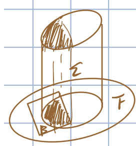
2 Conditional Expectation

2.1 Conditioning on an Event

Definition 2.1

For any integrable random variable ξ and any event $B \in \mathcal{F}$ such that $P(B) \neq 0$ the *conditional expectation* of ξ given B is defined by

$$E(\xi|B) = \frac{1}{P(B)} \int_B \xi \, dP.$$



Example 2.1

Three coins, 10p, 20p and 50p are tossed. The values of those coins that land heads up are added to work out the total amount ξ. What is the expected total amount ξ given that two coins have landed heads up?

Let B denote the event that two coins have landed heads up. We want to find $E(\xi|B)$. Clearly, B consists of three elements,

$$B = \{\mathrm{HHT}, \mathrm{HTH}, \mathrm{THH}\},$$

each having the same probability $\frac{1}{8}$. (Here H stands for heads and T for tails.) The corresponding values of ξ are

$$\begin{aligned} \xi(\mathrm{HHT}) &= 10 + 20 = 30, \\ \xi(\mathrm{HTH}) &= 10 + 50 = 60, \\ \xi(\mathrm{THH}) &= 20 + 50 = 70. \end{aligned}$$

Therefore

$$E(\xi|B) = \frac{1}{P(B)} \int_B \xi \, dP = \frac{1}{\frac{3}{8}} \left(\frac{30}{8} + \frac{60}{8} + \frac{70}{8} \right) = 53\tfrac{1}{3}.$$

Exercise 2.1

Show that $E(\xi|\Omega) = E(\xi)$.

Hint The definition of $E(\xi)$ involves an integral and so does the definition of $E(\xi|\Omega)$. How are these integrals related?

Solution 2.1

Since $P(\Omega) = 1$ and $\int_\Omega \xi \, dP = E(\xi)$,

$$E(\xi|\Omega) = \frac{1}{P(\Omega)} \int_\Omega \xi \, dP = E(\xi).$$

Exercise 2.2

Show that if

$$1_A(\omega) = \begin{cases} 1 & \text{for } \omega \in A \\ 0 & \text{for } \omega \notin A \end{cases}$$

(the *indicator function* of A), then

$$E(1_A|B) = P(A|B),$$

where

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

is the *conditional probability* of A given B.

Hint Write $\int_B 1_A \, dP$ as $P(A \cap B)$.

Solution 2.2

By Definition 2.1

$$\begin{aligned} E(1_A|B) &= \frac{1}{P(B)} \int_B 1_A \, dP \\ &= \frac{1}{P(B)} \int_{A \cap B} dP \\ &= \frac{P(A \cap B)}{P(B)} \\ &= P(A|B). \end{aligned}$$

2.2 Conditioning on a Discrete Random Variable η with possible values $y_1, y_2, \ldots$

= Conditioning on the events $\{\eta = y_n\}$

Definition 2.2

Let ξ be an integrable random variable and let η be a discrete random variable as above. Then the *conditional expectation* of ξ given η is defined to be a random variable $E(\xi|\eta)$ such that

$$E(\xi|\eta)(\omega) = E(\xi|\{\eta = y_n\}) \quad \text{if } \eta(\omega) = y_n$$

for any $n = 1, 2, \ldots$. variable event

Example 2.2

{HHH……TTT} 8种可能，古典概形

Three coins, 10p, 20p and 50p are tossed as in Example 2.1. What is the conditional expectation $E(\xi|\eta)$ of the total amount ξ shown by the three coins given the total amount η shown by the 10p and 20p coins only?

Clearly, η is a discrete random variable with four possible values: 0, 10, 20 and 30. We find the four corresponding conditional expectations in a similar way as in Example 2.1:

$$E(\xi|\{\eta = 0\}) = 25, \qquad E(\xi|\{\eta = 10\}) = 35,$$
$$E(\xi|\{\eta = 20\}) = 45, \qquad E(\xi|\{\eta = 30\}) = 55.$$

Therefore

$$E(\xi|\eta)(\omega) = \begin{cases} 25 & \text{if } \eta(\omega) = 0, \\ 35 & \text{if } \eta(\omega) = 10, \\ 45 & \text{if } \eta(\omega) = 20, \\ 55 & \text{if } \eta(\omega) = 30. \end{cases}$$

Example 2.3

Take $\Omega = [0, 1]$ with the σ-field of Borel sets and P the Lebesgue measure on $[0, 1]$. We shall find $E(\xi|\eta)$ for

$$\xi(x) = 2x^2, \qquad \eta(x) = \begin{cases} 1 & \text{if } x \in [0, \frac{1}{3}], \\ 2 & \text{if } x \in (\frac{1}{3}, \frac{2}{3}), \\ 0 & \text{if } x \in [\frac{2}{3}, 1]. \end{cases}$$

Clearly, η is discrete with three possible values 1, 2, 0. The corresponding events are

$$\{\eta = 1\} = [0, \tfrac{1}{3}],$$
$$\{\eta = 2\} = (\tfrac{1}{3}, \tfrac{2}{3}),$$
$$\{\eta = 0\} = [\tfrac{2}{3}, 1].$$

For $x \in [0, \frac{1}{3}]$

$$E(\xi|\eta)(x) = E(\xi|[0, \tfrac{1}{3}]) = \frac{1}{\frac{1}{3}} \int_0^{\frac{1}{3}} 2x^2 dx = \frac{2}{27}.$$

For $x \in (\frac{1}{3}, \frac{2}{3})$

$$E(\xi|\eta)(x) = E(\xi|(\tfrac{1}{3}, \tfrac{2}{3})) = \frac{1}{\frac{1}{3}} \int_{\frac{1}{3}}^{\frac{2}{3}} 2x^2 dx = \frac{14}{27}.$$

And for $x \in (\frac{2}{3}, 1]$

$$E(\xi|\eta)(x) = E(\xi|(\tfrac{2}{3}, 1]) = \frac{1}{\frac{1}{3}} \int_{\frac{2}{3}}^{1} 2x^2 dx = \frac{38}{27}.$$

The graph of $E(\xi|\eta)$ is shown in Figure 2.1 together with those of ξ and η.

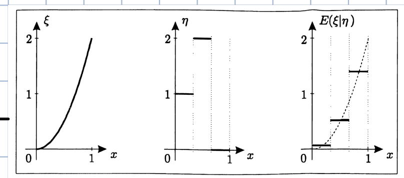


Figure 2.1. The graph of $E(\xi|\eta)$ in Example 2.3

Exercise 2.3

Show that if η is a constant function, then $E(\xi|\eta)$ is constant and equal to $E(\xi)$.

Hint The event $\{\eta = c\}$ must be $\emptyset$ or Ω for any $c \in \mathbb{R}$.

Proof: Since $\eta: \Omega \to R$ and η is a constant function

then $\exists c \in R$ s.t. for all $w \in \Omega$, $\eta(w) = c$

i.e. $\{\eta \in c\} = \Omega$

Thus, $E[\xi|\eta] = E[\xi|\{\eta = c\}] = E[\xi|\Omega] = E[\xi]$.

Q.E.D.

$$\begin{cases} \{\eta = c\} = \Omega, & \text{if } c = d \\ \{\eta = c\} = \phi, & \text{if } c \neq d \end{cases}$$

Solution 2.3

Since η is constant, it has only one value $c \in \mathbb{R}$, for which

$$\{\eta = c\} = \Omega.$$

Therefore $E(\xi|\eta)$ is constant and equal to

$$E(\xi|\eta)(\omega) = E(\xi|\{\eta = c\}) = E(\xi|\Omega) = E(\xi)$$

for each $\omega \in \Omega$. The last equality has been verified in Exercise 2.1.

Exercise 2.4

Show that

$$E(1_A|1_B)(\omega) = \begin{cases} P(A|B) & \text{if } \omega \in B \\ P(A|\Omega \setminus B) & \text{if } \omega \notin B \end{cases}$$

for any B such that $1 \neq P(B) \neq 0$.

Proof: $1_B = \begin{cases} 1, & w \in B \\ 0, & w \notin B \end{cases} \Rightarrow \{1_B = 1\} \in B, \ \{1_B = 0\} = \Omega \setminus B$

If $w \in B$, then $E(1_A|1_B)(w) = E(1_A|\{1_B = 1\}) = E(1_A|B) = P(A|B)$

If $w \notin B$, then $E(1_A|1_B)(w) = E(1_A|\{1_B = 0\}) = E(1_A|\Omega \setminus B) = P(A|\Omega \setminus B)$.

Solution 2.4

The indicator function 1_B takes two values 1 and 0. The sets on which these values are taken are

$$\{1_B = 1\} = B, \qquad \{1_B = 0\} = \Omega \setminus B.$$

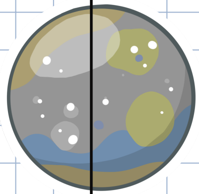
Thus, for $\omega \in B$

$$E(1_A|1_B)(\omega) = E(1_A|B) = P(A|B),$$

as in Exercise 2.2. Similarly, for $\omega \in \Omega \setminus B$

$$E(1_A|1_B)(\omega) = E(1_A|\Omega \setminus B) = P(A|\Omega \setminus B).$$

$$E[1_A|1_B] = \frac{1}{P(1_B)} \int_{1_B} 1_A dP$$
$$= \frac{1}{P(1_B = 0)} \int_{\{1_B = 0\}} 1_A dP$$
$$= \frac{1}{P(\Omega \setminus B)} \int_{\Omega \setminus B} 1_A dP$$
$$= \frac{P[A \cap (\Omega \setminus B)]}{P(\Omega \setminus B)}.$$



Exercise 2.5

Assuming that η is a discrete random variable, show that

$$E(E(\xi|\eta)) = E(\xi).$$

$$E(\xi|\eta) = \begin{cases} \cdots & \eta = y_1 \\ \vdots & \\ \cdots & \eta = y_n \end{cases}$$

Hint Observe that

$$\int_B E(\xi|\eta)\, dP = \int_B \xi\, dP$$

for any event B on which η is constant. The desired equality can be obtained by covering Ω by countably many disjoint events of this kind.

Proof: For any event B: $\int_B E(\xi|B)\,dP = \int_B \left(\frac{1}{P(B)}\int_B \xi\,dP\right)dP = \int_B \xi\,dP$.

Since η is discrete, it can equal to countably many values $\eta = y_1, y_2, \ldots$ Assume that $y_i \neq y_j$ if $i \neq j$, then

$\boxed{\text{the sets } \{\eta = y_1\}, \{\eta = y_2\}, \ldots \text{ are pairwise disjoint.}}$ #别忘

and $\{\eta = y_1\} \cup \{\eta = y_2\} \cup \cdots = \Omega$

Therefore,

$$\begin{aligned} E(E(\xi|\eta)) &= \int_\Omega E(\xi|\eta)\,dP \\ &= \sum_n \int_{\{\eta = y_n\}} E(\xi|\{\eta = y_n\})\,dP \\ &= \sum_n \int_{\{\eta = y_n\}} \frac{1}{P(\{\eta = y_n\})}\left(\int_{\{\eta = y_n\}} \xi\,dP\right)dP = \sum_n \int_{\{\eta = y_n\}} \xi\,dP \\ &= \int_\Omega \xi\,dP \\ &= E(\xi) \end{aligned}$$

先拆Ω

Thus, $E(E(\xi|\eta)) = E(\xi)$ when η is discrete

Q.E.D.

Solution 2.5

First we observe that

$$\int_B E(\xi|B)\,dP = \int_B \left(\frac{1}{P(B)}\int_B \xi\,dP\right)dP = \int_B \xi\,dP \tag{2.4}$$

for any event B.

Since η is discrete, it has countably many values $y_1, y_2, \ldots$. We can assume that these values are pairwise distinct, i.e. $y_i \neq y_j$ if $i \neq j$. The sets $\{\eta = y_1\}, \{\eta = y_2\}, \ldots$ are then pairwise disjoint and they cover the whole space Ω. Therefore, by (2.4)

$$\begin{aligned} E(E(\xi|\eta)) &= \int_\Omega E(\xi|\eta)\,dP \\ &= \sum_n \int_{\{\eta = y_n\}} E(\xi|\{\eta = y_n\})\,dP \\ &= \sum_n \int_{\{\eta = y_n\}} \xi\,dP \\ &= \int_\Omega \xi\,dP \\ &= E(\xi). \end{aligned}$$

Proposition 2.1

If ξ is an integrable random variable and η is a discrete random variable, then

1) $E(\xi|\eta)$ is $\sigma(\eta)$-measurable;
2) For any $A \in \sigma(\eta)$

$$\int_A E(\xi|\eta)\,dP = \int_A \xi\,dP. \tag{2.1}$$

Proof

Suppose that η has pairwise distinct values $y_1, y_2, \ldots$. Then the events

$$\{\eta = y_1\}, \{\eta = y_2\}, \ldots$$

1) Proof: Assume that $y_i \neq y_j$ for $i \neq j$, $i, j \in \mathbb{N}$.

then the events $\{\eta = y_1\}, \{\eta = y_2\}, \ldots$ are pairwise disjoint

and $\{\eta = y_1\} \cup \{\eta = y_2\} \cup \cdots = \Omega$.

