Chapter 3: Conditional Expectation

Recall that the function $X : \Omega \to \mathbb{R}^*$ is called a random variable if $X^{-1}(B) \in \mathcal{F}$ for every $B \in \mathcal{B}$, where $\mathcal{B}$ is the Borel σ-field. Note that to define a *r.v.* one only requires Ω and $\mathcal{F}$. The pair $(\Omega, \mathcal{F})$ is called a measurable space (A measurable space is simply the pair (set, σ-field)). Similarly, $(\mathbb{R}^*, \mathcal{B})$ is also a measurable space.

Definition 1. *Given two measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$, we say a function $f : \Omega_1 \to \Omega_2$ is $\mathcal{F}_1/\mathcal{F}_2$-measurable if $f^{-1}(E) \in \mathcal{F}_1$, $\forall E \in \mathcal{F}_2$.*

From the above definition, one can indeed say that a *r.v.* X defined on the probability space $(\Omega, \mathcal{F}, P)$ is a $\mathcal{F}/\mathcal{B}$-measurable function from Ω to $\mathbb{R}^*$. Since the co-domain of X is always $\mathbb{R}^*$ and the Borel σ-field $\mathcal{B}$ is the de-facto σ-field, we drop $\mathcal{B}$ from the definition and say that X is $\mathcal{F}$-measurable. In this chapter, when we say X is a *r.v.*, then it implicitly means that X is an $\mathcal{F}$-measurable *r.v.*

Definition 2. *A function $s : \Omega \to \mathbb{R}_+$ is called a non-negative, $\underline{\mathcal{G}\text{-measurable}}$ simple function with non-infinity coefficients if it has the form*

$$s = \sum_{i=1}^{n} a_i I_{A_i}, \text{ where } A_i \in \mathcal{G}, 0 \le a_i < \infty, 1 \le i \le n \text{ and } A_i \cap A_j = \emptyset,\ i \ne j.$$

The above definition is similar to the definition of simple functions from Chapter 2. However, the difference is that the individual sets A_i belong to the σ-field $\mathcal{G}$ in the above definition.

Definition 3. *Given a probability space $(\Omega, \mathcal{F}, P)$, we say $\mathcal{G}$ is a sub σ-field if $\mathcal{G}$ itself is a σ-field and $\mathcal{G} \subseteq \mathcal{F}$.*

A few observations:

Lemma 1. *Given a r.v. X (which implies that X is $\mathcal{F}$-measurable) and let $\mathcal{G}_1$ and $\mathcal{G}_2$ be two sub σ-fields with $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$. If X is $\mathcal{G}_1$-measurable, then X is also $\mathcal{G}_2$-measurable.*

Proof. Since X is $\mathcal{G}_1$-measurable, we have $X^{-1}(E) \in \mathcal{G}_1$, $\forall E \in \mathcal{B}$. Since $\mathcal{G}_1 \subseteq \mathcal{G}_2$, we have $X^{-1}(E) \in \mathcal{G}_2$, $\forall E \in \mathcal{B}$. Therefore, X is $\mathcal{G}_2$-measurable. □

The converse of the above statement is not true. For example, let $A, B \subsetneq \Omega$, $A \ne B \ne \emptyset$. Consider $\mathcal{G}_2 = \{A, B, A^c, B^c, A \cap B, A \cup B, A^c \cap B, A^c \cup B, A^c \cap B^c, A^c \cup B^c, A \cap B^c, A \cup B^c, \Omega, \emptyset\}$ and $\mathcal{G}_1 = \{B, B^c, \Omega, \emptyset\}$. Now note that I_A is $\mathcal{G}_2$-measurable, but not $\mathcal{G}_1$-measurable.

Lemma 2. *Let X and Y be two r.v.s which are integrable. They are $\mathcal{F}$-measurable by definition. Now let $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ-field. Then*

1. *If X and Y are $\mathcal{G}$-measurable and $\mathbb{E}[X\mathbb{I}_A] = \mathbb{E}[Y\mathbb{I}_A], \forall A \in \mathcal{G}$, then $X = Y$ a.s. The proof is similar to Theorem 13 of Chapter 2.*
2. *If X and Y are $\mathcal{G}$-measurable and $\mathbb{E}[X\mathbb{I}_A] \geq \mathbb{E}[Y\mathbb{I}_A], \forall A \in \mathcal{G}$, then $X \geq Y$ a.s. The proof is similar to Theorem 14 of Chapter 2.*
3. *Also, $X + Y$ and cY (where $c \in \mathbb{R}$) are $\mathcal{G}$-measurable and integrable r.v.s. The proof is similar to the proofs of Proposition 3 and Proposition 6 of Chapter 2.*
4. *If X is a non-negative r.v. which is $\mathcal{G}$-measurable, then there exists a sequence (s_n), s.t. $s_n \uparrow X$, where s_n are non-negative, $\underline{\mathcal{G}\text{-measurable}}$ simple functions with non-infinity coefficients. This means that for each $\omega \in \Omega$, we have $(s_n(\omega))$ is a monotonically increasing sequence and $\lim\limits_{n\to\infty} s_n(\omega) = X(\omega)$.*
5. *If X is a $\mathcal{G}$-measurable r.v. (not necessarily non-negative), then X can be decomposed as follows:*

$$X = X^+ - X^-,$$

where

$$X^+(\omega) = \begin{cases} X(\omega), & \text{if } X(\omega) \geq 0, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

and

$$X^-(\omega) = \begin{cases} -X(\omega), & \text{if } X(\omega) < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

We have similar decomposition in Section 4.3 of Chapter 2. The difference here is that X^+ and X^- in the above case are $\mathcal{G}$-measurable while those in Section 4.3 of Chapter 2 are $\mathcal{F}$-measurable.

