An Introduction to Financial Mathematics*

J†

1 Introduction

This colloquium is an introduction to interest rate models, or more generally fixed income models. However, to achieve the goal, the very first step is inevitably understanding basic knowledge of financial mathematics. By basic knowledge, we mean as follows:

(i) probability space

(ii) conditional expectation

(iii) Brownian motion

(iv) Itô's formula

(v) arbitrage opportunity

(vi) risk-neutral price

However, understanding the above concepts with rigor takes a good amount of dedication. Therefore, this lecture will often come with rough explanations. Interested readers may want to refer to He et al. (2018); Shreve (2004).

Before the details, here is one question for you. If you can answer the following question, this rudiment part may not be for you:

"For a fair value of derivative contracts, must you choose the risk-neutral price?"

The answer is "no". Sometime you must, but sometimes you don't have to. However, if you are a practitioner, there is no harm in choosing risk-neutral prices anyway. Well, does it sound so confusing? This lecture may help you answer the question clearly.

2 Probability Spaces

Most literature of mathematical finance begin with the following one magic sentence:

" Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space."

Then it is followed by something like this: "Now, let us equipped with a filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 0}$". Let's break the terms down together.

Definition 2.1. *A collection of Ω, denoted by $\mathcal{F}$, is called a σ-algebra of Ω if the following conditions are satisfied.*

*This lecture note can be downloaded in https://github.com/junbeoml22

†S & T in YSK. junbeoml22@gmail.com

(i) $\emptyset \in \mathcal{F}$

(ii) *if* $A \in \mathcal{F}$, *then* $A^c \in \mathcal{F}$

(iii) *if* $A_i \in \mathcal{F}$, $i \in \mathbb{N}$, *then* $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

σ-algebras is only for mathematical technicality. There is no financial meaning! (oh please don't try...). So to speak, we have to remove some bad sets from the power set. The main reason is that we take *the axiom of choice* (see Folland, 2013, p.20).

Definition 2.2. *$\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ is called a filtration of Ω if, for any $0 \leq s \leq t$, we have $\mathcal{F}_s \subseteq \mathcal{F}_t$, and each $\mathcal{F}_t$ is a filtration of Ω.*

Definition 2.3. *A σ-algebra generated by A is the smallest σ-algebra containing A.*

Problem 2.4. Prove that

$$\cap_{i=1}^{\infty} \mathcal{A}_i, \quad A \subseteq \mathcal{A}_i, \tag{2.1}$$

is the σ-algebra generated by A.

Example 2.5. Consider two tosses of an unfair coin such that $\mathbb{P}(\{H\}) = 1/3$ and $\mathbb{P}(\{T\}) = 2/3$. Then we have $\Omega = \{HH, HT, TH, TT\}$. The elements in Ω are called events. Now, denote the information at i-th toss by $\mathcal{F}_i$, $i = 0, 1, 2$. At first, we just do not know what will happen, so we have to say "well...by two tosses, it will HH or HT or TH or TT", which is simply a meaningless statement. Mathematically, this can be interpreted as $\{\Omega\} \approx \mathcal{F}$. But, for $\mathcal{F}_0$ to be a σ-algebra, we should set $\mathcal{F}_0 = \{\emptyset, \Omega\}$. Well done!

Let's turn to the next toss. After the first toss, we will know what the first toss is. Mathematically, this means that $\{\{HH, HT\}, \{TH, TT\}\} \subseteq \mathcal{F}_1$. Again for $\mathcal{F}_1$ to be σ-algebra,

$$\mathcal{F}_1 = \{\emptyset, \{HH, HT\}, \{TH, TT\}, \Omega\} \tag{2.2}$$

Likewise, at the second toss, $\mathcal{F}_2$ should contain

$$\{HH\}, \{HT\}, \{TH, \}, \{TT\}. \tag{2.3}$$

For $\mathcal{F}_2$ to be σ-algebra, it becomes the power set of Ω.

Notice that $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2$. This is describing the flow of information.

Definition 2.6. *Let $(\Omega, \mathcal{F})$ be a measurable space. A random variable $\xi : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable if $\{\omega \mid \xi(\omega) > a\} \in \mathcal{F}$ for any $a \in \mathbb{R}$.*

One example is step functions as in the next example.

