Chapter 2 (Foundations of Probability) 06.12.23

Note: I skip most of the standard measure-theoretic definitions as they should be familiar anyways...

2.1 Probability Spaces and Random Elements

Definition (Random Variables and Elements): A random variable (random vector) on a measurable space $(\Omega, \mathcal{F})$ is a $\mathcal{F}/\mathcal{B}(\mathbb{R})$-measurable function $X: \Omega \to \mathbb{R}$ (respectively $\mathcal{F}/\mathcal{B}(\mathbb{R}^k)$-measurable function $X: \Omega \to \mathbb{R}^k$). A random element between measurable spaces $(\Omega, \mathcal{F})$ and $(\mathcal{X}, \mathcal{G})$ is a $\mathcal{F}/\mathcal{G}$-measurable function $X: \Omega \to \mathcal{X}$.

2.2 σ-Algebras and Knowledge

Lemma (Factorization): Let $(\Omega, \mathcal{F})$, $(\mathcal{X}, \mathcal{G})$, $(\mathcal{Y}, \mathcal{H})$ be measurable spaces and let $X: \Omega \to \mathcal{X}$ and $Y: \Omega \to \mathcal{Y}$ be random elements. Suppose $(\mathcal{Y}, \mathcal{H})$ is a Borel space). Then Y is $\sigma(X)$-measurable (i.e., $\sigma(Y) \subseteq \sigma(X)$) if and only if there exists a $\mathcal{G}/\mathcal{H}$-measurable map $f: \mathcal{X} \to \mathcal{Y}$ such that $Y = f \circ X$.

Definition (Filtration, Adapted, Predictable): Let $(\Omega, \mathcal{F})$ be a measurable space, then a filtration is a sequence $(\mathcal{F}_t)_{t=0}^n$ of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_t \subseteq \mathcal{F}_{t+1}$ for all $t<n$. (Note that $n=\infty$ is allowed and $\mathcal{F}_\infty := \sigma(\bigcup_{t=0}^\infty \mathcal{F}_t)$.) A sequence of random variables $(X_t)_{t=1}^n$ is adapted to a filtration $\mathbb{F} = (\mathcal{F}_t)_{t=0}^n$ if X_t is $\mathcal{F}_t$-measurable. A sequence of random variables $(X_t)_{t=1}^n$ is $\mathbb{F}$-predictable if X_t is $\mathcal{F}_{t-1}$-measurable. A filtered probability space is a tuple $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space and $\mathbb{F}$ is a filtration of $\mathcal{F}$.

2.3 Conditional Probabilities

Definition (Conditional Probability): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A, B \in \mathcal{F}$ such that $\mathbb{P}(B) > 0$. Then the <u>conditional probability</u> $\mathbb{P}(A|B)$ is $\mathbb{P}(A|B) = \mathbb{P}(A \cap B) / \mathbb{P}(B)$.

Theorem (Bayes Rule): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A, B \in \mathcal{F}$ such that $\mathbb{P}(A) > 0$ and $\mathbb{P}(B) > 0$. Then:

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

2.4 Independence

Definition (Independence): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A collection $\mathcal{G} \subseteq \mathcal{F}$ is <u>pairwise independent</u> if for all $A, B \in \mathcal{G}$, $A \neq B$, A and B are independent. The events in $\mathcal{G}$ are said to be <u>mutually independent</u> (or just <u>independent</u>) if for any finite set $A_1, \dots, A_n \in \mathcal{F}$ of distinct events,

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \prod_{i=1}^{n} \mathbb{P}(A_i).$$

Two collections of events, $\mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{F}$, are <u>independent</u> if for all $A \in \mathcal{G}_1$ and all $B \in \mathcal{G}_2$, A and B are independent. Two random variables X and Y are <u>independent</u> if $\sigma(X)$ and $\sigma(Y)$ are independent. The notions of pairwise/mutual independence apply accordingly.

2.5 Integration and Expectation

Proposition (Independent Expectation): If X and Y are independent and either $\mathbb{E}[|X|], \mathbb{E}[|Y|] < \infty$ or $\mathbb{E}[|XY|] < \infty$, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Proposition (Tail Expectation): Let X be a nonnegative random variable. Then $\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X > x)\,dx$.

Definition ((Complementary) CDF): The <u>(complementary) cumulative distribution function</u> of a random variable X is ($x \mapsto \mathbb{P}(X > x)$) $F_X(x) = \mathbb{P}(X \leq x)$. Note that F_X is increasing, right-continuous, and $\lim_{x \to -\infty} F_X(x) = 0$ and $\lim_{x \to \infty} F_X(x) = 1$. The CDF captures all aspects of the distribution of X.

Proposition (Law Of The Unconcious Statistician – LOTUS): Let $X : \Omega \to \mathcal{X}$ be a random variable and let $\mathbb{P}_X$ be its push-forward. Let $g : \mathcal{X} \to \mathbb{R}$ be measurable, then

$$\mathbb{E}[g(X)] = \int g(x)\,d\mathbb{P}_X(x),$$

provided that either side exists.

2.6 Conditional Expectation

Definition (Conditional Expectation): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $X : \Omega \to \mathbb{R}$ be a random variable, and let $\mathcal{H} \subseteq \mathcal{F}$ be a sub-σ-algebra. The conditional expectation of X given $\mathcal{H}$ is denoted by $\mathbb{E}[X \mid \mathcal{H}]$ and defined to be any $\mathcal{H}$-measurable random variable on Ω such that for all $H \in \mathcal{H}$, $\int_H \mathbb{E}[X \mid \mathcal{H}]\,d\mathbb{P} = \int_H X\,d\mathbb{P}$. Given a random variable Y, the conditional expectation of X given Y is $\mathbb{E}[X \mid \sigma(Y)]$.

Theorem: The conditional expectation always exists and for two $f_1 : \Omega \to \mathcal{X}$, $f_2 : \Omega \to \mathcal{X}$, $f_1 = f_2$ almost surely.

Theorem (Properties of Conditional Expectation): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{F}$ be sub-σ-algebras and X, Y integrable random variables on $(\Omega, \mathcal{F}, \mathbb{P})$. Then:

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

(ii) $\mathbb{E}[1 \mid \mathcal{G}] = 1$ a.s.

(iii) $\mathbb{E}[X + Y \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{G}] + \mathbb{E}[Y \mid \mathcal{G}]$ a.s.

