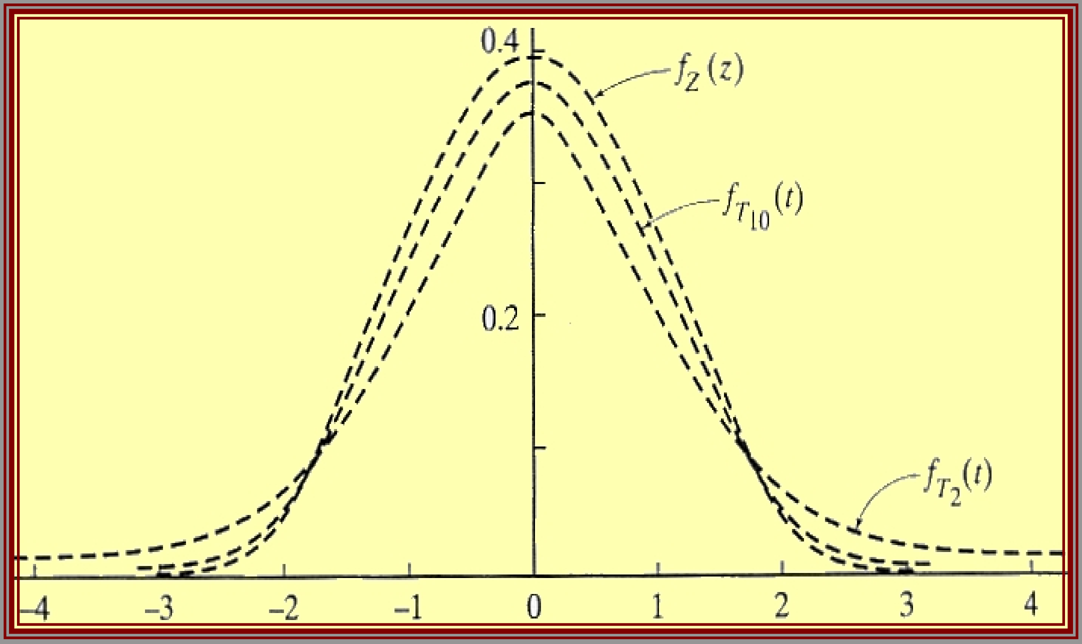
ISSN 1726-3328

JPSS

A comprehensive journal of probability and statistics
for theorists, methodologists, practitioners, teachers, and others



JOURNAL OF PROBABILITY AND STATISTICAL SCIENCE

Volume 18 Number 1 February 2020

ISSN 1726-3328

JPSS

Journal of Probability and Statistical Science

Volume 18 Number 1 February 2020

Table of Contents

Published by: Susan Rivers' Cultural Institute, Hsinchu City, Taiwan, ROC.

Journal of Probability and Statistical Science
18(1), 87-98, Feb. 2020

Elementary Mathematical Tools to Understand the Binomial Model in the Pricing of Financial Options

Krisada Khruachalee Winai Bodhisuwan Andrei Volodin
Kasetsart University *University of Regina*

ABSTRACT Financial mathematics is typically characterized as applying tools and techniques from mathematics to various topics in finance and financial markets. Special attention has been given to Financial Mathematics after the last financial crisis in 2007 and 2008 that resulted in the collapse of a number of large financial institutions in the world. Financial options are the most important products traded, and therefore the subject of pricing for options is of great importance for preventing future financial crises and making the work of financial institutions smooth and reliable. A main focus in economic analysis is devoted to a simultaneous combination of quantitative techniques provided by Financial Mathematics and the analytical analysis to derive option pricing by particular financial experts and advisors. Whenever this is possible, this will provide the best prediction for the value of the reward received. The most important and the easiest way to achieve these goals is to implement the famous Binomial Model, developed by Cox *et al.* [1] in 1973.

Keywords Binomial model; Conditional Expectation; $\mathcal{D}$-measurable; Financial mathematics; Martingale; Pricing options.

1. Introduction

While teaching a class on Financial Mathematics, the authors of this note mentioned that students have difficulties to understand some crucial Mathematical Tools such as Conditional Expectations and Martingales. Without this tools it is impossible to construct and explain the Binomial Model in the pricing of financial options. The main goal of this note is to provide very simple and elementary definitions, explanations, and derivations of these tools from the Elementary Probability level. Some of our derivations are based on the paper by Shiryaev *et al.*

□Received October 2019, revised December 2019, in final form January 2020.
□Krisada Khruachalee and Winai Bodhisuwan (corresponding author) are affiliated to the Department of Statistics at Kasetsart University, Chatuchak, Bangkok, Thailand; emails: krisada.khr@ku.th and fsciwnb@ku.ac.th. Andrei Volodin is a Professor in the Department of Mathematics and Statistics at University of Regina, Regina, Saskatchewan, Canada; email: andrei@uregina.ca.

 JPSS: ISSN 1726-3328

[3], but they rely on σ-algebras, while we do not. The notion of a partition is borrowed from the textbook by Shiryaev [2].

We provide all necessary mathematical tools from probability theory for derivation the Binomial Model for pricing options, while we do not discuss the Binomial Pricing Model itself. The special attention is given to such notions from probability theory as conditional expectation and martingale.

2. Conditional Expectation by Partitions

In the following let $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$ be a finite sample space.

