

Solutions

Chapter 2 Foundations of Probability

2.1 Since g is \mathcal{G}/\mathcal{H} -measurable, therefore $\forall C \in \mathcal{H}, \exists B = g^{-1}(C) \in \mathcal{G}$. Similarly, since f is \mathcal{F}/\mathcal{G} -measurable, $\forall B \in \mathcal{G}, \exists A = f^{-1}(B) \in \mathcal{F}$. Thus $\forall C \in \mathcal{H}, \exists A = f^{-1}(g^{-1}(C)) = (g \circ f)^{-1}(C) \in \mathcal{F}$ and the proof is complete.

2.2 We claim that $X = (X_1, X_2, \dots, X_n)$ is $\mathcal{F}/\mathcal{B}(\mathbb{R}^n)$ measurable. Define $a = (a_1, a_2, \dots, a_n)$ $b = (b_1, b_2, \dots, b_n)$ with $a, b \in \mathbb{R}^n$ where $a < b$. Since X_1, X_2, \dots, X_n is $\mathcal{F}/\mathcal{B}(\mathbb{R})$ measurable, therefore $\exists A_1 = X_1^{-1}((a_1, b_1)), A_2 = X_2^{-1}((a_2, b_2)), \dots, A_n = X_n^{-1}((a_n, b_n)) \in \mathcal{F}$. Let $A = A_1 \cap A_2 \cap \dots \cap A_n = \bigcap_{i=1}^n A_i$. It follows that $X^{-1}((a, b)) = \bigcap_{i=1}^n ((a, b)) = A \in \mathcal{F}$. Therefore X is $\mathcal{F}/\mathcal{B}(\mathbb{R}^n)$ measurable and X is random vector.

2.3

- (i) We need to show that Σ_X is closed under countable union. Let $U_i = X^{-1}(A_i), A_i \in \Sigma, i \in \mathbb{N}$. It follows that $\bigcup_{i=1}^{\infty} U_i = \bigcup_{i=1}^{\infty} X^{-1}(A_i) = X^{-1}(\bigcup_{i=1}^{\infty} A_i)$. Since $\bigcup_{i=1}^{\infty} A_i \in \Sigma$ (Σ is sigma algebra), $\bigcup_{i=1}^{\infty} U_i \in \Sigma_X$.
- (ii) We need to show that Σ_X is closed under set subtraction $-$. $\forall U_1, U_2 \in \Sigma_X, U_1 - U_2 = X^{-1}(A_1) - X^{-1}(A_2) = X^{-1}(A_1 - A_2)$. Since $A_1 - A_2 \in \Sigma$ (Σ is sigma algebra), $U_1 - U_2 \in \Sigma_X$.
- (iii) We need to show that Σ_X is closed to \mathcal{U} itself. Since $\mathcal{U} = X^{-1}(\mathcal{V})$ and $\mathcal{V} \in \Sigma$, it follows that $\mathcal{U} \in \Sigma_X$.

2.4

- (a) (i) We need to show that $\mathcal{F}|_A$ is closed under countable union. Let $X_1 = A \cap B_1, X_2 = A \cap B_2, \dots$ and $X' = \bigcup_{i=1}^{\infty} X_i$ and $B' = \bigcup_{i=1}^{\infty} B_i$ where $B_1, B_2, \dots \in \mathcal{F}$. Since \mathcal{F} is sigma algebra, $B' \in \mathcal{F}$. Furthermore, since $X' = \bigcup_{i=1}^{\infty} X_i = \bigcup_{i=1}^{\infty} A \cap B_i = A \cap \left(\bigcup_{i=1}^{\infty} B_i \right) = A \cap B'$, we can see that $X' \in \mathcal{F}|_A$.
- (ii) We need to show that $\mathcal{F}|_A$ is closed under set subtraction $-$. $\forall X_1, X_2 \in \mathcal{F}|_A, X_1 - X_2 = (A \cap B_1) - (A \cap B_2) = A \cap (B_1 - B_2)$. Since $B_1 - B_2 \in \mathcal{F}$ (\mathcal{F} is sigma algebra), it follows that $X_1 - X_2 \in \mathcal{F}|_A$.
- (iii) We need to show that Σ_X is closed to A itself. Since $\emptyset \in \mathcal{F}$, we have $\emptyset = A \cap \emptyset \in \mathcal{F}|_A$ and $A = \emptyset^C \in \mathcal{F}|_A$.
- (b) Let $P = \{A \cap B : B \in \mathcal{F}\}, Q = \{B : B \subset A, B \in \mathcal{F}\}$.
 - (i) We claim that $P \subset Q$. Let $X = A \cap B, B \in \mathcal{F}$. Since $A \in \mathcal{F}, X = A \cap B \in \mathcal{F}$. Furthermore, $X \in Q = \{B : B \subset A, B \in \mathcal{F}\}$.
 - (ii) We claim that $Q \subset P$. $\forall X \in Q$, we have $X \subset A$ and $X \in \mathcal{F}$, which means that $X = X \cap A$ and $X \in \mathcal{F}$. It follows that $X \in P$.
 - (iii) Take both (i)(ii) into consideration, we can see that $P = Q$.