6. *If X is a r.v. and $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous function, then $\phi(X)$ is a r.v. Specifically, if X is a $\mathcal{G}$-measurable r.v. and $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous function, then $\phi(X)$ is a $\mathcal{G}$-measurable r.v.*
7. *If X is a r.v. and $\phi : \mathbb{R} \to \mathbb{R}$ is a monotonically non-decreasing function (or a monotonically non-increasing function), then $\phi(X)$ is a r.v. Specifically, if X is a $\mathcal{G}$-measurable r.v. and $\phi : \mathbb{R} \to \mathbb{R}$ is a monotonically non-decreasing function (or a monotonically non-increasing function), then $\phi(X)$ is a $\mathcal{G}$-measurable r.v.*

Notation: We use the following notation in this chapter: For $A \in \mathcal{B}$, we let

$$\int_A X dP = \mathbb{E}[XI_A] \tag{3}$$

Therefore, with this notation, we have $\mathbb{E}[X] = \int X dP$.

Definition 4. *Given a sub σ-field $\mathcal{G} \subseteq \mathcal{F}$ and a r.v. X ($\mathcal{F}$-measurable by definition) which is integrable (i.e., $\mathbb{E}[|X|] < \infty$), then there exists a r.v. $\mathbb{E}[X|\mathcal{G}]$ which satisfies the following:*

1. *$\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable and integrable.*
2. $$\int_A X dP = \int_A \mathbb{E}[X|\mathcal{G}] \, dP, \qquad \forall A \in \mathcal{G}. \tag{4}$$

$\mathbb{E}[X|\mathcal{G}]$ is called the conditional expected value of X given $\mathcal{G}$.

For an integrable *r.v.* X, the proof of the existence of $\mathbb{E}[X|\mathcal{G}]$ is given in Page 445, [BPM][1]. A few important remarks about the above definition:

1. $\mathbb{E}[X|\mathcal{G}]$ is a $\mathcal{G}$-measurable *r.v.* This means that $\mathbb{E}[X|\mathcal{G}]$ is a function from Ω to $\mathbb{R}^*$. Further, $\mathcal{G}$-measurable implies that for every $E \in \mathcal{B}$, we have $\mathbb{E}[X|\mathcal{G}]^{-1}(E) \in \mathcal{G}$.
2. $\mathbb{E}[X|\mathcal{G}]$ is integrable, *i.e.*, $\int |\mathbb{E}[X|\mathcal{G}]| dP < \infty$.
3. $\mathbb{E}[X|\mathcal{G}]$ is only defined if X is integrable.
4. Let Y be any other integrable, $\mathcal{G}$-measurable *r.v.* which satisfies Definition 4, then Lemma 2(1) ensures that $Y = \mathbb{E}[X|\mathcal{G}]$ *a.s.* This implies that there might exist different $\mathcal{G}$-measurable *r.v.*s which satisfy Definition 4, however, they all are the same in the almost sure sense.

It is very hard in general to find $\mathbb{E}[X|\mathcal{G}]$. But in certain scenarios, one can deduce it. We illustrate those scenarios in the following theorem. The only tools we need to prove these claims are Lemmas 1 and 2.

Theorem 1. *Let X and Y be a r.v.s and are integrable. Then*

1. *If X is $\mathcal{G}$-measurable, then $\mathbb{E}[X|\mathcal{G}] = X$ a.s.*
2. *$\mathbb{E}[X+Y|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$ a.s. (Note that $\mathbb{E}[X+Y|\mathcal{G}]$ is well defined since $X+Y$ is integrable (follows from Proposition 6 of Chapter 2)).*
3. *For $c \in \mathbb{R}$, we have $\mathbb{E}[cX|\mathcal{G}] = c\mathbb{E}[X|\mathcal{G}]$ a.s. (Note that $\mathbb{E}[cX|\mathcal{G}]$ is well defined since cX is integrable (follows from Proposition 6 of Chapter 2)).*
4. *If $X \geq Y$ a.s., then $\mathbb{E}[X|\mathcal{G}] \geq \mathbb{E}[Y|\mathcal{G}]$ a.s.*
5. *$\big|\mathbb{E}[X|\mathcal{G}]\big| \leq \mathbb{E}\big[|X| \, \big|\mathcal{G}\big]$ a.s.*
6. *$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$.*