Definition 1.1 A class of subsets or events of the space Ω

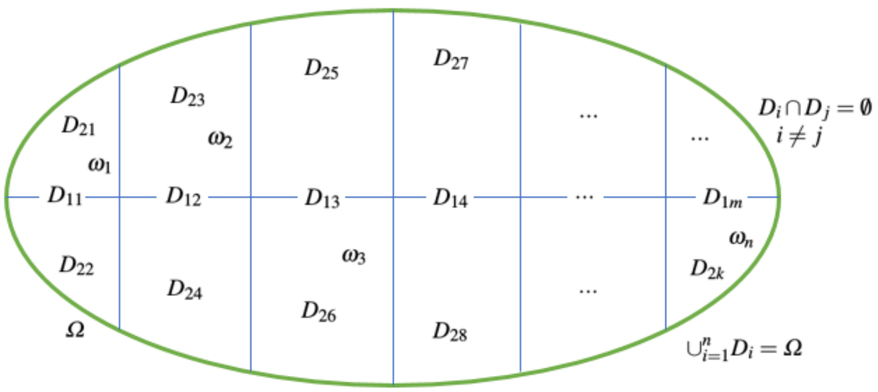
$$\mathcal{D}=\{D_1,D_2,\cdots,D_k\},\ D_i\subseteq\Omega,\ (i=1,\cdots,k)$$

is called a *partition* of the space Ω, if

1) The events D_i are disjoint (mutually exclusive), that is, $D_i\bigcap D_j=\varnothing,\ (i\neq j)$, and
2) $\bigcup_{i=1}^{k}D_i=\Omega.$

The sets D_i are called *atoms* of the partition $\mathcal{D}$. □

Suppose $\mathcal{D}_1=\{D_{11},D_{12},\cdots,D_{1m}\}$ and $\mathcal{D}_2=\{D_{21},D_{22},\cdots,D_{2k}\}$ are two partitions of Ω.



Definition 1.2 The partition $\mathcal{D}_2$ *majorates* the partition $\mathcal{D}_1$, or $\mathcal{D}_2$ is *more fine* than $\mathcal{D}_1$, (denoted $\mathcal{D}_1\preceq\mathcal{D}_2$), if each atom of the partition $\mathcal{D}_2$ is a subset of an atom of the partition $\mathcal{D}_1$. That is, for any $1\le q\le k$ there exists $p, 1\le p\le m$ such that $D_{2q}\subset D_{1p}$. □

Note that if partition $\mathcal{D}_2$ majorates partition $\mathcal{D}_1$, then $\mathcal{D}_2$ is simply a partition of atoms of partition $\mathcal{D}_1$. That is, for any $1\le p\le m$ there exists $J_p\subset\{1,2,\cdots,k\}$ such that $D_{1p}=\bigcup_{q\in J_p}D_{2q}$ and index sets J_p are disjoint and $\bigcup_{p=1}^{m}J_p=\{1,2,\cdots,k\},\ k\ge m.$

Also note that there is the "smallest" partition $\mathcal{D}_{\min}=\{\Omega\}$ and the "largest" partition $\mathcal{D}_{\max}=\{D_1=\{\omega_1\},D_2=\{\omega_2\},\cdots,D_n=\{\omega_n\}\}$ (recall that there are n elementary outcomes in $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$). Then for any partition $\mathcal{D}$ we have that $\mathcal{D}_{\min}\preceq\mathcal{D}\preceq\mathcal{D}_{\max}$.

Generally speaking, we can define a probability only on a partition $\mathcal{D}$ in the following way.

Definition 1.3 Let $\mathcal{D} = \{D_1, D_2, \cdots, D_k\}$ be a partition of space Ω. A *probability* **P** on the partition $\mathcal{D}$ is a function of $\mathcal{D}$ into an interval [0, 1], such that the following conditions are satisfied:

1) $\mathbf{P}(D_i) \geq 0, (i = 1, \cdots, k)$ and 2) $\sum_{i=1}^{k} \mathbf{P}(D_i) = 1.$ □

We can extend the notion of probability via summation on disjoint unions on the class of sets $\mathcal{F}$ which consists of all possible unions of elements of the partition $\mathcal{D}$, that is, if $A \in \mathcal{F}$, then $A = \bigcup_{i \in I} D_i$, where I is any collection of indexes from 1 to k (including $I = \varnothing$). If $A = \bigcup_{i \in I} D_i$, then we define the probability of the event A as $\mathbf{P}(A) = \sum_{i \in I} \mathbf{P}(D_i)$.

In the following we assume that the probability function **P** is defined on the partition $\mathcal{D}_{\max}$. In this case, for any $A \subset \Omega$ we then have $\mathbf{P}(A) = \sum_{\omega \in A} \mathbf{P}(\omega)$.

Definition 1.4 Let $\mathcal{D}_1 = \{D_{11}, D_{12}, \cdots, D_{1m}\}$ and $\mathcal{D}_2 = \{D_{21}, D_{22}, \cdots, D_{2k}\}$ be two partitions of Ω. We say that partitions $\mathcal{D}_1$ and $\mathcal{D}_2$ are *independent* if

$$\mathbf{P}(D_{1i} \cap D_{2j}) = \mathbf{P}(D_{1i})\mathbf{P}(D_{2j}), \quad \text{for all } 1 \leq i \leq m, \ 1 \leq j \leq k.$$ □

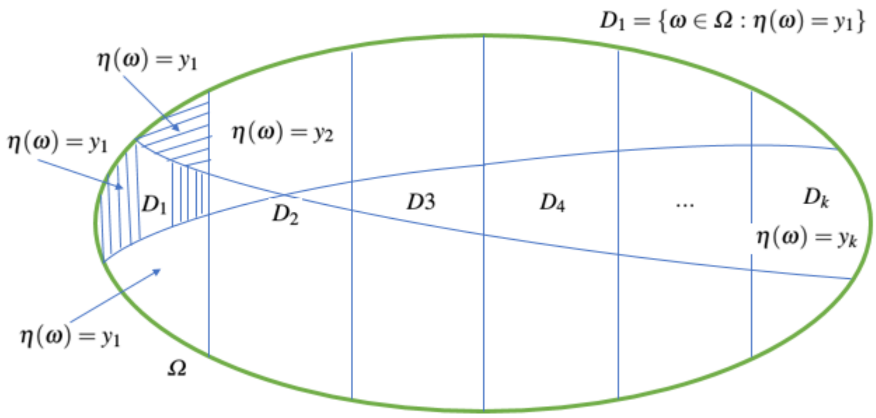
Having defined probability on a partition, we now introduce the important concept of a random variable.

