Group Assignment 1

Introduction to Data Science H23

Elise Hammarström Theodora Moldovan Ella Schmidtobreick
Georgios Tsouderos Finn Vaughankraska

January 3, 2024

All group members attempted the proofs/problems individually before meeting. After discussing, we finalized the problems and each of us chose a problem to write up in LaTeX. Lemma 1.14 and the "tower property" were written by Ella, Lemma 2.8 was written by Theodora, Finn and Georgios did property 4 of Lemma 2.18, and Exercise 2.59 was written by Elise.

1 Proof Lemma 1.14

We want to prove the following Lemma:

Lemma 1. *Given a probability triple* $(\Omega, \mathcal{F}, \mathbb{P})$ *then for* $A \in \mathcal{F}$ *with* $\mathbb{P}(A) \neq 0$,

$$\mathbb{P}(\cdot \mid A) : \mathcal{F} \to [0, 1]$$

is a probability measure as in Definition 1.10 over $(\Omega, \mathcal{F})$.

Proof. Let $B \in \mathcal{F}$ be an arbitrary event. By definition 1.13. the following is holds:

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

Show axiom 1 "Something happens":

$$\mathbb{P}(\Omega \mid A) \overset{1.13}{=} \frac{\mathbb{P}(\Omega \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$$

Show axiom 2 "Addition rule": Let $B, C \in \mathcal{F}$ be arbitrary sets with $B \cap C = \emptyset$. Then, the following holds:

$$\begin{aligned}\mathbb{P}(B \cap C \mid A) &\overset{1.13}{=} \frac{\mathbb{P}([B \cup C] \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}([A \cap B] \cup [A \cap C])}{\mathbb{P}(A)} \overset{1.10}{=} \frac{\mathbb{P}(A \cap B) + \mathbb{P}(A \cap C)}{\mathbb{P}(A)} \\ &= \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} + \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(A)} \overset{1.13}{=} \mathbb{P}(B \mid A) + \mathbb{P}(C \mid A)\end{aligned}$$

Remark that from $B \cap C = \emptyset$ it directly follows, that $[A \cap B] \cap [A \cap C] = \emptyset$. □

2 Proof Lemma 2.8

We want to prove the following Lemma:

Lemma 2. *Given a probability triple* $(\Omega, \mathcal{F}, \mathbb{P})$ *and an event* $A \in \mathcal{F}$, *the following properties hold:*

1. $\mathbb{1}_A = 1 - \mathbb{1}_{A^c}$ (*complementation behaves like the probability*) (1)
2. $\mathbb{1}_{A \cap B} = \mathbb{1}_A \cdot \mathbb{1}_B$ (*intersection becomes product*) (2)
3. $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \cdot \mathbb{1}_B$ (*union becomes addition - intersection*) (3)

Proof. To prove that each property holds, we evaluate all cases for each property.

1. Complementation behaves like the probability: $\mathbb{1}_A = 1 - \mathbb{1}_{A^c}$

Case 1: If $\omega \in A$, then $\mathbb{1}_{A(\omega)} = 1$ and $\mathbb{1}_{A^c(\omega)} = 0$. Thus, $\mathbb{1}_{A(\omega)} = 1 - \mathbb{1}_{A^c(\omega)} = 1 - 0 = 1$.

Case 2: If $\omega \notin A$, then $\mathbb{1}_{A(\omega)} = 0$ and $\mathbb{1}_{A^c(\omega)} = 1$. Thus, $\mathbb{1}_{A(\omega)} = 1 - \mathbb{1}_{A^c(\omega)} = 1 - 1 = 0$.

2. Intersection becomes product: $\mathbb{1}_{A\cap B} = \mathbb{1}_A \cdot \mathbb{1}_B$

Case 1: If $\omega \in A \cap B$, then $\mathbb{1}_{A(\omega)} = 1$ and $\mathbb{1}_{B(\omega)} = 1$. Thus, $\mathbb{1}_{A(\omega)} * \mathbb{1}_{B(\omega)} = 1 * 1 = 1 = \mathbb{1}_{A\cap B}$

Case 2a: If $\omega \notin A \cap B$, then $\mathbb{1}_{A(\omega)} = 0$ and $\mathbb{1}_{B(\omega)} = 0$. Thus, $\mathbb{1}_{A(\omega)} * \mathbb{1}_{B(\omega)} = 0 * 0 = 0 = \mathbb{1}_{A\cap B}$
Case 2b: If $\omega \notin A \cap B$, then $\mathbb{1}_{A(\omega)} = 1$ and $\mathbb{1}_{B(\omega)} = 0$. Thus, $\mathbb{1}_{A(\omega)} * \mathbb{1}_{B(\omega)} = 1 * 0 = 0 = \mathbb{1}_{A\cap B}$
Case 2c: If $\omega \notin A \cap B$, then $\mathbb{1}_{A(\omega)} = 0$ and $\mathbb{1}_{B(\omega)} = 1$. Thus, $\mathbb{1}_{A(\omega)} * \mathbb{1}_{B(\omega)} = 0 * 1 = 0 = \mathbb{1}_{A\cap B}$

3. Union becomes addition - intersection: $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \cdot \mathbb{1}_B$

Case 1a: If $\omega \in A \cup B$ and $\omega \notin A \cap B$, then $\mathbb{1}_{A(\omega)} = 1$ or $\mathbb{1}_{B(\omega)} = 1$ and $\mathbb{1}_A \cdot \mathbb{1}_B = 0$, since one of the factors is 0. Thus, $\mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \cdot \mathbb{1}_B = 1 + 0 - 0 = 1 = \mathbb{1}_{A\cup B(\omega)}$