Example 2.7. Let $(\Omega, \mathcal{F})$ be a measurable space and

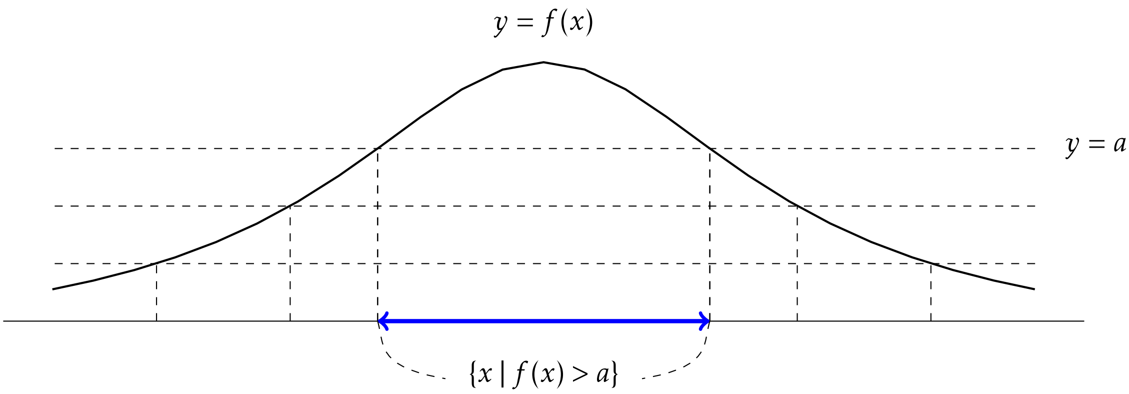
$$F(\omega) = \sum_{i=1}^{\infty} a_i \mathbb{1}_{A_i}, \quad A_i \in \mathcal{F}. \tag{2.4}$$

Then F is $\mathcal{F}$-measurable.

In addition, a stochasitc process keeping measurable for a filtration is called *adapted*.

Definition 2.8. *Let $\mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}$ be a filtration of Ω. A stochastic process X is $\mathbb{F}$-adapted if X_t is $\mathcal{F}_t$-measurable for any $t \geq 0$.*

So to speak, if you know the value of a random variable ξ by the information $\mathcal{F}$, we call the variable $\mathcal{F}$-measurable. Mathematically, measurability is the minimal requirement for Lebesgue integral. The key idea in Lebesgue integral is chopping functions horizontally first.



In this case, the bottom of the rectangle must be a measurable set. The problem is that σ-algebra was given first. In the other way around, we can define σ-algebra for a given function.

Definition 2.9. *Let $f : \Omega \to \mathbb{R}$. The σ-algebra generated by f is the smallest σ-algebra so that f is measurable. Then, we denote it by $\sigma(f)$.*

Example 2.10. We do the coin thing again and consider

$$\xi(w) = \mathbb{1}_{\{HH,HT\}}(\omega) + 2\mathbb{1}_{\{TH,TT\}}(\omega) \tag{2.5}$$

$\sigma(\xi)$ must contain $\{HH, HT\}$ and $\{TH, TT\}$. Therefore, for it to be σ-algebra, it should be $\sigma(\xi) = \mathcal{F}_1$ in Example 2.5.

Now, we are ready to discuss *conditional expectation* of a modern version. Let us recall the classical one first:

$$\mathbb{E}[X \mid Y = y] = \sum_i x_i \frac{\mathbb{P}(X = x_i, Y = y)}{\mathbb{P}(Y = y)}. \tag{2.6}$$

Under this definition, we can not consider cases that $\mathbb{P}(Y = y) = 0$. So instead, you may be tempted to define it with respect to level sets of Y, e.g., something like

$$\{y - \varepsilon < Y < y + \varepsilon\}. \tag{2.7}$$

This idea naturally leads us to define conditional expectation by $\sigma(Y)$, namely a σ-algebra.