Definition 1.5 A function η from the sample space into the real line, that is,

$$\eta : \Omega \to (-\infty, \infty),$$

is called a *random variable*. □

Definition 1.6 Let η be a random variable taking on values $y_1, y_2, \cdots, y_k$. Denote by $D_j = \{\omega : \eta(\omega) = y_j\}$. Then the partition $\mathcal{D}_\eta = \{D_1, D_2, \cdots, D_k\}$ is called a *partition generated by random variable* η.



□

In the same way we can introduce the partition $\mathcal{D}_{\eta_1, \eta_2, \cdots, \eta_m}$ generated by a family of random variables $\eta_1, \eta_2, \cdots, \eta_m$. This partition consists of the atoms

$$D_{y_1, y_2, \cdots, y_m} = \{\omega : \eta_1(\omega) = y_1, \eta_2(\omega) = y_2, \cdots, \eta_m(\omega) = y_m\}.$$

Note that if a random variable η is given, then in the following we will consider only probabilities **P** defined on the partition $\mathcal{D}_\eta$ such that

$$\mathbf{P}(\eta = y_j) > 0, \ j = 1, 2, \cdots, k.$$

There is one random variable that plays a special role in the following

Definition 1.7 Let A be an event, that is, $A \subset \Omega$. The random variable I_A is called an *indicator function* if it takes only two values:

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}.$$

We note that the partition generated by the indicator function I_A has the simple structure: $\mathcal{D}_{I_A} = \{A, A^c\}$. □

Definition 1.8 Let $\mathcal{D} = \{D_1, D_2, \cdots, D_k\}$ be a partition, then random variable $\eta = \eta(\omega)$ is *measurable with respect to the partition* $\mathcal{D}$, or $\mathcal{D}$*-measurable*, if $\mathcal{D}_\eta \preceq \mathcal{D}$. This means that the random variable takes constant values on atoms of the partition $\mathcal{D}$ and hence the random variable can be written in the form

$$\eta(\omega) = \sum_{i=1}^{k} y_i \cdot I_{D_i}(\omega),$$

where some of the scales y_i may be equal. □

Sometimes it is more convenient for us to consider only distinct y_i. Then we can collect all D_i for which random variable η takes the same values. In this case the random variable η can be represented as

$$\eta = \sum_{j=1}^{l} y_j \cdot I_{A_j}, \quad A_j = \{\omega : \eta(\omega) = y_j = \textstyle\bigcup_{i \in J_j} D_i\},$$

where the index sets J_j are disjoint and $\bigcup_{j=1}^{l} J_j = \{1, 2, \cdots, k\}$, $l \le k$.

Definition 1.9 Let $\xi = \xi(\omega)$ be a random variable taking values $x_1, x_2, \cdots, x_l$:

$$\xi = \sum_{j=1}^{l} x_j \cdot I_{A_j}, \quad A_j = \{\omega : \xi(\omega) = x_j\}.$$

The *mathematical expectation* (unconditional) of the random variable ξ is defined as

$$\mathbf{E}(\xi) = \sum_{j=1}^{l} x_j \mathbf{P}(A_j).$$ □

We now state and prove a basic result concerning the expectation of the indicator function.

Fact: $\mathbf{E}(I_A) = \mathbf{P}(A)$.

Proof. $\mathbf{E} = 0 \cdot \mathbf{P}(I_A = 0) + 1 \cdot \mathbf{P}(I_A = 1) = \mathbf{P}(A)$. □

Proposition 1.1 Let $\xi = \xi(\omega)$ and $\eta = \eta(\omega)$ be two random variables such that ξ is measurable with respect to the partition $\mathcal{D}_\eta$ generated by the random variable η. Then there exists a function f such that $\xi = f(\eta)$.

Proof. Let $y_1, y_2, \cdots, y_k$ be the values taken by the random variable η. Then the partition $\mathcal{D}_\eta$ consists of atoms

$$D_j = \{\omega : \eta(\omega) = y_j\},\ 1 \le j \le k \quad \text{and} \quad \eta(\omega) = \sum_{j=1}^{k} y_j \cdot I_{D_j}(\omega).$$

Since ξ is measurable with respect to the partition $\mathcal{D}_\eta$, it can be written as

$$\xi(\omega) = \sum_{j=1}^{k} x_j \cdot I_{D_j}(\omega).$$

The construction of the function f now is quite simple. Let $f(y_j) = x_j,\ 1 \le j \le k.$ Obviously, $\xi(\omega) = f[\eta(\omega)]$ by the definition of the function f. The proposition is proved. □

Proposition 1.1 has a simple generalization. Namely, let $\xi = \xi(\omega)$ and $\eta_1(\omega),\ \eta_2(\omega),\ \cdots, \eta_m(\omega)$ be random variables such that ξ is measurable by the partition $\mathcal{D}_{\eta_1,\eta_2,\cdots,\eta_m}$ generated by the random variables $\eta_1, \eta_2, \cdots, \eta_m$. Then there exists a function f such that $\xi = f(\eta_1, \eta_2, \cdots, \eta_m)$.