2.5

- (a) Clearly $\sigma(\mathcal{G})$ should be the intersection of all σ -algebras that contain \mathcal{G} . Formally speaking, let $\mathcal{K} = \{\mathcal{F} | \mathcal{F} \text{ is a } \sigma\text{-algebra and contains } \mathcal{G}\}$. Then $\bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$ contains exactly those sets that are in every σ -algebra that contains \mathcal{G} . Given its existence, we only need to prove that $\bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$ is the smallest σ -algebra that contains \mathcal{G} .

First we show $\bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$ is a σ -algebra. Since \mathcal{F} is a σ -algebra and therefore $\Omega \in \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$, it follows that $\Omega \in \bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$. Next, for any $A \in \bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$, $A^c \in \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$. Since they are all σ -algebras, $A^c \in \bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$. Hence $A^c \in \bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$. Finally, for any $\{A_i\}_i \subset \bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$, $\{A_i\}_i \subset \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$. Since they are all σ -algebras, $\bigcup_i A_i \in \mathcal{F}$ for all $\mathcal{F} \in \mathcal{K}$. Hence $\bigcup_i A_i \in \bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$.

It is quite obvious that $\bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F}$ is the smallest one as $\bigcap_{\mathcal{F} \in \mathcal{K}} \mathcal{F} \subseteq \mathcal{F}'$ for all $\mathcal{F}' \in \mathcal{K}$.

- (b) We first introduce a useful lemma: the map X is \mathcal{F}/\mathcal{G} -measurable if and only if $\sigma(X) \subseteq \mathcal{F}$, where $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{G}\}$ is the σ -algebra generated by X . With this lemma, the main idea to prove X is $\mathcal{F}/\sigma(\mathcal{G})$ -measurable is to show that $\sigma(X) = \{X^{-1}(A) : A \in \sigma(\mathcal{G})\} \subseteq \mathcal{F}$.

Let $X^{-1}(\mathcal{G}) = \{X^{-1}(A) : A \in \mathcal{G}\}$. Clearly we have $X^{-1}(\mathcal{G}) \subseteq \mathcal{F}$. $\sigma(X^{-1}(\mathcal{G}))$ is the smallest σ -algebra that contains $X^{-1}(\mathcal{G})$. And we know \mathcal{F} is a σ -algebra that contains $X^{-1}(\mathcal{G})$. According to the result of the previous question, $\sigma(X^{-1}(\mathcal{G})) \subseteq \mathcal{F}$. Furthermore, $\sigma(X^{-1}(\mathcal{G})) = X^{-1}(\sigma(\mathcal{G})) = \{X^{-1}(A) : A \in \sigma(\mathcal{G})\} = \sigma(X)$. Hence $\sigma(X) \subseteq \mathcal{F}$.

Readers can further refer to the penultimate paragraph in Page 16, where the author provides a general idea to check whether a map is measurable.

- (c) The idea is to show $\forall B \in \mathfrak{B}(\mathbb{R}), \mathbb{I}\{A\}^{-1}(B) \in \mathcal{F}$.

If $\{0, 1\} \in B$, $\mathbb{I}\{A\}^{-1}(B) = \Omega \in \mathcal{F}$. If $\{0\} \in B$, $\mathbb{I}\{A\}^{-1}(B) = A^c \in \mathcal{F}$. If $\{1\} \in B$, $\mathbb{I}\{A\}^{-1}(B) = A \in \mathcal{F}$. If $\{0, 1\} \cap B = \emptyset$, $\mathbb{I}\{A\}^{-1}(B) = \emptyset \in \mathcal{F}$.

2.6 As the hint suggests, Y is not $\sigma(X)$ -measurable under such conditions since $Y^{-1}((0, 1)) = (0, 1) \notin \sigma(X)$, where $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{G}\} = \{\emptyset, \mathbb{R}\}$.