(iv) $\mathbb{E}[XY \mid \mathcal{G}] = Y\,\mathbb{E}[X \mid \mathcal{G}]$ a.s. if $\mathbb{E}[XY]$ exists and Y is $\mathcal{G}$-measurable.

(v) If $\mathcal{G}_1 \subseteq \mathcal{G}_2$, then $\mathbb{E}[X \mid \mathcal{G}_1] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}_2] \mid \mathcal{G}_1]$ a.s.

(vi) If $\sigma(X)$ is independent of $\mathcal{G}_2$ given $\mathcal{G}_1$, then $\mathbb{E}[X \mid \sigma(\mathcal{G}_1 \cup \mathcal{G}_2)] = \mathbb{E}[X \mid \mathcal{G}_1]$ a.s.

(vii) If $\mathcal{G} = \{\emptyset, \Omega\}$, then $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$ a.s.

Definition (Conditional Independence): Two event systems $\mathcal{A}$ and $\mathcal{B}$ are independent given a σ-algebra $\mathcal{F}$ if for all $A \in \mathcal{A}$ and all $B \in \mathcal{B}$, $\mathbb{P}(A \cap B \mid \mathcal{F}) = \mathbb{P}(A \mid \mathcal{F})\,\mathbb{P}(B \mid \mathcal{F})$ a.s.

2.7 Notes

Definition (Absolutely Continuous): Let $\mathbb{P}$ and $\mathbb{Q}$ be measures over $(\Omega, \mathcal{F})$. Then $\mathbb{P}$ is <u>absolutely continuous</u> w.r.t. $\mathbb{Q}$ if for all $A \in \mathcal{F}$, $\mathbb{Q}(A) = 0 \Rightarrow \mathbb{P}(A) = 0$. We also say $\mathbb{Q}$ dominates $\mathbb{P}$ or write $\mathbb{P} \ll \mathbb{Q}$.

Theorem (Radon-Nikodym Derivative): Let $\mathbb{P}$ and $\mathbb{Q}$ be measures over $(\Omega, \mathcal{F})$ and suppose $\mathbb{Q}$ is σ-finite. Then the <u>density</u>, or <u>Radon-Nikodym derivative</u>, of $\mathbb{P}$ w.r.t. $\mathbb{Q}$, denoted $d\mathbb{P}/d\mathbb{Q}$, exists if and only if $\mathbb{P} \ll \mathbb{Q}$. The Radon-Nikodym derivative is the function $d\mathbb{P}/d\mathbb{Q} : \Omega \to \mathbb{R}$ such that

$$\mathbb{P}(A) = \int_A \frac{d\mathbb{P}}{d\mathbb{Q}}\, d\mathbb{Q} \qquad \text{for all } A \in \mathcal{F}.$$

It is unique up to a $\mathbb{Q}$-null set.

Proposition (Change-Of-Measure): Let $\mathbb{P}$ and $\mathbb{Q}$ be measures over $(\Omega, \mathcal{F})$ such that $d\mathbb{P}/d\mathbb{Q}$ exists. Let X be a $\mathbb{P}$-integrable random variable, then $\int X \, d\mathbb{P} = \int X \, {}^{d\mathbb{P}}/_{d\mathbb{Q}} \, d\mathbb{Q}$.

Proposition (Radon-Nikodym Chain Rule): Let $\mathbb{P}, \mathbb{Q}, \mathbb{S}$ be measures with $\mathbb{P} \ll \mathbb{Q} \ll \mathbb{S}$. Then

$$\frac{d\mathbb{P}}{d\mathbb{S}} = \frac{d\mathbb{P}}{d\mathbb{Q}} \frac{d\mathbb{Q}}{d\mathbb{S}}.$$

Definition (Support): Let X be a topological space and let μ be a measure over $(X, \mathcal{B}(X))$, then the <u>support</u> of X is

$$\operatorname{Supp}(x) := \{x \in X \mid \mu(U) > 0 \text{ for all neighborhoods } U \text{ of } x\}.$$

2.9 <u>Exercises</u> 07.12.23

(1) *Proof.* Let $\mathcal{F}, \mathcal{G}, \mathcal{H}$ be σ-algebras and let f and g be $\mathcal{F}/\mathcal{G}$- and $\mathcal{G}/\mathcal{H}$-measurable, respectively. Consider $g \circ f$. Let $A \in \mathcal{H}$, then $g^{-1}(A) \in \mathcal{G}$ and $f^{-1}(g^{-1}(A)) \in \mathcal{F}$. Thus, as $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$, $g \circ f$ is $\mathcal{F}/\mathcal{H}$-measurable.

$\square$

(2) *Proof.* Let $(\Omega, \mathcal{F})$ be a measurable space and let $X_1, \dots, X_n : \Omega \to \mathbb{R}$ be random variables. Define $X : \Omega \to \mathbb{R}^n$ by $X = (X_1, \dots, X_n)$. Let $A \subseteq \mathbb{R}^n$ be an open rectangle, then there are open intervals $A_1, \dots, A_n \subseteq \mathbb{R}$ with $A = A_1 \times \dots \times A_n$. But then

$$X^{-1}(A) = \bigcap_{i=1}^{n} X_i^{-1}(A_i)$$

and as for each i, $X_i^{-1}(A_i) \in \mathcal{F}$, also $X^{-1}(A) \in \mathcal{F}$. Hence, X is measurable as $\mathcal{B}(\mathbb{R}^n)$ is generated by open rectangles.

$\square$

(3) *Proof.* Let $\mathcal{U}$ be a set and let $(\mathcal{V}, \Sigma)$ be a measurable space. Let $X : \mathcal{U} \to \mathcal{V}$ be a function. We show that $\mathcal{F} = \{X^{-1}(A) \mid A \in \Sigma\}$ is a σ-algebra over $\mathcal{U}$. First, as $X^{-1}(\mathcal{V}) = \mathcal{U}$ and $\mathcal{V} \in \Sigma$, also $\mathcal{U} \in \mathcal{F}$. Now let $B \in \mathcal{F}$. Then there is an $A \in \Sigma$ with $X^{-1}(A) = B$. But then also $\mathcal{V} \setminus A \in \Sigma$ and $X^{-1}(\mathcal{V} \setminus A) = \mathcal{U} \setminus X^{-1}(A) = \mathcal{U} \setminus B \in \mathcal{F}$. Finally, let $(B_i) \subseteq \mathcal{F}$. Then for each i there is an $A_i \in \Sigma$ with $X(A_i) = B_i$. Moreover, $\bigcup A_i \in \Sigma$ and

$$X^{-1}\left(\bigcup A_i\right) = \bigcup X^{-1}(A_i) = \bigcup B_i \in \mathcal{F}.$$

Hence, $\mathcal{F}$ is a σ-algebra over $\mathcal{U}$.