Definition 2.11. *[conditional expectation] $\xi^{\mathcal{F}}$ is the conditional expectation of ξ on $\mathcal{F}$ if*

(i) $\xi^{\mathcal{F}}$ *is $\mathcal{F}$-measurable,*

(ii) $\int_A \xi \, d\mathbb{P} = \int_A \xi^{\mathcal{F}} \, d\mathbb{P}$, *for any $A \in \mathcal{F}$.*

Often, we denote $\xi^{\mathcal{F}} = \mathbb{E}[\xi \mid \mathcal{F}]$.

Example 2.12. Now, we do the coin thing again with a fair coin. Ω, $\mathcal{F}_0$,$\mathcal{F}_1$, $\mathcal{F}_2$ are same as in the above examples. Consider

$$\xi(\omega) := \mathbb{1}_{\{HH,HT\}}(\omega) + 2\mathbb{1}_{\{TH,TT\}}(\omega), \tag{2.8}$$

$$\phi(\omega) := \mathbb{1}_{\{HH\}}(\omega) + 2\mathbb{1}_{\{HT\}}(\omega) + 3\mathbb{1}_{\{TH\}}(\omega) + 4\mathbb{1}_{\{TT\}}(\omega). \tag{2.9}$$

Then we will find $\mathbb{E}[\phi \mid \xi = x]$ (the classical version) and $\mathbb{E}[\phi \mid \sigma(\xi)]$ (the modern version). Then, we will see the connection between the two:

$$\mathbb{E}[\phi \mid \xi = x] = \mathbb{E}[\phi \mid \sigma(\xi)](\xi^{-1}(x)). \tag{2.10}$$

Recall that we consider a fair coin toss such that

$$\mathbb{P}(\{HH\}) = \frac{1}{4},\ \mathbb{P}(\{HT\}) = \frac{1}{4},\ \mathbb{P}(\{TH\}) = \frac{1}{4},\ \mathbb{P}(\{TT\}) = \frac{1}{4}. \tag{2.11}$$

Then it is straightforward to calculate (classical) conditional expectations, for example,

$$\begin{aligned}\mathbb{P}(\phi = 2 \mid \xi = 1) =& \frac{\mathbb{P}(\phi = 2, \xi = 1)}{\mathbb{P}(\xi = 1)} \\ =& \frac{\mathbb{P}(\{HT\})}{\mathbb{P}(\{HH,HT\})} = 1/4. \end{aligned} \tag{2.12}$$

Then, by following the classical definition (2.6), we can obtain

$$\mathbb{E}[\phi \mid \xi = 1] = 1.5 \tag{2.13}$$

$$\mathbb{E}[\phi \mid \xi = 2] = 3.5. \tag{2.14}$$

Now, we turn to the modern one. First, it is easy to get

$$\sigma(\xi) = \{\emptyset, \Omega, \{HH,HT\}, \{TH,TT\}\} \tag{2.15}$$

To satisfy the first condition in definition 2.11, we need to begin with guseeing a measurable random variable. Actually, we just now have already done it. Let us consider $\phi^{\sigma(\xi)}: \Omega \to \mathbb{R}$ as

$$\phi^{\sigma(\xi)}(\omega) = 1.5\mathbb{1}_{\{HH,HT\}}(\omega) + 3.5\mathbb{1}_{TH,TT}(\omega). \tag{2.16}$$

It is obviously $\delta(\xi)$-measurable. To check (ii) in Definition 2.11, choose $\{HH,HT\} \in \sigma(\xi)$. Then

$$\begin{aligned}\int_{\{HH,HT\}} \phi(\omega)\mathbb{P}(\omega) &= \int_{\Omega} \phi(\omega)\mathbb{1}_{\{HH,HT\}}(\omega)\mathbb{P}(\omega) \\ &= \int_{\Omega} \Big[\mathbb{1}_{\{HH\}}(\omega) + 2\mathbb{1}_{\{HT\}}(\omega)\Big]\mathbb{P}(\omega) \\ &= \mathbb{P}(\{HH\}) + 2\mathbb{P}(\{HT\}) = 1/4 + 2/4 = 0.75 \end{aligned} \tag{2.17}$$