Definition 1.11 Let $\mathcal{D}_\xi$ and $\mathcal{D}_\eta$ be two partitions. We say that partitions $\mathcal{D}_\xi$ and $\mathcal{D}_\eta$ are independent, if for all $A \in \mathcal{D}_\xi$ and $B \in \mathcal{D}_\eta$ we have that $\mathbf{P}(AB) = \mathbf{P}(A)\mathbf{P}(B)$. □

Definition 1.12 Two random variables $\xi = \xi(\omega)$ and $\eta = \eta(\omega)$ are said to be *independent* if the partitions $\mathcal{D}_\xi$ and $\mathcal{D}_\eta$ are independent. □

Let $A \subset \Omega$ and let $\mathbf{P}(A \mid D_i)$ be the conditional probability of event A by event D_i. That is,

$$\mathbf{P}(A \mid D_i) = \frac{\mathbf{P}(A \cap D_i)}{\mathbf{P}(D_i)}.$$

The collection of conditional probabilities $\{\mathbf{P}(A \mid D_i)\}_{i=1}^{k}$ defines a random variable given by

$$\mathbf{P}(A \mid \mathcal{D})(\omega) = \sum_{i=1}^{k} \mathbf{P}(A \mid D_i) I_{D_i}(\omega), \tag{1}$$

which takes the values $\mathbf{P}(A \mid D_i)$ on the atoms D_i of the partition $\mathcal{D}$. The random variable (1) is called *the conditional probability* of the event A with respect to the partition $\mathcal{D}$.

The following are some important properties of the conditional probability with respect to the partition.

1°. If two events A and B in $\mathcal{F}$ are disjoint, then $\mathbf{P}(A \cup B \mid \mathcal{D}) = \mathbf{P}(A \mid \mathcal{D}) + \mathbf{P}(B \mid \mathcal{D})$.

2°. Recall that $\mathcal{D}_{\min} = \{\Omega\}$ is the minimal partition. Then $\mathbf{P}(A \mid \mathcal{D}_{\min}) = \mathbf{P}(A)$. For the random variable $\mathbf{P}(A|\mathcal{D})$ we can define an expectation as follow.

$$\mathbf{E}[\mathbf{P}(A|\mathcal{D})]=\mathbf{E}\left(\sum_{i=1}^{k}\mathbf{P}(A|D_i)I_{D_i}(\omega)\right)=\sum_{i=1}^{k}\mathbf{P}(A|D_i)\mathbf{P}(D_i)=\sum_{i=1}^{k}\mathbf{P}(A\cap D_i)=\mathbf{P}(A).$$

Therefore we have obtained the formula of Total Probability in a modified form:

3°. $\mathbf{E}[\mathbf{P}(A|\mathcal{D})]=\mathbf{P}(A).$

Definition 1.13 Conditional probability $\mathbf{P}(A|\mathcal{D}_\eta)$ is called the *conditional probability of event A with respect to random variable* η and denoted by $\mathbf{P}(A|\eta)$. Denote

$$\mathbf{P}(A|\eta=y_j)=\mathbf{P}(A|D_j),\text{ where } D_j=\{\omega:\eta(\omega)=y_j\}.$$

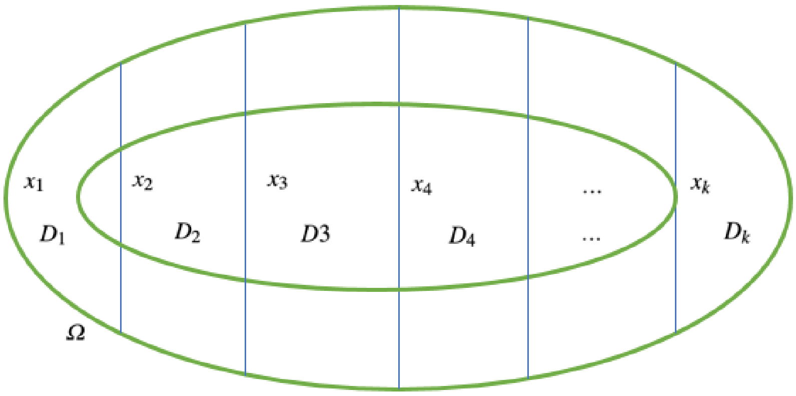
In the same way we can introduce the conditional probability of an event A with respect to a family of random variables $\eta_1,\eta_2,\cdots,\eta_m$. The partition here is generated by the atoms

$$D_{y_1,y_2,\cdots,y_m}=\{\omega:\eta_1(\omega)=y_1,\eta_2(\omega)=y_2,\cdots,\eta_m(\omega)=y_m\}.$$

This conditional probability is then denoted by $\mathbf{P}(A|\eta_1,\eta_2,\cdots,\eta_m)$. □

Definition 1.14 The *conditional mathematical expectation of the random variable* ξ by the partition $\mathcal{D}$ with respect to the formula:

$$\mathbf{E}(\xi|\mathcal{D})=\sum_{j=1}^{l}x_j\mathbf{P}(A_j|\mathcal{D}). \tag{2}$$



If the partition $\mathcal{D}$ is generated by the random variables $\eta_1,\eta_2,\cdots,\eta_k$, then the conditional mathematical expectation $\mathbf{E}(\xi|\mathcal{D}_{\eta_1,\eta_2,\cdots,\eta_k})$ is called the *conditional mathematical expectation of a random variable* ξ *with respect to* $\eta_1,\eta_2,\cdots,\eta_k$ and is denoted by $\mathbf{E}(\xi|\eta_1,\eta_2,\cdots,\eta_k)$. □

According to this definition, the conditional mathematical expectation of a random variable is also a random variable, see (2). Now we approach this notion from a different perspective. Define the conditional mathematical expectation $\mathbf{E}(\xi|D_i)$ of random variable ξ with respect to an event D_i by the formula

$$\mathbf{E}(\xi|D_i)=\sum_{j=1}^{l}x_j\mathbf{P}(A_j|D_i)=\frac{1}{\mathbf{P}(D_i)}\mathbf{E}(\xi\cdot I_{D_i}).$$

Now we let

$$\mathbf{E}(\xi \mid \mathcal{D}) = \sum_{i=1}^{k} [\mathbf{E}(\xi \mid D_i) \cdot I_{D_i}(\omega)] = \sum_{i=1}^{k} \left(\frac{\mathbf{E}(\xi \cdot I_{D_i})}{\mathbf{P}(D_i)} \cdot I_{D_i}(\omega) \right). \tag{3}$$

Note the value of the conditional mathematical expectation of a random variable with respect to the partition does not depend on the representation of the random variable. Moreover, for its calculation we can follow scheme (3) and we will obtain the same result.