As the σ-field $\sigma(\eta)$ is generated by these events

$\Rightarrow$ any set $A \in \sigma(\eta)$ is a countable union of the sets of the form $\{\eta = y_i\}$.

As $E[\xi|\eta]$ is constant on each set $\{\eta = y_i\}$, $i \in \mathbb{N}$

and $E[\xi|\eta](\omega) = E[\xi|\{\eta = y_i\}]$ if $\eta(\omega) = y_i$ $\forall i \in \mathbb{N}$.

Thus, $\forall c \in \mathbb{R}$, $\{E[\xi|\eta] = c\} = \{\eta = y_i\} \in \sigma(\eta)$

if $c = E[\xi|\{\eta = y_i\}]$ for some $i \in \mathbb{N}$.

or $\{E[\xi|\eta] = c\} = \phi \in \sigma(\eta)$ if $c \neq E[\xi|\{\eta = y_i\}]$ $\forall i \in \mathbb{N}$.

Thus, $E[\xi|\eta]$ is $\sigma(\eta)$-measurable.

Q.E.D.

What I Want

2) Proof: $\forall i \in \mathbb{N}$,

$$\int_{\{\eta = y_i\}} E[\xi|\eta](\omega)P(d\omega) \overset{\text{Def. 2.2}}{=} \int_{\{\eta = y_i\}} E[\xi|\{\eta = y_i\}](\omega)P(d\omega)$$

$$\overset{\text{Def 2.1}}{=} \int_{\{\eta = y_i\}} \frac{1}{P(\{\eta = y_i\})}\int_{\{\eta = y_i\}} \xi(\omega)P(d\omega)$$

$$= \int_{\{\eta = y_i\}} \xi(\omega)P(d\omega).$$

As $\{\eta = y_1\}, \{\eta = y_2\}, \ldots$ are pairwise disjoint and cover Ω,

then each $A \in \sigma(\eta)$ is a countable union of the events of the form $\{\eta = y_i\}$, thus,

$$\int_A E[\xi|\eta](\omega)P(d\omega) = \int_A \xi(\omega)P(d\omega).$$

Q.E.D.

Proof. Suppose that η has pairwise distinct values $y_1, y_2, \ldots$. Then, the events

$$\{\eta = y_1\},\ \{\eta = y_2\},\ \ldots$$

are pairwise disjoint and cover Ω. The σ-field $\sigma(\eta)$ is generated by these events, in fact that every $A \in \sigma(\eta)$ is a countable union of the sets of the form $\{\eta = y_i\}$. Note that $\mathbb{E}[\xi \mid \eta]$ is constant on each of these sets of the form $\{\eta = y_i\}$.

$\{\eta = y_i\} \in \sigma(\eta)$

In fact,

$$\mathbb{E}[\xi \mid \eta](\omega) = \mathbb{E}[\xi \mid \{\eta = y_i\}] \quad \text{if} \quad \eta(\omega) = y_i$$

for any $i = 1, 2, \ldots$ by Definition 2.2. Thus, for any $c \in \mathbb{R}$ we have

$$\{\mathbb{E}[\xi \mid \eta] = c\} = \{\eta = y_i\} \in \sigma(\eta), \text{ if } c = \mathbb{E}[\xi \mid \{\eta = y_i\}] \text{ for some } i,$$

or

$$\{\mathbb{E}[\xi \mid \eta] = c\} = \emptyset \in \sigma(\eta), \text{ if } c \neq \mathbb{E}[\xi \mid \{\eta = y_i\}] \text{ for all } i.$$

Therefore, it must be $\sigma(\eta)$-measurable.

For each i, according to Definitions 2.1 and 2.2, we have

$$\int_{\{\eta=y_i\}} \mathbb{E}[\xi \mid \eta](\omega)\mathbb{P}(\mathrm{d}\omega) \overset{2.2}{=} \int_{\{\eta=y_i\}} \mathbb{E}[\xi \mid \{\eta = y_i\}](\omega)\mathbb{P}(\mathrm{d}\omega) \overset{2.1}{=} \int_{\{\eta=y_i\}} \xi(\omega)\mathbb{P}(\mathrm{d}\omega).$$

Since each $A \in \sigma(\eta)$ is a countable union of sets of the form $\{\eta = y_i\}$, which are pairwise disjoint, it follows that

$$\int_A \mathbb{E}[\xi \mid \eta](\omega)\mathbb{P}(\mathrm{d}\omega) = \int_A \xi(\omega)\mathbb{P}(\mathrm{d}\omega),$$

as required. □

2.3 Conditioning on an Arbitrary Random Variable

Definition 2.3

Let ξ be an integrable random variable and let η be an arbitrary random variable. Then the *conditional expectation* of ξ given η is defined to be a random variable $E(\xi|\eta)$ such that

1) $E(\xi|\eta)$ is $\sigma(\eta)$-measurable; $\Rightarrow$

2) For any $A \in \sigma(\eta)$

$$\int_A E(\xi|\eta)\, dP = \int_A \xi\, dP.$$

1) $\mathbb{E}[\xi \mid \eta]$ is a $\sigma(\eta)$-measurable random variable, i.e., according to Doob-Dynkin Lemma,

$$\mathbb{E}[\xi \mid \eta] = \varphi(\eta);$$

Remark 2.1

We can also define the *conditional probability* of an event $A \in \mathcal{F}$ given η by

$$P(A|\eta) = E(1_A|\eta),$$

where 1_A is the indicator function of A.

Remark. We recall that two random variables $\xi = \eta$ a.s. if and only if $\mathbb{P}(\xi = \eta) = 1$, or equivalently, $\mathbb{P}(\xi \neq \eta) = 0$.

Do the conditions of Definition 2.3 characterize $E(\xi|\eta)$ uniquely? The lemma below implies that $E(\xi|\eta)$ is defined to within equality on a set of full measure. Namely,

Uniques

$$\text{if } \xi = \xi' \text{ a.s., then } E(\xi|\eta) = E(\xi'|\eta) \text{ a.s.} \tag{2.2}$$

$\#\ E[\xi|\eta] = E[\xi'|\eta] \not\Rightarrow \xi = \xi'$

The existence of $E(\xi|\eta)$ will be discussed later in this chapter.

Lemma 2.1

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$. If ξ is a $\mathcal{G}$-measurable random variable and for any $B \in \mathcal{G}$

$$\int_B \xi \, dP = 0,$$

then $\xi = 0$ a.s. $P(\xi=0)=1$ i.e. $\forall c>0$, $P(\{\xi \geq c\}) = P(\{\xi \leq -c\}) = 0$

Proof

Observe that $P\{\xi \geq \varepsilon\} = 0$ for any $\varepsilon > 0$ because

$$0 \leq \varepsilon P\{\xi \geq \varepsilon\} = \int_{\{\xi \geq \varepsilon\}} \varepsilon \, dP \leq \int_{\{\xi \geq \varepsilon\}} \xi \, dP = 0.$$

The last equality holds, since $\{\xi \geq \varepsilon\} \in \mathcal{G}$. Similarly, $P\{\xi \leq -\varepsilon\} = 0$ for any $\varepsilon > 0$. As a consequence,

$$P\{-\varepsilon < \xi < \varepsilon\} = 1$$

for any $\varepsilon > 0$.

Let us put

$$A_n = \{-\tfrac{1}{n} < \xi < \tfrac{1}{n}\}.$$

Then $P(A_n) = 1$ and

$$\{\xi = 0\} = \bigcap_{n=1}^{\infty} A_n.$$

Because the A_n form a contracting sequence of events, it follows that

$$P\{\xi = 0\} = \lim_{n\to\infty} P(A_n) = 1,$$

as required. □

Proof. Since ξ is G-measurable ($\xi: \Omega \to \mathbb{R}$) then $\forall c>0$, $\underbrace{\{\xi \geq c\}}_{B} \in G$

i.e. $\{w \mid \xi(w) \geq c\} \in G$.

$\Rightarrow \boxed{0 \leq cP(\xi \geq c)} = c\int_{\{\xi \geq c\}} 1 \, P(dw)$

$= \int_{\{\xi \geq c\}} c \, P(dw) \leq \int_{\{\xi \geq c\}} \xi(w) P(dw) = 0$

$\Rightarrow P(\xi \geq c) = 0$

Similarly, $\boxed{0 \geq -cP(\xi \leq -c)} = \int_{\{\xi \leq -c\}} -c \, P(dw) \geq \int_{\{\xi \leq -c\}} \xi(w) P(dw) = 0$

$\Rightarrow P(\xi \leq -c) = 0$

Thus, $P(-c < \xi < c) = 1 - P(\xi \leq -c) - P(\xi \geq c) = 1$

Recall c is arbitrary positive real number, then

for any $n \in \mathbb{N}$ denote $A_n = \{-\frac{1}{n} < \xi < \frac{1}{n}\}$, $P(A_n) = 1$

On the other hand, $\{\xi = 0\} = \bigcap_{n=1}^{\infty} A_n$

Thus, $P(\{\xi=0\}) = \lim_{n\to\infty} P(A_n) = 1$ as $A_1, A_2 \dots$ is a contracting sequence of events.

Q.E.D.

One way to show the existence of $E[\xi|\eta]$:

by finding out the exact formula of $E[\xi|\eta] = \varphi(\eta)$.

欲证

Show the Existence of $E[\xi|\eta]$, which satisfies the conditions 1) and 2) of Def. 2.3.

已知

Example 2.4

Take $\Omega = [0,1]$ with the σ-field of Borel sets and P the Lebesgue measure on $[0,1]$. We shall find $E(\xi|\eta)$ for

$$\xi(x) = 2x^2, \qquad \eta(x) = \begin{cases} 2 & \text{if } x \in [0, \frac{1}{2}), \\ x & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

Here η is no longer discrete and the general Definition 2.3 should be used.

First we need to describe the σ-field $\sigma(\eta)$. For any Borel set $B \subset [\frac{1}{2}, 1]$ we have

$$B = \{\eta \in B\} \in \sigma(\eta)$$

and

$$[0, \tfrac{1}{2}) \cup B = \{\eta \in B\} \cup \{\eta = 2\} \in \sigma(\eta).$$

In fact sets of these two kinds exhaust all elements of $\sigma(\eta)$. The inverse image $\{\eta \in C\}$ of any Borel set $C \subset \mathbb{R}$ is of the first kind if $2 \notin C$ and of the second kind if $2 \in C$.

If $E(\xi|\eta)$ is to be $\sigma(\eta)$-measurable, it must be constant on $[0, \frac{1}{2})$ because η is. If for any $x \in [0, \frac{1}{2})$

$$\begin{aligned} E(\xi|\eta)(x) &= E(\xi|[0, \tfrac{1}{2})) \\ &= \frac{1}{P([0, \frac{1}{2}))} \int_{[0, \frac{1}{2})} \xi(x)\, dx \\ &= \frac{1}{\frac{1}{2}} \int_0^{\frac{1}{2}} 2x^2\, dx \\ &= \frac{1}{6}, \end{aligned}$$

then

$$\int_{[0, \frac{1}{2})} E(\xi|\eta)(x)\, dx = \int_{[0, \frac{1}{2})} \xi(x)\, dx,$$

i.e. condition 2) of Definition 2.3 will be satisfied for $A = [0, \frac{1}{2})$.

Moreover, ~~if~~ $E(\xi|\eta) = \xi$ on $[\frac{1}{2}, 1]$, then of course

on $[\frac{1}{2},1]$, $\xi(x) = \xi(\eta) = 2x^2$

then ξ is $\sigma(\eta)$ measurable

$\Rightarrow E[\xi|\eta] = \xi$

$$\int_B E(\xi|\eta)(x)\, dx = \int_B \xi(x)\, dx$$

for any Borel set $B \subset [\frac{1}{2}, 1]$.

Therefore, we have found that

$$E(\xi|\eta)(x) = \begin{cases} \frac{1}{6} & \text{if } x \in [0, \frac{1}{2}), \\ 2x^2 & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

Because every element of $\sigma(\eta)$ is of the form B or $[0, \frac{1}{2}) \cup B$, where $B \subset [\frac{1}{2}, 1]$ is a Borel set, it follows immediately that conditions 1) and 2) of Definition 2.3 are satisfied. The graph of $E(\xi|\eta)$ is presented in Figure 2.2 along with those of ξ and η.