For the other part,

$$\begin{aligned}\int_{\{HH,HT\}} \phi^{\sigma(\xi)}(\omega)\mathbb{P}(\omega) &= \int_{\{HH,HT\}} \Big[1.5\mathbb{1}_{\{HH,HT\}}(\omega) + 3.5\mathbb{1}_{\{TH,TT\}}(\omega)\Big]\mathbb{P}(\omega) \\ &= \int_{\{HH,HT\}} 1.5\mathbb{1}_{\{HH,HT\}}(\omega)\mathbb{P}(\omega) \\ &= 1.5\mathbb{P}(\{HH,HT\}) + 2\mathbb{P}(\{HT\}) = 1.5/2 = 0.75 \end{aligned} \tag{2.18}$$

Likewise, we can check it with $\{TH,TT\}$. Hence, $\mathbb{E}[\phi \mid \sigma(\xi)] = \xi^{\sigma(\xi)}$. Finally, notice that we can still recover the classical one by the following relationship:

$$\begin{aligned}\mathbb{E}[\phi \mid \xi = 1] =& \mathbb{E}[\phi \mid \sigma(\xi)](\xi^{-1}(1)), \\ \mathbb{E}[\phi \mid \xi = 2] =& \mathbb{E}[\phi \mid \sigma(\xi)](\xi^{-1}(2)). \end{aligned} \tag{2.19}$$

Before moving on, let us ask some questions.

1. Given that you know the price of a stock today, what is the probability whether it rains tomorrow?

2. Given that you know the price of a stock today, what is the expectation of the stock price today?

3. Recall that a conditional expectation is a random variable, e.g., the expectation of a stock price given the information of day-1 is not known on day-0. Under the info of day-0, what is the expectation of the random variable (the expectation of conditional expectation)?

The above statements can be cast into the following theorem.

Theorem 2.13. *i. If ξ is independent with $\mathcal{F}$, $\mathbb{E}[\xi \mid \mathcal{F}] = \mathbb{E}[\xi]$.*

ii. If ξ is $\mathcal{F}$-measurable, then $\mathbb{E}[\xi \mid \mathcal{F}] = \xi$.

iii. Let $\mathcal{F} \subseteq \mathcal{G}$. Then $\mathbb{E}[\mathbb{E}[\xi \mid \mathcal{G}] \mid \mathcal{F}] = \mathbb{E}[\xi \mid \mathcal{F}]$.

Problem 2.14. Prove theorem 2.13.

Definition 2.15. *$(X_t)_{t\geq 0}$ is a sub-martingale (resp. super-martingale) if $X_t \in \mathbb{L}^1$ for any $t \geq 0$ and*

$$\mathbb{E}(X_s \mid \mathcal{F}_t) \geq X_t, \quad t \leq s \quad (\textit{resp. } \mathbb{E}(X_s \mid \mathcal{F}_t) \leq X_t, \quad t \leq s). \tag{2.20}$$

If $(X_t)_{t\geq 0}$ are both a sub-martingale and a super-martingale, then it is called a martingale.

Note that when we call some a stochastic process martingale, it is involved with the probability and filtration. Therefore, it is preferred by theorists to denote $(\mathbb{F}, \mathbb{P})$-martingale. Often, we will omit it for simplicity if there is no possibility for confusion. However, when we deal with change of measure, i.e., Girsanov's theorem, we will explicitly denote the filtration and probability.

3 Brownian Motions and their Stochastic Integration

Definition 3.1. *Let $(\Omega, \mathbb{F}, \mathcal{F}, \mathbb{P})$ be a filtered probability space.*
Let $(W_t)_{t\geq 0}$ be a stochastic process such that

- *W is $\mathbb{F}$-adapted*
- *$W_t - W_s \sim N(0, t-s)$*
- *W_t is continuous in t a.s*
- *$W_0 = 0$.*

Then $(W_t)_{t\geq 0}$ is called the standard Brownian motion.

Problem 3.2. Calculate $\mathbb{E}[W_t]$, $\mathbb{E}[W_t^2]$, and $\mathbb{E}[W_t^3]$.

