned}
I(\alpha h+\beta k) &= \sum_{(i,j)\in H}(\alpha Z_{i-1}+\beta U_{j-1})(B(v_{ij})-B(v_{ij})\\
&= \alpha\sum_{(i,j)\in H} Z_{i-1}(B(v_{ij})-B(v_{ij})\\
&+ \beta\sum_{(i,j)\in H} U_{j-1})(B(v_{ij})-B(v_{ij})\\
&= \alpha\sum_{i=1}^{p} Z_{i-1}\sum_{j\in[1,q]:(i,j)\in H}(B(v_{ij})-B(v_{ij})\\
&+ \beta\sum_{j=1}^{q} U_{j-1}\sum_{i\in[1,q]:(i,j)\in H}(B(v_{ij})-B(v_{ij})\\
&= \alpha I(h)+\beta I(k),
\end{aligned}$$

where we used Formulas (2.8)–(2.10) above for both h and k. $\square$

4. Expectation and Variance of $I(h)$. Let us use the representation (2.1) of a function $h\in\mathcal{C}_{pm}$. We have

$$h^2=\sum_{i=1}^{p} Z_{i-1}^2 1_{]t_{i-1},t_i]},$$

since the additional terms of the form

$$Z_{i-1}Z_{j-1}1_{]t_{i-1},t_i]}1_{]t_{j-1},t_j]}, 1\leq i\neq j\leq p$$

are zero from the product of indicators functions based on disjoints supports. So se have

$$\|h\|^2_{L^2([a,b]\times\Omega)} = \mathbb{E}\biggl(\int_a^b h^2(t,\omega)\ \ dt\biggr) = \sum_{i=1}^{p} \mathbb{E}(Z^2_{i-1})(t_i - t_{i-1}),$$

that is

$$\|h\|_{L^2([a,b]\times\Omega)} = \sum_{i=1}^{p} \mathbb{E}(Z^2_{i-1})(t_i - t_{i-1}). \tag{2.12}$$

But we have

$$\begin{aligned}
&I(h)^2 = \sum_{i=1}^{p} Z^{(}_{i-1} B(t_i) - B(t_{i-1}))^2\\
&+2 \sum_{1\leq i<j\leq} Z_{i-1}Z_{j-1}(B(t_i) - B(t_{i-1}))(B(t_j) - B(t_{j-1}))\\
&=: A + 2B.
\end{aligned}$$

But conditioning each term in B associated with $(i < j)$ by $\mathcal{F}_{t_{j-1}}$ and by using the $\mathcal{F}_{t_{j-1}}$-measurability of Z_{i-1} in Line 1 and the projection rule $\mathbb{E}(\circ/\mathcal{B}_1) = \mathbb{E}(\mathbb{E}(\circ/\mathcal{B}_2)/\mathcal{B}_1)$ for sub-σ-algebras $\mathcal{B}_1 \subset \mathcal{B}_2$ of $\mathcal{A}$ in Line L3, and the fact that $Z_{j-1}(B(t_i){-}B(t_{i-1}))$ is $\mathcal{F}_{t_{j-1}}$-measurable in Line L4 and finally the fact that $(B(t_j){-}B(t_{j-1})$ is independent of $\mathcal{F}_{t_{j-1}}$ in Line L5, we get