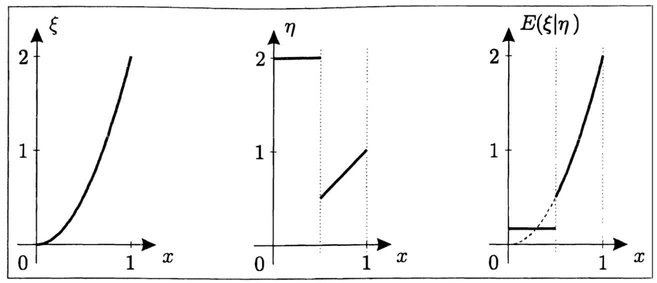


Figure 2.2. The graph of $E(\xi|\eta)$ in Example 2.4

Exercise 2.6

Let $\Omega = [0,1]$ with Lebesgue measure as in Example 2.4. Find the conditional expectation $E(\xi|\eta)$ if

$$\xi(x) = 2x^2, \qquad \eta(x) = 1 - |2x - 1|.$$

Hint First describe the σ-field generated by η. Observe that η is symmetric about $\frac{1}{2}$. What does it tell you about the sets in $\sigma(\eta)$? What does it tell you about $E(\xi|\eta)$ if it is to be $\sigma(\eta)$-measurable? Does it need to be symmetric as well? For any A in $\sigma(\eta)$ try to transform $\int_A \xi\, dP$ to make the integrand symmetric.

Solution 2.6

First we need to describe the σ-field $\sigma(\eta)$ generated by η. Observe that η is symmetric about $\frac{1}{2}$,

$$\eta(x) = \eta(1-x)$$

for any $x \in [0,1]$. We claim that $\sigma(\eta)$ consists of all Borel sets $A \subset [0,1]$ symmetric about $\frac{1}{2}$, i.e. such that

$$A = 1 - A. \tag{2.5}$$

Indeed, if A is such a set, then $A = \{\eta \in B\}$, where

$$B = \{2x : x \in A \cap [0, \tfrac{1}{2}]\}$$

is a Borel set, so $A \in \sigma(\eta)$. On the other hand, if $A \in \sigma(\eta)$, then there is a Borel set B in $\mathbb{R}$ such that $A = \{\eta \in B\}$. Then

$$\begin{aligned} x \in A &\Leftrightarrow \eta(x) \in B \\ &\Leftrightarrow \eta(1-x) \in B \\ &\Leftrightarrow 1 - x \in A, \end{aligned}$$

so A satisfies (2.5).

We are ready to find $E(\xi|\eta)$. If it is to be $\sigma(\eta)$-measurable, it must be symmetric about $\frac{1}{2}$, i.e.

$$E(\xi|\eta)(x) = E(\xi|\eta)(1-x)$$

for any $x \in [0,1]$. For any $A \in \sigma(\eta)$ we shall transform the integral below so as to make the integrand symmetric about $\frac{1}{2}$:

$$\begin{aligned} \int_A 2x^2\, dx &= \int_A x^2\, dx + \int_A x^2\, dx \\ &= \int_A x^2\, dx + \int_{1-A} (1-x)^2\, dx \\ &= \int_A x^2\, dx + \int_A (1-x)^2\, dx \\ &= \int_A \left(x^2 + (1-x)^2\right) dx. \end{aligned}$$

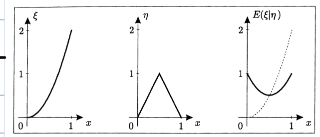


Figure 2.3. The graph of $E(\xi|\eta)$ in Exercise 2.6

It follows that

$$E(\xi|\eta)(x) = x^2 + (1-x)^2.$$

The graphs of ξ, η and $E(\xi|\eta)$ are shown in Figure 2.3.

Soln: $\eta(x) = 1-|2x-1| = \begin{cases} -2x+2, & x\in[\frac{1}{2},+\infty) \\ 2x, & x\in(-\infty,\frac{1}{2}) \end{cases}$

Observed that η is symmetric about $\frac{1}{2}$,

$\Rightarrow \eta(x) = \eta(1-x), \ \forall x\in[0,1]$.

Claim that $\sigma(\eta)$ consists of all Borel sets $A\subset[0,1]$ symmetric about $\frac{1}{2}$ s.t. $A = 1-A$

Proof: If A is indeed such a set then $A = \{\eta\in B\}$ where $B = \{2x : x\in A\cap[0,\frac{1}{2}]\}$ is a Borel set.

$\Rightarrow A = \{\eta\in B\}\in\sigma(\eta)$

As $A\in\sigma(\eta)$ there there's a Borel set B in $\mathbb{R}$ s.t.

$$A = \{\eta\in B\} \Rightarrow x\in A \Leftrightarrow \eta(x)\in B$$
$$\Leftrightarrow \eta(1-x)\in B$$
$$\Leftrightarrow 1-x\in A$$
$$\Rightarrow A = 1-A$$

Q.E.D.

As η is symmetric about $\frac{1}{2}$ then if $E[\xi|\eta]$ is about to be $\sigma(\eta)$-measurable then it must be symmetric about $\frac{1}{2}$.

i.e. $\boxed{E[\xi|\eta](x) = E[\xi|\eta](1-x) \quad \forall x\in[0,1].}$

For any $A\in\sigma(\eta)$,

$$\int_A \xi(x)dx = \int_A 2x^2dx = \int_A x^2dx + \int_A x^2dx$$
$$= \int_A x^2dx + \boxed{\int_{1-A}(1-x)^2dx}$$
$$= \int_A x^2dx + \int_A (1-x)^2dx$$
$$= \int_A (x^2+(1-x)^2)dx$$

and by Def. $\int_A \xi(x)dx = \int_A E(\xi|\eta)dx$

Thus, $E[\xi|\eta] = x^2+(1-x)^2 = \dfrac{\xi(x)+\xi(1-x)}{2}$ (Ex. 2.15) □.

$2x^2 \neq x^2+(1-x)^2$

$(\xi \neq E[\xi|\eta])$

即使积分相等.

Exercise 2.7

Let Ω be the unit square $[0,1]\times[0,1]$ with the σ-field of Borel sets and P the Lebesgue measure on $[0,1]\times[0,1]$. Suppose that ξ and η are random variables on Ω with joint density

$$f_{\xi,\eta}(x,y) = x+y$$

for any $x,y\in[0,1]$, and $f_{\xi,\eta}(x,y) = 0$ otherwise. Show that

$$E(\xi|\eta) = \frac{2+3\eta}{3+6\eta}. = F(\eta) \text{ a Borel Function}$$

Hint It suffices (why?) to show that for any Borel set B

$$\int_{\{\eta\in B\}}\xi\,dP = \int_{\{\eta\in B\}}\frac{2+3\eta}{3+6\eta}\,dP.$$

Try to express each side of this equality as an integral over the square $[0,1]\times[0,1]$ using the joint density $f_{\xi,\eta}(x,y)$.

Solution 2.7

Since

$$\{\eta\in B\} = [0,1]\times B$$

for any Borel set B, we have

$$\int_{\{\eta\in B\}}\xi\,dP = \int_B\int_{\mathbb{R}} x f_{\xi,\eta}(x,y)\,dx\,dy$$
$$= \int_B\left(\int_{[0,1]} x(x+y)\,dx\right)dy$$
$$= \int_B\left(\frac{1}{3}+\frac{1}{2}y\right)dy$$

and

$$\int_{\{\eta\in B\}}\frac{2+3\eta}{3+6\eta}\,dP = \int_B\int_{\mathbb{R}}\frac{2+3y}{3+6y}f_{\xi,\eta}(x,y)\,dx\,dy$$
$$= \int_B\frac{2+3y}{3+6y}\left(\int_{[0,1]}(x+y)\,dx\right)dy$$
$$= \int_B\left(\frac{1}{3}+\frac{1}{2}y\right)dy.$$

Because each event in $\sigma(\eta)$ is of the form $\{\eta\in B\}$ for some Borel set B, this means that condition 2) of Definition 2.3 is satisfied. The random variable $\frac{2+3\eta}{3+6\eta}$ is $\sigma(\eta)$-measurable, so condition 1) holds too. It follows that

$$E(\xi|\eta) = \frac{2+3\eta}{3+6\eta}.$$

Proof: Since $\{\eta\in B\} = [0,1]\times B$ for any Borel set B (on $\mathbb{R}$) and $\int_{\{\eta\in B\}}\frac{2+3\eta}{3+6\eta}dP = \int_B\int_{\mathbb{R}}\frac{2+3y}{3+6y}f_{\xi,\eta}(x,y)\,dxdy$

$$= \int_B\frac{2+3y}{3+6y}\left[\int_0^1(x+y)dx\right]dy$$
$$= \int_B\frac{2+3y}{3+6y}\left[\frac{1}{2}x^2+xy\right]_0^1dy$$
$$= \int_B\frac{2+3y}{3+6y}\times\frac{1+2y}{2}dy$$
$$= \int_B\frac{2+3y}{6}dy$$
$$= \int_B\left(\frac{1}{3}+\frac{1}{2}y\right)dy$$

then $\int_{\{\eta\in B\}}\xi\,dp = \int_B\int_{\mathbb{R}} x f_{\xi,\eta}(x,y)\,dxdy$

$$= \int_B\left(\int_{[0,1]}x(x+y)dx\right)dy$$
$$= \int_B\int_0^1 x^2+xy\,dx\,dy$$
$$= \int_B\left[\frac{1}{3}x^3+\frac{1}{2}x^2y\right]_0^1dy$$
$$= \int_B\left[\frac{1}{3}+\frac{1}{2}y\right]dy$$

$\int_{\{\eta\in B\}}dp = \int_B\int_{\mathbb{R}}xf(x,y)dx\,dy$

Thus, $\int_{\{\eta\in B\}}\xi\,dp = \int_{\{\eta\in B\}}\frac{2+3\eta}{3+6\eta}dp$

as $\{\eta\in B\}\in\sigma(\eta)$ then by Def. 2.3, $\frac{2+3\eta}{3+6\eta} = E(\xi|\eta)$.

Q.E.D.

$\boxed{\{\eta\in B\} = \{(\xi,\eta)\in[0,1]\times B\}.}$

Exercise 2.8

Let Ω be the unit square $[0,1] \times [0,1]$ with Lebesgue measure as in Exercise 2.7. Find $E(\xi|\eta)$ if ξ and η are random variables on Ω with joint density

$$f_{\xi,\eta}(x,y) = \frac{3}{2}\left(x^2+y^2\right)$$

for any $x, y \in [0,1]$, and $f_{\xi,\eta}(x,y) = 0$ otherwise.

Hint This is slightly harder than Exercise 2.7 because here we have to derive a formula for the conditional expectation. Study the solution to Exercise 2.7 to find a way of obtaining such a formula.

Solution 2.8

We are looking for a Borel function $F : \mathbb{R} \to \mathbb{R}$ such that for any Borel set B

$$\int_{\{\eta\in B\}} \xi \, dP = \int_{\{\eta\in B\}} F(\eta)\, dP. \tag{2.6}$$

Then $E(\xi|\eta) = F(\eta)$ by Definition 2.3.

We shall transform both integrals above using the joint density $f_{\xi,\eta}(x,y)$ in much the same way as in the solution to Exercise 2.7, except that here we do not know the exact form of $F(y)$. Namely,

$$\begin{aligned}\int_{\{\eta\in B\}} \xi\, dP &= \int_B \int_{\mathbb{R}} x\, f_{\xi,\eta}(x,y)\, dx\, dy \\ &= \frac{3}{2}\int_B \left(\int_{[0,1]} x\left(x^2+y^2\right) dx\right) dy \\ &= \frac{3}{2}\int_B \left(\frac{1}{4}+\frac{1}{2}y^2\right) dy\end{aligned}$$

and

$$\begin{aligned}\int_{\{\eta\in B\}} F(\eta)\, dP &= \int_B \int_{\mathbb{R}} F(y)\, f_{\xi,\eta}(x,y)\, dx\, dy \\ &= \frac{3}{2}\int_B F(y) \left(\int_{[0,1]} \left(x^2+y^2\right) dx\right) dy \\ &= \frac{3}{2}\int_B F(y)\left(\frac{1}{3}+y^2\right) dy.\end{aligned}$$

Then, (2.6) will hold for any Borel set B if

$$F(y) = \frac{\frac{1}{4}+\frac{1}{2}y^2}{\frac{1}{3}+y^2} = \frac{3+6y^2}{4+12y^2}.$$

It follows that

$$E(\xi|\eta) = F(\eta) = \frac{3+6\eta^2}{4+12\eta^2}.$$

Soln: To find $E[\xi|\eta]$, by Def. 2.3 we have to find a Borel function $F:\cancel{[0,1]}\,\mathbb{R}\to\mathbb{R}$ s.t. for any Borel set B ~~on [0,1]~~. (1+2)

$$\int_{\{\eta\in B\}} \xi\, dP = \int_{\{\eta\in B\}} F(\eta)\, dP$$

So $E[\xi|\eta] = F(\eta)$

As
$$\begin{aligned}\int_{\{\eta\in B\}} \xi\, dP &= \int_B\int_{\cancel{[0,1]}\,\mathbb{R}} x f_{\xi\eta}(x,y)\,dx\,dy \\ &= \int_B\int_0^1 x\cdot\frac{3}{2}(x^2+y^2)\,dx\,dy \\ &= \frac{3}{2}\int_B\left[\frac{1}{4}x^4+\frac{1}{2}x^2y^2\right]_0^1 dy \\ &= \frac{3}{2}\int_B\left(\frac{1}{4}+\frac{1}{2}y^2\right)dy\end{aligned}$$

and
$$\begin{aligned}\boxed{\int_{\{\eta\in B\}} F(\eta)\, dP = \int_B\int_{\cancel{[0,1]}\,\mathbb{R}} F(y) f_{\xi\eta}(x,y)\,dx\,dy} \\ &= \frac{3}{2}\int_B\int_0^1 F(y)\cdot(x^2+y^2)\,dx\,dy \\ &= \frac{3}{2}\int_B F(y)\left[\frac{1}{3}x^3+xy^2\right]_0^1 dy \\ &= \frac{3}{2}\int_B F(y)\left(\frac{1}{3}+y^2\right)dy\end{aligned}$$

If F is indeed a such function then

$$\frac{1}{4}+\frac{1}{2}y^2 = F(y)\left(\frac{1}{3}+y^2\right)$$

$$F(y) = \frac{\frac{1+2y^2}{4}}{\frac{1+3y^2}{3}} = \frac{3+6y^2}{4+12y^2}$$

$$\Rightarrow \quad E[\xi|\eta] = F(\eta) = \frac{3+6\eta^2}{4+12\eta^2}$$

□.