2.7 First we have $\mathbb{P}(\Omega | B) = \frac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$. Then, for all $A \in \mathcal{F}$, $\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \geq 0$. Next, for all $A \in \mathcal{F}$, $\mathbb{P}(A^c | B) = \frac{\mathbb{P}(A^c \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}((\Omega - A) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B) - \mathbb{P}(A \cap B)}{\mathbb{P}(B)} = 1 - \mathbb{P}(A | B)$. Finally, for all countable collections of disjoint sets $\{A_i\}_i$ with $A_i \in \mathcal{F}$ for all i , we have $\mathbb{P}(\bigcup_i A_i | B) = \frac{\mathbb{P}((\bigcup_i A_i) \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(\bigcup_i (A_i \cap B))}{\mathbb{P}(B)} = \sum_i \frac{\mathbb{P}(A_i \cap B)}{\mathbb{P}(B)} = \sum_i \mathbb{P}(A_i | B)$.

2.8 With the definition of conditional probability, we have $\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$.

2.9

- (a) There are 36 possible events:

1. $X_1 < 2, X_2 = \text{even}$:

$$X_1 = 1, X_2 = 2$$

$$X_1 = 1, X_2 = 4$$

$$X_1 = 1, X_2 = 6$$

$$\text{So, } \mathbb{P}(X_1 < 2, X_2 = \text{even}) = \frac{3}{36} = \frac{1}{12}$$

$$\text{and } \mathbb{P}(X_1 < 2) = \frac{1}{3}, \mathbb{P}(X_2 = \text{even}) = \frac{1}{2}$$

So, $\mathbb{P}(X_1 < 2, X_2 = \text{even}) = \mathbb{P}(X_1 < 2) * \mathbb{P}(X_2 = \text{even})$. According to the definition of independent event, two events are independent.

- (b) Prove $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

2.10

- (a) Empty sets and complete sets are independent of any event:

$$P(A \cap \Omega) = P(A) = 1 * P(A) = P(\Omega) * P(A)$$

$$P(A \cap \phi) = P(\phi) = 0 = P(\phi) * P(A)$$

- (b) Prove when $P(A) = 0$ or 1 , A is independent of any event: for any $B \in \Omega$

$$P(A) \in \{0, 1\}$$

$$\text{When } P(A) = 1, P(A^c \cap B) \leq P(A^c) = 1 - P(A) = 0,$$

$$\text{we have } P(A \cap B) = P(A \cap B) + P(A^c \cap B) = P(B) = P(A)P(B)$$

$$\text{When } P(A) = 0, \text{ we have } P(A \cap B) \leq P(A) = 0 = P(A)P(B)$$

- (c) $P(A^c \cap A) = P(A)P(A^c)$

$$\text{we have } 0 = P(A)(1 - P(A)) \Rightarrow P(A) \in \{0, 1\}$$

- (d) $P(A \cap A) = P(A)P(A), P(A) = 0, 1$

- (e) $\Omega = (1, 1), (1, 0), (0, 1), (0, 0)$

$$\text{Just verify that each case is independent : } P(A = 1, B = 1) = \frac{1}{4} = \frac{1}{2} * \frac{1}{2} = P(A)P(B)$$

$$P(A = 1, B = 0) = \frac{1}{4} = \frac{1}{2} * \frac{1}{2} = P(A)P(B)$$

$$P(A = 0, B = 1) = \frac{1}{4} = \frac{1}{2} * \frac{1}{2} = P(A)P(B)$$

$$P(A = 0, B = 0) = \frac{1}{4} = \frac{1}{2} * \frac{1}{2} = P(A)P(B)$$

- (f) $P(X_1 \leq 2) = 2/3$

$$P(X_1 = X_2) = 3/9 = 1/3$$

$$P(X_1 \leq 2, X_1 = X_2) = P(X_1 = X_2 = 1) + P(X_1 = X_2 = 2) = 1/9 + 1/9 = 2/9$$

$$\text{So, } P(X_1 \leq 2, X_1 = X_2) = P(X_1 = X_2)P(X_1 \leq 2)$$

- (g) Necessity : $\frac{|A \cap B|}{n} = P(A \cap B) = P(A)P(B) = \frac{|A|}{n} \frac{|B|}{n}$

$$\Rightarrow |A \cap B| * n = |A||B|$$

$$\text{Sufficiency : } |A \cap B| * n = |A||B| \Rightarrow \frac{|A|}{n} \frac{|B|}{n} = \frac{|A \cap B|}{n}$$

$$\Rightarrow P(A \cap B) = P(A)P(B)$$

- (h) $|A \cap B| * n = |A||B|, \Rightarrow n \leq |A| \leq n \Rightarrow A \neq \phi \text{ or } \Omega$