Proof. 1. Since X is $\mathcal{G}$-measurable, X satisfies condition 1 of Definition 4. Also, X satisfies condition 2 of Definition 4 trivially. Therefore, $\mathbb{E}[X|\mathcal{G}] = X$ *a.s.*

[1] Patrick Billingsley, Probability and Measure

2. Since $\mathbb{E}[X|\mathcal{G}]$ and $\mathbb{E}[Y|\mathcal{G}]$ are $\mathcal{G}$-measurable and integrable (by definition of conditional expectation), we have $\mathbb{E}[X|\mathcal{G}]+\mathbb{E}[Y|\mathcal{G}]$ is $\mathcal{G}$-measurable and integrable (follows from the Lemma 2(3)). Now, for $A \in \mathcal{G}$, we have

$$\begin{aligned}\int_A \mathbb{E}[X+Y|\mathcal{G}]\, dP &= \int_A (X+Y)\, dP = \int_A X dP + \int_A Y dP \\ &= \int_A \mathbb{E}[X|\mathcal{G}]\, dP + \int_A \mathbb{E}[Y|\mathcal{G}]\, dP \\ &= \int_A (\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}])\, dP.\end{aligned}$$

Now since the above equality is true for every $A \in \mathcal{G}$ and $\mathbb{E}[X+Y|\mathcal{G}]$ and $\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$ are $\mathcal{G}$-measurable $r.v.$s, the claim follows from Lemma 2(1).

3. Since $\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable and integrable (by definition of conditional expectation), we have $c\mathbb{E}[X|\mathcal{G}]$ is $\mathcal{G}$-measurable and integrable (follows from the Lemma 2(3)). Now, for $A \in \mathcal{G}$, we have

$$\begin{aligned}\int_A \mathbb{E}[cX|\mathcal{G}]\, dP = \int_A cX dP = c\int_A X dP &= c\int_A \mathbb{E}[X|\mathcal{G}]\, dP \\ &= \int_A c\mathbb{E}[X|\mathcal{G}])dP.\end{aligned}$$

Now since the above equality is true for every $A \in \mathcal{G}$ and $\mathbb{E}[cX|\mathcal{G}]$ and $c\mathbb{E}[X|\mathcal{G}]$ are $\mathcal{G}$-measurable $r.v.$s, the claim follows from Lemma 2(1).

4. For $A \in \mathcal{G}$, we have

$$\begin{aligned}\int_A \mathbb{E}[X|\mathcal{G}]\, dP &= \int_A X dP \text{ (follows from the def of cond exp)} \\ &\geq \int_A Y dP \text{ (follows since } X \geq Y \ a.s. \text{ and from Prop 6(3))} \\ &= \int_A \mathbb{E}[Y|\mathcal{G}]\, dP \text{ (follows from the def of cond exp).}\end{aligned}$$

Now since the above equality is true for every $A \in \mathcal{G}$ and $\mathbb{E}[X|\mathcal{G}]$ and $\mathbb{E}[Y|\mathcal{G}]$ are $\mathcal{G}$-measurable $r.v.$s, the claim follows from Lemma 2(2).

5. Note that $X \leq |X|$. Therefore, from part (4) of the theorem, we have

$$\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}\big[|X|\ \big|\mathcal{G}\big] \ \ a.s. \tag{5}$$

Also, $-X \leq |X|$. Therefore, again from part (4) of the theorem, we have

$$\begin{aligned}&\mathbb{E}\big[|X|\ \big|\mathcal{G}\big] \geq \mathbb{E}[-X|\mathcal{G}] = -\mathbb{E}[X|\mathcal{G}]. \\ &\Rightarrow \mathbb{E}[X|\mathcal{G}] \geq -\mathbb{E}\big[|X|\ \big|\mathcal{G}\big] \ \ a.s. \end{aligned} \tag{6}$$

Hence, from Theorem 8(4) of Chapter 1 and Eqs. (5-6), we have

$$\big|\mathbb{E}[X|\mathcal{G}]\,\big| \leq \mathbb{E}\big[|X|\ \big|\mathcal{G}\big] \ \ a.s.$$

6. Indeed,

$$\begin{aligned}\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}\right]\right] &= \int_\Omega \mathbb{E}\left[X|\mathcal{G}\right] dP \text{ (by notation)} \\ &= \int_\Omega X dP \text{ (by the def of conditional expectation)} \\ &= \mathbb{E}\left[X\right] \text{ (by notation).}\end{aligned}$$

□

Theorem 2. *Let (X_n) be a sequence of r.v.s. and there exists an integrable r.v. Y s.t. $|X_n| \leq Y$ a.s. for every n. If $\lim_{n\to\infty} X_n = X^*$ a.s., then $\lim_{n\to\infty} \mathbb{E}\left[X_n|\mathcal{G}\right] = \mathbb{E}\left[X^*|\mathcal{G}\right]$ a.s.*

Proof. First, we will verify that all the terms involved in the theorem statement are well-defined. Note the since $|X_n| \leq Y$ *a.s.*, we have $\mathbb{E}\left[|X_n|\right] \leq \mathbb{E}\left[Y\right]$ which implies that X_n is integrable. Further, since $\lim_{n\to\infty} X_n = X^*$ *a.s.* we have $|X^*| \leq Y$ *a.s.* Therefore, by the same argument as above, we have $\mathbb{E}\left[|X^*|\right] < \infty$, *i.e.*, X^* is integrable. Thus $\mathbb{E}\left[X_n|\mathcal{G}\right]$ and $\mathbb{E}\left[X^*|\mathcal{G}\right]$ are well-defined.

