

MA400. September Introductory Course
(Financial Mathematics, Risk & Stochastics)
P5. Conditional Expectation

1. We assume that an underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is fixed.
2. Recall from Chapter P2 where we defined random variables that σ -algebras are mathematical models for information. Suppose that we are interested in a random variable X , say the price of a certain stock at some future date. Consider a situation where the information provided by a σ -algebra $\mathcal{G} \subseteq \mathcal{F}$ becomes available to us. For instance, \mathcal{G} can identify with the information $\sigma(Z)$ that is associated with the observation of another random variable Z . The conditional expectation $\mathbb{E}[X | \mathcal{G}]$ is the expected value of X if we are given the information \mathcal{G} .

The conditional expectation $\mathbb{E}[X | \mathcal{G}]$ is a random variable that is \mathcal{G} -measurable, namely,

$$\sigma(\mathbb{E}[X | \mathcal{G}]) \subseteq \mathcal{G}.$$

Indeed, once we possess the information \mathcal{G} and we have updated our beliefs about the likelihood of random events, the conditional expectation should be simply the “updated” expectation of X . To appreciate more this point, we consider the following “extreme” cases.

- Suppose that the information \mathcal{G} contains the “knowledge” of the actual value of X , i.e., X is \mathcal{G} -measurable ($\sigma(X) \subseteq \mathcal{G}$). In this case, $\mathbb{E}[X | \mathcal{G}] = X$ (see 18.(iii) below), which reflects the idea that “knowledge” of \mathcal{G} implies “knowledge” of the actual value of X .
 - Suppose that X is independent of the information \mathcal{G} . In this case, $\mathbb{E}[X | \mathcal{G}] = \mathbb{E}[X]$ (see 18.(xi) below), which conveys the idea that “knowledge” of \mathcal{G} provides no information about X .
 - The trivial σ -algebra $\{\Omega, \emptyset\}$ can be viewed as a model for “absence of information”: we can interpret Ω as the event that “something occurs” and \emptyset as the event that “nothing happens”. In this context, the result $\mathbb{E}[X | \{\Omega, \emptyset\}] = \mathbb{E}[X]$ (see 18.(i) below), which expresses expectation as a “special case” of conditional expectation, is only the “expected” one.
3. Given an event $B \in \mathcal{F}$, recall that $\mathbb{E}[\mathbf{1}_B] = \mathbb{P}(B)$. In view of this observation, it is natural to identify the conditional probability $\mathbb{P}(B | \mathcal{G})$ of B given a σ -algebra $\mathcal{G} \subseteq \mathcal{F}$ with the conditional expectation $\mathbb{E}[\mathbf{1}_B | \mathcal{G}]$ of the random variable $\mathbf{1}_B$ given the σ -algebra \mathcal{G} (see Definition 17 below).
 4. It is important to keep in mind that conditional expectation and conditional probability are random variables: **probability theory is concerned with the future!**

5. In what follows, we start from elementary considerations and we end up with the general mathematical Definitions 15 and 17 of conditional expectation and conditional probability.

Elementary conditional probability

6. Given an event $B \in \mathcal{F}$, $\mathbb{P}(B)$ quantifies our views on how likely it is for the event B to occur. Now, suppose that we have been informed that chance outcomes are restricted within an event $A \in \mathcal{F}$. In other words, suppose that somebody informs us that all likely to happen events are subsets of A , and all events that are subsets of A^c are impossible to occur.

How should we modify our views, namely our probability measure, to account for this scenario? Let us denote by $\mathbb{P}(B \mid A)$ our modified belief on the likelihood of the event $B \in \mathcal{F}$ given the knowledge that A has occurred. Since the only new information that we possess is that chance outcomes are restricted within the event A , it is natural to postulate that $\mathbb{P}(B \mid A)$ should be proportional to $\mathbb{P}(B \cap A)$, namely

$$\mathbb{P}(B \mid A) \sim \mathbb{P}(B \cap A). \quad (1)$$

However, our beliefs should “add up” to 1, so that we have a proper probability measure. This means that we should impose the requirement that $\mathbb{P}(\Omega \mid A) = 1$. Since $\mathbb{P}(\Omega \cap A) = \mathbb{P}(A)$, we conclude that we should scale the right hand side of (1) by $1/\mathbb{P}(A)$ to obtain

$$\mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}. \quad (2)$$

(Of course, this formula makes sense only if $\mathbb{P}(A) > 0$.)