Exercise 2.9

Let Ω be the unit disc $\{(x,y) : x^2+y^2 \le 1\}$ with the σ-field of Borel sets and P the Lebesgue measure on the disc normalized so that $P(\Omega) = 1$, i.e.

$$P(A) = \frac{1}{\pi}\iint_A dx\, dy$$

for any Borel set $A \subset \Omega$. Suppose that ξ and η are the projections onto the x and y axes,

$$\xi(x,y) = x, \qquad \eta(x,y) = y$$

for any $(x,y) \in \Omega$. Find $E(\xi^2|\eta)$.

Hint What is the joint density of ξ and η? Use this density to transform the integral

$$\int_{\{\eta\in B\}} \xi^2\, dP$$

for an arbitrary Borel set B so that the integrand becomes a function of η. How is this function of η related to $E(\xi^2|\eta)$?

Solution 2.9

We are looking for a Borel function $F : \mathbb{R} \to \mathbb{R}$ such that for any Borel set $B \subset \mathbb{R}$

$$\int_{\{\eta\in B\}} \xi^2\, dP = \int_{\{\eta\in B\}} F(\eta)\, dP. \tag{2.7}$$

Then, by Definition 2.3 we shall have $E(\xi^2|\eta) = F(\eta)$.

Let us transform both sides of (2.7). To do so we observe that the random variables ξ and η have uniform joint distribution over the unit disc $\Omega = \{(x,y) : x^2+y^2 \le 1\}$, with density

$$f_{\xi,\eta}(x,y) = \frac{1}{\pi}$$

if $x^2+y^2 \le 1$, and $f_{\xi,\eta}(x,y) = 0$ otherwise. It follows that

$$\begin{aligned}\int_{\{\eta\in B\}} \xi^2\, dP &= \int_B\int_{\mathbb{R}} x^2 f_{\xi,\eta}(x,y)\, dx\, dy \\ &= \frac{1}{\pi}\int_B\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x^2\, dx\, dy \\ &= \frac{2}{3\pi}\int_B \left(1-y^2\right)^{3/2} dy\end{aligned}$$

and

$$\begin{aligned}\int_{\{\eta\in B\}} F(\eta)\, dP &= \int_B\int_{\mathbb{R}} F(y)\, f_{\xi,\eta}(x,y)\, dx\, dy \\ &= \frac{1}{\pi}\int_B F(y)\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} dx\, dy \\ &= \frac{2}{\pi}\int_B F(y)\left(1-y^2\right)^{1/2} dy.\end{aligned}$$

If (2.7) is to be satisfied for all Borel sets B, then

$$F(y) = \frac{1}{3}\left(1-y^2\right).$$

This means that

$$E(\xi^2|\eta)(x,y) = F(\eta(x,y)) = F(y) = \frac{1}{3}\left(1-y^2\right)$$

for any (x,y) in Ω.

Proof: If there's a Borel function $F:\mathbb{R}\to\mathbb{R}$ s.t. for any Borel set $B\subset\mathbb{R}$, $\int_{\{\eta\in B\}}\xi^2\,dP = \int_{\{\eta\in B\}}F(\eta)\,dP$

then by Def. 2.3 we can define $E[\xi^2|\eta] = F(\eta)$.

As $P(A) = \iint_A f(x,y)\,dx\,dy = \iint_A \frac{1}{\pi}\,dx\,dy$, so $f(x,y) = \begin{cases}\frac{1}{\pi}, & \text{if } x^2+y^2\le 1\\ 0, & \text{otherwise.}\end{cases}$

and ξ,η have uniform joint distribution over Ω.

As
$$\begin{aligned}\int_{\{\eta\in B\}}\xi^2\,dP &= \int_B\int_{\mathbb{R}} x^2 f_{\xi\eta}(x,y)\,dx\,dy \\ &= \int_B\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} x^2\cdot\frac{1}{\pi}\,dx\,dy \\ &= \frac{1}{\pi}\int_B \frac{1}{3}x^3\Big|_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}\,dy \\ &= \frac{1}{3\pi}\int_B 2\cdot(1-y^2)^{\frac{3}{2}}\,dy \\ &= \frac{2}{3\pi}\int_B (1-y^2)^{\frac{3}{2}}\,dy\end{aligned}$$

and
$$\begin{aligned}\boxed{\int_{\{\eta\in B\}}F(y)\,dP = \int_B\int_{\mathbb{R}}F(y)\cdot\frac{1}{\pi}\,dx\,dy} \\ &= \int_B\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} F(y)\cdot\frac{1}{\pi}\,dx\,dy \\ &= \frac{1}{\pi}\int_B F(y)\cdot 2(1-y^2)^{\frac{1}{2}}\,dy \\ &= \frac{2}{\pi}\int_B F(y)(1-y^2)^{\frac{1}{2}}\,dy\end{aligned}$$

$$\Rightarrow \quad \frac{2}{\pi}F(y)(1-y^2)^{\frac{1}{2}} = \frac{2}{3\pi}(1-y^2)^{\frac{3}{2}}$$

$$F(y) = \frac{1}{3}(1-y^2)$$

$\Rightarrow E[\xi^2|\eta] = F(\eta) = F(\eta(x,y)) = F(y) = \frac{1}{3}(1-y^2)$ for any $x,y\in\Omega$.

□.

2.4 Conditioning on a σ-Field

Proposition 2.2

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let η and η' be two random variables such that $\sigma(\eta) = \sigma(\eta')$. Then

$$\mathbb{E}[\xi \mid \eta] = \mathbb{E}[\xi \mid \eta'] \quad \text{a.s.}$$

holds for any integrable random variable ξ.

What truly matters is the $\sigma(\eta)$ generated by η, but not the actual value of η.

Proof

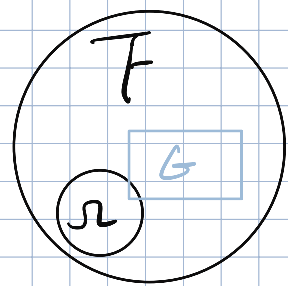
This is an immediate consequence of Lemma 2.1. □

Definition 2.4

Let ξ be an integrable random variable on a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$. Then the *conditional expectation* of ξ given $\mathcal{G}$ is defined to be a random variable $E(\xi|\mathcal{G})$ such that

1) $E(\xi|\mathcal{G})$ is $\mathcal{G}$-measurable;
2) For any $A \in \mathcal{G}$

$$\int_A E(\xi|\mathcal{G})\, dP = \int_A \xi\, dP. \tag{2.3}$$



Remark 2.2

The *conditional probability* of an event $A \in \mathcal{F}$ given a σ-field $\mathcal{G}$ can be defined by

$$P(A|\mathcal{G}) = E(1_A|\mathcal{G}),$$

where 1_A is the indicator function of A.

The notion of conditional expectation with respect to a σ-field extends conditioning on a random variable η in the sense that

$$E(\xi|\sigma(\eta)) = E(\xi|\eta),$$

where $\sigma(\eta)$ is the σ-field generated by η.

Proposition 2.3

$E(\xi|\mathcal{G})$ exists and is unique in the sense that if $\xi = \xi'$ a.s., then $E(\xi|\mathcal{G}) = E(\xi'|\mathcal{G})$ a.s.

Proof

Existence and uniqueness follow, respectively, from Theorem 2.1 below and Lemma 2.1. □

Existence:

Theorem 2.1 (Radon–Nikodym)

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$. Then for any random variable ξ there exists a $\mathcal{G}$-measurable random variable ζ such that

$$\int_A \xi \, dP = \int_A \zeta \, dP$$

for each $A \in \mathcal{G}$.

The Radon–Nikodym theorem is important from a theoretical point of view. However, in practice there are usually other ways of establishing the existence of conditional expectation, for example, by finding an explicit formula, as in the examples and exercises in the previous section. The proof of the Radon–Nikodym theorem is beyond the scope of this course and is omitted.

Exercise 2.10

Show that if $\mathcal{G} = \{\emptyset, \Omega\}$, then $E(\xi|\mathcal{G}) = E(\xi)$ a.s.

Hint What random variables are $\mathcal{G}$-measurable if $\mathcal{G} = \{\emptyset, \Omega\}$?

2.10 Show that if $\mathcal{G} = \{\emptyset, \Omega\}$, then $\mathbb{E}[\xi \mid \mathcal{G}] = \mathbb{E}[\xi]$.

Proof. Firstly, we claim that a random variable ζ is $\mathcal{G}$-measurable if and only if ζ is a constant random variable, that is, there exists a constant c such that $\zeta(\omega) = c$ for all $\omega \in \Omega$.

In fact, if there exists a constant c such that $\zeta(\omega) = c$ for all $\omega \in \Omega$, then for any $B \in \mathcal{B}(\mathbb{R})$, we have that $\{\zeta \in B\} = \Omega$ whenever $c \in B$, and $\{\zeta \in B\} = \emptyset$ whenever $c \notin B$, and so that ζ is $\mathcal{G}$-measurable; if ζ has at least two values $c_1 \neq c_2$, that is, there exist $\omega_1, \omega_2 \in \Omega$ such that $\zeta(\omega_1) = c_1$ and $\zeta(\omega_2) = c_2$. Then, letting $B \in \mathcal{B}(\mathbb{R})$ such that $c_1 \in B$ and $c_2 \notin B$, we have that $\{\zeta \in B\} \neq \emptyset$ since $\omega_1 \in \{\zeta \in B\}$, and $\{\zeta \in B\} \neq \Omega$ since $\omega_2 \notin \{\zeta \in B\}$. This implies that $\{\zeta \in B\} \notin \mathcal{G}$, and so that ζ is not $\mathcal{G}$-measurable.

Now, according to Definition 2.4, we can assume that $\mathbb{E}[\xi \mid \mathcal{G}](\omega) = c$ for all $\omega \in \Omega$, where c is a constant. Again according to Definition 2.4, we have

$$c = c\mathbb{P}(\Omega) = \int_\Omega c\mathbb{P}(d\omega) = \boxed{\int_\Omega \mathbb{E}[\xi \mid \mathcal{G}](\omega)\mathbb{P}(d\omega) = \int_\Omega \xi(\omega)\mathbb{P}(d\omega)} = \mathbb{E}[\xi].$$

Def. 2.4

Thus, we conclude that $\mathbb{E}[\xi \mid \mathcal{G}] = \mathbb{E}[\xi]$. □

Solution 2.10

If $\mathcal{G} = \{\emptyset, \Omega\}$, then any constant random variable is $\mathcal{G}$-measurable. Since

"和"

$$\int_\Omega \xi \, dP = E(\xi) = \int_\Omega E(\xi) \, dP$$

不是 $G=\phi$ or $G=\Omega$

and

$$\int_\emptyset \xi \, dP = 0 = \int_\emptyset E(\xi) \, dP,$$

it follows that $E(\xi|\mathcal{G}) = E(\xi)$ a.s., as required.

Exercise 2.11

Show that if ξ is $\mathcal{G}$-measurable, then $E(\xi|\mathcal{G}) = \xi$ a.s.

Hint The conditions of Definition 2.4 are trivially satisfied by ξ if ξ is $\mathcal{G}$-measurable.