$$
\begin{aligned}
&\mathbb{E}(B)\\
&= \sum_{1\leq i<j\leq} \mathbb{E}\Bigg(\mathbb{E}\Bigg(Z_{i-1}Z_{j-1}(B(t_i)-B(t_{i-1}))(B(t_j)-B(t_{j-1}))/\mathcal{F}_{t_{i-1}}\Bigg)\Bigg)\\
&= \sum_{1\leq i<j\leq} \mathbb{E}\Bigg(Z_{i-1}\mathbb{E}\Bigg(Z_{j-1}(B(t_i)-B(t_{i-1}))(B(t_j)-B(t_{j-1}))/\mathcal{F}_{t_{i-1}}\Bigg)\Bigg) \quad (L2)\\
&= \sum_{1\leq i<j\leq} \mathbb{E}\Bigg(Z_{i-1}\mathbb{E}\Bigg(\mathbb{E}\Bigg(Z_{j-1}(B(t_i)-B(t_{i-1}))(B(t_j)-B(t_{j-1}))/\mathcal{F}_{t_{j-1}}\Bigg)/\mathcal{F}_{t_{i-1}}\Bigg)\Bigg) \quad (L3)\\
&= \sum_{1\leq i<j\leq} \mathbb{E}\Bigg(Z_{i-1}\mathbb{E}\Bigg(Z_{j-1}(B(t_i)-B(t_{i-1}))\mathbb{E}\Bigg((B(t_j)-B(t_{j-1}))/\mathcal{F}_{t_{j-1}}\Bigg)/\mathcal{F}_{t_{i-1}}\Bigg)\Bigg) \quad (L4)\\
&= \sum_{1\leq i<j\leq} \mathbb{E}\Bigg(Z_{i-1}\mathbb{E}\Bigg(Z_{j-1}(B(t_i)-B(t_{i-1}))\mathbb{E}\Bigg(B(t_j)-B(t_{j-1}))\Bigg)/\mathcal{F}_{t_{i-1}}\Bigg)\Bigg) \quad (L5)\\
&= 0
\end{aligned}
$$

since the quantity in the most interior parentheses

$$\mathbb{E}\Bigg(B(t_j)-B(t_{j-1}))\Bigg)/\mathcal{F}_{t_{i-1}}\Bigg)$$

is zero.

Next we have

$$\begin{aligned}
\mathbb{E}(A) &= \mathbb{E}\biggl(\sum_{i=1}^{p} Z_{i-1}^{(}B(t_i)-B(t_{i-1}))^2\biggr)\\
&= \sum_{i=1}^{p}\mathbb{E}\biggl(\mathbb{E}\biggl(Z_{i-1}^2(B(t_i)-B(t_{i-1}))^2/\mathcal{F}_{t_{i-1}}\biggr)\biggr)\\
&= \sum_{i=1}^{p}\mathbb{E}\biggl(Z_{i-1}^2\mathbb{E}\biggl((B(t_i)-B(t_{i-1}))^2/\mathcal{F}_{t_{i-1}}\biggr)\biggr)\\
&= \sum_{i=1}^{p}\mathbb{E}\biggl(Z_{i-1}^2\mathbb{E}\biggl((B(t_i)-B(t_{i-1}))^2\biggr)\biggr)\\
&= \sum_{i=1}^{p}\mathbb{E}(Z_{i-1}^2)(t_i-t_{i-1}).
\end{aligned}$$

By comparing with Formula (2.12), we get

$$\|I(h)\|_{L^2(\Omega)} = \|h\|_{L^2([a,b]\times\Omega)}\,, \tag{2.13}$$

that is, in other terms

$$\mathbb{E}I(h)=0 \quad \textit{and} \quad \mathbb{V}ar I(h) = \|h\|^2_{L^2([a,b]\times\Omega)}\,. \tag{2.14}$$

In conclusion the operator from I from $\mathcal{C}_{pm}$ to $L^2(\Omega)$ is an isometric linear mapping. Now, let us justify the extension to $\mathcal{C}_0$.

Let us consider two $(f_n)_{n\leq 1}$ and $(g_n)_{n\leq 1}$ in $\mathcal{C}_{pm}([a,b]$ converging to f in $L^2([a,b]\times\Omega)$, as $n\rightarrow 0$. Let continue our proofs through the following steps.

5. The sequence $(I(f_n)_{n\geq 1}$ **has a limit in** $L^2(\Omega)$. By Formulas (2.12)-(2.13), it is clear that $(I(f_n)_)n \geq 1$ is in $L^2(\Omega)$. It is enough to show it is Cauchy in $L^2(\Omega)$ to reach the conclusion. By linearity and Formula (2.13), we have

$$\begin{aligned}
&\|I(f_n) - I(f_m)\|^2_{L^2(\Omega)} = \|I(f_n) - I(f_m)\|^2_{L^2(\Omega)} \\
&= \|I(f_n - f_m)\|^2_{L^2([a,b]\times\Omega)} \\
&= \mathbb{E}\biggl(\int a^b (f_n(t) - f_m(t))^2 \ dt\biggr).
\end{aligned}$$