We can check that the function $\mathbb{P}(\cdot \mid A) : \mathcal{F} \rightarrow [0, 1]$ defined (2) is indeed a probability measure on (Ω, \mathcal{F}) [*Exercise*].

7. *Law of total probability.* Suppose that the events $A_1, A_2, \dots, A_n \in \mathcal{F}$ form a partition of Ω i.e., $A_i \cap A_j = \emptyset$ for all $i \neq j$, and $\bigcup_{i=1}^n A_i = \Omega$. Given an event $B \in \mathcal{F}$, the events $B \cap A_1, B \cap A_2, \dots, B \cap A_n$ are pairwise disjoint and $\bigcup_{i=1}^n B \cap A_i = B$. Therefore, the additivity property of a probability measure and (2) imply

$$\begin{aligned} \mathbb{P}(B) &= \mathbb{P}(B \cap A_1) + \mathbb{P}(B \cap A_2) + \dots + \mathbb{P}(B \cap A_n) \\ &= \mathbb{P}(B \mid A_1)\mathbb{P}(A_1) + \mathbb{P}(B \mid A_2)\mathbb{P}(A_2) + \dots + \mathbb{P}(B \mid A_n)\mathbb{P}(A_n). \end{aligned} \quad (3)$$

8. *Bayes' theorem.* Suppose that the events $A_1, A_2, \dots, A_n \in \mathcal{F}$ form a partition of Ω . Given an event $B \in \mathcal{F}$ such that $\mathbb{P}(B) > 0$, and any $k = 1, 2, \dots, n$, (2) and the law of total probability (3) imply

$$\begin{aligned}\mathbb{P}(A_k | B) &= \frac{\mathbb{P}(A_k \cap B)}{\mathbb{P}(B)} \\ &= \frac{\mathbb{P}(B | A_k)\mathbb{P}(A_k)}{\mathbb{P}(B | A_1)\mathbb{P}(A_1) + \mathbb{P}(B | A_2)\mathbb{P}(A_2) + \dots + \mathbb{P}(B | A_n)\mathbb{P}(A_n)}.\end{aligned}\quad (4)$$

9. Conditional probabilities defined as in (2) have an *a posteriori* character: we have been informed and we know that event A has occurred. How should we develop our theory to account for a *prior to observation* perspective? In other words, suppose that we anticipate an observation that will inform us on whether A or A^c occurs. How should we modify our views to account for this situation?

Given the arguments in Paragraph 6 above, the natural answer is to set

$$\begin{aligned}\mathbb{P}(B | \text{"observation of } A \text{ or } A^c\text{"}) &= \begin{cases} \mathbb{P}(B | A) & \text{if } A \text{ occurs,} \\ \mathbb{P}(B | A^c) & \text{if } A^c \text{ occurs,} \end{cases} \\ &= \mathbb{P}(B | A)\mathbf{1}_A + \mathbb{P}(B | A^c)\mathbf{1}_{A^c},\end{aligned}\quad (5)$$

provided, of course, that $0 < \mathbb{P}(A) < 1$. Observe that our views on how likely it is for the event B to occur have now become a simple *random variable*. Given any sample $\omega \in \Omega$, the conditional probability of the event B takes the value $\mathbb{P}(B | A)$ if $\omega \in A$ and takes the value $\mathbb{P}(B | A^c)$ if $\omega \in A^c$.