Solution 2.11

Because the trivial identity

$$\int_A \xi \, dP = \int_A \xi \, dP$$

holds for any $A \in \mathcal{G}$ and ξ is $\mathcal{G}$-measurable, it follows that $E(\xi|\mathcal{G}) = \xi$ a.s.

Exercise 2.12

Show that if $B \in \mathcal{G}$, then

$$E(E(\xi|\mathcal{G})|B) = E(\xi|B).$$

Hint The conditional expectation on either side of the equality involves an integral over B. How are these integrals related to one another?

Solution 2.12

By Definition 2.3

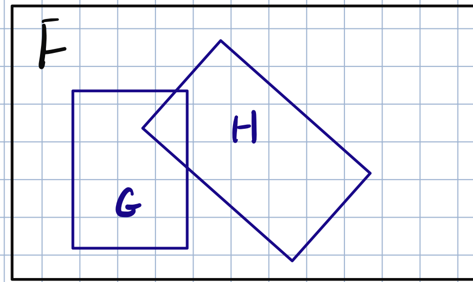
$$\int_B E(\xi|\mathcal{G}) \, dP = \int_B \xi \, dP,$$

since $B \in \mathcal{G}$. It follows that

$$\begin{aligned} E(E(\xi|\mathcal{G})|B) &= \frac{1}{P(B)} \int_B E(\xi|\mathcal{G}) \, dP \\ &= \frac{1}{P(B)} \int_B \xi \, dP \\ &= E(\xi|B). \end{aligned}$$

2.5 General Properties

Here a, b are arbitrary real numbers, ξ, ζ are integrable random variables on a probability space $(\Omega, \mathcal{F}, P)$ and $\mathcal{G}, \mathcal{H}$ are σ-fields on Ω contained in $\mathcal{F}$. In 3) we also assume that the product $\xi\zeta$ is integrable. All equalities and the inequalities in 6) hold P-a.s.



$\Omega \overset{\mathcal{G}}{\exists} \in G/H/\cdots \in F$

Proposition 2.4

Conditional expectation has the following properties:

1) $E(a\xi + b\zeta|\mathcal{G}) = aE(\xi|\mathcal{G}) + bE(\zeta|\mathcal{G})$ (*linearity*);

Proof

1) For any $B \in \mathcal{G}$

$$\begin{aligned}\int_B (aE(\xi|\mathcal{G}) + bE(\zeta|\mathcal{G}))\, dP &= a\int_B E(\xi|\mathcal{G})\, dP + b\int_B E(\zeta|\mathcal{G})\, dP \\ &= a\int_B \xi\, dP + b\int_B \zeta\, dP \\ &= \int_B (a\xi + b\zeta)\, dP.\end{aligned}$$

By uniqueness this proves the desired equality.

by Def. 2.4 $\int_B E[a\xi+b\zeta|G]dP = \int_B a\xi+b\zeta\, dP$ $\Rightarrow E[a\xi+b\zeta|G] = aE[\xi|G] + bE[\zeta|G]$.

$\int_A (r.v.)_1\, dP = \int_A (r.v.)_2\, dP \;\cancel{\Rightarrow}\; (r.v.)_1 = (r.v.)_2$.

2) $E(E(\xi|\mathcal{G})) = E(\xi)$;

2) This follows by putting $A = \Omega$ in the condition 2) of Definition 2.4. Also, the Property 2) is a special case of Property 5) when $\mathcal{H} = \{\emptyset, \Omega\}$.

$A \in G$, $G \in F$, $\Omega \in F$

$\Rightarrow$ A could be Ω

Proof: As $E[\xi|G]$ is a random variable s.t. $E[\xi|G]$ is G-measurable and for any $A \in G$, $A = \Omega$

$$\int_A E[\xi|G]\,dP = \int_A \xi\, dP.$$

Thus, for any $A \in G$, $A = \Omega$

$$E[E[\xi|G]] = \int_A E[\xi|G]\,dP = \int_A \xi\, dP = E[\xi]$$

$\square$.

3) $E(\xi\zeta|\mathcal{G}) = \xi E(\zeta|\mathcal{G})$ if ξ is $\mathcal{G}$-measurable (*taking out what is known*);

3) We first verify the result for $\xi = 1_A$, where $A \in \mathcal{G}$. In this case, for any $B \in \mathcal{G}$, we have

$$\begin{aligned}\int_B 1_A(\omega)\mathbb{E}[\zeta \mid \mathcal{G}](\omega)\mathbb{P}(d\omega) &= \int_{A\cap B} \mathbb{E}[\zeta \mid \mathcal{G}](\omega)\mathbb{P}(d\omega) \\ &= \int_{A\cap B} \zeta(\omega)\mathbb{P}(d\omega) \quad \text{as } (A\cap B) \in G \\ &= \int_B 1_A(\omega)\zeta(\omega)\mathbb{P}(d\omega).\end{aligned}$$

$\int_A \xi\left(\int_A \zeta\, dP\right) dP \neq \int_A \xi\zeta\, dP$.

e.g. $\sum_i \xi_i (\sum_i \zeta_i) = \sum_i\sum_j \xi_i\zeta_j \neq \sum_i \xi_i\zeta_i$

Proof: Firstly, verify the result for $\xi = 1_A$, where $A \in G$.

then for any $B \in G$, $\int_B \xi E[\zeta|G](\omega) P(d\omega)$

$= \int_B 1_A(\omega) E[\zeta|G](\omega) P(d\omega)$

necessity of 1_A $= \int_{A\cap B} E[\zeta|G](\omega) P(d\omega)$

by Def. 2.4 $= \int_{A\cap B} \zeta(\omega) P(d\omega)$ as $A\cap B \in G$

necessity of 1_A $= \int_B 1_A \zeta\, dP$

$= \int_B \xi\zeta\, dP$

By uniqueness it follows that

$$1_A \mathbb{E}[\zeta \mid \mathcal{G}] = \mathbb{E}[1_A \zeta \mid \mathcal{G}].$$

Similarly, we obtain the result for a $\mathcal{G}$-measurable step function,

$$\xi = \sum_{j=1}^{m} a_j 1_{A_j},$$

where $A_j \in \mathcal{G}$ for $j = 1, \ldots, m$. Finally, the result in the general case follows by approximating ξ by $\mathcal{G}$-measurable step function.

By uniqueness, $1_A E[Z|G] = E[1_A Z|P]$

Similarly, obtain the result for a G-measurable step function

$\xi = \sum_{j=1}^{m} a_j 1_{A_j}$ where $A_i \in G$, $j = 1, 2, \ldots, m$

Finally, obtain the result for the general case by approximating ξ by a G-measurable step function.

Q.E.D.

4) $E(\xi|\mathcal{G}) = E(\xi)$ if ξ is independent of $\mathcal{G}$ (*an independent condition drops out*); 在random variable解决

Proof: If ξ is independent of G, then $\boxed{\text{for any } B \in G,}$

then the r.v. ξ and 1_B are independent for any $B \in G$. → r.v.

$$\int_B E[\xi]\, dP = E[\xi] E[1_B]$$
$$= E[\xi 1_B] \quad \text{By Proposition 1.1}$$
$$= \int_B \xi\, dP$$

Thus, $E[\xi | G] = E[\xi]$.

Q.E.D.

4) Since ξ is independent of $\mathcal{G}$, the random variables ξ and 1_B are independent for any $B \in \mathcal{G}$. It follows by Proposition 1.1 (independent random variables are uncorrelated) that

$= E(\xi)\int_B dP = E(\xi) P(B)$

$$\int_B E(\xi)\, dP = E(\xi) E(1_B) \quad \text{as } 1_B \text{ is an indicator func. of } B$$
$$= E(\xi 1_B) = \int_G \xi 1_B\, dP$$
$$= \int_B \xi\, dP,$$

which proves the assertion. Fro $\int_B E[\xi|P]\, dP = \int_B \xi\, dP$ for any $B \in G$

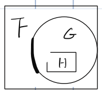
Proposition 1.1

If two integrable random variables $\xi, \eta : \Omega \to \mathbb{R}$ are independent, then they are *uncorrelated*, i.e.

$$E(\xi\eta) = E(\xi)E(\eta),$$

provided that the product $\xi\eta$ is also integrable. If $\xi_1, \ldots, \xi_n : \Omega \to \mathbb{R}$ are

5) $E(E(\xi|\mathcal{G})|\mathcal{H}) = E(\xi|\mathcal{H})$ if $\mathcal{H} \subset \mathcal{G}$ (*tower property*);



5) By Definition 2.4, we have

$$\int_B \mathbb{E}[\xi \mid \mathcal{G}](\omega) \mathbb{P}(d\omega) = \int_B \xi(\omega) \mathbb{P}(d\omega), \quad \forall B \in \mathcal{G},$$
$$\int_B \mathbb{E}[\xi \mid \mathcal{H}](\omega) \mathbb{P}(d\omega) = \int_B \xi(\omega) \mathbb{P}(d\omega), \quad \forall B \in \mathcal{H}.$$

Because $\mathcal{H} \subset \mathcal{G}$ it follows that

$$\int_B \mathbb{E}[\xi \mid \mathcal{G}](\omega) \mathbb{P}(d\omega) = \int_B \mathbb{E}[\xi \mid \mathcal{H}](\omega) \mathbb{P}(d\omega)$$

for every $B \in \mathcal{H}$. Applying Definition 2.4 once again, we obtain

$$\mathbb{E}\Big[\mathbb{E}[\xi \mid \mathcal{G}] \mid \mathcal{H}\Big] = \mathbb{E}[\xi \mid \mathcal{H}].$$

Proof: As $E[E(\xi|G)|H]$ and $E(\xi|H)$ are H-measurable

so $\forall B \in H$, $\int_B E[E(\xi|G)|H]\,dP = \int_B E(\xi|G)\,dP$

and $\int_B E[\xi|H]\,dP = \int_B \xi\,dP$

as $H \subset G$ so $B \in G$, $\Rightarrow \int_B E(\xi|G)\,dP = \int_B \xi\,dP$

Thus, $E[\xi|B] = E[E(\xi|G)|H]$ $\square$.

6) If $\xi \geq 0$, then $E(\xi|\mathcal{G}) \geq 0$ (*positivity*).

6) For any n we put

$$A_n = \left\{ E(\xi|\mathcal{G}) \leq -\frac{1}{n} \right\}.$$

\# Then $A_n \in \mathcal{G}$. If $\xi \geq 0$ a.s., then

$$0 \leq \int_{A_n} \xi\,dP = \int_{A_n} E(\xi|\mathcal{G})\,dP \leq -\frac{1}{n}P(A_n), \to 0$$

which means that $P(A_n) = 0$. Because

$$\{E(\xi|\mathcal{G}) < 0\} = \bigcup_{n=1}^{\infty} A_n$$

it follows that

$$P\{E(\xi|\mathcal{G}) < 0\} = 0, = \lim_{n\to\infty} P(A_n)$$

completing the proof. $\square$

Proof: Let $A_n = \{E[\xi|G] \leq -\frac{1}{n}\}$, $\forall n \in N$ then as $E[\xi|G]$ is G-measurable so is A_n $\boxed{\text{i.e. } A_n \in G.}$

$\boxed{\text{If } \xi \geq 0 \text{ then } 0 \leq \int_{A_n} \xi\,dp = \int_{A_n} E[\xi|G]\,dp} \leq -\frac{1}{n}P(A_n)$

→ 不能直接由 $\int_{A_n} E[\xi|G]dp > 0$ 得 $E[\xi|G] > 0$

因为 $A_n \in G$；但可由 $\xi \geq 0$ 得 $\int_{A_n} \xi dp \geq 0$，由积分性质

$\Rightarrow P(A_n) = 0$ （忽略 $P(A_n) < 0$）. $\forall n \in N$.

As A_n is an expanding sequence, so

$$0 = \bigcup_{n=1}^{\infty} P(A_n) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} P(\{E[\xi|n] \leq -\frac{1}{n}\}) = P(\{E[\xi|G] < 0\})$$

\#

i.e. $P(\{E[\xi|G] < 0\}) = 0$

Thus, $E[\xi|G] \geq 0$ Q.E.D.

\# 注意取交还是并.

Proposition 2.4

Assume ξ and ζ are integrable random variables. The conditional expectation possesses the following properties:

1) If a and b are two constants, then

$$\mathbb{E}[a\xi + b\zeta \mid \mathcal{G}] = a\mathbb{E}[\xi \mid \mathcal{G}] + b\mathbb{E}[\zeta \mid \mathcal{G}];$$

2) $\mathbb{E}[\mathbb{E}[\xi \mid \mathcal{G}]] = \mathbb{E}[\xi]$; $\mathcal{H} = \{\emptyset, \Omega\}$

3) If ξ is $\mathcal{G}$-measurable, then $\mathbb{E}[\xi \mid \mathcal{G}] = \xi$; and

$$\mathbb{E}[\xi\zeta \mid \mathcal{G}] = \xi\mathbb{E}[\zeta \mid \mathcal{G}];$$

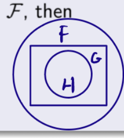
when $\xi\zeta$ is integrable.