Theorem 3.3. *(i) Brownian motions are $(\mathbb{F}, \mathbb{P})$-martingale.*

(ii) Let $(W_t)_{t\geq 0}$ be a Borwnian Motion. Then, $(e^{\sigma W_t - \sigma^2 t/2})_{t\geq 0}$ is a martingale for any $\sigma \in \mathbb{R}$.

Proof. (i)

$$
\begin{aligned}
\mathbb{E}[W_t|\mathcal{F}_s] =& \mathbb{E}[W_t - W_s + W_s|\mathcal{F}_s] \\
=& W_s + \mathbb{E}[W_t - W_s|\mathcal{F}_s] && \text{(Measurability)} \\
=& W_s + \mathbb{E}[W_t - W_s] = W_s. && \text{(Independence)}
\end{aligned} \tag{3.1}
$$

(ii)

$$
\begin{aligned}
\mathbb{E}\left[\exp\left\{\sigma W_t - \frac{\sigma^2}{2}t\right\}\middle|\mathcal{F}_s\right] =& \mathbb{E}\left[\exp\left\{\sigma(W_t - W_s + W_s) - \frac{\sigma^2}{2}t\right\}\middle|\mathcal{F}_s\right] \\
=& \exp\left\{-\frac{\sigma^2}{2}t\right\}\mathbb{E}\left[\exp\left\{\sigma(W_t - W_s + W_s)\right\}\middle|\mathcal{F}_s\right] \\
=& \exp\left\{\sigma W_s - \frac{\sigma^2}{2}t\right\}\mathbb{E}\left[\exp\left\{\sigma(W_t - W_s)\right\}\middle|\mathcal{F}_s\right] && \text{(Measurability)} \\
=& \exp\left\{\sigma W_s - \frac{\sigma^2}{2}t\right\}\mathbb{E}\left[\exp\left\{\sigma(W_t - W_s)\right\}\middle|\mathcal{F}_s\right] \\
=& \exp\left\{\sigma W_s - \frac{\sigma^2}{2}t\right\}\mathbb{E}\left[\exp\left\{\sigma(W_t - W_s)\right\}\right] && \text{(Independence)} \\
=& \exp\left\{\sigma W_s - \frac{\sigma^2}{2}t\right\}\exp\{\mathbb{V}[\sigma(W_t - W_s)]\} && \text{(why?)} \\
=& \exp\left\{\sigma W_s - \frac{\sigma^2}{2}t\right\}\exp\{\sigma^2(t-s)/2\} = \exp\left\{\sigma W_s - \frac{\sigma^2}{2}s\right\}
\end{aligned} \tag{3.2}
$$

∎

The second item in the above theorem will be used to prove Girsanov's theorem.

Definition 3.4.

$$
\int_0^t \xi_s \, \mathrm{d}W_s = \lim \sum_i \xi_{t_i}(W_{t_{i+1}} - W_t). \tag{3.3}
$$

Theorem 3.5. *(1)* $\int_0^t \xi_s \, \mathrm{d}W_s$ *is a local-martingale.*

(2) If ξ *is a deterministic process,*

$$
\int_0^t \xi_s \, \mathrm{d}W_s \sim N\left(0, \int_0^t \xi_s^2 \, \mathrm{d}s\right) \tag{3.4}
$$

Theorem 3.6 (Itô's formual)**.** *Let* $F \in C^2(\mathbb{R})$ *and* X *follows*

$$
\mathrm{d}X_t = a(t, X_t)\,\mathrm{d}t + b(t, X_t)\,\mathrm{d}W_t. \tag{3.5}
$$

Then we have

$$
F(X_t) - F(X_0) = \int_0^t \left[F'(X_s)a(t, X_s) + \boldsymbol{F''(X_s)\frac{b^2(s, X_s)}{2}}\right]\mathrm{d}s + \int_0^t F'(X_s)b(s, X_s)\,\mathrm{d}W_s. \tag{3.6}
$$

The formula is often put in the following formal representation:

$$dF(X_t) = \left[F'(X_s)a(t,X_t) + F''(X_t)\frac{b^2(t,X_t)}{2}\right]dt + F'(X_t)b(t,X_t)\,dW_t \tag{3.7}$$