10. Given events $A, B \in \mathcal{F}$ such that $0 < \mathbb{P}(A) < 1$, the random variable Y defined by

$$Y = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}\mathbf{1}_A + \frac{\mathbb{P}(A^c \cap B)}{\mathbb{P}(A^c)}\mathbf{1}_{A^c}$$

is the conditional probability of B given the σ -algebra $\{\emptyset, \Omega, A, A^c\}$. We denote this conditional probability by

$$\mathbb{P}(B | \{\emptyset, \Omega, A, A^c\}) = \mathbb{E}[\mathbf{1}_B | \{\emptyset, \Omega, A, A^c\}].$$

We note that this random variable Y has the following properties:

- (i) Y is $\{\emptyset, \Omega, A, A^c\}$ -measurable.
- (ii) $\mathbb{E}[|Y|] < \infty$.
- (iii) $\mathbb{E}[\mathbf{1}_C Y] = \mathbb{E}[\mathbf{1}_C X]$ for all $C \in \{\emptyset, \Omega, A, A^c\}$. Indeed, given any $C \in \{\emptyset, \Omega, A, A^c\}$,

$$\begin{aligned}
 \mathbb{E}[\mathbf{1}_C Y] &= \mathbb{E} \left[\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \mathbf{1}_C \mathbf{1}_A + \frac{\mathbb{P}(A^c \cap B)}{\mathbb{P}(A^c)} \mathbf{1}_C \mathbf{1}_{A^c} \right] \\
 &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \mathbb{P}(A \cap C) + \frac{\mathbb{P}(A^c \cap B)}{\mathbb{P}(A^c)} \mathbb{P}(A^c \cap C) \\
 &= \begin{cases} \mathbb{P}(\emptyset \cap B), & \text{if } C = \emptyset, \\ \mathbb{P}(\Omega \cap B), & \text{if } C = \Omega, \\ \mathbb{P}(A \cap B), & \text{if } C = A, \\ \mathbb{P}(A^c \cap B), & \text{if } C = A^c, \end{cases} \\
 &= \mathbb{E}[\mathbf{1}_C \mathbf{1}_B].
 \end{aligned}$$

Conditional expectation of a simple random variable given another simple random variable

11. Consider two simple random variables X and Z and suppose that

$$X = \sum_{i=1}^n x_i \mathbf{1}_{\{X=x_i\}} \quad \text{and} \quad Z = \sum_{j=1}^m z_j \mathbf{1}_{\{Z=z_j\}},$$

for some distinct x_1, \dots, x_n and z_1, \dots, z_m . Also, assume that $\mathbb{P}(Z = z_j) > 0$ for all $j = 1, \dots, m$.

We observe that the σ -algebra $\sigma(Z)$ is easy to describe: it consists of all possible unions of sets in the family $\{\{Z = z_1\}, \dots, \{Z = z_m\}\}$, namely,

$$\sigma(Z) = \left\{ \bigcup_{k \in J} \{Z = z_k\} \mid J \subseteq \{1, \dots, m\} \right\}, \quad (6)$$

with the convention that

$$\bigcup_{k \in \emptyset} \{Z = z_k\} = \emptyset.$$

12. Suppose that we have made an “experiment” that has informed us about the actual value of Z . In particular, suppose that we have been given the information that the actual value of the random variable Z is z_j , for some $j = 1, \dots, m$. In this context where we know that the event $\{Z = z_j\}$ has occurred, we should revise our probabilities from $\mathbb{P}(\cdot)$ to $\mathbb{P}(\cdot \mid Z = z_j)$. Furthermore, we should revise the expectation of X from $\mathbb{E}[X]$ to

$$\mathbb{E}[X \mid Z = z_j] = \sum_{i=1}^n x_i \mathbb{P}(X = x_i \mid Z = z_j).$$

This conditional expectation, *the conditional expectation of X given that the random variable Z is equal to z_j* , which is a real number, has an *a posteriori* character: we have been informed that the actual value of Z is z_j .

Prior to observation, namely, before we observe the value of the random variable Z , it is natural to define *the conditional expectation of X given the information we can obtain from the observation of the random variable Z* by

$$\mathbb{E}[X \mid \sigma(Z)] \equiv \mathbb{E}[X \mid Z] = \sum_{j=1}^m \mathbb{E}[X \mid Z = z_j] \mathbf{1}_{\{Z=z_j\}},$$

which is a random variable.