$\square$

(4) Let $(\Omega, \mathcal{F})$ be a measurable space, let $A \subseteq \Omega$, and define $\mathcal{F}|_A = \{A \cap B \mid B \in \mathcal{F}\}$.

(a) *Proof.* We show that $(A, \mathcal{F}|_A)$ is a measurable space. Clearly, $A \in \mathcal{F}|_A$ as $\Omega \in \mathcal{F}$ and $A \cap \Omega = A$. Now let $C \in \mathcal{F}|_A$ and $B \in \mathcal{F}$ such that $C = A \cap B$. Then also $\Omega \setminus B \in \mathcal{F}$ and

$$A \cap (\Omega \setminus B) = (A \cap \Omega) \setminus (A \cap B) = A \setminus C,$$

so $A \setminus C \in \mathcal{F}|_A$. Moreover, let $(C_i) \subseteq \mathcal{F}|_A$ and $(B_i) \in \mathcal{F}|_A$ such that $C_i = A \cap B_i$ for all i. But then also $\bigcup B_i \in \mathcal{F}$ and

$$A \cap \left(\bigcup B_i\right) = \bigcup (A \cap B_i) = \bigcup C_i,$$

so $\bigcup C_i \in \mathcal{F}|_A$. Hence, $(A, \mathcal{F}|_A)$ is a measurable space.

$\square$

(4) (b) *Proof.* Suppose $A \in \mathcal{F}$. "$\subseteq$": Let $C \in \mathcal{F}|_A$. Then there is a $B \in \mathcal{F}$ with $C = A \cap B$. Thus, as $A, B \in \mathcal{F}$ also $C \in \mathcal{F}$, and $C \subseteq A$ is clear. Hence, $C \in \{\tilde{B} \mid \tilde{B} \in \mathcal{F}, \tilde{B} \subseteq A\}$.
"$\supseteq$": Let $C \in \{\tilde{B} \mid \tilde{B} \in \mathcal{F}, \tilde{B} \subseteq A\}$. Then $C = A \cap C$ and thus, $C \in \mathcal{F}|_A$. Therefore, $\mathcal{F}|_A = \{B \mid B \in \mathcal{F}, B \subseteq A\}$.

$\square$

(5) Let $\mathcal{G} \subseteq 2^\Omega$ be nonempty and let $\sigma(\mathcal{G})$ be the smallest σ-algebra such that $\mathcal{G} \subseteq \sigma(\mathcal{G})$.

(a) *Proof.* We show that

$$\sigma(\mathcal{G}) = \bigcap_{\mathcal{H} \in \tilde{\mathcal{H}}} \mathcal{H},$$

where $\tilde{\mathcal{H}}$ is the set of all σ-algebras over Ω that contain $\mathcal{G}$. We ~~first~~ show that this is actually a σ-algebra, that it contains $\mathcal{G}$, and finally that it is the smallest. Clearly, $\Omega \in \sigma(\mathcal{G})$ as $\Omega \in \mathcal{H}$ for all $\mathcal{H} \in \tilde{\mathcal{H}}$. Let $A \in \sigma(\mathcal{G})$. Then $A \in \mathcal{H}$ for all $\mathcal{H} \in \tilde{\mathcal{H}}$ and thus $\Omega \setminus A \in \mathcal{H}$ for all $\mathcal{H} \in \tilde{\mathcal{H}}$. Hence, also $\Omega \setminus A \in \sigma(\mathcal{G})$. Finally, let $(A_i) \subseteq \sigma(\mathcal{G})$. Then $A_i \in \mathcal{H}$ for all $\mathcal{H} \in \tilde{\mathcal{H}}$ and all i and therefore also $\bigcup A_i \in \mathcal{H}$ for all $\mathcal{H} \in \tilde{\mathcal{H}}$. Thus, $\bigcup A_i \in \sigma(\mathcal{G})$ and $\sigma(\mathcal{G})$ is a σ-algebra. Moreover, as $\mathcal{G} \subseteq \mathcal{H}$ for all $\mathcal{H} \in \tilde{\mathcal{H}}$, also $\mathcal{G} \subseteq \sigma(\mathcal{G})$. Finally, let $\mathcal{F} \subseteq 2^\Omega$ be any σ-algebra with $\mathcal{G} \subseteq \mathcal{F}$. Then $\mathcal{F} \in \tilde{\mathcal{H}}$ and thus $\mathcal{F} \supseteq \sigma(\mathcal{G})$.

$\square$

(b) *Proof.* Let $(\Omega', \mathcal{F})$ be a measurable space and let $X : \Omega' \to \Omega$ be $\mathcal{F}/\mathcal{G}$-measurable. Now let $A \in \mathcal{G}$, then $X^{-1}(\Omega \setminus A) = \Omega' \setminus X^{-1}(A) \in \mathcal{F}$ as $X^{-1}(A) \in \sigma(\mathcal{G})$. Now let $A_1, A_2, \ldots \in \mathcal{G}$. Then $X^{-1}(\bigcup A_i) = \bigcup X^{-1}(A_i) \in \mathcal{F}$ as $X^{-1}(A_i) \in \mathcal{F}$ for each A_i. Hence, as every $A \in \sigma(\mathcal{G})$ can be constructed from countable unions and complements of elements of $\mathcal{G}$, we have $X^{-1}(A) \in \mathcal{F}$ for all $A \in \sigma(\mathcal{G})$ and thus X is $\mathcal{F}/\sigma(\mathcal{G})$-measurable.

$\square$

(c) *Proof.* Let $\mathcal{F}$ be a σ-algebra and let $\mathbb{1}_A$ be the indicator function for some $A \in \mathcal{F}$. Let B be some subset of the codomain of $\mathbb{1}_A$. Then:

$$\mathbb{1}_A^{-1}(B) = \begin{cases} \emptyset & \text{if } 0, 1 \notin B \\ A & \text{if } 0 \notin B,\ 1 \in B \\ \Omega \setminus A & \text{if } 0 \in B,\ 1 \notin B \\ \Omega & \text{if } 0, 1 \in B \end{cases}$$

In all cases, $\mathbb{1}_A^{-1}(B) \in \mathcal{F}$ as $A \in \mathcal{F}$, so $\mathbb{1}_A$ is indeed measurable. In fact, it is measurable if <u>and only if</u> $A \in \mathcal{F}$.