For $n \in \mathbb{N}$, let's define,

$$Z_n = \sup_{k\geq n} |X_k - X^*|. \tag{7}$$

Note that Z_n is a *r.v.* (follows from Proposition 7 of Chapter 2). The intuition behind defining Z_n is as follows: Note that the goal of the proof is to show that $\lim_{n\to\infty} \left|\mathbb{E}\left[X_n|\mathcal{G}\right] - \mathbb{E}\left[X^*|\mathcal{G}\right]\right| = 0$ *a.s.* To achieve that we consider the following:

$$\begin{aligned}0 \leq \left|\mathbb{E}\left[X_n|\mathcal{G}\right] - \mathbb{E}\left[X^*|\mathcal{G}\right]\right| &\leq \left|\mathbb{E}\left[X_n - X^*|\mathcal{G}\right]\right| \text{ (follows from Theorem 1(2))} \\ &\leq \mathbb{E}\left[|X_n - X^*|\big|\mathcal{G}\right] \text{ (follows from Theorem 1(5))} \\ &\leq \mathbb{E}\left[Z_n|\mathcal{G}\right]. \end{aligned} \tag{8}$$

Now by applying squeeze theorem to the above inequality (assuming that the limit on the right side of the inequality exists), we get

$$0 \leq \lim_{n\to\infty} \left|\mathbb{E}\left[X_n|\mathcal{G}\right] - \mathbb{E}\left[X^*|\mathcal{G}\right]\right| \leq \lim_{n\to\infty} \mathbb{E}\left[Z_n|\mathcal{G}\right]. \tag{9}$$

Goal: If we could show that $\lim_{n\to\infty} \mathbb{E}\left[Z_n|\mathcal{G}\right] = 0$ *a.s.*, then we are done.

• First we will analyze the properties of the sequence (Z_n). It is easy to verify that

$$Z_n \geq 0 \text{ } a.s. \text{ and } Z_{n+1} \leq Z_n \text{ } a.s., \quad \forall n \in \mathbb{N}. \tag{10}$$

Therefore (Z_n) is a monotonically decreasing sequence of non-negative *r.v.*s. and $\lim_{n\to\infty} Z_n = 0$ *a.s.* (this follows since $\lim_{n\to\infty} X_n = X^*$). Also, for $n \in \mathbb{N}$, we have

$$|Z_n| = \sup_{k \geq n} |X_k - X^*| \leq \sup_{k \geq n} |X_k| + |X^*| \leq 2|Y|. \tag{11}$$

This further implies that

$$\mathbb{E}\left[|Z_n|\right] \leq 2\mathbb{E}\left[|Y|\right] < \infty. \tag{12}$$

Hence, Z_n is integrable, $\forall n$. Therefore,

$$\lim_{n\to\infty} \mathbb{E}\left[Z_n\right] = 0 \; a.s. \text{ (follows from dominated convergence theorem).} \tag{13}$$

• Now we will analyze the properties of the sequence $(\mathbb{E}\left[Z_n|\mathcal{G}\right])$. From Eq. (12), it follows that Z_n is integrable and therefore, $\mathbb{E}\left[Z_n|\mathcal{G}\right]$ exists.

Note that from Eq. (10) and Theorem 1(4), we have

$$0 \leq \mathbb{E}\left[Z_{n+1}|\mathcal{G}\right] \leq \mathbb{E}\left[Z_n|\mathcal{G}\right] \; a.s., \forall n. \tag{14}$$

Therefore $(\mathbb{E}\left[Z_n|\mathcal{G}\right])$ is a monotonically decreasing sequence of non-negative $r.v.$s. and therefore $\lim_{n\to\infty} \mathbb{E}\left[Z_n|\mathcal{G}\right]$ exists almost surely. Let

$$Z^* = \lim_{n\to\infty} \mathbb{E}\left[Z_n|\mathcal{G}\right]. \tag{15}$$

Since $\mathbb{E}\left[Z_n|\mathcal{G}\right] \geq 0 \; a.s.$, we have $Z^* \geq 0 \; a.s.$ Now we will apply dominated convergence theorem (DCT) on the sequence $(\mathbb{E}\left[Z_n|\mathcal{G}\right])$. To apply that we need to satisfy the hypothesis of DCT, *i.e.*, the sequence is bounded by an integrable $r.v.$ In our case, we have from Eq. (12) and Theorem 1(4),

$$\mathbb{E}\Big[|Z_n| \; \Big|\mathcal{G}\Big] \leq 2\mathbb{E}\Big[|Y| \; \Big|\mathcal{G}\Big]. \tag{16}$$