- (j) Let's take a counter example: roll a die and set the A event to $\{1, 2, 3\}$, B event is set to $\{1, 2, 4\}$, C event is set to $\{1, 4, 5, 6\}$

$$P(X_1 X_2 X_3) = \frac{1}{6}$$

$$P(X_1)P(X_2)P(X_3) = (1/2) * (1/2) * (2/3) = 1/6$$

$$\text{while } P(X_1 \cap X_2) = 1/3 \neq \frac{1}{2} * \frac{1}{2}$$

2.11

- (a) $X: \Omega \rightarrow x$

Because X, Y are independent equivalent to $\sigma(X), \sigma(Y)$ are independent;

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

$$P(\phi \cap A) = P(\phi) = 0 = P(\phi)P(A)$$

$$P(\Omega \cap A) = P(A) = P(\Omega)P(A)$$

(b) We know that $P(X = x) = 1$

$$P(X = x|Y) = \frac{P((X = x) \cap Y)}{P(Y)} = 1 = P(X = x)$$

$$P(X \neq x|Y) = 1 - P(X = x|Y) = 0 = P(X \neq x)$$

(c) Notice the relation: $P(A) = P(X(A) = 1)$

$$P(B) = P(X(B) = 1)$$

$$P(A \cap B) = P(X(A \cap B) = 1)$$

The first two formulas follow the definition. Let's prove the third equation:

$$P(X(A \cap B) = 1) = P(X(A) + X(B) - X(A \cup B) = 1)$$

Let's discuss $X(A), X(B), X(A \cup B)$:

$X(A)$	$X(B)$	$X(A \cup B)$	$X(A) + X(B) - X(A \cup B)$
1	1	1	1
1	0	1	0
0	1	1	0
0	0	0	0

We can see that, $P(X(A \cap B) = 1) = P(X(A) + X(B) - X(A \cup B) = 1)$, this is only one case of the first row of the table.

that is $P(X(A \cap B) = 1) = P(X(A) = 1, X(B) = 1) = P(A \cap B)$

that is $P(X(A \cap B) = 1) = P(A \cap B)$

So, $P(A \cap B) = P(A)P(B)$ is equivalent to $P(X(A \cap B) = 1) = P(X(A) = 1)P(X(B) = 1)$

(d) A_i pairwise i $\Leftrightarrow I\{A_i\}$ pairwise i

$$\text{mutual i} \Leftrightarrow P(\bigcap_i A_i) = \prod_i P(A_i)$$

$$\Leftrightarrow P(\bigcap_{i \in K^I} A_i \cap \bigcap_{i \in K^I} A_i^c)$$

$$= \prod_{i \in K^I} P(A_i) \prod_{i \in K^I} P(A_i^c)$$

$$\Leftrightarrow \{\phi, \Omega, A_i, A_i^c\} \text{ mutual independent}$$

$$\Leftrightarrow \sigma(I\{w \in A_i\}) \text{ mutual i}$$

$$\Leftrightarrow I\{w \in A_i\} \text{ mutual i}$$

2.12 X integrable $|X|$ integrable

(a) For any $A \in \mathcal{B}(R) \Rightarrow A$ is open,

so, $f^{-1}(A)$ is open, so $f^{-1}(A) \in \mathcal{B}(R)$

(b) X is known to be a random variable, $f(x) = |x|$ continuous.

r.v. X is $\mathbf{F}/\mathbf{B}(R)$ -measurable

$\Rightarrow |X|$ is $\mathbf{B}(R)/\mathbf{B}(R)$ -measurable

$\Rightarrow |X|$ is $\mathbf{F}/\mathbf{B}(R)$ -measurable

$\Rightarrow |X|$ is r.v.

From (a)(b), X integrable $\Leftrightarrow |X|$ integrable.

2.14

(a) Assume $\forall i, X_i$ is simple function.