By using the C_2-inequality in the later line, we have

$$\begin{aligned}
&\|I(f_n) - I(f_m)\|^2_{L^2(\Omega)} \leq 2\mathbb{E}\biggl(\int a^b (f_n(t) - f_(t))^2 \ dt \\
&+\mathbb{E}\biggl(\int a^b (f_m(t) - f_(t))^2 \ dt\biggr) \\
&= 2\biggl(\|f - f_n\|^2_{L^2([a,b]\times\Omega)} + \|f - f_n\|^2_{L^2([a,b]\times\Omega)}\biggr) \\
&\to 0 \ \ as \ \ (n,m) \to (+\infty, +\infty). \ \square
\end{aligned}$$

6. The sequence $(I(f_n))_{n\geq 1}$ **and** $(I(g_n))_{n\geq 1}$ **have the same limit in** $L^2(\Omega)$. By the previous result, we already know that both sequences have a limit in $L^2(\Omega)$. But, by using the same techniques, we have

$$\|I(f_n) - I(g_n)\|^2_{L^2(\Omega)} = \|I(f_n - g_n)\|^2_{L^2(\Omega)}$$
$$= \|I(f_n - g_n)\|^2_{L^2([a,b]\times\Omega)}$$
$$= \mathbb{E}\biggl(\int a^b (f_n(t) - g_n(t))^2 \ \ dt\biggr)$$
$$\leq 2\mathbb{E}\biggl(\int a^b (f_n(t) - f_(t))^2 \ \ dt$$
$$+\mathbb{E}\biggl(\int a^b (g_n(t) - f_(t))^2 \ \ dt\biggr)$$
$$= 2\biggl(\|f_n - f\|^2_{L^2([a,b]\times\Omega)} + \|g_n - f\|^2_{L^2([a,b]\times\Omega)}\biggr)$$
$$\to 0 \ \ as \ \ n \to +\infty. \ \square$$

This shows that the definition of $I(f)$ is consistent, meaning it does not depend on the sequence converging to f in $L^2([a,b]\times\Omega)$.

7. I **is a linear isometric mapping from** $\mathcal{C}_0$ **to** $L^2([a,b])$**.** The linearity immediate as an extension of the linearity of I is linear on $\mathcal{C}_{pm}$ by means of limits. Since the convergence in any L^2 implies that of the norm, we fix $f \in \mathcal{C}_0$ and use a sequence $(f_n)_{n\leq 1}$ and $(g_n)_{n\leq 1}$ in $\mathcal{C}_{pm}([a,b]$ converging to f in $L^2([a,b]\times\Omega)$, as $n \to 0$. So, the three facts

$$(1) \ \ f_n \to_{L^2([a,b]\times\Omega)} f, \ \ (2) \ \ I(f_n) \to_{L^2(\Omega)} I(f),$$

and

$$(3)\, \|I(f_n)\|_{L^2(\Omega)} = \|f_n\|_{L^2([a,b]\times\Omega)}\,, \quad n \geq 0$$

together imply that

$$\forall f \in \mathcal{C}_{pmsi}, \quad \|I(f)\|_{L^2(\Omega)} = \|I(f)\|_{L^2([a,b]\times\Omega)}\,.$$

3. The Itô Integral

1. Summary of the construction.

2. Examples. From the proof of Lemma 8.3, we have that if f is such that $\mathbb{E}(f(\circ, s)f(\circ, t))$ is continuous in $(s,t) \in [a,b]^2$, we may use a known approximating sequence of the form

$$f_n(\circ, t) = f(\circ, a) + \sum_{i=1}^{n} f(t_{i-1}, \omega) 1_{]t_{i-1}, \ \ t_i]}$$

with

$$\max_{1 \leq j \leq n} (t_{j,n} - t_{j-1,n}) \to 0.$$

Finally,

$$I(f_n) = \sum_{i=1}^{n} f(t_{i-1}, \omega)(B(t_i) - B(t_{i-1})) \to_{L^2(\Omega)} I(f) = \int_a^b \ f \ \ dB_t. \tag{3.1}$$

Formula 3.1 offers a practical method to find the first examples which will be used to get more integrals based on properties to come later.