13. The random variable

$$Y = \sum_{j=1}^m \mathbb{E}[X \mid Z = z_j] \mathbf{1}_{\{Z=z_j\}},$$

namely, the conditional expectation $\mathbb{E}[X \mid \sigma(Z)] \equiv \mathbb{E}[X \mid Z]$ of X given the σ -algebra $\sigma(Z)$ has the following properties:

- (i) Y is $\sigma(Z)$ -measurable because Y is constant on the event $\{Z = z_j\}$ for all $j = 1, \dots, m$.
- (ii) $\mathbb{E}[|Y|] < \infty$ because $|Y|$ can take only m possible values.
- (iii) $\mathbb{E}[\mathbf{1}_C Y] = \mathbb{E}[\mathbf{1}_C X]$ for all $C \in \sigma(Z)$. To see this claim, we consider any event

$C \in \sigma(Z)$ and we use (6) to calculate

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_C Y] &= \mathbb{E} \left[\left(\sum_{k \in J} \mathbf{1}_{\{Z=z_k\}} \right) \left(\sum_{j=1}^m \sum_{i=1}^n x_i \frac{\mathbb{P}(X = x_i, Z = z_j)}{\mathbb{P}(Z = z_j)} \mathbf{1}_{\{Z=z_j\}} \right) \right] \\
&= \mathbb{E} \left[\sum_{i=1}^n x_i \sum_{k \in J} \sum_{j=1}^m \frac{\mathbb{P}(X = x_i, Z = z_j)}{\mathbb{P}(Z = z_j)} \mathbf{1}_{\{Z=z_j\}} \mathbf{1}_{\{Z=z_k\}} \right] \\
&= \mathbb{E} \left[\sum_{i=1}^n x_i \sum_{k \in J} \frac{\mathbb{P}(X = x_i, Z = z_k)}{\mathbb{P}(Z = z_k)} \mathbf{1}_{\{Z=z_k\}} \right] \\
&= \sum_{i=1}^n x_i \sum_{k \in J} \mathbb{P}(X = x_i, Z = z_k) \\
&= \sum_{i=1}^n x_i \sum_{k \in J} \mathbb{E} [\mathbf{1}_{\{X=x_i\} \cap \{Z=z_k\}}] \\
&= \sum_{i=1}^n x_i \sum_{k \in J} \mathbb{E} [\mathbf{1}_{\{X=x_i\}} \mathbf{1}_{\{Z=z_k\}}] \\
&= \sum_{i=1}^n x_i \mathbb{E} \left[\mathbf{1}_{\{X=x_i\}} \sum_{k \in J} \mathbf{1}_{\{Z=z_k\}} \right] \\
&= \sum_{i=1}^n x_i \mathbb{E} [\mathbf{1}_{\{X=x_i\}} \mathbf{1}_C] \\
&= \mathbb{E} \left[\mathbf{1}_C \sum_{i=1}^n x_i \mathbf{1}_{\{X=x_i\}} \right] \\
&= \mathbb{E} [\mathbf{1}_C X].
\end{aligned}$$

Conditional expectation of a continuous random variable given another continuous random variable

14. Suppose that X and Z are continuous random variables with joint probability density function f_{XZ} , so that

$$f_Z(z) = \int_{-\infty}^{\infty} f_{XZ}(x, z) dx$$

is the probability density function of Z , and assume that

$$\mathbb{E}[|X|] = \int_{-\infty}^{\infty} |x| f_X(x) dx = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x| f_{XZ}(x, z) dx dz < \infty.$$

We define the *conditional probability density function* of X given Z by

$$f_{X|Z}(x|z) = \begin{cases} f_{XZ}(x, z)/f_Z(z), & \text{if } f_Z(z) \neq 0, \\ 0, & \text{if } f_Z(z) = 0. \end{cases}$$

The random variable Y

$$Y = \int_{-\infty}^{\infty} x f_{X|Z}(x|Z) dx$$

is the conditional expectation $\mathbb{E}[X | \sigma(Z)] \equiv \mathbb{E}[X | Z]$ of X given the σ -algebra $\sigma(Z)$ and has the following properties:

- (i) Y is $\sigma(Z)$ -measurable.
- (ii) $\mathbb{E}[|Y|] < \infty$. Indeed, we note that $x \mapsto |x|$ is a convex function and we use Jensen's inequality to calculate

$$\begin{aligned} \mathbb{E}[|Y|] &= \int_{-\infty}^{\infty} \left| \int_{-\infty}^{\infty} x f_{X|Z}(x|z) dx \right| f_Z(z) dz \\ &\leq \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} |x| f_{X|Z}(x|z) dx \right) f_Z(z) dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x| f_{X|Z}(x|z) f_Z(z) dx dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |x| f_{XZ}(x, z) dx dz \\ &= \mathbb{E}[|X|] < \infty. \end{aligned}$$

- (iii) $\mathbb{E}[\mathbf{1}_C Y] = \mathbb{E}[\mathbf{1}_C X]$ for all $C \in \sigma(Z)$. To see this claim, we first note, given any event $C \in \sigma(Z)$, there exists $A \in \mathcal{B}(\mathbb{R})$ such that

$$C = \{\omega \in \Omega \mid Z(\omega) \in A\}.$$

In view of this observation, we calculate

$$\begin{aligned}
\mathbb{E}[\mathbf{1}_C Y] &= \mathbb{E}\left[\mathbf{1}_{\{Z \in A\}} \left(\int_{-\infty}^{\infty} x f_{X|Z}(x|Z) dx\right)\right] \\
&= \int_{-\infty}^{\infty} \mathbf{1}_{\{z \in A\}} \left(\int_{-\infty}^{\infty} x f_{X|Z}(x|z) dx\right) f_Z(z) dz \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{1}_{\{z \in A\}} x f_{X|Z}(x|z) f_Z(z) dx dz \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathbf{1}_{\{z \in A\}} x f_{XZ}(x, z) dx dz \\
&= \mathbb{E}[\mathbf{1}_{\{Z \in A\}} X] \\
&= \mathbb{E}[\mathbf{1}_C X].
\end{aligned}$$

Definitions and existence

15. *Definition.* Consider a random variable X such that $\mathbb{E}[|X|] < \infty$, and let $\mathcal{G} \subseteq \mathcal{F}$ be a σ -algebra on Ω . The conditional expectation $\mathbb{E}[X \mid \mathcal{G}]$ of the random variable X given the σ -algebra \mathcal{G} is any random variable Y such that

- (i) Y is \mathcal{G} -measurable,
- (ii) $\mathbb{E}[|Y|] < \infty$, and
- (iii) for every event $C \in \mathcal{G}$,

$$\mathbb{E}[\mathbf{1}_C Y] = \mathbb{E}[\mathbf{1}_C X].$$

We say that a random variable Y with the properties (i)–(iii) is *a version of the conditional expectation* $\mathbb{E}[X \mid \mathcal{G}]$ of X given \mathcal{G} , and we write $Y = \mathbb{E}[X \mid \mathcal{G}]$, \mathbb{P} -a.s..

16. *Theorem.* Consider a random variable X such that $\mathbb{E}[|X|] < \infty$, and let $\mathcal{G} \subseteq \mathcal{F}$ be a σ -algebra. There exists a random variable Y having properties (i)–(iii) in Definition 15.

Furthermore, Y is unique in the sense that, if \tilde{Y} is another random variables satisfying the required properties, then $\tilde{Y} = Y$, \mathbb{P} -a.s..

17. *Definition.* Consider an event $B \in \mathcal{F}$, and let $\mathcal{G} \subseteq \mathcal{F}$ be a σ -algebra. The conditional probability of B given \mathcal{G} is the *random variable* defined by

$$\mathbb{P}(B \mid \mathcal{G}) = \mathbb{E}[\mathbf{1}_B \mid \mathcal{G}].$$

Properties of conditional expectation

18. *Theorem.* In the following list of properties of conditional expectation, we assume that all random variables are in \mathcal{L}^1 , and that $\mathcal{G}, \mathcal{H} \subseteq \mathcal{F}$ are σ -algebras on Ω .