Proposition 2.4 (Continuous)

4) If ξ is independent of $\mathcal{G}$, then $\mathbb{E}[\xi \mid \mathcal{G}] = \mathbb{E}[\xi]$;

5) If $\mathcal{H}$ is a sub-σ-field of $\mathcal{F}$ such that $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, then

$$\mathbb{E}\Big[\mathbb{E}[\xi \mid \mathcal{G}] \mid \mathcal{H}\Big] = \mathbb{E}[\xi \mid \mathcal{H}];$$



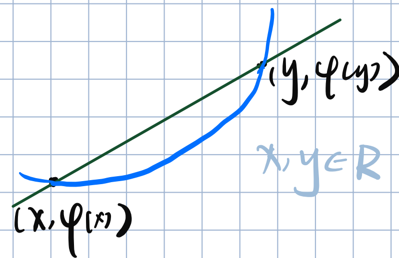
6) If $\xi \geq 0$ a.s., then $\mathbb{E}[\xi \mid \mathcal{G}] \geq 0$ a.s..

5) 叠加 conditional expection 从外往内读，
取 restricted area 的交集

The next theorem, which will be stated without proof, involves the notion of a convex function, such as $\max(1, x)$ or e^{-x}, for example. In this course the theorem will be used mainly for $|x|$, which is also a convex function. In general, we call a function $\varphi : \mathbb{R} \to \mathbb{R}$ *convex* if for any $x, y \in \mathbb{R}$ and any $\lambda \in [0,1]$

$$\varphi(\lambda x + (1-\lambda) y) \leq \lambda\varphi(x) + (1-\lambda)\varphi(y).$$

This condition means that the graph of φ lies below the cord connecting the points with coordinates $(x, \varphi(x))$ and $(y, \varphi(y))$.



Theorem 2.2 (Jensen's Inequality)

Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function and let ξ be an integrable random variable on a probability space $(\Omega, \mathcal{F}, P)$ such that $\varphi(\xi)$ is also integrable. Then

$$\varphi(E(\xi|\mathcal{G})) \leq E(\varphi(\xi)|\mathcal{G}) \quad \text{a.s.}$$

for any σ-field $\mathcal{G}$ on Ω contained in $\mathcal{F}$.

积分的凸函数值 ≤ 凸函数值的积分

一般形式 [編輯]

簡森不等式

延森不等式可以用測度論或機率論的語言給出。這兩種方式都表明同一個很一般的結果。

測度論的版本 [編輯]

probability measure

假設μ是集合Ω的正測度，使得$\mu(\Omega) = 1$。若g是勒貝格可積的實值函數，而φ是在g的值域上定義的凸函數，則

$$\varphi\left(\int_\Omega g\, d\mu\right) \leq \int_\Omega \varphi \circ g\, d\mu$$

機率論的版本 [編輯]

以機率論的名詞，μ是個機率測度。函數g換作實值隨機變數X（就純數學而言，兩者沒有分別）。在Ω空間上，任何函數相對於機率測度μ的積分就成了期望值。這不等式就說，若φ是任一凸函數，則

$$\varphi(E(X)) \leq E(\varphi(X))$$

Jasen gap =

$$E[\varphi(\xi)] - \varphi(E(\xi)) \geq 0$$

Applications and special cases [edit]

Form involving a probability density function [edit]

Suppose Ω is a measurable subset of the real line and $f(x)$ is a non-negative function such that

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

In probabilistic language, f is a probability density function.

Then Jensen's inequality becomes the following statement about convex integrals:

If g is any real-valued measurable function and φ is convex over the range of g, then

$$\varphi\left(\int_{-\infty}^{\infty} g(x)f(x)\,dx\right) \le \int_{-\infty}^{\infty} \varphi(g(x))f(x)\,dx.$$

If $g(x) = x$, then this form of the inequality reduces to a commonly used special case:

$$\varphi\left(\int_{-\infty}^{\infty} x\,f(x)\,dx\right) \le \int_{-\infty}^{\infty} \varphi(x)\,f(x)\,dx.$$

This is applied in Variational Bayesian methods.

Example: even moments of a random variable [edit]

If $g(x) = x^{2n}$, and X is a random variable, then g is convex as

$$\frac{d^2 g}{dx^2}(x) = 2n(2n-1)x^{2n-2} \ge 0 \quad \forall\, x \in \mathbb{R}$$

and so

$$g(\mathrm{E}[X]) = (\mathrm{E}[X])^{2n} \le \mathrm{E}[X^{2n}].$$

In particular, if some even moment $2n$ of X is finite, X has a finite mean. An extension of this argument shows X has finite moments of every order $l \in \mathbb{N}$ dividing n.

Alternative finite form [edit]

Let $\Omega = \{x_1, \ldots x_n\}$, and take μ to be the counting measure on Ω, then the general form reduces to a statement about sums:

$$\varphi\left(\sum_{i=1}^{n} g(x_i)\lambda_i\right) \le \sum_{i=1}^{n} \varphi(g(x_i))\lambda_i,$$

provided that $\lambda_i \ge 0$ and

$$\lambda_1 + \cdots + \lambda_n = 1.$$

There is also an infinite discrete form.

Statistical physics [edit]

Jensen's inequality is of particular importance in statistical physics when the convex function is an exponential, giving:

$$e^{\mathrm{E}[X]} \le \mathrm{E}\left[e^X\right],$$

where the expected values are with respect to some probability distribution in the random variable X.

Proof: Let $\varphi(x) = e^x$ in $\varphi(\mathrm{E}[X]) \le \mathrm{E}[\varphi(X)]$.

2.6 Various Exercises on Conditional Expectation

Exercise 2.13

Mrs. Jones has made a steak and kidney pie for her two sons. Eating more than a half of it will give indigestion to anyone. While she is away having tea with a neighbour, the older son helps himself to a piece of the pie. Then the younger son comes and has a piece of what is left by his brother. We assume that the size of each of the two pieces eaten by Mrs. Jones' sons is random and uniformly distributed over what is currently available. What is the expected size of the remaining piece given that neither son gets indigestion?

Hint All possible outcomes can be represented by pairs of numbers, the portions of the pie consumed by the two sons. Therefore Ω can be chosen as a subset of the plane. Observe that the older son is restricted only by the size of the pie, while the younger one is restricted by what is left by his brother. This will determine the shape of Ω. Next introduce a probability measure on Ω consistent with the conditions of the exercise. This can be done by means of a suitable density over Ω. Now you are in a position to compute the probability that neither son will get indigestion. What is the corresponding subset of Ω? Finally, define a random variable on Ω representing the portion of the pie left by the sons and compute the conditional expectation.

Solution 2.13

The whole pie will be represented by the interval $[0,1]$. Let $x \in [0,1]$ be the portion consumed by the older son. Then $[0, 1-x]$ will be available to the younger one, who takes a portion of size $y \in [0, 1-x]$. The set of all possible outcomes is

$$\Omega = \{(x,y) : x, y \ge 0, x + y \le 1\}.$$

The event that neither of Mrs. Jones' sons will get indigestion is

$$A = \left\{(x,y) \in \Omega : x, y < \frac{1}{2}\right\}.$$

These sets are shown in Figure 2.4. If x is uniformly distributed over $[0,1]$ and y is uniformly distributed over $[0, 1-x]$, then the probability measure P over Ω with density

$$f(x,y) = \frac{1}{1-x}$$

will describe the joint distribution of outcomes (x,y), see Figure 2.5.

Now we are in a position to compute

$$P(A) = \int_A f(x,y)\,dx\,dy = \int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} \frac{1}{1-x}\,dx\,dy = \ln\sqrt{2}.$$

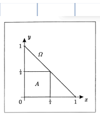


Figure 2.4. The sets Ω and A in Exercise 2.13

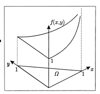


Figure 2.5. The density $f(x,y)$ in Exercise 2.13

The random variable

$$\xi(x,y) = 1 - x - y$$

defined on Ω represents the size of the portion left by Mrs. Jones' sons. Finally, we find that

$$E(\xi|A) = \frac{1}{P(A)}\int_A (1-x-y)\,f(x,y)\,dx\,dy = \frac{1}{\ln\sqrt{2}}\int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} \frac{1-x-y}{1-x}\,dx\,dy = \frac{1-\ln\sqrt{2}}{\ln 4}.$$

Soln: Represent the whole pie by interval $[0,1]$, Let $x \in [0,1]$ be the proportion eaten by the older. Then $y \in [0, 1-x]$ is the proportion avaliable to the the younger,

The set of all possible outcomes is

$$\Omega = \{(x,y) : x, y \ge 0,\ x+y \le 1\}$$

The event neither of the sons get indegestion is

$$A = \{(x,y) \in \Omega : x, y \le \tfrac{1}{2}\}$$

As X, Y are uniformly distributed on $[0,1]$ and $[0,1-x]$ respectively, then $f(x,y) = \frac{1}{1} \times \frac{1}{1-x} = \frac{1}{1-x}$ will describe the distribution of outcomes (x,y).

当 $x, y \le \frac{1}{2}$ 时
且 indigestion 时二者
a.s. 独立

$$\Rightarrow P(A) = \iint_A f(x,y)\,dx\,dy = \int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} \frac{1}{1-x}\,dx\,dy = \ln\sqrt{2}$$

$$\Rightarrow E[1-x-y \mid A] = \frac{1}{P(A)}\int_A (1-x-y)\cdot f(x,y)\,dx\,dy$$

别忘!

$$= \frac{1}{\ln\sqrt{2}}\int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}} (1-x-y)\cdot\frac{1}{1-x}\,dx\,dy$$

$$= \frac{1-\ln\sqrt{2}}{\ln 4}.$$

□.

Exercise 2.14

As a probability space take $\Omega = [0, 1)$ with the σ-field of Borel sets and the Lebesgue measure on $[0, 1)$. Find $E(\xi|\eta)$ if

$$\xi(x) = 2x^2, \qquad \eta(x) = \begin{cases} 2x & \text{for } 0 \le x < \frac{1}{2}, \\ 2x - 1 & \text{for } \frac{1}{2} \le x < 1. \end{cases}$$

Hint What do events in $\sigma(\eta)$ look like? What do $\sigma(\eta)$-measurable random variables look like? If you devise a neat way of describing these, it will make the task of finding $E(\xi|\eta)$ much easier. You will need to transform the integrals in condition 2) of Definition 2.3 to find a formula for the conditional expectation.

Exercise 2.14

As a probability space take $\Omega = [0, 1)$ with the σ-field of Borel sets and the Lebesgue measure on $[0, 1)$. Find $E(\xi|\eta)$ if

$$\xi(x) = 2x^2, \qquad \eta(x) = \begin{cases} 2x & \text{for } 0 \le x < \frac{1}{2}, \\ 2x - 1 & \text{for } \frac{1}{2} \le x < 1. \end{cases}$$

Hint What do events in $\sigma(\eta)$ look like? What do $\sigma(\eta)$-measurable random variables look like? If you devise a neat way of describing these, it will make the task of finding $E(\xi|\eta)$ much easier. You will need to transform the integrals in condition 2) of Definition 2.3 to find a formula for the conditional expectation.

Solu: Observerd that if $x \in [\frac{1}{2}, 1)$ then $\eta(x) = \eta(x - \frac{1}{2})$

so for any Borel set $B \in [\frac{1}{2}, 1)$, $B = B - \frac{1}{2}$

As $\sigma(\eta)$ is a compositron of all such B and $B - \frac{1}{2}$,

then for any $A \in \sigma(\eta)$, $A = A - \frac{1}{2}$

$$\int_A \xi \, dp = \int_A 2x^2 \, dx = \int_A x^2 dx + \int_A x^2 dx$$
$$= \int_A x^2 dx + \int_{A-\frac{1}{2}} (x - \tfrac{1}{2})^2 dx$$
$$= \int_A x^2 dx + \int_A (x - \tfrac{1}{2})^2 dx$$
$$= \int_A [x^2 + (x - \tfrac{1}{2})^2] \, dx$$

Since $E[\xi|\eta]$ is $\sigma(\eta)$-measurable so $\int_A E[\xi|\eta] dp = \int_A \xi \, dp$

$\Rightarrow E[\xi|\eta] = x^2 + (x - \frac{1}{2})^2 = 2x^2 - x + \frac{1}{4}$ for $x \in [\frac{1}{2}, 1)$

Similarly, for any Borel Set $B \subset [0, \frac{1}{2})$, $B = B + \frac{1}{2}$

$\Rightarrow E[\xi|\eta] = x^2 + (x + \frac{1}{2})^2 = 2x^2 + x + \frac{1}{4}$ for $x \in [0, \frac{1}{2})$

Thus, $E[\xi|\eta] = \begin{cases} x^2 + (x + \frac{1}{2})^2, & x \in [0, \frac{1}{2}) \\ x^2 + (x - \frac{1}{2})^2, & x \in [\frac{1}{2}, 1). \end{cases}$ $\square$.