4. Computations

In the proof of Lemma 8.3, in the Appendix (**??**), its proved that : for $f \in \mathcal{C}_0$ such that the function $\mathbb{E}(f(\circ,s)f(\circ,t))$ is continuous in $(s,t) \in [a,b]^2$, we may chose as approximating sequences as follows. Let $n \geq 1$ and let $t_0 = a < t_{n,1} < \cdots < t_{n,n} = b$ be a partition of $[a,b]$ such that, as $n \rightarrow +\infty$,

$$\max_{1 \leq j \leq n} (t_{j,n} - t_{j-1,n}) \rightarrow 0.$$

we take for each $n \geq 2$,

$$f_n(\circ,t) = f(\circ,a) + \sum_{i=1}^{n} f(t_{i-1},\omega) 1_{]t_{i-1},\ \ t_i]}.$$

As a consequence, for

$$Z_n = \sum_{i=1}^{p} f_{t-i-1}(B(t_i) - B(t_{i-1}))$$

converges to

$$I(f) = \int_a^b f_t \ \ dB_t$$

in L^2. We are going to use this to compute several stochastic integrals.

(1) - f(t)=B(t).

$$\int_a^b B_t \ \ dB_t = \frac{B(b)^2 - B(a)^2 - (b-a)}{2}. \tag{4.1}$$

(2) - $f(t) = B^2(t)$.

$$\int_a^b B_t^2 \ \ dB_t = \frac{B(b)^3 - B(a)^3 - \int_a^b B(t) \ \ dt}{2}. \tag{4.2}$$

where $\int_a^b B(t) \ \ dt$ a path-wise Riemann integral of the Brownian movement.

Important Remark. By no way, these integrals are could be interpreted as Rieman-stieljes ones.

APPENDIX .

Proof of Lemma 8.3.

We follows the proof of Kuo (2000) and consider three cases.

Case 1: The function $\mathbb{E}(f(\circ,s)f(\circ,t))$ is continuous in $(s,t) \in [a,b]^2$. Let $n \geq 1$ and let $t_0 = a < t_{n,1} < \cdots < t_{n,n} = b$ be a partition of $[a,b]$ such that, as $n \to +\infty$,

$$\max_{1 \leq j \leq n} (t_{j,n} - t_{j-1,n}) \to 0.$$

We define for each $n \geq 2$,

$$f_n(\circ,t) = f(\circ,a) + \sum_{i=1}^{n} f(t_{i-1},\omega) 1_{]t_{i-1},\ t_i]}.$$

Since f is $\mathcal{F}_t$-adapted, it becomes clear that that $Z_{i-1}(\omega) = f(t_{i-1},\omega)$ is $\mathcal{F}_{t_{i-1}}$ measurable. Now, for each $t \in [a,b]$, $f_n(\circ,t)$ is $\mathcal{F}_t$-measurable since for any $B \in \mathcal{B}(\mathbb{R})$, we have (we denote by $f_{n,t}$ and f_t partial function for t fixed),

$$\begin{aligned}
(f_{n,t} \in B) &= (f_{n,t} \in B) \cap (t = a) + \sum_{i=1}^{n-1} (f_{n,t} \in B) \cap (t \in]t_{i-1},\ t_i) \\
&= (f_a \in B) \cap (t = a) + \sum_{i=1}^{n-1} (f_{t_{i-1}} \in B) \cap (t \in]t_{i-1},\ t_i]) \in \mathcal{F}_{t_{i-1}} \\
&\subset \mathcal{F}_t. \ \square
\end{aligned}$$

By the assumption of this case, for $x \in [a,b]$ fixed, as $y \to x$, that is $(x,y) \to (x,x)$ and as $(y,y) \to (x,x)$,

$$\begin{aligned}
&\mathbb{E}|f(x) - f(y)|^2 \\
&= \mathbb{E}(f^2(x) + f^2(y) - 2f(x)f(y))|^2 \\
&\to \mathbb{E}(f^2(x) + f^2(x) - 2f(x)f(x))|^2 \\
&= 0.
\end{aligned}$$