Also, $2\mathbb{E}\Big[|Y| \; \Big|\mathcal{G}\Big]$ is integrable (by the definition of condition expectation). Therefore, $2\mathbb{E}\Big[|Y| \; \Big|\mathcal{G}\Big]$ is the required bound required for the hypothesis of DCT. Now by appealing to DCT, we obtain the following:

$$\begin{aligned}
\mathbb{E}\left[Z^*\right] &= \int Z^* dP \text{ (by notation)} \\
&= \lim_{n\to\infty} \int \mathbb{E}\left[Z_n|\mathcal{G}\right] dP \text{ (by DCT on the seq. } (\mathbb{E}\left[Z_n|\mathcal{G}\right])) \\
&= \lim_{n\to\infty} \int Z_n dP \text{ (by def. of conditional expectation)} \\
&= \lim_{n\to\infty} \mathbb{E}\left[Z_n\right] \text{ (by notation)} \\
&= 0 \text{ (from Eq. (13)).}
\end{aligned}$$

This implies that $Z^* = 0 \; a.s.$ (follows from the fact that $Z^* \geq 0$ and Theorem 12 in Chapter 2 (take $A = \Omega$ in Theorem 12 of Chapter 2)). Therefore,

$$\lim_{n\to\infty} \mathbb{E}\left[Z_n|\mathcal{G}\right] = Z^* = 0 \; a.s.$$

Thus the claim follows (see goal). □

Theorem 3. *Let $\mathcal{G}_1, \mathcal{G}_2$ be sub σ-fields with $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$. Let X be an integrable r.v.. Then*

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}_1\right]|\mathcal{G}_2\right] = \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}_2\right]|\mathcal{G}_1\right] = \mathbb{E}\left[X|\mathcal{G}_2\right]. \tag{17}$$

Proof. We will first prove that $\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}_1\right]|\mathcal{G}_2\right] = \mathbb{E}\left[X|\mathcal{G}_2\right]$. For that note that $\mathbb{E}\left[X|\mathcal{G}_1\right]$ is a $\mathcal{G}_1$-measurable *r.v.* (follows from the definition of conditional expectation). Since $\mathcal{G}_1 \subseteq \mathcal{G}_2$, we have $\mathbb{E}\left[X|\mathcal{G}_1\right]$ is also $\mathcal{G}_2$-measurable (follows from Lemma 1). Therefore,

$$\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}_1\right]|\mathcal{G}_2\right] = \mathbb{E}\left[X|\mathcal{G}_2\right]. \tag{18}$$

The above claim follows from Theorem 1(1).

Now we will prove the second part that $\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}_2\right]|\mathcal{G}_1\right] = \mathbb{E}\left[X|\mathcal{G}_2\right]$. For that we make use of Lemma 2(1). Now consider, for $A \in \mathcal{G}_1$, we have

$$\begin{aligned}
\int_A \mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}_2\right]|\mathcal{G}_1\right] dP &= \int_A \mathbb{E}\left[X|\mathcal{G}_2\right] dP \text{ (follows from the def. of } \mathbb{E}\left[\cdot|\mathcal{G}_1\right]) \\
&= \int_A X dP \text{ (follows since } A \in \mathcal{G}_1 \subseteq \mathcal{G}_2 \text{ and def. of } \mathbb{E}\left[\cdot|\mathcal{G}_2\right]) \\
&= \int_A \mathbb{E}\left[X|\mathcal{G}_1\right] dP \text{ (follows since } A \in \mathcal{G}_1 \text{ and def. of } \mathbb{E}\left[\cdot|\mathcal{G}_1\right])
\end{aligned}$$

Since $A \in \mathcal{G}_1$ is chosen arbitrarily, the above equality holds $\forall A \in \mathcal{G}_1$. Also, $\mathbb{E}\left[\mathbb{E}\left[X|\mathcal{G}_2\right]|\mathcal{G}_1\right]$ and $\mathbb{E}\left[X|\mathcal{G}_1\right]$ are $\mathcal{G}_1$-measurable *r.v.*s. Therefore, the claim follows from Lemma 2(1). $\square$

Theorem 4. *Let X, Y be r.v.s. with X and XY are integrable. Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub σ-field. Then*

$$\mathbb{E}\left[XY|\mathcal{G}\right] = Y\mathbb{E}\left[X|\mathcal{G}\right] \ a.s.$$

Proof. We consider different types of $\mathcal{G}$-measurable Y here:

Case 1: Y is a $\mathcal{G}$-measurable indicator function, *i.e.*, $Y = I_B$, where $B \in \mathcal{G}$:

Now, for $A \in \mathcal{G}$, we have $A \cap B \in \mathcal{G}$ (since $A, B \in \mathcal{G}$) and

$$\begin{aligned}
\int_A \mathbb{E}\left[XY|\mathcal{G}\right] dP &= \int_A \mathbb{E}\left[XI_B|\mathcal{G}\right] dP \\
&= \int_A XI_B dP \text{ (follows from the def. of conditional expectation)} \\
&= \int_{A\cap B} X dP \\
&= \int_{A\cap B} \mathbb{E}\left[X|\mathcal{G}\right] dP \text{ (since } A \cap B \in \mathcal{G} \text{ and the def. of cond. exp.)} \\
&= \int_A I_B \mathbb{E}\left[X|\mathcal{G}\right] dP = \int_A Y\mathbb{E}\left[X|\mathcal{G}\right] dP.
\end{aligned}$$