Solution 2.14

The σ-field $\sigma(\eta)$ generated by η consists of sets of the form $B \cup (B + \frac{1}{2})$ for some Borel set $B \subset [0, \frac{1}{2})$. Thus, we are looking for a $\sigma(\eta)$-measurable random

for any $\sigma(\eta)$-measurable r.v. α, $\alpha(x) = \alpha(x + \frac{1}{2})$

variable ζ such that for each Borel set $B \subset [0, \frac{1}{2})$

$$\int_{B \cup (B + \frac{1}{2})} \xi(x) \, dx = \int_{B \cup (B + \frac{1}{2})} \zeta(x) \, dx. \tag{2.8}$$

Then $E(\xi|\eta) = \zeta$ by Definition 2.3.

Transforming the integral on the left-hand side, we obtain

$$\begin{aligned} \int_{B \cup (B + \frac{1}{2})} \xi(x) \, dx &= \int_B 2x^2 \, dx + \int_{B + \frac{1}{2}} 2x^2 \, dx \\ &= \int_B 2x^2 \, dx + \int_B 2\left(x + \tfrac{1}{2}\right)^2 dx \\ &= 2 \int_B \left(x^2 + \left(x + \tfrac{1}{2}\right)^2\right) dx. \end{aligned}$$

For ζ to be $\sigma(\eta)$-measurable it must satisfy

$$\zeta(x) = \zeta\left(x + \tfrac{1}{2}\right) \tag{2.9}$$

for each $x \in [0, \frac{1}{2})$. Then

$$\begin{aligned} \int_{B \cup (B + \frac{1}{2})} \zeta(x) \, dP &= \int_B \zeta(x) \, dx + \int_{B + \frac{1}{2}} \zeta(x) \, dx \\ &= \int_B \zeta(x) \, dx + \int_B \zeta\left(x + \tfrac{1}{2}\right) dx \\ &= \int_B \zeta(x) \, dx + \int_B \zeta(x) \, dx \\ &= 2 \int_B \zeta(x) \, dx. \end{aligned}$$

If (2.8) is to hold for any Borel set $B \subset [0, \frac{1}{2})$, then

$$\zeta(x) = x^2 + \left(x + \tfrac{1}{2}\right)^2$$

for each $x \in [0, \frac{1}{2})$. The values of $\zeta(x)$ for $x \in [\frac{1}{2}, 1)$ can be obtained from (2.9). It follows that

$$E(\xi|\eta)(x) = \zeta(x) = \begin{cases} x^2 + \left(x + \frac{1}{2}\right)^2 & \text{for } 0 \le x < \frac{1}{2}, \\ \left(x - \frac{1}{2}\right)^2 + x^2 & \text{for } \frac{1}{2} \le x < 1. \end{cases}$$

The graphs of ξ, η and $E(\xi|\eta)$ are shown in Figure 2.6.

Exercise 2.15

Take $\Omega = [0, 1]$ with the σ-field of Borel sets and P the Lebesgue measure on $[0, 1]$. Let

$$\eta(x) = x(1 - x) \Leftrightarrow \eta(x) = \eta(1 - x)$$

for $x \in [0, 1]$. Show that

$$E(\xi|\eta)(x) = \frac{\xi(x) + \xi(1 - x)}{2}$$

for any $x \in [0, 1]$.

Proof. Since $\eta(x) = \eta(1-x)$, so the σ-field $\sigma(\eta)$ consists of Borel sets $B \subset [0,1]$ s.t. $B = 1 - B$ where $1 - B = \{1 - x : x \in B\}$.

事件x的Probability measure

"事件~~结果~~相同"即B事件$x \in B$，$\eta(x)$的值与$1-B$中的事件$(1-x)$的$\eta(1-x)$值相同.

For any such B, ($B \in \sigma(\eta)$ and $B \subset [0,1]$):

$$\int_B \xi(x) \, dp = \frac{1}{2} \int_B \xi(x) dx + \frac{1}{2} \int_B \xi(x) dx$$
$$= \frac{1}{2} \int_B \xi(x) dx + \frac{1}{2} \int_{1-B} \xi(1-x) dx \quad (x \in 1-B\text{时}, 1-x\text{才等价于原B中的}x)$$
$$= \frac{1}{2} \int_B \xi(x) dx + \frac{1}{2} \int_B \xi(1-x) dx \quad 1 - B = B$$
$$= \int_B \frac{\xi(x) + \xi(1-x)}{2} dx$$

As $\frac{\xi(x) + \xi(1-x)}{2}$ is $\sigma(\eta)$-measurable so by Def. 2.3, $\int_B E[\xi|\eta] dp = \int_B \xi \, dp$.

$E[\xi|\eta] = \frac{\xi(x) + \xi(1-x)}{2}$ Q.E.D.

Solution 2.15

Since $\eta(x) = \eta(1 - x)$, the σ-field $\sigma(\eta)$ consists of Borel sets $B \subset [0, 1]$ such that

$$\{x : x \in B\} = B = 1 - B, = \{1 - x : x \in B\}$$

where $1 - B = \{1 - x : x \in B\}$. For any such B

$$\begin{aligned} \int_B \xi(x) \, dx &= \frac{1}{2} \int_B \xi(x) \, dx + \frac{1}{2} \int_B \xi(x) \, dx \\ &= \frac{1}{2} \int_B \xi(x) \, dx + \frac{1}{2} \int_{1-B} \xi(1 - x) \, dx \\ &= \frac{1}{2} \int_B \xi(x) \, dx + \frac{1}{2} \int_B \xi(1 - x) \, dx \\ &= \int_B \frac{\xi(x) + \xi(1 - x)}{2} \, dx. \end{aligned}$$

Because $\frac{1}{2}(\xi(x) + \xi(1 - x))$ is $\sigma(\eta)$-measurable, it follows that

$$E(\xi|\eta)(x) = \frac{\xi(x) + \xi(1 - x)}{2}.$$

Exercise 2.16

Let ξ, η be integrable random variables with joint density $f_{\xi,\eta}(x, y)$. Show that

$$E(\xi|\eta) = \frac{\int_{\mathbb{R}} x \, f_{\xi,\eta}(x, \eta) \, dx}{\int_{\mathbb{R}} f_{\xi,\eta}(x, \eta) \, dx}.$$

Hint Study the solutions to Exercises 2.7 and 2.8.

Solution 2.16

We are looking for a Borel function $F(y)$ such that

$$\int_{\{\eta\in B\}} \xi \, dP = \int_{\{\eta\in B\}} F(\eta)\, dP$$

for $\boxed{\text{any Borel set } B \text{ in } \mathbb{R}.}$ Because $F(\eta)$ is $\sigma(\eta)$-measurable and each event in $\sigma(\eta)$ can be written as $\{\eta \in B\}$ for some Borel set B, this will mean that

$E(\xi|\eta) = F(\eta)$

Let us transform the two integrals above using the joint density of ξ and η:

$$\begin{aligned}\int_{\{\eta\in B\}} \xi \, dP &= \int_B \int_{\mathbb{R}} x\, f_{\xi,\eta}(x,y)\, dx\, dy \\ &= \int_B \left(\int_{\mathbb{R}} x\, f_{\xi,\eta}(x,y)\, dx \right) dy\end{aligned}$$

and

$$\begin{aligned}\int_{\{\eta\in B\}} F(\eta)\, dP &= \int_B \int_{\mathbb{R}} F(y)\, f_{\xi,\eta}(x,y)\, dx\, dy \\ &= \int_B F(y) \left(\int_{\mathbb{R}} f_{\xi,\eta}(x,y)\, dx \right) dy.\end{aligned}$$

If these two integrals are to be equal for each Borel set B, then

$$F(y) = \frac{\int_{\mathbb{R}} x\, f_{\xi,\eta}(x,y)\, dx}{\int_{\mathbb{R}} f_{\xi,\eta}(x,y)\, dx}.$$

It follows that

$$E(\xi|\eta) = F(\eta) = \frac{\int_{\mathbb{R}} x\, f_{\xi,\eta}(x,\eta)\, dx}{\int_{\mathbb{R}} f_{\xi,\eta}(x,\eta)\, dx}.$$

Remark 2.3. Denote

$$f_{\xi,\eta}(x \mid y) = \frac{f_{\xi,\eta}(x,y)}{f_\eta(y)},$$

where

$$f_\eta(y) = \int_{\mathbb{R}} f_{\xi,\eta}(x,y) \mathrm{d}x$$

is the marginal density of η. Then, by the result of Exercise 2.16, we have

$$\mathbb{E}[\xi \mid \eta = y] = \int_{\mathbb{R}} x f_{\xi,\eta}(x \mid y)\, \mathrm{d}x.$$

We call $f_{\xi,\eta}(x \mid y)$ the conditional density of ξ given $\eta = y$.

5. Integration of Random Variables

A random variable ξ is said to be *integrable* if

$$\int_\Omega |\xi(\omega)| \mathbb{P}(\mathrm{d}\omega) < \infty.$$

Then, the expectation, or the mean, of ξ exists, and it is given by

$$\mathbb{E}[\xi] = \int_\Omega \xi(\omega)\mathbb{P}(\mathrm{d}\omega) = \int_{-\infty}^{\infty} x \mathrm{d}F_\xi(x).$$

If ξ is continuous, and $f_\xi(x)$ is the density

$$\mathbb{E}[\xi] = \int_{-\infty}^{\infty} x f_\xi(x)\mathrm{d}x,$$

and if ξ is discrete, and $f_\xi(x)$ is the probability distribution,

$$\mathbb{E}[\xi] = \sum_{i=1}^{\infty} x_i f_\xi(x_i).$$

Some useful formulae: If $h(\xi)$ is integrable, and $A \in \mathcal{F}$, then

换元

$$\begin{aligned} \mathbb{E}[h(\xi)] &= \int_\Omega h(\xi(\omega))\mathbb{P}(\mathrm{d}\omega) = \int_{-\infty}^{\infty} h(x)\mathrm{d}F_\xi(x). \\ \mathbb{P}(A) &= \mathbb{E}[1_A] = \int_\Omega 1_A(\omega)\mathbb{P}(\mathrm{d}\omega) = \int_A \mathbb{P}(\mathrm{d}\omega). \end{aligned}$$

Let $x := \xi(\omega)$ then $x \in R$

$$\int_\Omega h(\xi(\omega))\, dP(\omega)$$
$$= \int_R h(x)\, dP(\xi^{-1}(x))$$
$$= \int_R h(x)\, dP(\{\xi \in R\})$$
$$= \int_R h(x)\, dF(\xi \le x)$$
$$= \int_R h(x)\, dF_\xi(x) \quad \forall x \in R.$$

Basic properties of expectation:

(1) If $c_1, \ldots, c_n$ are constants, then

$$\mathbb{E}\left[\sum_{i=1}^{n} c_i \xi_i\right] = \sum_{i=1}^{n} c_i \mathbb{E}[\xi_i].$$

(2) If $\xi \ge 0$, then $\mathbb{E}[\xi] \ge 0$; and if $\xi = 0$, then $\mathbb{E}[\xi] = 0$.

(3) If $\xi \equiv c$ for some $c \in \mathbb{R}$, then $\mathbb{E}[\xi] = c$.

(4) If $A \in \mathcal{F}$, then $\mathbb{E}[1_A] = \mathbb{P}(A)$.

(5) **Monotone convergence theorem**. If $\{\xi_n\}$ is a sequence of random variables with $0 \le \xi_n \le \xi_{n+1}$ for all n, and there is an integrable random variable ξ such that $\xi_n \to \xi$ as $n \to \infty$, then

$$\lim_{n\to\infty} \mathbb{E}[\xi_n] = \mathbb{E}\left[\lim_{n\to\infty} \xi_n\right] = \mathbb{E}[\xi]. \qquad (*)$$

例有"上确界"

(6) **Lebesgue's dominated convergence theorem**. If $\{\xi_n\}$ is a sequence of random variables such that $\xi_n \to \xi$ as $n \to \infty$, and there exists an integrable random variable η such that $|\xi_n| \le |\eta|$ holds for all n, then $(*)$ holds. 例有上界

(7) **Fatou's lemma**. Assume that $\{\xi_n\}$ is a sequence of nonnegative random variables, and there is an integrable random variable η such that $\xi_n \le \eta$ for all n. Then

$$\mathbb{E}\left[\liminf_n \xi_n\right] \le \liminf_n \mathbb{E}[\xi_n] \le \limsup_n \mathbb{E}[\xi_n] \le \mathbb{E}\left[\limsup_n \xi_n\right]$$

6. Conditional Probability

Conditional probability: For any events $A, B \in \mathcal{F}$ such that $\mathbb{P}(B) \neq 0$, the *conditional probability of A given B* is given by

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Total probability formula I: Let $A \in \mathcal{F}$ and let $B_1, B_2, \dots$ be a sequence of pairwise disjoint events such that $B_1 \cup B_2 \cup \dots = \Omega$ and $\mathbb{P}(B_i) \neq 0$ for any i. Then

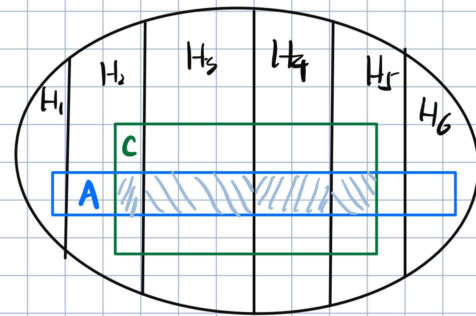
$$\mathbb{P}(A) = \sum_i \mathbb{P}(A \mid B_i)\mathbb{P}(B_i).$$

→ 用之前务必证 pairwise disjoint 和 $\bigcup_n B_i = \Omega$.