Now we present the fundamental properties of the conditional expectation.

1) Linearity: $\mathbf{E}(a \cdot \xi + b \cdot \eta \mid \mathcal{D}) = a \cdot \mathbf{E}(\xi \mid \mathcal{D}) + b \cdot \mathbf{E}(\eta \mid \mathcal{D}),$ where a and b are arbitrary constants.

Proof. Let

$$\xi = \sum_{j=1}^{L} (x_j \cdot I_{A_j}) \quad \text{and} \quad \eta = \sum_{i=1}^{k} (y_i \cdot I_{D_i}).$$

Then

$$a \cdot \xi + b \cdot \eta = a \cdot \sum_{j=1}^{L} (x_j \cdot I_{A_j}) + b \cdot \sum_{i=1}^{k} (y_i \cdot I_{D_i}) = \sum_{ij} [(a \cdot x_j + b \cdot y_i) \cdot I_{A_j \cap D_i}].$$

Then $\mathbf{E}(a \cdot \xi + b \cdot \eta) = \sum_{ij} (a \cdot x_j + b \cdot y_i) + \mathbf{P}(A_j \cap D_i),$ and

$$\begin{aligned}\mathbf{E}(a \cdot \xi + b \cdot \eta \mid \mathcal{D}) &= \sum_{ij} (a \cdot x_j + b \cdot y_i) + \mathbf{P}(A_j \cap D_i \mid \mathcal{D}) \\ &= \sum_{j} [a \cdot x_j \cdot \mathbf{P}(A_j \mid \mathcal{D})] + \sum_{i} [b \cdot y_i \cdot \mathbf{P}(D_i \mid \mathcal{D})] \\ &= a \cdot \sum_{j} [x_j \cdot \mathbf{P}(A_j \mid \mathcal{D})] + b \cdot \sum_{i} [y_i \cdot \mathbf{P}(D_i \mid \mathcal{D})] = a \cdot \mathbf{E}(\xi \mid \mathcal{D}) + b \cdot \mathbf{E}(\eta \mid \mathcal{D}). \ \square\end{aligned}$$

2) $\mathbf{E}(\xi \mid \mathcal{D}_{\min}) = \mathbf{E}(\xi).$

Proof. Recall that $\mathcal{D}_{\min}$ consists of only one set Ω, that is, $\mathcal{D}_{\min} = \{\Omega\}$.

$$\mathbf{E}(\xi \mid \mathcal{D}_{\min}) = \sum_{i=1}^{k} \left(\frac{\mathbf{E}(\xi \cdot I_{D_i})}{\mathbf{P}(D_i)} \cdot I_{D_i} \right) = \frac{\mathbf{E}(\xi \cdot \Omega)}{\mathbf{P}(\Omega)} \cdot I_{\Omega} = \mathbf{E}(\xi). \qquad \square$$

3) Constants: $\mathbf{E}(c \mid \mathcal{D}) = c,$ where c is a constant.

4) If $\xi = I_A,$ then $\mathbf{E}(\xi \mid \mathcal{D}) = \mathbf{P}(A \mid \mathcal{D}).$

The next property generalizes the Law of Total Probability.

5) Expectation Law: $\mathbf{E}[\mathbf{E}(\xi \mid \mathcal{D})] = \mathbf{E}(\xi).$

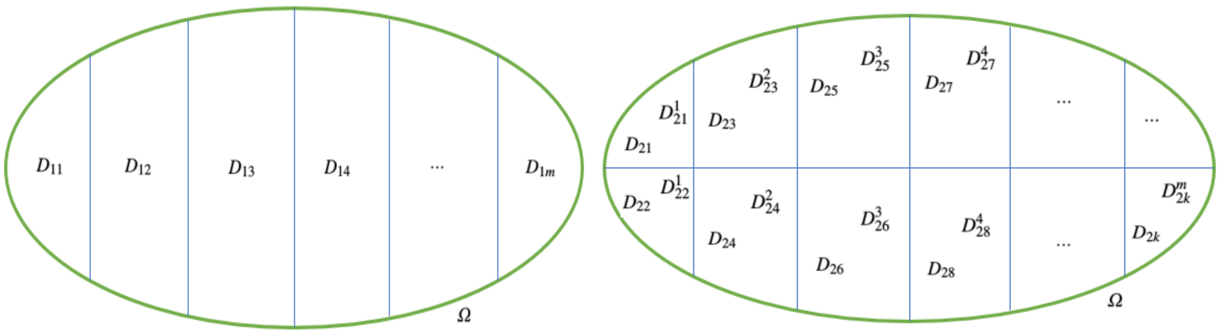
Proof. It is enough to observe what follows:

$$\mathbf{E}[\mathbf{E}(\xi \mid \mathcal{D})] = \mathbf{E}\left(\sum_{j=1}^{L} x_j \mathbf{P}(A_j \mid \mathcal{D}) \right) = \sum_{j=1}^{L} x_j \mathbf{E}[\mathbf{P}(A_j \mid \mathcal{D})] = \sum_{j=1}^{L} x_j \mathbf{P}(A_j)] = \mathbf{E}(\xi). \qquad \square$$

6) Tower Property: Let $\mathcal{D}_1 \preceq \mathcal{D}_2$. Then $\mathbf{E}[\mathbf{E}(\xi \mid \mathcal{D}_2) \mid \mathcal{D}_1] = \mathbf{E}(\xi \mid \mathcal{D}_1)$.