The above theorem is a stochastic version of fundamental theorem of calculus. If we do not have the middle term

$$\int_0^t F''(X_s)\frac{b^2(s,X_s)}{2}\,ds, \tag{3.8}$$

(3.6) can be rewritten as

$$F(X_t) - F(X_0) = \int_0^t F'(X_s)\,dX_s. \tag{3.9}$$

Notice the analogy between (3.9) and the classical version of fundamental theorem of calculus:

$$g(b) - g(a) = \int_a^b g'(x)\,dx. \tag{3.10}$$

However, we need an additional term (3.8), since Brownian motions move so hard as much as unrealistic. Actually, Brownian motions are not a realistic object at all. Bluementhal's 0-1 Law explains how hard Brownian motions move.

Theorem 3.7 (Bluementhal's 0-1 Law)**.** *Let $\mathbb{F}$ be the filtration generated by a Brownian motion $(W_t)_{t\geq 0}$. Then $\mathcal{F}_{0+}$ is trivial. More precisely, for any $A \in \mathcal{F}_{0+}$, we have either*

$$\mathbb{P}(A) = 1 \quad \textit{or} \quad \mathbb{P}(A) = 0. \tag{3.11}$$

Corollary 3.8. *Let $\tau := \inf\{t \geq 0 \mid W_t > 0\}$ and $\nu := \inf\{t \geq 0 \mid W_t < 0\}$. Then*

$$\mathbb{P}(\tau = 0) = \mathbb{P}(\nu = 0) = 1. \tag{3.12}$$

By corollary 3.8, for any interval $(0,\epsilon)$ $\epsilon > 0$, there is zero of Brownian motions. However, Brownian motion are Markov Processes, i.e., historical path does not affect its next move. Therefore, as soon as it hits the zero, it becomes to have Bluementhal's 0-1 property, i.e., it hits another zero again immediately. In other words, there is no interval that Brownian motions are monotonic. This is a reason why we have the adjustment term (3.8). The remaining step to get (3.8) is only a matter of calculation. The next lemma is used for the calculation.

Lemma 3.9 (Quadratic Variation)**.** *Let $(W_t)_{t\geq 0}$ be a Brownian motion and $\{\Pi\}$ be partitions of $[0,T]$. Then*

$$\lim_{\|\Pi\|\to 0}\sum_i (W_{t_{i+1}} - W_{t_i})^2 \to T \quad \textit{in} \quad \mathbb{L}^2. \tag{3.13}$$

Example 3.10. Find $\mathbb{E}[W_t^4]$.
Sol. Put $dW_t = dW_t$. You can put $a(t,x) = 0$ and $b(t,x) = 1$ in (3.5). Let $F(x) = x^4$ and we apply Itô's formual:

$$dW_t^4 = 6W_t^2\,dt + 4W_t^3\,dW_t. \tag{3.14}$$

Then, it follows that

$$W_t^4 - W_0^4 = \int_0^t 6W_s^2 \,\mathrm{d}s + \int_0^t 4W_s^3 \,\mathrm{d}W_s. \tag{3.15}$$

Therefore, we have

$$\begin{aligned} \mathbb{E}[W_t^4] &= \int_0^t 6\mathbb{E}[W_s^2]\,\mathrm{d}s && (3.16)\\ &= \int_0^t 6s\,\mathrm{d}s && (3.17)\\ &= 3t^2. && (3.18) \end{aligned}$$

In what follows, we will use the very popular recipe:

$$\begin{aligned} \mathrm{d}W_t\,\mathrm{d}W_t &= \mathrm{d}t.\\ \mathrm{d}W_t\,\mathrm{d}t &= 0.\\ \mathrm{d}t\,\mathrm{d}t &= 0. \end{aligned}$$

References

Folland, G. B. (2013). *Real analysis: modern techniques and their applications*. John Wiley & Sons.

He, S.-w., Wang, J.-g., & Yan, J.-a. (2018). *Semimartingale theory and stochastic calculus*. Routledge.

Shreve, S. E. (2004). *Stochastic calculus for finance ii: Continuous-time models* (Vol. 11). Springer Science & Business Media.