- (i) $\mathbb{E}[X \mid \{\Omega, \emptyset\}] = \mathbb{E}[X]$.
- (ii) If Y is a version of $\mathbb{E}[X \mid \mathcal{G}]$, then $\mathbb{E}[Y] = \mathbb{E}[X]$.
- (iii) If X is \mathcal{G} -measurable, then $\mathbb{E}[X \mid \mathcal{G}] = X$, \mathbb{P} -a.s..
- (iv) (*Linearity*) Given constants $a_1, a_2 \in \mathbb{R}$, and random variables X_1, X_2 ,

$$\mathbb{E}[a_1 X_1 + a_2 X_2 \mid \mathcal{G}] = a_1 \mathbb{E}[X_1 \mid \mathcal{G}] + a_2 \mathbb{E}[X_2 \mid \mathcal{G}], \quad \mathbb{P}\text{-a.s..}$$

Precisely, if Y_1 is a version of $\mathbb{E}[X_1 \mid \mathcal{G}]$ and Y_2 is a version of $\mathbb{E}[X_2 \mid \mathcal{G}]$, then $a_1 Y_1 + a_2 Y_2$ is a version of $\mathbb{E}[a_1 X_1 + a_2 X_2 \mid \mathcal{G}]$.

- (v) (*Conditional monotone convergence theorem*) If (X_n) is an increasing sequence of positive random variables (i.e., $0 \leq X_1 \leq X_2 \leq \dots \leq X_n \leq \dots$) converging to the random variable X , \mathbb{P} -a.s., then

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}] = \mathbb{E}[X | \mathcal{G}], \quad \mathbb{P}\text{-a.s.}$$

- (vi) (*Conditional dominated convergence theorem*) If (X_n) is a sequence of random variables that converges to a random variable X , \mathbb{P} -a.s., and is such that $|X_n| \leq Z$, \mathbb{P} -a.s., for all $n \geq 1$, for some random variable Z , then

$$\lim_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}] = \mathbb{E}[X | \mathcal{G}], \quad \mathbb{P}\text{-a.s.}$$

- (vii) (*Conditional Fatou's lemma*) If (X_n) is a sequence of random variables such that $X_n \geq Z$, \mathbb{P} -a.s., for all $n \geq 1$, for some random variable Z , then

$$\mathbb{E} \left[\liminf_{n \rightarrow \infty} X_n | \mathcal{G} \right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}], \quad \mathbb{P}\text{-a.s.}$$

Similarly, if (X_n) is a sequence of random variables such that $X_n \leq Z$, \mathbb{P} -a.s., for all $n \geq 1$, for some random variable Z , then

$$\mathbb{E} \left[\limsup_{n \rightarrow \infty} X_n | \mathcal{G} \right] \geq \limsup_{n \rightarrow \infty} \mathbb{E}[X_n | \mathcal{G}], \quad \mathbb{P}\text{-a.s.}$$

- (viii) (*Conditional Jensen's inequality*) Given a random variable X and a convex function $g : \mathbb{R} \rightarrow \mathbb{R}$,

$$\mathbb{E}[g(X) | \mathcal{G}] \geq g(\mathbb{E}[X | \mathcal{G}]), \quad \mathbb{P}\text{-a.s.}$$

- (ix) (*Tower property*) If $\mathcal{H} \subseteq \mathcal{G}$, then

$$\mathbb{E}[\mathbb{E}[X | \mathcal{G}] | \mathcal{H}] = \mathbb{E}[X | \mathcal{H}], \quad \mathbb{P}\text{-a.s.}$$

- (x) (*"Taking out what is known"*) If Z is \mathcal{G} -measurable, then

$$\mathbb{E}[ZX | \mathcal{G}] = Z\mathbb{E}[X | \mathcal{G}], \quad \mathbb{P}\text{-a.s.}$$

- (xi) (*Independence*) If \mathcal{H} is independent of $\sigma(\sigma(X) \cup \mathcal{G})$, then

$$\mathbb{E}[X | \sigma(\mathcal{G} \cup \mathcal{H})] = \mathbb{E}[X | \mathcal{G}], \quad \mathbb{P}\text{-a.s.}$$

In particular, if $\sigma(X)$ and \mathcal{H} are independent σ -algebras,

$$\mathbb{E}[X | \mathcal{H}] = \mathbb{E}[X], \quad \mathbb{P}\text{-a.s.}$$