Proof. Let $\mathcal{D}_1 = \{D_{11}, D_{12}, \cdots, D_{1m}\}$ and $\mathcal{D}_2 = \{D_{21}, D_{22}, \cdots, D_{2k}\}$. Following Definition 1.4, since $\mathcal{D}_1 \preceq \mathcal{D}_2$, for any $1 \le p \le m$ there exists $J_p \subset \{1, 2, \cdots, k\}$ such that $D_{1p} = \bigcup_{q \in J_p} D_{2q}$ and index sets J_p are disjoint and $\bigcup_{p=1}^{m} J_p = \{1, 2, \cdots, k\}$, $k \ge m$. According to this we specially label or rearrange the atoms from the partition $\mathcal{D}_2$ taking into consideration as to which atom from the partition $\mathcal{D}_1$ they correspond. Namely, we attach an additional label p, so it becomes D_{2q}^p for atoms from $\mathcal{D}_2$, such that

$$D_{1p} = \bigcup_{q \in J_p} D_{2q}^p, \ 1 \le p \le m.$$



Note that in our “previous” notation for any event A,

$$\mathbf{P}(A \mid \mathcal{D}_2) = \sum_{q=1}^{k} [\mathbf{P}(A \mid D_{2q}) \cdot I_{D_{2q}}],$$

while with our “new” notation,

$$\mathbf{P}(A \mid \mathcal{D}_2) = \sum_{p=1}^{m} \sum_{q \in J_p} [\mathbf{P}(A \mid D_{2q}^p) \cdot I_{D_{2q}^p}].$$

The main advantage of our “new” notation that we will use in this proof is that

$$\mathbf{P}(D_{2q}^i \mid D_{1p}) = \begin{cases} 0, & i \ne p, \\ \dfrac{\mathbf{P}(D_{2q}^p)}{\mathbf{P}(D_{1p})}, & i = p. \end{cases}$$

Now, if $\xi = \sum_{j=1}^{l} x_j \cdot I_{A_j}$, then $\mathbf{E}(\xi \mid \mathcal{D}_2) = \sum_{j=1}^{l} x_j \cdot \mathbf{P}(A_j \mid \mathcal{D}_2)$ and

$$\mathbf{E}[\mathbf{E}(\xi \mid \mathcal{D}_2) \mid \mathcal{D}_1] = \mathbf{E}\left(\sum_{j=1}^{l} x_j \cdot \mathbf{P}(A_j \mid \mathcal{D}_2) \mid \mathcal{D}_1 \right) = \sum_{j=1}^{l} x_j \mathbf{E}[\mathbf{P}(A_j \mid \mathcal{D}_2) \mid \mathcal{D}_1].$$

Hence it is sufficient to show that $\mathbf{E}[\mathbf{P}(A_j \mid \mathcal{D}_2) \mid \mathcal{D}_1] = \mathbf{P}(A_j \mid \mathcal{D}_1)$. As we already mentioned above,

$$\mathbf{P}(A_j \mid \mathcal{D}_2) = \sum_{p=1}^{m} \sum_{q \in J_p} [\mathbf{P}(A_j \mid D_{2q}^p) \cdot I_{D_{2q}^p}].$$

and hence

$$\mathbf{E}[\mathbf{P}(A_j \mid \mathcal{D}_2) \mid \mathcal{D}_1]$$

$$= \sum_{p=1}^{m} \sum\nolimits_{q \in J_p} [\mathbf{P}(A_j \mid D_{2q}^{p}) \cdot \mathbf{P}(D_{2q}^{p} \mid \mathcal{D}_1)]$$

$$= \sum_{p=1}^{m} \sum\nolimits_{q \in J_p} \mathbf{P}(A_j \mid D_{2q}^{p}) \cdot \left(\sum_{i=1}^{m} \mathbf{P}(D_{2q}^{p} \mid D_{1i}) \cdot I_{D_{1i}} \right)$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \sum_{p=1}^{m} \sum\nolimits_{q \in J_p} [\mathbf{P}(A_j \mid D_{2q}^{p}) \cdot \mathbf{P}(D_{2q}^{p} \mid D_{1i})]$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \left(\sum_{p=1, p \neq i}^{m} \sum\nolimits_{q \in J_p} [\mathbf{P}(A_j \mid D_{2q}^{p}) \cdot \mathbf{P}(D_{2q}^{p} \mid D_{1i})] + \sum\nolimits_{q \in J_i} [\mathbf{P}(A_j \mid D_{2q}^{i}) \cdot \mathbf{P}(D_{2q}^{i} \mid D_{1i})] \right)$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \left(\sum_{p=1, p \neq i}^{m} \sum\nolimits_{q \in J_p} [\mathbf{P}(A_j \mid D_{2q}^{p}) \cdot 0] + \sum\nolimits_{q \in J_i} \mathbf{P}(A_j \mid D_{2q}^{i}) \cdot \frac{\mathbf{P}(D_{2q}^{i})}{\mathbf{P}(D_{1i})} \right)$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \cdot \sum\nolimits_{q \in J_i} \mathbf{P}(A_j \mid D_{2q}^{i}) \cdot \frac{\mathbf{P}(D_{2q}^{i})}{\mathbf{P}(D_{1i})}$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \cdot \sum\nolimits_{q \in J_i} \frac{\mathbf{P}(A_j \cap D_{2q}^{i})}{\mathbf{P}(D_{2q}^{i})} \cdot \frac{\mathbf{P}(D_{2q}^{i})}{\mathbf{P}(D_{1i})}$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \cdot \sum\nolimits_{q \in J_i} \frac{\mathbf{P}(A_j \cap D_{2q}^{i})}{\mathbf{P}(D_{1i})}$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \cdot \frac{1}{\mathbf{P}(D_{1i})} \sum\nolimits_{q \in J_i} \mathbf{P}(A_j \cap D_{2q}^{i})$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \cdot \frac{1}{\mathbf{P}(D_{1i})} \mathbf{P}(A_j \cap D_{1i}) \qquad (\text{ since } \sum\nolimits_{q \in J_i} D_{2q}^{i} = D_{1i})$$