Also for $x \in [a, b]$, $f_n(x) = f(t_{j(x),n})$, with $t_{j(x),n} \to x$, as $n \to +\infty$. Applying the result before, we have

$$\mathbb{E}|f(x) - f_n(x)|^2 \to 0 \ \ as \ \ n \to +\infty,$$

and so for any $\eta > 0$, by the second triangle inequality, for large values of n,

$$(\mathbb{E}f_n(x)^2)^{1/2} \leq \eta + (\mathbb{E}f_n(x)^2)^{1/2}.$$

But, by Minkowski Inequality in $L^2([a, b], \lambda)$ and by the boundedness the continuous function $\mathbb{E}(f(\circ, s)f(\circ, t))$ on the compact set $[a, b]^2$ and by the later formula, for large values on n,

$$\begin{aligned}
(\mathbb{E}|f(x) - f_n(x)|^2)^{1/2} \ &\leq \ (\mathbb{E}f(x)^2)^{1/2} + (\mathbb{E}f_n(x)^2)^{1/2} \\
&\leq \ \eta + \sup_{s \in [a,b]} (\mathbb{E}f(s)^2)^{1/2} \in \mathbb{R}_+
\end{aligned}$$

Now let us apply the Fatou-Lebesgue Theorem with respect to the finite Lebesgue measure $\lambda_{[a,b]}$ to have, as $n \to +\infty$,

$$\|f-f_n\|^2_{L^2([a,b]\times\Omega,\mathcal{B}([a,b])\otimes\mathcal{A})} = \int_{[a,b]} \mathbb{E}|f(x)-f_n(x)|^2 \ \ d\lambda_{[a,b]} \rightarrow 0. \ \square$$

Case 2: f is bounded (right-continuous). Let F be the *cdf* of the standard distribution function and consider the Lebesgue-Stieljes integration ($d\lambda_F = dF$). We define $n \geq 1$, for $t \in [a,b]$, for $\omega \in \Omega$, the measurable function $x \rightarrow f(t - x/n, , \omega)$. It is bounded and thus, is integrable with respect to the finite F-Lebesgue-Stieltjes measure. So the following function is well defined

$$\begin{aligned} g_n(t,,\omega) &= \int_{\mathbb{R}} 1_{[0,n(t-a)]} f(t-x/n) \ \ dF(x) \\ &= \int_0^{n(t-a)} f(t-x/n) \ \ dF(x) \end{aligned}$$

Let us fix n Let us put, for $p \geq 1$, $x_j = (i/p)n(t-a)$, $j \in [0,p]$

$$g_{n,m}(t,\omega) = f(t - x_0/n, \omega) + \sum_{j=1}^{p} f(t - x_j/n, \omega) 1_{]x_{j-1}, \ \ x_j]}.$$

Is clear that $g_{n,m}(\circ, t)$ is $\mathcal{F}$-adapted. Also, for each fixed ω, $g_{n,m}(t,\omega)$ is bounded and right-continuity, it converges $f(t-x/n, \omega)$ on $[0, n(t-a)]$ as $p \rightarrow +\infty$ and by the dominated convergence theorem, as $p \rightarrow +\infty$:

$$\int_0^{n(t-a)} g_{n,p}(t-x/n) \ \ dF(x) \rightarrow g_n(t,\omega),$$

and by the dominated Lebesgue theorem based the boundedness $g_{n,p}$ and g_m and on the finiteness of the measure λ_F, we have as $p \rightarrow +\infty$

$$\|g_n - f_{n,p}\|^2_{L^2([a,b]\times\Omega)} = \int_0^{n(t-a)} \mathbb{E}|g_{n,p}(x) - g_n(x)|^2 dF(x) \to 0. \quad (S21)$$

Here, we may consider the integral as a Lebesgue-Stieltjes one and proceed to a change of variable to have

$$g_n(t,\omega) = ne^{nt}\int_t^a e^{-u} f(u,\omega) \ \ du$$

and next for $(s,t) \in [a,b]$,

$$\mathbb{E}(g_n(t,\omega)g_n(\omega,s)) = n^2 e^{n(t+s)} \int_t^a \int_s^a e^{-(u+v)} f(u,\omega)f(v,\omega) \ \ du \ \ dv.$$