$\square$

(6) TODO

(7) *Proof.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $B \in \mathcal{F}$, $\mathbb{P}(B) > 0$. Consider $\mathbb{Q} : \mathcal{F} \to \mathbb{R}$, $\mathbb{Q}(A) = \mathbb{P}(A \mid B)$, where $\mathbb{P}(A \mid B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$. We show that $\mathbb{Q}$ is a probability measure over $(\Omega, \mathcal{F})$. First, clearly

$$\mathbb{Q}(\emptyset) = \mathbb{P}(\emptyset \mid B) = \mathbb{P}(\emptyset \cap B)/\mathbb{P}(B) = 0/\mathbb{P}(B) = 0.$$

Moreover,

$$\mathbb{Q}(\Omega) = \mathbb{P}(\Omega \cap B)/\mathbb{P}(B) = \mathbb{P}(B)/\mathbb{P}(B) = 1.$$

Now let $(A_i) \subseteq \mathcal{F}$ with $A_i \cap A_j = \emptyset$ for $i \neq j$. Then also $(A_i \cap B) \cap (A_j \cap B) = \emptyset$ for $i \neq j$ and thus,

$$\mathbb{Q}(\bigcup A_i) = \mathbb{P}((\bigcup A_i) \cap B)/\mathbb{P}(B) = \mathbb{P}(\bigcup (A_i \cap B))/\mathbb{P}(B)$$

$$= \left(\sum \mathbb{P}(A_i \cap B)\right)/\mathbb{P}(B) = \sum \mathbb{Q}(A_i).$$

Hence, $\mathbb{Q}$ is a probability measure. (Nonnegativity is trivial.)

$\square$

(8) *Proof.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A, B \in \mathcal{F}$, $\mathbb{P}(A), \mathbb{P}(B) > 0$. Then $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$ and $\mathbb{P}(B|A) = \mathbb{P}(A \cap B)/\mathbb{P}(A)$. Hence,

$$\mathbb{P}(A|B)\,\mathbb{P}(B) = \mathbb{P}(B|A)\,\mathbb{P}(A).$$

Dividing both sides by $\mathbb{P}(B)$ yields Bayes' rule:

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

$\square$

(9) (a) Let A and B denote the event sets for '$X_1 < 2$' and 'X_2 is even', respectively. That is,

$$A = \{1\} \times \{1, 2, 3, 4, 5, 6\},$$

$$B = \{1, 2, 3, 4, 5, 6\} \times \{2, 4, 6\}.$$

We have $\mathbb{P}(A) = 6/6^2 = 1/6$ and $\mathbb{P}(B) = 18/6^2 = 1/2$. Moreover,

$$A \cap B = \{1\} \times \{2, 4, 6\},$$

and thus $\mathbb{P}(A \cap B) = 1/12 = 1/6 \cdot 1/2 = \mathbb{P}(A)\,\mathbb{P}(B)$. That is, A and B are independent.

(b) Let $A \in \sigma(X_1)$ and $B \in \sigma(X_2)$. We have that for any set $C \subseteq \{1, 2, 3, 4, 5, 6\}$, $X_1^{-1}(C) = C \times \{1, 2, 3, 4, 5, 6\}$, and similar for X_2. Hence, there are $A', B' \subseteq \{1, 2, 3, 4, 5, 6\}$ such that $A = A' \times \{1, \dots, 6\}$ and $B = \{1, \dots, 6\} \times B'$. Thus, $\mathbb{P}(A) = 6|A'|/6^2 = |A'|/6$ and $\mathbb{P}(B) = |B'|/6$. Moreover, $A \cap B = A' \times B'$ and $\mathbb{P}(A \cap B) = |A'| \cdot |B'|/6^2$. Thus, we have $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$.

10.12.23

(10) (a) *Proof.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. As $\emptyset \cap \Omega = \emptyset$ and $\mathbb{P}(\emptyset) = 0$, we have

$$\mathbb{P}(\emptyset \cap \Omega) = \mathbb{P}(\emptyset) = 0 = \mathbb{P}(\emptyset)\mathbb{P}(\Omega),$$

so $\emptyset$ and Ω are independent.

$\square$

(b) *Proof.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A \in \mathcal{F}$ such that $\mathbb{P}(A) \in \{0, 1\}$. Suppose $\mathbb{P}(A) = 0$. Let $B \in \mathcal{F}$. Then $\mathbb{P}(A)\mathbb{P}(B) = 0\mathbb{P}(B) = 0$. We thus need to show that $\mathbb{P}(A \cap B) = 0$. As $A \cap B \subseteq A$, we have

$$\mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0,$$

and thus, as $\mathbb{P}$ is nonnegative, $\mathbb{P}(A \cap B) = 0$ and therefore A and B are independent. Now suppose $\mathbb{P}(A) = 1$. Let $B \in \mathcal{F}$. Then $\mathbb{P}(A)\mathbb{P}(B) = 1\mathbb{P}(B) = \mathbb{P}(B)$, so we need to show that $\mathbb{P}(A \cap B) = \mathbb{P}(B)$. First, as $A \cap B \subseteq B$, we have $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$. Suppose A and B are disjoint, i.e., $A \cap B = \emptyset$. Then

$$1 \geq \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) = 1 + \mathbb{P}(B),$$

so $\mathbb{P}(B) = 0 = \mathbb{P}(\emptyset) = \mathbb{P}(A \cap B)$. Suppose $A \cap B \neq \emptyset$. Then we can write A as

$$A = (A \cap B) \cup (A \setminus B)$$

where $(A \cap B) \cap (A \setminus B) = \emptyset$, and therefore,

$$\mathbb{P}(A) = \mathbb{P}((A \cap B) \cup (A \setminus B)) = \mathbb{P}(A \cap B) + \mathbb{P}(A \setminus B)$$

$$= \mathbb{P}(A \cap B) + \mathbb{P}(A) - \mathbb{P}(B). \quad (\text{As } \mathbb{P}(A) = 1 < \infty.)$$

Subtracting and adding $\mathbb{P}(A)$ and $\mathbb{P}(B)$, respectively, yields $\mathbb{P}(A \cap B) = \mathbb{P}(B)$.