Since $A \in \mathcal{G}$ is chosen arbitrarily, the above equality follows $\forall A \in \mathcal{G}$. Now by appealing to Lemma 2(1), we have

$$\mathbb{E}\left[XY|\mathcal{G}\right] = Y\mathbb{E}\left[X|\mathcal{G}\right] \;\; a.s. \tag{19}$$

Case 2: Y is a $\mathcal{G}$-measurable non-negative simple function with non-infinity coefficients, *i.e.*, $Y = \sum_{i=1}^{n} b_i I_{B_i}$, where $b_i \in \mathbb{R}$ and $B_i \in \mathcal{G}$:

Now for $A \in \mathcal{G}$, we have $A \cap B_i \in \mathcal{G}$, $1 \leq i \leq n$ (since $A, B_i \in \mathcal{G}$) and

$$\begin{aligned}
\int_A \mathbb{E}\left[XY|\mathcal{G}\right] dP &= \int_A \mathbb{E}\left[X \sum_{i=1}^{n} b_i I_{B_i} \middle| \mathcal{G}\right] dP = \int_A b_i \sum_{i=1}^{n} \mathbb{E}\left[XI_{B_i}|\mathcal{G}\right] dP \\
&= \sum_{i=1}^{n} b_i \int_A \mathbb{E}\left[XI_{B_i}|\mathcal{G}\right] dP \text{ (follows from Theorem 1(2))} \\
&= \sum_{i=1}^{n} b_i \int_A XI_{B_i} dP = \sum_{i=1}^{n} b_i \int_{A\cap B_i} X dP = \sum_{i=1}^{n} b_i \int_{A\cap B_i} \mathbb{E}\left[X|\mathcal{G}\right] dP \\
&= \sum_{i=1}^{n} b_i \int_A I_{B_i} \mathbb{E}\left[X|\mathcal{G}\right] dP = \int_A \sum_{i=1}^{n} b_i I_{B_i} \mathbb{E}\left[X|\mathcal{G}\right] dP \\
&= \int_A \mathbb{E}\left[X|\mathcal{G}\right] \sum_{i=1}^{n} b_i I_{B_i} dP = \int_A Y\mathbb{E}\left[X|\mathcal{G}\right] dP.
\end{aligned}$$

Since $A \in \mathcal{G}$ is chosen arbitrarily, the above equality follows $\forall A \in \mathcal{G}$. Now by appealing to Lemma 2(1), we have

$$\mathbb{E}\left[XY|\mathcal{G}\right] = Y\mathbb{E}\left[X|\mathcal{G}\right] \;\; a.s. \tag{20}$$

Case 3: Y is a $\mathcal{G}$-measurable, non-negative $r.v.$:

We make use of Theorem 2 and Case 2 to prove this case.
Since $Y \geq 0$, there exists a (s_n), $s.t.$ $s_n \uparrow Y$, where s_n are non-negative, $\underline{\mathcal{G}\text{-measurable}}$ simple functions with non-infinity coefficients (Lemma 2(4)). Since $(s_n) \uparrow Y$, we have $\lim_{n\to\infty} Xs_n = XY$ and $|Xs_n| \leq |XY|$. Therefore, from Theorem 2, we have

$$\lim_{n\to\infty} \mathbb{E}\left[Xs_n|\mathcal{G}\right] = \mathbb{E}\left[XY|\mathcal{G}\right]. \tag{21}$$

Now from Case 2, we have

$$\mathbb{E}\left[Xs_n|\mathcal{G}\right] = s_n\mathbb{E}\left[X|\mathcal{G}\right]. \tag{22}$$

From Eqs (21-22) and the fact that $s_n \uparrow Y$, we have

$$\mathbb{E}\left[XY|\mathcal{G}\right] = Y\mathbb{E}\left[X|\mathcal{G}\right]. \tag{23}$$

Case 4: Y is a $\mathcal{G}$-measurable $r.v.$ (not necessarily non-negative):

Since Y is a $\mathcal{G}$-measurable $r.v.$, we have the following decomposition for Y (Lemma 2(5)):

$$Y = Y^+ - Y^- \text{ and } XY = XY^+ - XY^- \ a.s. \tag{24}$$

Note that Y^+ and Y^- are non-negative, $\mathcal{G}$-measurable $r.v.$s and $|XY^+| \leq |XY|$ and $|XY^-| \leq |XY|$ which implies that XY^+ and XY^- are integrable. Hence, by Case 3, we have

$$\mathbb{E}\left[XY^+|\mathcal{G}\right] = Y^+\mathbb{E}\left[X|\mathcal{G}\right] \text{ and } \mathbb{E}\left[XY^-|\mathcal{G}\right] = Y^-\mathbb{E}\left[X|\mathcal{G}\right]. \tag{25}$$