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E} \left[\sum_{i=1}^n X_i \right] \\
&= \mathbb{E} \left[\sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} \mathbb{I}_{A_{i,j}} \{\omega\} \right] \\
&= \int_{\Omega} \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} \mathbb{I}_{A_{i,j}} \{\omega\} d\mathbb{P}(\omega) \\
&= \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} \int_{\Omega} \mathbb{I}_{A_{i,j}} \{\omega\} d\mathbb{P}(\omega) \\
&= \sum_{i=1}^n \sum_{j=1}^m \alpha_{i,j} \mathbb{P}(A_{i,j}) \\
&= \sum_{i=1}^n \mathbb{E}[X_i]
\end{aligned}$$

(b) Assume $\forall i, X_i$ is non-negative random variable.

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E} \left[\sum_{i=1}^n X_i \right] \\
&= \sup \left\{ \int_{\Omega} h d\mathbb{P} : h \text{ is simple and } 0 \leq h \leq X = \sum_{i=1}^n X_i \right\} \\
&= \sum_{i=1}^n \sup \left\{ \int_{\Omega} h_i d\mathbb{P} : h_i \text{ is simple and } 0 \leq h_i \leq X_i \right\} \\
&= \sum_{i=1}^n \mathbb{E}[X_i]
\end{aligned}$$

(c) Assume $\forall i, X_i$ is arbitrary random variable.

$$\begin{aligned}
\mathbb{E}[X] &= \mathbb{E} \left[\sum_{i=1}^n X_i \right] \\
&= \mathbb{E} \left[\sum_{i=1}^n (X_i^+ - X_i^-) \right] \\
&= \mathbb{E} \left[\sum_{i=1}^n X_i^+ \right] - \mathbb{E} \left[\sum_{i=1}^n X_i^- \right] \\
&= \sum_{i=1}^n \mathbb{E}[X_i^+] - \sum_{i=1}^n \mathbb{E}[X_i^-] \\
&= \sum_{i=1}^n (\mathbb{E}[X_i^+] - \mathbb{E}[X_i^-]) \\
&= \sum_{i=1}^n \mathbb{E}[X_i]
\end{aligned}$$

(a) Assume X is simple function.

$$\begin{aligned}
\mathbb{E}[cX] &= \mathbb{E}\left[c \sum_{i=1}^n \alpha_i \mathbb{I}_{A_i}\{\omega\}\right] \\
&= \int_{\Omega} c \sum_{i=1}^n \alpha_i \mathbb{I}_{A_i}\{\omega\} d\mathbb{P}(\omega) \\
&= c \int_{\Omega} \sum_{i=1}^n \alpha_i \mathbb{I}_{A_i}\{\omega\} d\mathbb{P}(\omega) \\
&= c\mathbb{E}[X]
\end{aligned}$$

(b) Assume X is non-negative random variable.

$$\begin{aligned}
\mathbb{E}[cX] &= \sup \left\{ \int_{\Omega} h d\mathbb{P} : h \text{ is simple and } 0 \leq h \leq cX \right\} \\
&= c \sup \left\{ \int_{\Omega} h' d\mathbb{P} : h' \text{ is simple and } 0 \leq h' \leq X \right\} \\
&= c\mathbb{E}[X]
\end{aligned}$$

(c) Assume X is arbitrary random variable.

(i) $c \geq 0$

$$\begin{aligned}
\mathbb{E}[cX] &= \mathbb{E}[(cX)^+] - \mathbb{E}[(cX)^-] \\
&= \mathbb{E}[c(X)^+] - \mathbb{E}[c(X)^-] \\
&= c\mathbb{E}[(X)^+] - c\mathbb{E}[(X)^-] \\
&= c\mathbb{E}[X]
\end{aligned}$$

(ii) $c < 0$

By definition, we have

$$\begin{aligned}
(cX)^+ &= cX \mathbb{I}\{cX > 0\} \\
&= cX \mathbb{I}\{X < 0\} \text{ (since } c < 0) \\
&= (-c)(-X) \mathbb{I}\{X < 0\} \\
&= (-c)(X)^-
\end{aligned}$$

Along the similar line, we have

$$\begin{aligned}
(cX)^- &= -cX \mathbb{I}\{cX < 0\} \\
&= -cX \mathbb{I}\{X > 0\} \\
&= -c(X)^+
\end{aligned}$$

Now we can see that

$$\begin{aligned}
\mathbb{E}[cX] &= \mathbb{E}[(cX)^+] - \mathbb{E}[(cX)^-] \\
&= \mathbb{E}[(-c)(X)^-] - \mathbb{E}[-c(X)^+] \\
&= -c\mathbb{E}[(X)^-] + c\mathbb{E}[(X)^+] \\
&= c\mathbb{E}[X]
\end{aligned}$$