$\square$

(c) Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A \in \mathcal{F}$ be independent of its complement, $\Omega \setminus A$. Then A is trivial.

Proof. Suppose $\mathbb{P}(A \cap (\Omega \setminus A)) = \mathbb{P}(A)\,\mathbb{P}(\Omega \setminus A)$. On the left hand side, we have

$$\mathbb{P}(A \cap (\Omega \setminus A)) = \mathbb{P}(\emptyset) = 0,$$

so $\mathbb{P}(A)\,\mathbb{P}(\Omega \setminus A) = 0$, too. Thus, either $\mathbb{P}(A) = 0$ or $\mathbb{P}(\Omega \setminus A) = 0$ and due to $\mathbb{P}(\Omega \setminus A) = 1 - \mathbb{P}(A)$, we have $\mathbb{P}(A) \in \{0, 1\}$.

$\square$

(d) We can conclude that A is trivial.

Proof. Let $A \in \mathcal{F}$ be independent of itself. Then $\mathbb{P}(A) = (\mathbb{P}(A))^2$, so $\mathbb{P}(A) \in \{0, 1\}$.

$\square$

(e) Let $\Omega = \{h, t\}^2$ and let $X_i : \Omega \to \{h, t\}$, $X_i(\omega) = \omega_i$ be the coins for $i \in \{1, 2\}$. Let $\mathcal{F} = 2^\Omega$ and $\mathbb{P} : \mathcal{F} \to \mathbb{R}$, $\mathbb{P}(A) = |A|/4$ for all $A \in \mathcal{F}$. Then we have the following independent pairs:

A	B
$\{h\} \times \{h, t\}$	$\{h, t\} \times \{h\}$
$\{t\} \times \{h, t\}$	$\{h, t\} \times \{t\}$
$\emptyset$	any set
Ω	any set
any set	$\emptyset$
any set	Ω

(f) Let $X_i : \Omega \to \{1, 2, 3\}$ be the random variables for $i \in \{1, 2\}$. Consider $A = \{X_1 \leq 2\}$ and $B = \{X_1 = X_2\}$. Then

$$A = \{1, 2\} \times \{1, 2, 3\},$$

$$B = \{(1,1), (2,2), (3,3)\},$$

and $\mathbb{P}(A) = 2 \cdot 3/9 = 6/9 = 2/3$, $\mathbb{P}(B) = 1/3$. Moreover, with $A \cap B = \{(1,1), (2,2)\}$, we have $\mathbb{P}(A \cap B) = 2/9$. Hence, as $\mathbb{P}(A)\mathbb{P}(B) = 2/3 \cdot 1/3 = 2/9$, A and B are independent.

$\square$

(g) Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an n-element, ~~finite~~ uniform probability space. That is, $\mathbb{P}(A) = |A|/n$ where $A \in \mathcal{F} = 2^{\Omega}$. '$\Leftarrow$': Let $A, B \in \mathcal{F}$ such that $n|A \cap B| = |A| \cdot |B|$. Then

$$\mathbb{P}(A \cap B) = \frac{1}{n}|A \cap B| = \frac{1}{n^2} n|A \cap B|$$

$$= \frac{1}{n}|A| \cdot \frac{1}{n}|B| = \mathbb{P}(A)\mathbb{P}(B),$$

so A and B are independent. '$\Rightarrow$': Let $A, B \in \mathcal{F}$ such that $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. That is,

$$\frac{1}{n}|A \cap B| = \mathbb{P}(A \cap B) = \frac{1}{n^2}|A| \cdot |B| = \mathbb{P}(A)\mathbb{P}(B).$$

Multiplying both sides by n yields the desired relation.

$\square$

(h) Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an n-element, uniform probability space. That is, $|\Omega| = n$, $\mathbb{P}(A) = |A|/n$, where $A \in \mathcal{F} = 2^{\Omega}$. Suppose n is prime. Let $A, B \in \mathcal{F}$ be independent, then $n|A \cap B| = |A| \cdot |B|$. Suppose, for contradiction, that both A and B are nontrivial. That is, $0 < |A|, |B| < n$. Thus, as $|A| \cdot |B| = n|A \cap B|$ and $|A| \cdot |B| > 0$, also $|A \cap B| > 0$. Let $k = |A|$, $\ell = |B|$, and $m = |A \cap B|$, all of which are natural numbers. We have $k\ell = nm$. Moreover, $m \leq k, \ell < n$. Let $\tilde{k} = k/\gcd(k,m)$ and $\tilde{m} = m/\gcd(k,m)$. Then $\tilde{k}$ and $\tilde{m}$ are coprime and $\tilde{k}\ell = n\tilde{m}$. However, this implies that $\tilde{k} | n$ as $\tilde{k} \nmid \tilde{m}$. ↯ But n is prime. Hence, A and B cannot be independent if they are both nontrivial.

$\square$

(i) TODO

(j) No. Let A, B be dependent events and let $C = \emptyset$. Then

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\emptyset) = 0 = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

(11) (a) Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be a constant random variable. Then $\sigma(X) = \{\emptyset, \Omega\}$, and both $\emptyset$ and Ω are independent of every event.

$\square$

(b) Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be constant a.s. Then for all $A \in \sigma(X)$, we have $\mathbb{P}(A) \in \{0, 1\}$ and thus every A is trivial. Hence, independent of every event.