By the continuity of the indefinite integral (in (s,t)), $\mathbb{E}(g_n(t,\omega)g_n(s,\omega))$ is continuous in (s,t). So, for each fixed $n \geq 1$, we may apply the result in the first case to get a sequence of step functions $(f_{n,p})_{p\geq 1}$ such that as $p \to +\infty$:

$$\|g_n - f_{n,p}\|_{L^2([a,b]\times\Omega)}. \quad (S22)$$

Finally, by combining Formulas (S21), (S22) and the $C-2$ inequality, we get

$$\|f - f_{n,p}\|^2_{L^2([a,b]\times\Omega)} \leq 2(\|f - g_n\|^2_{L^2([a,b]\times\Omega)} + \|g_n - f_{n,p}\|^2_{L^2([a,b]\times\Omega)}), \quad (S23)$$

where we let $p \to +\infty$ first and $n \to +\infty$ next. $\square$

Case 3. General case. For $n \geq 1$ fixed, for ω fixed, we consider

$$g_n(t) = f(t,\omega)1_{(|f(\omega,t)|)}, \quad t \in [a,b].$$

The function $g_n(t)$ is $\mathcal{F}$-adapted and bounded. By the monotone convergence theorem, we have

$$\|f_n - g_n\|^2_{L^2([a,b]\times\Omega)} = \int 1_{]n, \quad +\infty]}(f(t,\omega))|f(t,\omega)|^2 d(\lambda) \otimes \mathbb{P}) \rightarrow 0, \quad (S31)$$

since the sequence of integrands $(h_n)_{n\geq 1}$ satisfies

$$\forall n \geq 1, \;\; 0 \leq h_n \leq f^2 \;\; integrable, \;\; h_n \nearrow 0 \;\; as \;\; n \rightarrow +\infty,$$

From there, we can apply the second step to each g_n and conclude by the same method. ■.

Conclusions and Perspectives

5. Achievements

This dissertation focused on the fundamental tools of stochastic modelling :

(1) The theory of Martingales (discrete) and application.

(2) The foundation of stochastic processes through the Kolmogorov Existence process.

(3) The detailed study of two very important stochastic processes.

(3a) The Brownian process

(3a) The Poison process.

(4) The introduction to Stochastic calculus.

Although its large volume and its deepness, we were be able to achieve it. But the full scope proposed by professeur Lo included :

(5) The Stochastic differential equations

(6) The continuous-times Martingales.

Due to time limitation and certainly the limit of the scope of a master degree, we tried to reach at least the parts 1 to 3.

This leads to consider Parts (5) and (6) as immediate perspectives.

6. Perspectives

After we complete Part 5 and 6, we would be in a advantageous to go deeply in Stochastics in Finance, through the important book of Shiryaev A. N. (1999) Essentials of Stochastic Finance : Facts, Models, Theory.

Bibliography

Patrick Billingsley (1995). *Probability and Measure*. Wiley. Third Edition.

Kai Lai Chung (1974). *A Course in Probability Theory*. Academic Press. New-York.

(2000). *Introduction to Stochastic Integration*. Springer.

Lo, G. S. (2019). Introduction to Stochastic processes. Gaton Berger University.

Lo, G.S.(2018). *Measure Theory and Integration by and for the learner*. SPAS Editions. Saint-Louis, Calgary, Abuja. Doi : http://dx.doi.org/10.16929/sbs/2016.0005, ISBN : 978-2-9559183-5-7.

Lo, G.S.(2018). *Mathematical Foundation of Probability Theory*. SPAS Books Series. Saint-Louis, Senegal - Calgary, Canada. Doi : http://dx.doi.org/10.16929/sbs/2016.0008. Arxiv : arxiv.org/pdf/1808.01713

Lo G.S., K. T. A. Ngom M. and Diallo M.(2018). Weak Convergence (IIA) - Functional and Random Aspects of the Univariate Extreme Value Theory. Arxiv : 1810.01625

Michel Loève (1997). *Probability Theory I*. Springer Verlag. Fourth Edition.

Neveu J.(1975). *Discrete Parameters Martingales*. Elsevier.

Neveu J.(1965). *Mathematical Foundation of the Calculus of probability*. HOLDEN-DAYD series in Probability and Statistics. San Francisco, London, Amsterdam (Translated by Amiel Feinstein from French)

Taylor H.M and Karlin S.(1987). *Introduction to Stochastic Modelling*. Third Edition. Academic Press.