Therefore,

$$\begin{aligned}
\mathbb{E}\left[XY|\mathcal{G}\right] &= \mathbb{E}\left[XY^+|\mathcal{G}\right] - \mathbb{E}\left[XY^-|\mathcal{G}\right] \\
&= Y^+\mathbb{E}\left[X|\mathcal{G}\right] - Y^-\mathbb{E}\left[X|\mathcal{G}\right] \\
&= (Y^+ - Y^-)\mathbb{E}\left[X|\mathcal{G}\right] \\
&= Y\mathbb{E}\left[X|\mathcal{G}\right].
\end{aligned}$$

□

Lemma 3. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a convex function. Then ϕ is continuous and right hand derivative of ϕ exists for every $x \in \mathbb{R}$. Further, the right hand derivative of ϕ (denoted as ϕ'_+) is a monotonically non-decreasing function of x. Also, ϕ satisfies*

$$\phi(x) \geq \phi(\bar{x}) + \phi'_+(\bar{x})(x - \bar{x}) \tag{26}$$

Lemma 4. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be a continuous function (or) a monotonically non-decreasing function. Then ϕ takes bounded sets to bounded sets,* i.e., *if $A \subset \mathbb{R}$ is bounded (there exists $0 < K < \infty$ s.t. $|x| \leq K$, $\forall x \in A$), then $\phi(A)$ is also bounded, where $\phi(A) = \{\phi(x) : x \in A\}$,* i.e., *there exists a $K' > 0$ s.t. $|\phi(x)| \leq K'$, $\forall x \in A$ or simply one can write as $|\phi(A)| \leq K'$.*

Lemma 5. *If $\phi : \mathbb{R} \to \mathbb{R}$ is a continuous function and $\lim_{n\to\infty} x_n = x$, then $\lim_{n\to\infty} \phi(x_n) = \phi(x)$.*

Proof. **Case 1: $\mathbb{E}\left[X|\mathcal{G}\right]$ is bounded $a.s.$:**

$\mathbb{E}\left[X|\mathcal{G}\right]$ is bounded $a.s.$ means that there exists a $0 \leq K < \infty$ and $N \in \mathcal{G}$ $s.t.$ $|\mathbb{E}\left[X|\mathcal{G}\right](\omega)| \leq K$, $\forall\omega \in N$ and $P(N^c) = 0$ (follows from the definition of almost surely). This has a few implications:

1. Since ϕ is continuous and $[-K, K]$ is a bounded set, there exists a $0 \leq K' < \infty$ $s.t.$ $|\phi([-K, K])| \leq K'$ (follows from Lemma 4). Now since $|\mathbb{E}\left[X|\mathcal{G}\right](\omega)| \leq K$, $\forall\omega \in N$, we have $\mathbb{E}\left[X|\mathcal{G}\right](\omega) \in [-K, K]$, $\forall\omega \in N$. Therefore, $|\phi\left(\mathbb{E}\left[X|\mathcal{G}\right](\omega)\right)| \leq K'$, $\forall\omega \in N$ with $P(N^c) = 0$. This means that $|\phi\left(\mathbb{E}\left[X|\mathcal{G}\right]\right)| \leq K'$ $a.s.$ which further implies that $\mathbb{E}\left[|\phi\left(\mathbb{E}\left[X|\mathcal{G}\right]\right)|\right] \leq K'$. Thus $\phi\left(\mathbb{E}\left[X|\mathcal{G}\right]\right)$ is integrable.

2. Since ϕ'_+ is a monotonically non-decreasing function, one can similarly show that $\phi'_+(\mathbb{E}[X|\mathcal{G}])$ is integrable.

Now observe that since ϕ is a convex function, we obtain the following from Lemma 3:

$$\phi(\mathbb{E}[X|\mathcal{G}]) + \phi'_+(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G}) \leq \phi(X). \tag{27}$$

First we will state some observations:

(F1) $\phi(\mathbb{E}[X|\mathcal{G}])$ is $\mathcal{G}$-measurable (follows from Lemma 1(6)) and integrable (from above observation). Therefore, $\mathbb{E}\left[\phi(\mathbb{E}[X|\mathcal{G}])\middle|\mathcal{G}\right] = \phi(\mathbb{E}[X|\mathcal{G}])$ (follows from Theorem 1(1)).

(F2) $\phi'_+(\mathbb{E}[X|\mathcal{G}])$ is $\mathcal{G}$-measurable (follows from Lemma 1(7)) and integrable (from above observation).