$\square$

(c) *Proof.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A, B \in \mathcal{F}$ be events. Let $X = \mathbb{1}_A$ and $Y = \mathbb{1}_B$. 'If:' Suppose X and Y are independent. Then $\sigma(X)$ and $\sigma(Y)$ are independent. As $A \in \sigma(X)$ and $B \in \sigma(Y)$, A and B are independent. 'Only if:' Suppose A and B are independent. We have

$$\sigma(X) = \{\emptyset, \Omega, A, \Omega \setminus A\},$$

$$\sigma(Y) = \{\emptyset, \Omega, B, \Omega \setminus B\}.$$

Let $\tilde{A} \in \sigma(X)$ and $\tilde{B} \in \sigma(Y)$. If $\tilde{A} \in \{\emptyset, \Omega\}$ or $\tilde{B} \in \{\emptyset, \Omega\}$, things are trivial. If $\tilde{A} = A$ and $\tilde{B} = B$, $\tilde{A}$ and $\tilde{B}$ are independent by assumption. If $\tilde{A} = \Omega \setminus A$ and $\tilde{B} = B$, we have

$$\mathbb{P}(\tilde{A}) = 1 - \mathbb{P}(A) \qquad \mathbb{P}(\tilde{B}) = \mathbb{P}(B)$$

$$\begin{aligned}
\mathbb{P}(\tilde{A} \cap \tilde{B}) &= \mathbb{P}((\Omega \setminus A) \cap B) = \mathbb{P}(B \setminus (A \cap B)) \\
&= \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) \\
&= (1 - \mathbb{P}(A))\mathbb{P}(B) = \mathbb{P}(\tilde{A})\mathbb{P}(\tilde{B}),
\end{aligned}$$

so $\tilde{A}$ and $\tilde{B}$ are independent. If $\tilde{A} = A$ and $\tilde{B} = \Omega \setminus B$, the argument is analogous. If $\tilde{A} = \Omega \setminus A$ and $\tilde{B} = \Omega \setminus B$, we have $\mathbb{P}(\tilde{A}) = 1 - \mathbb{P}(A)$ and $\mathbb{P}(\tilde{B}) = 1 - \mathbb{P}(B)$, and

$$\begin{aligned}
\mathbb{P}(\tilde{A} \cap \tilde{B}) &= \mathbb{P}((\Omega \setminus A) \cap (\Omega \setminus B)) \\
&= \mathbb{P}((\Omega \cap (\Omega \setminus B)) \setminus (A \cap (\Omega \setminus B))) \\
&= \mathbb{P}((\Omega \setminus B) \setminus (A \setminus (A \cap B))) \\
&= 1 - \mathbb{P}(B) - (\mathbb{P}(A) - \mathbb{P}(A \cap B)) \\
&= 1 - \mathbb{P}(B) - \mathbb{P}(A) + \mathbb{P}(A)\mathbb{P}(B) \\
&= (1 - \mathbb{P}(A))(1 - \mathbb{P}(B)) = \mathbb{P}(\tilde{A})\mathbb{P}(\tilde{B})
\end{aligned}$$

$\square$

(d) Let $(A_i)_{i=1}^{n}$ be n events and let $X_i = \mathbb{1}_{A_i}$ be indicators for all $i = 1, \dots, n$. Then (A_i) are pairwise/mutually independent if and only if (X_i) are pairwise/mutually independent.

The proof is straightforward but tedious.

(12) *Proof.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $X, Y : \Omega \to \mathbb{R}$ be random variables such that $X \le Y$. Let X^+, X^-, Y^+, Y^- be the nonnegative and nonpositive parts of X and Y. Then $0 \le X^+ \le Y^-$ and $X^- \ge Y^- \ge 0$. Hence,

$$\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-] \le \mathbb{E}[Y^+] - \mathbb{E}[Y^-] = \mathbb{E}[Y].$$

$\square$

(13) (a) *Proof.* Let $f : \mathbb{R} \to \mathbb{R}$ be continuous. Let $O \subseteq \mathbb{R}$ be open and recall that the Borel algebra is generated by the open sets. As f is continuous, $f^{-1}(O)$ is also open and thus measurable. Hence, f is measurable.

$\square$

(b) *Proof.* Let $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a random variable and consider $|\cdot| : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. As $|\cdot|$ is continuous, it is measurable. As $|X| = |\cdot| \circ X$, $|X|$ is also measurable.

$\square$

(c) *Proof.* Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be a random variable. 'If:' Suppose that $|X|$ is integrable. Then $\mathbb{E}[X^+] \le \mathbb{E}[|X|] < \infty$ and $\mathbb{E}[X^-] \le \mathbb{E}[|X^-|] < \infty$, so X is integrable. 'Only if:' Suppose that X is integrable. Then

$$\mathbb{E}[|X|] = \underbrace{\mathbb{E}[X^+]}_{<\infty} + \underbrace{\mathbb{E}[X^-]}_{<\infty} < \infty,$$

so $|X|$ is integrable.

$\square$

(74) ~~TODO~~ $\rightarrow$ pg. 22

(75) Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let (X_i) be a (possibly infinite) sequence on that space. Assume that $\mathbb{E}[X_i] < \infty$ for all i and $\mathbb{E}[\sum_i |X_i|] < \infty$. Let $\underbrace{X = \sum_i X_i}_{\text{assume } X \text{ exists}}$, and $\tilde{X} = \sum_i |X_i|$. Let $X_n = \sum_{i=1}^n X_i$. Then

$$|X_n| \leq \sum_{i=1}^n |X_i| \leq \sum_i |X_i| = \tilde{X},$$

so (X_n) is dominated by the integrable function $\tilde{X}$. Hence, by the dominated convergence theorem,

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_i X_i\right] = \sum_i \mathbb{E}[X_i].$$

$\square$

(76) Proof. Let X be an integrable random variable and let $c \in \mathbb{R}$. Suppose X is simple. Then $\mathbb{E}[cX] = c\,\mathbb{E}[X]$ is trivial. Suppose X is nonnegative. Then

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

Now let X be any integrable random variable. Then $\mathbb{E}[cX] = \mathbb{E}[cX^+] - \mathbb{E}[cX^-] = c(\mathbb{E}[X^+] - \mathbb{E}[X^-]) = c\,\mathbb{E}[X]$.