Total probability formula II: If $H_k \in \mathcal{F}$, $\mathbb{P}(H_k \cap H_l) = 0$ for $k \neq l$, and $\mathbb{P}(\bigcup_k H_k) = 1$, then

$$\mathbb{P}(A \mid C) = \sum_k \mathbb{P}(A \mid C \cap H_k)\mathbb{P}(H_k \mid C).$$

$$= \sum_k \frac{P(A \cap (C \cap H_k))}{P(C \cap H_k)} \cdot \frac{P(H_k \cap C)}{P(C)} = \sum_k \frac{P(A \cap (C \cap H_k))}{P(C)}$$



7. Conditional Expectation for Given Random Variable

Let ξ be an integrable random variable and let η be an arbitrary random variable.

According to the definition, in order to find $\mathbb{E}[\xi \mid \eta]$, the conditional expectation of ξ given η, we need to find a random variable ζ such that

1) ζ is $\sigma(\eta)$-measurable, $\Rightarrow Z = E[\xi \mid \eta]$ exists (1)
2) for any $A \in \sigma(\eta)$,

$$\int_A \zeta(\omega)\mathbb{P}(\mathrm{d}\omega) = \int_A \xi(\omega)\mathbb{P}(\mathrm{d}\omega).$$

We can show that such random variable ζ is unique (2) (by Lemma 2.1). Thus, this random variable ζ is the conditional expectation of ξ given η, i.e. $\zeta = \mathbb{E}[\xi \mid \eta]$.

g(η) 和 E[ξ|η] 都是 σ(η) measurable 的 Borel function.

Proposition.

For any Borel measurable function $g(x)$,

$$\mathbb{E}\big[(\xi - g(\eta))^2\big] \geq \mathbb{E}\big[(\xi - \mathbb{E}[\xi \mid \eta])^2\big],$$

i.e. $\mathbb{E}[\xi \mid \eta]$ is the best estimate that one can make of ξ from knowledge of η when η is known. We say that $\mathbb{E}[\xi \mid \eta]$ is the regression of ξ on η.

Proof. Let $\zeta = \mathbb{E}[\xi \mid \eta]$. Then, we have

$$\begin{aligned}\mathbb{E}\big[(\xi - g(\eta))^2 \mid \eta\big] &= \mathbb{E}\big[(\xi - \zeta + \zeta - g(\eta))^2 \mid \eta\big] \\ &= \mathbb{E}\big[(\xi - \zeta)^2 \mid \eta\big] + \mathbb{E}\big[(\zeta - g(\zeta))^2 \mid \eta\big] \\ &\quad + 2\mathbb{E}\big[(\xi - \zeta)(\zeta - g(\eta)) \mid \eta\big].\end{aligned}$$

$E(\zeta - g(\eta) \mid \eta) = \zeta - g(\eta)$, $E(\eta \mid \eta) = E(\eta)$

Since η and $\zeta - g(\eta)$ are $\sigma(\eta)$-measurable, we get

$$\mathbb{E}\big[(\xi - \zeta)(\zeta - g(\eta)) \mid \eta\big] = (\zeta - g(\zeta))\big(\mathbb{E}[\xi \mid \eta] - \zeta\big) = 0,$$

and hence, we have

$$\begin{aligned}\mathbb{E}\big[(\xi - g(\eta))^2 \mid \eta\big] &= \mathbb{E}\big[(\xi - \zeta)^2 \mid \eta\big] + \mathbb{E}\big[(\zeta - g(\eta))^2 \mid \eta\big] \\ &\geq \mathbb{E}\big[(\xi - \zeta)^2 \mid \eta\big].\end{aligned}$$

Taking the expectations, we obtain the inequality in the proposition. □

* ζ-g(η) is σ(η) measurable

但 (ζ-g(η))² may not be so.

因为 $L^2 < L^1$

→ ∃φ s.t. φ(η)=E[ξ|η] as E[ξ|η] is σ(η)-measurable.

According to the Doob-Dynkin lemma, the problem finding $\mathbb{E}[\xi \mid \eta]$, the conditional expectation of ξ given η, becomes to find a Borel function $\varphi(x)$ such that, for any $A \in \sigma(\eta)$,

$$\int_A \varphi(\eta)(\omega)\mathbb{P}(\mathrm{d}\omega) = \int_A \xi(\omega)\mathbb{P}(\mathrm{d}\omega). \qquad (*)$$

Then we have that $\mathbb{E}\big[\xi \mid \eta\big] = \varphi(\eta)$.

However, there is no general formula or procedure to find such a function $\varphi(x)$. Chapter 2 gives some methods or ideas, in some special cases, to find the Borel function $\varphi(x)$.

- Let ξ and η are two random variables. If η is $\sigma(\xi)$-measurable, then there exists a Borel function $f : \mathbb{R} \longrightarrow \mathbb{R}$ such that $\eta = f(\xi)$. (Doob-Dynkin Lemma)

(1) If $\xi = \varphi(\eta)$ for a Borel function $\varphi(x)$, then ξ is $\sigma(\eta)$-measurable, and the quality $(*)$ of course holds for all $A \in \sigma(\eta)$. This means that $\mathbb{E}\big[\xi \mid \eta\big] = \xi$.

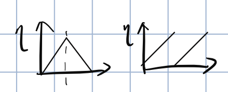
For instance, in Example 2.4, for the interval $[\frac{1}{2}, 1]$, we have that $\xi = 2\eta^2$ (that is, $\xi = \varphi(\eta)$), and so that, we take $\mathbb{E}\big[\xi \mid \eta\big] = \xi$ in the interval $[\frac{1}{2}, 1]$.

(2) If we well know the construction or the feature of $\sigma(\eta)$, we can try to guess the function $\varphi(\eta)$ since it is $\sigma(\eta)$-measurable, and then we try to show that it satisfies the equation $(*)$.

In Exercise 2.6, we use this idea.

由 σ(η) 所表现出的 pattern 猜 E[ξ|η].

e.g.



(3) If we know the joint density $f(x, y)$ of ξ and η, we have the formula to find $\varphi(\eta)$:

$$\varphi(\eta) = \frac{\int_{\mathbb{R}} x f(x, \eta)\, \mathrm{d}x}{\int_{\mathbb{R}} f(x, \eta)\mathrm{d}x},$$

which is given in Exercise 2.16, and we can also verify that the function $\varphi(\eta)$ satisfies the equation $(*)$.

In Exercise 2.7 and Exercise 2.16. we use this method.

(4) If η is discrete, then we can use the formula:

$$\mathbb{E}\big[\xi \mid \{\eta = y_i\}\big] = \frac{1}{\mathbb{P}(\eta = y_i)} \int_{\{\eta = y_i\}} \xi(\omega)\mathbb{P}(\mathrm{d}\omega),$$

which is given in Definition 2.1, and use Definition 2.2 to find $\mathbb{E}\big[\xi \mid \eta\big]$.

We use this procedure in Example 2.2.

8. Conditional Expectation for Given σ-Field

Conditional expectation for given σ-field: Let ξ be an integrable random variable, and let $\mathcal{G}$ be a sub-σ-field of $\mathcal{F}$. Then the *conditional expectation* of ξ given $\mathcal{G}$ is defined to be a random variable $\mathbb{E}[\xi \mid \mathcal{G}]$ such that

1) $\mathbb{E}[\xi \mid \mathcal{G}]$ is $\mathcal{G}$-measurable,
2) for any $A \in \mathcal{G}$,

$$\int_A \mathbb{E}\big[\xi \mid \mathcal{G}\big](\omega)\mathbb{P}(\mathrm{d}\omega) = \int_A \xi(\omega)\mathbb{P}(\mathrm{d}\omega).$$

Remarks: (1) $\mathbb{E}\big[\xi \mid \sigma(\eta)\big] = \mathbb{E}\big[\xi \mid \eta\big]$. (2) The existence of conditional expectations: Random-Nikodym theorem. (3) The uniqueness of conditional expectations: Lemma 2.1 and Proposition 2.3.

Basic properties: Assume ξ and ζ are integrable random variables. The conditional expectation possesses the following properties:

1) If a and b are two constants, then

$$\mathbb{E}\big[a\xi + b\zeta \mid \mathcal{G}\big] = a\mathbb{E}\big[\xi \mid \mathcal{G}\big] + b\mathbb{E}\big[\zeta \mid \mathcal{G}\big];$$

2) $\mathbb{E}\big[\mathbb{E}[\xi \mid \mathcal{G}]\big] = \mathbb{E}[\xi]$;
3) If ξ is $\mathcal{G}$-measurable, then $\mathbb{E}[\xi \mid \mathcal{G}] = \xi$; and

$$\mathbb{E}\big[\xi\zeta \mid \mathcal{G}\big] = \xi\mathbb{E}\big[\zeta \mid \mathcal{G}\big];$$

when $\xi\zeta$ is integrable.

4) If ξ is independent of $\mathcal{G}$, then $\mathbb{E}\big[\xi \mid \mathcal{G}\big] = \mathbb{E}[\xi]$;
5) If $\mathcal{H}$ is a sub-σ-field of $\mathcal{F}$ such that $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$, then

$$\mathbb{E}\Big[\mathbb{E}\big[\xi \mid \mathcal{G}\big] \mid \mathcal{H}\Big] = \mathbb{E}\big[\xi \mid \mathcal{H}\big];$$

6) If $\xi \geq 0$, then $\mathbb{E}[\xi \mid \mathcal{G}] \geq 0$.
7) (Jensen's Inequality). If $\varphi(x)$ is convex, then

$$\varphi\big(\mathbb{E}[\xi \mid \mathcal{G}]\big) \leq \mathbb{E}\big[\varphi(\xi) \mid \mathcal{G}\big].$$

Here we also have following facts:

- If $\xi_1, \xi_2, \dots$ is a martingale with respect to a filtration $\mathcal{F}_1, \mathcal{F}_2, \dots$, then

$$\mathbb{E}[\xi_{k+1} \mid \mathcal{F}_k] = \xi_k,$$

  holds for all k (Chapter 3 and Chapter 4).
- If $\xi_0, \xi_1, \xi_2, \dots$ is a Markov chain, then

$$\mathbb{P}\big(\xi_{k+1} = s \mid \mathcal{G}_k\big) = \mathbb{P}\big(\xi_{k+1} = s \mid \xi_k\big),$$

  where $\mathcal{G}_k = \sigma(\xi_0, \xi_1, \dots, \xi_k)$ (Chapter 5).

- Each event in $\sigma(\eta)$ can be written as $\{\eta \in B\}$ for some Borel set $B \in \mathcal{B}(\mathbb{R})$.
  - a combination of sets in the form $\{\eta = y_i\}$.

- $\sigma(\eta)$ consists of sets $B \subset [0,1]$ s.t. $B = 1 - B$ where $1-B = \{1-x : x \in B\}$.
  - of the form $B \cup (B + \frac{1}{2})$ for some Borel set $B \subset [0, \frac{1}{2})$
  - $A \subset [0,1]$ symmetric about $\frac{1}{2}$. i.e. $A = 1 - A$.

- If A is indeed such set then $A = \{\eta \in B\}$ where

  $B = \{2x : x \in A \cap [0, \frac{1}{2})\}$ is a Borel set in $\mathcal{B}(\mathbb{R})$, so $A = \sigma(\eta)$.

  On the other hand, if $A \in \sigma(\eta)$ then there's a B in $\mathbb{R}$ s.t.

  $A = \{\eta \in B\}$ then

$$
\begin{aligned}
x \in A &\iff \eta(x) \in B \\
&\iff \eta(1-x) \in B \\
&\iff (1-x) \in A
\end{aligned}
$$

  so $A = 1 - A$.