$$= \sum_{i=1}^{m} I_{D_{1i}} \cdot \mathbf{P}(A_j \mid D_{1i}) = \mathbf{P}(A_j \mid \mathcal{D}_1). \qquad \square$$

7) Suppose the random variable η is measurable with respect the partition $\mathcal{D}$. Then

$$\mathbf{E}(\eta \mid \mathcal{D}) = \eta.$$

Proof. Because η is $\mathcal{D}$-*measurable,* we have $\eta = \sum_{i=1}^{k} y_i \cdot I_{D_i}$, where $\mathcal{D}_i \preceq \mathcal{D}$. Then

$$\mathbf{E}(\eta \mid \mathcal{D}) = \sum_{i=1}^{k} (\eta \mid \mathcal{D}_i) \cdot I_{D_i} = \sum_{i=1}^{k} \mathbf{E}\left(\sum_{j=1}^{L} y_i \cdot I_{D_i} \mid \mathcal{D}_i \right) \cdot I_{D_i}$$

$$= \sum_{i=1}^{k}\sum_{j=1}^{L}\frac{\mathbf{E}(y_i \cdot I_{D_j} \cdot I_{D_i})}{\mathbf{P}(D_i)} \cdot I_{D_i}$$

$$= \sum_{i=1}^{k}\frac{\mathbf{E}(y_i \cdot I_{D_i})}{\mathbf{P}(D_i)} \cdot I_{D_i} = \sum_{i=1}^{k} y_i \frac{\mathbf{E}(I_{D_i})}{\mathbf{P}(D_i)} \cdot I_{D_i} = \eta. \qquad \square$$

8) Independence Law: If the random variables ξ and η are independent, then

$$\mathbf{E}(\xi \mid \eta) = \mathbf{E}(\xi).$$

Proof. Let $\eta = \sum_{i=1}^{k} y_i \cdot I_{D_i}$. Thus the partition generated by η is $\{D_1, D_2, \cdots, D_k\}$. Further suppose $\xi = \sum_{j=1}^{L} x_j \cdot I_{A_j}$. Thus the partition generated by ξ is $\{A_1, A_2, \cdots, A_L\}$. Since ξ and η are independent, this means that their partitions $\mathcal{D}_i$ and A_j are independent, that is,

$$\mathbf{P}(D_i \cap A_j) = \mathbf{P}(D_i)\mathbf{P}(A_j),\ 1 \le i \le k,\ 1 \le j \le L.$$

Equivalently, $\mathbf{P}(D_i \mid A_j) = \mathbf{P}(D_i)$ or $\mathbf{P}(A_j \mid D_i) = \mathbf{P}(A_j)$. Re-writing in terms of expectations we can express this independence as

$$\mathbf{E}(I_{D_i} \cdot I_{A_j}) = \mathbf{E}(I_{D_i}) \cdot \mathbf{E}(I_{A_j}) = 0 \cdot \mathbf{P}(D_i \cap A_j)^c + 1 \cdot \mathbf{P}(D_i \cap A_j) = \mathbf{P}(D_i \cap A_j) = \mathbf{P}(D_i)\mathbf{P}(A_j).$$

Then $\mathbf{E}(\xi \mid \eta) = \sum_{i=1}^{k} \frac{\mathbf{E}(\xi \cdot I_{D_i})}{\mathbf{P}(D_i)} \cdot I_{D_i}$, where

$$\mathbf{E}(\xi \cdot I_{D_i}) = \mathbf{E}\left(\sum_{j=1}^{L} x_j \cdot I_{A_j}\right) \cdot I_{D_i} = \sum_{j=1}^{L}[x_j \cdot \mathbf{E}(I_{D_i} \cdot I_{A_j})]$$

$$= \sum_{j=1}^{L}[x_j \mathbf{E}(I_{D_i}) \cdot \mathbf{E}(I_{A_j})] = \sum_{j=1}^{L}[x_j \mathbf{P}(D_i) \cdot \mathbf{P}(A_j)].$$

Returning to our original equation,

$$\mathbf{E}(\xi \mid \eta) = \sum_{i=1}^{k}\sum_{j=1}^{L}\{[x_j \mathbf{P}(D_i) \cdot \mathbf{P}(A_j)] \cdot I_{D_i}\} = \left(\sum_{j=1}^{L}[x_j \mathbf{P}(A_j)]\right)\left(\sum_{i=1}^{k}\mathbf{P}(D_i)\right) = \mathbf{E}(\xi). \qquad \square$$

9) $\mathbf{E}(\eta \mid \eta) = \eta.$

10) Stability: If the random variable η is $\mathcal{D}$-measurable, then $\mathbf{E}(\eta \cdot \xi \mid \mathcal{D}) = \eta \cdot \mathbf{E}(\xi \mid \mathcal{D})$.

Proof. Let $\xi = \sum_{j=1}^{L} x_j \cdot I_{A_j}$ and $\eta = \sum_{i=1}^{k} y_i \cdot I_{D_i}$. Then $\xi \cdot \eta = \sum_{j=1}^{L}\sum_{i=1}^{k}(x_j y_j \cdot I_{A_j D_i})$.