$\square$

(17) *Proof.* Let $X = \mathbb{1}_A$ and $Y = \mathbb{1}_B$ be indicator function and suppose they are independent. Then A and B are independent and we have

$$\mathbb{E}[X\ Y] = \int \mathbb{1}_A \mathbb{1}_B \, d\mathbb{P} = \int \mathbb{1}_{A \cap B} \, d\mathbb{P}$$

$$= \mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = \mathbb{E}[X]\mathbb{E}[Y].$$

Now let $X = \sum_i \alpha_i \mathbb{1}_{A_i}$ and $Y = \sum_i \beta_i \mathbb{1}_{B_i}$ be simple functions and that they are independent. Suppose, w.l.o.g. that $A_i \cap A_j = \emptyset$ for $i \neq j$ and $B_i \cap B_j = \emptyset$ for $i \neq j$. Moreover, we have that all A_i and B_j are independent. Hence,

$$\mathbb{E}[X\ Y] = \int \left(\sum_i \alpha_i \mathbb{1}_{A_i}\right)\left(\sum_j \beta_j \mathbb{1}_{B_j}\right) d\mathbb{P}$$

$$= \int \sum_{ij} \alpha_i \beta_j \mathbb{1}_{A_i \cap B_j} \, d\mathbb{P}$$

$$= \sum_{ij} \alpha_i \beta_j \mathbb{P}(A_i \cap B_j)$$

$$= \sum_{ij} \alpha_i \beta_j \mathbb{P}(A_i)\mathbb{P}(B_j)$$

$$= \left(\sum_i \alpha_i \mathbb{P}(A_i)\right)\left(\sum_j \beta_j \mathbb{P}(B_j)\right)$$

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

We still need to show that all A_i, B_j are independent. Let A_i and B_j be arbitrary sets from X and Y, respectively. Then $A_i \in \sigma(X)$ and $B_j \in \sigma(Y)$ as $A_i = X^{-1}(\{\alpha_i\})$ and $B_j = Y^{-1}(\{\beta_j\})$. Thus, as X and Y are independent, A_i and B_j are.

TODO

(18) ~~Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$ be sub-σ-algebras of $\mathcal{F}$. Let X be a random variable. We show that~~

$$\cancel{\mathbb{E}[X \mid \mathcal{G}_2] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}_1] \mid \mathcal{G}_2] \quad \text{a.s.}}$$

~~Let $f_2 = \mathbb{E}[X \mid \mathcal{G}_2]$, $f_1 = \mathbb{E}[X \mid \mathcal{G}_1]$, and $f_{12} = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}_1] \mid \mathcal{G}_2]$ be $\mathcal{G}_2$-, $\mathcal{G}_1$, and $\mathcal{G}_2$-measurable functions representing the conditional expectations. Note that they are unique a.s.~~

$$\cancel{\int_{G_2} f_2 \, d\mathbb{P} = \int_{G_2} X \, d\mathbb{P} \qquad \int_{G_1} f_1 \, d\mathbb{P} = \int_{G_1} X \, d\mathbb{P}}$$

$$\cancel{\int_{G_2} f_{12} \, d\mathbb{P} = \int_{G_2} f_1 \, d\mathbb{P}}$$

~~for all $G_1 \in \mathcal{G}_1$ and all $G_2 \in \mathcal{G}_2$. As $\mathcal{G}_1 \subseteq \mathcal{G}_2$, we also have~~

$$\cancel{\int_{G_1} f_2 \, d\mathbb{P} = \int_{G_1} X \, d\mathbb{P}}$$

~~for all $G_1 \in \mathcal{G}_1$ and therefore, $f_1 = f_2$ a.s. over $\mathcal{G}_1$.~~

~~TODO~~ $\rightarrow$ see pg. 27

(19) Consider $\Omega = \{-1, 1\}$ $\mathcal{F} = 2^\Omega$, and $\mathbb{P}(A) = |A|/2$ for $A \in \mathcal{F}$. Let $X, Y : \Omega \to \mathbb{R}$ be given as $X(\omega) = \omega$ and $Y(\omega) = -\omega$. Then $\mathbb{E}[X] = \mathbb{E}[Y] = 0$, but $\mathbb{E}[XY] = 1$.

(20) Let $X \geq 0$ be a nonnegative random variable. Then

$$\mathbb{E}[X] = \int_{[0,\infty)} \mathbb{P}(X > x)\, \lambda(dx)$$

Proof. First, note that we can write

$$X(\omega) = \int_{[0,\infty)} \mathbb{1}_{[0, X(\omega))}(x)\, \lambda(dx)$$

as

$$\int_{[0,\infty)} \mathbb{1}_{[0, X(\omega))}(x)\, \lambda(dx) = \int_{[0, X(\omega))} d\lambda = \lambda([0, X(\omega))) = X(\omega).$$

Plugging this into $\mathbb{E}[X]$, we get

$$\begin{aligned}
\mathbb{E}[X] = \int_\Omega X \, d\mathbb{P} &= \int_\Omega \left(\int_{[0,\infty)} \mathbb{1}_{[0, X(\omega)]}(x)\, \lambda(dx) \right) \mathbb{P}(d\omega) \\
&= \int_{[0,\infty)} \left(\int_\Omega \mathbb{1}_{[0, X(\omega)]}(x)\, \mathbb{P}(d\omega) \right) \lambda(dx) \\
&= \int_{[0,\infty)} \left(\int_\Omega \mathbb{1}\{\omega \mid X(\omega) > x\}(\omega)\, \mathbb{P}(d\omega) \right) \lambda(dx) \\
&= \int_{[0,\infty)} \mathbb{P}(X > x)\, \lambda(dx)
\end{aligned}$$

where we used $\mathbb{1}_{[0, X(\omega))}(x) = \mathbb{1}\{\omega \mid X(\omega) > x\}(x)$ as x is in the interval $[0, X(\omega))$ iff $X(\omega) > x$. We can swap the integrals due to Fubini-Tonelli.

$\square$

(Q7) Proof. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, let $\mathcal{G}, \mathcal{G}_1, \mathcal{G}_2 \subseteq \mathcal{F}$ be sub-σ-algebras of $\mathcal{F}$, and let X and Y be integrable random variables.

(i) Let $X \geq 0$. Suppose that $\mathbb{P}(\mathbb{E}[X \mid \mathcal{G}] < 0) > 0$. Then there is an $\epsilon < 0$ and a measurable $A \in \mathcal{G}$ such that $\mathbb{E}[X \mid \mathcal{G}] \leq \epsilon$ on A. But then

$$\int_A \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} \leq \int_A \epsilon \, d\mathbb{P} = \epsilon \, \mathbb{P}(A) < 0$$

while

$$\int_A \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} = \int_A X \, d\mathbb{P} \geq 0. \text{ ↯}$$

Hence, $\mathbb{E}[X \mid \mathcal{G}] \geq 0$ a.s.

(ii) Let $G \in \mathcal{G}$. Clearly, 1 is $\mathcal{G}$-measurable. We have

$$\int_G \mathbb{E}[1 \mid \mathcal{G}] \, d\mathbb{P} = \int_G 1 \, d\mathbb{P}$$

by definition and also, clearly $\int_G 1 \, d\mathbb{P} = \int_G 1 \, d\mathbb{P}$. Hence, by Theorem 2.77, $\mathbb{E}[1 \mid \mathcal{G}] = 1$ a.s.