Now, by applying $\mathbb{E}[\cdot|\mathcal{G}]$ on both sides of the inequality (27) (inequality still holds after applying by Theorem 1(4)), we get

$$\begin{aligned}
&\mathbb{E}\left[\phi(\mathbb{E}[X|\mathcal{G}]) + \phi'_+(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G})\,\middle|\mathcal{G}\right] \leq \mathbb{E}[\phi(X)|\mathcal{G}] \\
&\Rightarrow \underbrace{\mathbb{E}\left[\phi(\mathbb{E}[X|\mathcal{G}])\middle|\mathcal{G}\right]}_{||\ \text{(F1)}} + \underbrace{\mathbb{E}\left[\phi'_+(\mathbb{E}[X|\mathcal{G}])(X - \mathbb{E}[X|\mathcal{G})\,\middle|\mathcal{G}\right]}_{||\ \text{(F2) and Theorem 4}} \leq \mathbb{E}[\phi(X)|\mathcal{G}] \\
&\Rightarrow \quad \phi(\mathbb{E}[X|\mathcal{G}]) + \phi'_+(\mathbb{E}[X|\mathcal{G}])\underbrace{\mathbb{E}\left[(X - \mathbb{E}[X|\mathcal{G})\,\middle|\mathcal{G}\right]}_{=0} \leq \mathbb{E}[\phi(X)|\mathcal{G}] \\
&\Rightarrow \phi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\phi(X)|\mathcal{G}].
\end{aligned} \tag{28}$$

Case 2: $\mathbb{E}[X|\mathcal{G}]$ is not necessarily bounded:

Let $G_n = \{\omega \in \Omega : |\mathbb{E}[X|\mathcal{G}](\omega)| \leq n\}$. Since $\mathbb{E}[X|\mathcal{G}]$ is a $\mathcal{G}-$measurable *r.v.*, we have $G_n \in \mathcal{G}$ and therefore, I_{G_n} is $\mathcal{G}$-measurable. So,

$$\begin{aligned}
\left|\mathbb{E}[I_{G_n}X|\mathcal{G}]\right| &= \left|I_{G_n}\mathbb{E}[X|\mathcal{G}]\right| \text{ (follows from Theorem 4)} \\
&\leq n \text{ (follows from the definition of } G_n\text{)}.
\end{aligned} \tag{29}$$

Hence, $\mathbb{E}[I_{G_n}X|\mathcal{G}]$ is a bounded *r.v.* Therefore, by Case 1, we have

$$\phi(\mathbb{E}[I_{G_n}X|\mathcal{G}]) \leq \mathbb{E}[\phi(I_{G_n}X)|\mathcal{G}]. \tag{30}$$

Since $\mathbb{E}[I_{G_n}X|\mathcal{G}] = I_{G_n}\mathbb{E}[X|\mathcal{G}]$ (follows from Theorem 4), we obtain from Eq. (30) the following:

$$\begin{aligned}
\phi(I_{G_n}\mathbb{E}[X|\mathcal{G}]) &\leq \mathbb{E}[\phi(I_{G_n}X)|\mathcal{G}] \\
&= \mathbb{E}\left[I_{G_n}\phi(X) + I_{G_n^c}\phi(0)|\mathcal{G}\right] \\
&= I_{G_n}\mathbb{E}[\phi(X)|\mathcal{G}] + I_{G_n^c}\phi(0).
\end{aligned} \tag{31}$$

Now, note that since $\mathbb{E}[X|\mathcal{G}]$ is integrable, we have $\left|\mathbb{E}[X|\mathcal{G}]\right| < \infty$ *a.s.* (follows from Theorem 10 of Chapter 2). Hence, $\lim_{n\to\infty} I_{G_n}\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}]$ *a.s.* Therefore,

$$\lim_{n\to\infty} \phi\left(I_{G_n}\mathbb{E}[X|\mathcal{G}]\right) = \phi\left(\mathbb{E}[X|\mathcal{G}]\right) \ a.s. \tag{32}$$

The above equality follows from Lemma 5.

Also, since $\left|\mathbb{E}[X|\mathcal{G}]\right| < \infty$ *a.s.*, we have $\lim_{n\to\infty} I_{G_n} = I_\Omega$ *a.s.* and therefore $\lim_{n\to\infty} I_{G_n^c} = I_\emptyset$ *a.s.* ($\emptyset$ is the empty set). Hence,

$$\lim_{n\to\infty} I_{G_n}\mathbb{E}[\phi(X)|\mathcal{G}] = \mathbb{E}[\phi(X)|\mathcal{G}] \ a.s. \text{ and } \lim_{n\to\infty} I_{G_n^c}\phi(0) = 0 \ a.s. \tag{33}$$

Now we can apply $\lim_{n\to\infty}$ to both sides of the inequality of Eq. (31) since the individual limits exist (Eqs. (32-33)). Thus, by applying $\lim_{n\to\infty}$ to both sides of the inequality of Eq. (31), we obtain

$$\begin{aligned}
\phi\left(\mathbb{E}[X|\mathcal{G}]\right) = \lim_{n\to\infty} \phi\left(I_{G_n}\mathbb{E}[X|\mathcal{G}]\right) &\leq \lim_{n\to\infty} \big(I_{G_n}\mathbb{E}[\phi(X)|\mathcal{G}] + I_{G_n^c}\phi(0)\big) \\
&= \lim_{n\to\infty} I_{G_n}\mathbb{E}[\phi(X)|\mathcal{G}] + \lim_{n\to\infty} I_{G_n^c}\phi(0) \\
&= \mathbb{E}[\phi(X)|\mathcal{G}]
\end{aligned}$$

□