(a) Assume $X = \sum_{i=1}^n \alpha_i \mathbb{I}_{A_i} \{\omega\}$, $Y = \sum_{j=1}^m \beta_j \mathbb{I}_{B_j} \{\omega\}$ are simple functions.

$$\begin{aligned}
\mathbb{E}[XY] &= \mathbb{E}\left[\sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \mathbb{I}_{A_i} \{\omega\} \mathbb{I}_{B_j} \{\omega\}\right] \\
&= \int_{\Omega} \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \mathbb{I}_{A_i} \{\omega\} \mathbb{I}_{B_j} \{\omega\} d\mathbb{P}(\omega) \\
&= \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \mathbb{P}(A_i \cap B_j) \\
&= \sum_{i=1}^n \sum_{j=1}^m \alpha_i \beta_j \mathbb{P}(A_i) \mathbb{P}(B_j) \quad (\text{by the definition of independence}) \\
&= \left(\sum_{i=1}^n \alpha_i \mathbb{P}(A_i)\right) \left(\sum_{j=1}^m \beta_j \mathbb{P}(B_j)\right) \\
&= \mathbb{E}[X] \mathbb{E}[Y]
\end{aligned}$$

(b) Assume X, Y are non-negative random variables.

$$\begin{aligned}
\mathbb{E}[XY] &= \sup \{ \mathbb{E}[h] : h \text{ is simple and } 0 \leq h \leq XY \} \\
&= \sup \{ \mathbb{E}[h_1 h_2] : h_1, h_2 \text{ are simple and } 0 \leq h_1 \leq X, 0 \leq h_2 \leq Y \} \\
&= \sup \{ \mathbb{E}[h_1] \mathbb{E}[h_2] : h_1, h_2 \text{ are simple and } 0 \leq h_1 \leq X, 0 \leq h_2 \leq Y \} \\
&= \sup \{ \mathbb{E}[h_1] : h_1 \text{ is simple and } 0 \leq h_1 \leq X \} \cdot \sup \{ \mathbb{E}[h_2] : h_2 \text{ is simple and } 0 \leq h_2 \leq Y \} \\
&= \mathbb{E}[X] \mathbb{E}[Y]
\end{aligned}$$

(c) Assume X, Y are arbitrary random variables.

$$\begin{aligned}
\mathbb{E}[XY] &= \mathbb{E}[(X^+ - X^-)(Y^+ - Y^-)] \\
&= \mathbb{E}[X^+ Y^+ - X^+ Y^- - X^- Y^+ + X^- Y^-] \\
&= \mathbb{E}[X^+] \mathbb{E}[Y^+] - \mathbb{E}[X^+] \mathbb{E}[Y^-] - \mathbb{E}[X^-] \mathbb{E}[Y^+] + \mathbb{E}[X^-] \mathbb{E}[Y^-] \\
&= (\mathbb{E}[X^+] - \mathbb{E}[X^-])(\mathbb{E}[Y^+] - \mathbb{E}[Y^-]) \\
&= \mathbb{E}[X] \mathbb{E}[Y]
\end{aligned}$$

2.17 Before proving Ex.2.17, we need to make minor changes to the definition of conditional expectation and give a small lemma.

Definition 1. Assume $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space. $\mathcal{G} \subset \mathcal{F}$ is a sub- σ -algebra of \mathcal{F} . $X : \Omega \rightarrow \mathbb{R}$ is a random variable. The conditional expectation of X given \mathcal{G} is denoted by any random variable Y which satisfies the following 2 properties:

- Y is \mathcal{G} -measurable
- $\forall A \in \mathcal{G}$,

$$\int_A Y d\mathbb{P} = \int_A X d\mathbb{P}$$

Formally, we denote Y by notation $\mathbb{E}[X|\mathcal{G}]$.

Lemma 1. If X is \mathcal{G} -measurable, then $\mathbb{E}[X|\mathcal{G}] = X$ holds a.s.

Proof. Since X is \mathcal{G} -measurable, property1 holds. And property2 holds trivially. \square

We can now handily prove Ex.2.17. Since $\mathbb{E}[X|\mathcal{G}_1]$ is \mathcal{G}_1 -measurable and $\mathcal{G}_1 \subset \mathcal{G}_2$, we can see that $\mathbb{E}[X|\mathcal{G}_1]$ is \mathcal{G}_2 -measurable. By Lemma 1, $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]|\mathcal{G}_2] = \mathbb{E}[X|\mathcal{G}_1]$ holds almost surely.