(iii) Let $G \in \mathcal{G}$, then

$$\begin{aligned}
\int_G \mathbb{E}[X + Y \mid \mathcal{G}] \, d\mathbb{P} &\overset{(*)}{=} \int_G (X + Y) \, d\mathbb{P} \\
&\overset{(+)}{=} \int_G X \, d\mathbb{P} + \int_G Y \, d\mathbb{P} \\
&\overset{(*)}{=} \int_G \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} + \int_G \mathbb{E}[Y \mid \mathcal{G}] \, d\mathbb{P} \\
&\overset{(+)}{=} \int_G \left(\mathbb{E}[X \mid \mathcal{G}] + \mathbb{E}[Y \mid \mathcal{G}] \right) d\mathbb{P},
\end{aligned}$$

where $(*)$ is by definition and $(+)$ is due to linearity. Hence, by Theorem 2.77, we have $\mathbb{E}[X + Y \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{G}] + \mathbb{E}[Y \mid \mathcal{G}]$ a.s. as both $\mathbb{E}[X + Y \mid \mathcal{G}]$ and $\mathbb{E}[X \mid \mathcal{G}] + \mathbb{E}[Y \mid \mathcal{G}]$ are $\mathcal{G}$-measurable.

$\longrightarrow$

(iv) Let Y be $\mathcal{G}$-measurable and suppose $\mathbb{E}[XY]$ exists. Then $Y\mathbb{E}[X|\mathcal{G}]$ is also $\mathcal{G}$-measurable. Let $G \in \mathcal{G}$.

TODO

(v) Let $\mathcal{G}_1 \subseteq \mathcal{G}_2$. Consider $\mathbb{E}[X|\mathcal{G}_1]$ and $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1]$, then

$$\int_G \mathbb{E}[X|\mathcal{G}_1]\, d\mathbb{P} = \int_G X\, d\mathbb{P} \qquad \text{and}$$
$$\int_G \mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1]\, d\mathbb{P} = \int_G \mathbb{E}[X|\mathcal{G}_2]\, d\mathbb{P} \overset{(*)}{=} \int_G X\, d\mathbb{P}$$

for all $G \in \mathcal{G}_1$ where $(*)$ is due to $\mathcal{G}_1 \subseteq \mathcal{G}_2$. Hence, by Theorem 2.77, $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1]$ a.s.

(vii) Let $\mathcal{G} = \{\emptyset, \Omega\}$. We have

$$\int_\emptyset \mathbb{E}[X]\, d\mathbb{P} = 0 = \int_\emptyset X\, d\mathbb{P},$$
$$\int_\Omega \mathbb{E}[X]\, d\mathbb{P} = \mathbb{E}[X] = \int_\Omega X\, d\mathbb{P},$$

and thus $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ a.s.

(vi) Suppose $\sigma(X)$ is independent of $\mathcal{G}_2$ given $\mathcal{G}_1$, i.e., for all $A \in \sigma(X)$ and all $B \in \mathcal{G}_2$,

$$\mathbb{P}(A \cap B|\mathcal{G}_1) = \mathbb{P}(A|\mathcal{G}_1)\, \mathbb{P}(B|\mathcal{G}_1),$$

where $\mathbb{P}(A \cap B|\mathcal{G}_1) = \mathbb{E}[\mathbb{1}_{A\cap B}|\mathcal{G}_1]$, $\mathbb{P}(A|\mathcal{G}_1) = \mathbb{E}[\mathbb{1}_A|\mathcal{G}_1]$, and $\mathbb{P}(B|\mathcal{G}_1) = \mathbb{E}[\mathbb{1}_B|\mathcal{G}_1]$. We need to show that

$$\int_G \mathbb{E}[X|\mathcal{G}_1]\, d\mathbb{P} = \int_G X\, d\mathbb{P}$$

holds for all $G \in \sigma(\mathcal{G}_1 \cup \mathcal{G}_2)$.

TODO

(14) Yes, the definition of the Lebesgue integral can be extended to allow infinite values (for nonnegative random variables).

Definition (Lebesgue-integral for arbitrary nonnegative random variables): Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a measure space and let $X : \Omega \to [0, \infty)$ be a nonnegative measurable function. We define the integral over X as follows:

$$\int_\Omega X \, d\mathbb{P} = \sup \left\{ \int_\Omega h \, d\mathbb{P} \mid h \text{ simple and } 0 \leq h \leq X \right\}$$

If the supremum is not defined, set $\int_\Omega X \, d\mathbb{P} = \infty$.

Clearly, still $\int_\Omega \mathbb{1}_A \, d\mathbb{P} = \mathbb{P}(A)$ for measurable A as $\mathbb{P}(A) < \infty$. For linearity, let X_1, X_2 be nonnegative measurable functions and let $\alpha_1, \alpha_2 \in \mathbb{R}^+$. If $\int X_1 \, d\mathbb{P} < \infty$ and $\int X_2 \, d\mathbb{P} < \infty$, we have linearity as usual. Suppose $\int X_1 \, d\mathbb{P} = \infty$. Then

$$\alpha_1 \int_\Omega X_1 \, d\mathbb{P} + \alpha_2 \int_\Omega X_2 \, d\mathbb{P} = \infty + \alpha_2 \int_\Omega X_2 \, d\mathbb{P} = \infty$$

and

$$\begin{aligned}
&\int_\Omega (\alpha_1 X_1 + \alpha_2 X_2) \, d\mathbb{P} \\
&\quad = \sup \left\{ \int_\Omega h \, d\mathbb{P} \mid h \text{ simple and } 0 \leq h \leq \alpha_1 X_1 + \alpha_2 X_2 \right\} \\
&\quad \geq \sup \left\{ \int_\Omega h \, d\mathbb{P} \mid h \text{ simple and } 0 \leq h \leq X_1 \right\} \\
&\quad = \int_\Omega X_1 \, d\mathbb{P} = \infty,
\end{aligned}$$

and therefore $\int_\Omega (\alpha_1 X_1 + \alpha_2 X_2) \, d\mathbb{P} = \infty$ and we have linearity. (The case for $\int X_2 \, d\mathbb{P} = \infty$ is analogous.)

Thus, the definition can be extended consistently (for nonnegative random variables).