Hence

$$\mathbf{E}(\eta \cdot \xi \mid \mathcal{D}) = \sum_{j=1}^{L}\sum_{i=1}^{k}[x_j y_j \cdot \mathbf{P}(A_j D_i \mid \mathcal{D})]$$

$$= \sum_{j=1}^{L}\sum_{i=1}^{k}(x_j y_j)\sum_{m=1}^{k}[\mathbf{P}(A_j D_i \mid D_m) \cdot I_{D_m}]$$

$$= \sum_{j=1}^{L}\sum_{i=1}^{k}[x_j y_j \cdot \mathbf{P}(A_j D_i \mid D_i)\cdot I_{D_i}] = \sum_{j=1}^{L}\sum_{i=1}^{k}[x_j y_j \cdot \mathbf{P}(A_j \mid D_i)\cdot I_{D_i}].$$

On the other hand, since $I_{D_i}^2 = I_{D_i}$, $I_{D_i} I_{D_m} = 0\ (i \neq m)$, we have

$$\eta \cdot \mathbf{E}(\xi \mid \mathcal{D}) = \left(\sum_{i=1}^{k} y_j I_{D_i}\right)\left(\sum_{j=1}^{L} x_j \mathbf{P}(A_j \mid \mathcal{D})\right)$$

$$= \left(\sum_{i=1}^{k} y_j I_{D_i}\right)\cdot \sum_{m=1}^{k}\left(\sum_{j=1}^{L} x_j \mathbf{P}(A_j \mid D_m)\right)\cdot I_{D_m}$$

$$= \sum_{j=1}^{L}\sum_{i=1}^{k}[x_j y_j \cdot \mathbf{P}(A_j \mid D_i)\cdot I_{D_i}]. \qquad \square$$

3. Martingales

Let $(\Omega, \mathcal{D}, \mathbf{P})$ be a finite probability space, and let $\mathcal{D}_1 \preceq \mathcal{D}_2 \preceq \cdots \preceq \mathcal{D}_n$ be a sequence of partitions of Ω.

Definition 2.1 A sequence of random variables $\xi_1, \xi_2, \cdots, \xi_n$ is called a *martingale* (with respect to the sequence of partitions $\mathcal{D}_1 \preceq \mathcal{D}_2 \preceq \cdots \preceq \mathcal{D}_n$), if

1) ξ_k is $\mathcal{D}_k$-measurable, for each $1 \le k \le n$; and
2) $\mathbf{E}(\xi_{k+1} \mid \mathcal{D}_k) = \xi_k$, $1 \le k \le n-1$.

We denote such a martingale by $\xi = (\xi_k, \mathcal{D}_k)_{k=1}^{n}$. In the particular case, when $\mathcal{D}_k = \mathcal{D}_{\xi_1,\xi_2,\cdots,\xi_k}$, we say that the sequence $\xi = (\xi_k)$ is a martingale without specifying the sequence of partitions associated. □

From the definition of a martingale it follows immediately that the mathematical expectation $\mathbf{E}(\xi_k)$ is constant for all k: $\mathbf{E}(\xi_k) = \mathbf{E}(\xi_1)$. We now consider a classical example illustrating a basic example of martingales.

Example 2.1 Let $\eta_1, \eta_2, \cdots, \eta_n$ be independent identically distributed random variables with a Bernoulli distribution $\mathbf{P}(\eta_k = -1) = \mathbf{P}(\eta_k = 1) = 1/2$. Define

$$S_k = \eta_1 + \eta_2 + \cdots + \eta_k \qquad \text{and} \qquad \mathcal{D}_k = \mathcal{D}_{\eta_1,\eta_2,\cdots,\eta_k}.$$

The structure of the partitions associated with $\mathcal{D}_k$ is simple:

$$\mathcal{D}_1 = \{D^+, D^-\},$$

where $D^+ = \{\omega : \eta_1 = 1\}$, $D^- = \{\omega : \eta_1 = -1\}$; and

$$\mathcal{D}_2 = \{D^{++}, D^{+-}, D^{-+}, D^{--}\},$$

where $D^{++} = \{\omega : \eta_1 = 1, \eta_2 = 1\}$, $D^{--} = \{\omega : \eta_1 = -1, \eta_2 = -1\}$, and so on. Since

$$\mathcal{D}_{\eta_1,\eta_2,\cdots,\eta_k} = \mathcal{D}_{S_1,S_2,\cdots,S_k},$$

each of the random variables S_k is $\mathcal{D}_k$-measurable. Now we show that the second assumption of a martingale is also hold. Observe that

$$\mathbf{E}(S_{k+1} \mid \mathcal{D}_k) = \mathbf{E}(S_k + \eta_{k+1} \mid \mathcal{D}_k) = \mathbf{E}(S_k \mid \mathcal{D}_k) + \mathbf{E}(\eta_{k+1} \mid \mathcal{D}_k) = S_k + \mathbf{E}(\eta_{k+1}) = S_k.$$

References

[1] Cox, J. C., Ross, S. A., and Rubinstein, M. (1979). Option pricing: a simplified approach, *Journal of Financial Economics*, **7**(3), 229-263.

[2] Shiryaev, A. N. (2016). Probability, 3rd ed., translated from the fourth (2007) Russian edition by R. P. Boas and D. M. Chibisov. Graduate Texts in Mathematics, 95. Springer, New York.

[3] Shiryaev, A. N., Kabanov, Yu. M., Kramkov, D. O., and Mel'nikov, A. V. (1995). Toward a theory of pricing options of European and American types, I. Discrete time (Russian), translated from Teor. Veroyatnost. i Primenen. (1994), **39**(1), 23-79, *Theory of Probability and its Applications* (1994), **39**(1), 14-60.

 JPSS: ISSN 1726-3328