Case 1b: If $\omega \in A \cup B$ and $\omega \in A \cap B$, then $\mathbb{1}_{A(\omega)} = 1$ and $\mathbb{1}_{B(\omega)} = 1$ as well as $\mathbb{1}_A \cdot \mathbb{1}_B = 1$. Thus, $\mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \cdot \mathbb{1}_B = 1 + 1 - 1 = 1 = \mathbb{1}_{A\cup B(\omega)}$

Case 2: If $\omega \notin A \cup B$, then $\mathbb{1}_{A(\omega)} = 0$ and $\mathbb{1}_{B(\omega)} = 0$. Thus, $\mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \cdot \mathbb{1}_B = 0 + 0 - 0 = 0 = \mathbb{1}_{A\cup B(\omega)}$.

□

3 Proof property 4 of Theorem 2.18

Theorem 2.18.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability triple, and let X be an $\mathbb{R}$-valued continuous random variable. Then the following holds:

Property 4:

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

Proof. To prove the equation above, we use property 3 of Theorem 2.18. The following holds:

$$\int_a^b f(u)\,du = F(b) - F(a)$$

Then

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{b\to\infty} F(X = b) - \lim_{a\to-\infty} F(X = a)$$

Then by continuity, $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$. Let's consider two number sequences $\{x_n\}$ and $\{y_n\}$ such that $\{x_n\}$ is decreasing and $x_n \to -\infty$, while $\{y_n\}$ is increasing and $y_n \to \infty$. Take $A_n = \{\xi < x_n\}$ and $B_n = \{\xi < y_n\}$. Since x_n tends monotonically to $-\infty$, the sequence of sets A_n decreases monotonically to $\bigcap A_n = \emptyset$.

For any sequence of events $A_1, A_2, A_3, \ldots$ that is decreasing (i.e., $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$) and their intersection is the empty set (i.e., $\bigcap_{n=1}^{\infty} A_n = \emptyset$), the probability of these events also behaves in a continuous manner, i.e., $\lim_{n\to\infty} P(A_n) = 0$.

Then, by continuity, $P(A_n) \to 0$ as $n \to \infty$, which is the same as $\lim_{n\to\infty} F(x_n) = 0$. This and the monotonicity of $F(x)$ imply that $\lim_{x\to-\infty} F(x) = 0$. Since the sequence $\{y_n\}$ tends monotonically to ∞, the sequence of sets B_n increases to $\bigcup B_n = \Omega$. Thus, the sequence of sets increases to the entire sample space. As a result, $P(B_n) \to 1$. This implies that $\lim_{n\to\infty} F(y_n) = 1$, $\lim_{x\to\infty} F(x) = 1$. Thus,

$$\int_{-\infty}^{\infty} f(x)\,dx = \lim_{b\to\infty} F(X = b) - \lim_{a\to-\infty} F(X = a) = 1 - 0 = 1$$

$\square$

4 Solution Exercise 2.59

X and Y are two independent fair coin tosses (with a two-sided coin), i.e. $X, Y \sim Bernoulli(1/2)$.

$$\text{A coin toss} = \begin{cases} 1 & \text{Heads} \\ 0 & \text{Tails} \end{cases} \tag{4}$$

$$\text{We let } Z = X + Y \tag{5}$$

We want to find the PMF of $Z = X + Y$ given X. As we have two independent coin tosses, X and Y, the probability for each outcome of Heads/Tails is 1/2 respectively:

$$\mathbb{P}(x = 0) = 1/2$$
$$\mathbb{P}(x = 1) = 1/2$$
$$\mathbb{P}(y = 0) = 1/2$$
$$\mathbb{P}(y = 1) = 1/2$$

When X is known, we have two cases:
Case 1: $X = 0$ (tails)
If $X = 0$, there are two possibilities of Y, and therefore also two possibilities of Z.

$$\text{If } X = 0 \text{ and } Y = 0 \Rightarrow Z = 0 \tag{6}$$

$$\text{If } X = 0 \text{ and } Y = 1 \Rightarrow Z = 1 \tag{7}$$

With the probabilities as follows:

$$\mathbb{P}(Z = 0|X = 0) = \frac{\mathbb{P}(Y = 0, X = 0)}{\mathbb{P}(X = 0)} = \frac{\mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 0)}{\mathbb{P}(X = 0)} = \mathbb{P}(Y = 0) = 1/2 \tag{8}$$

$$\mathbb{P}(Z = 1|X = 0) = \frac{\mathbb{P}(Y = 1, X = 0)}{\mathbb{P}(X = 0)} = \frac{\mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 1)}{\mathbb{P}(X = 0)} = \mathbb{P}(Y = 1) = 1/2 \tag{9}$$

Case 2: $X = 1$ (heads)
If $X = 1$, there are two possibilities of Y, and therefore also two possibilities of Z.

$$\text{If } X = 1 \text{ and } Y = 0 \Rightarrow Z = 1 \tag{10}$$

$$\text{If } X = 1 \text{ and } Y = 1 \Rightarrow Z = 2 \tag{11}$$

With the probabilities as follows:

$$\mathbb{P}(Z = 1|X = 1) = \frac{\mathbb{P}(Y = 0, X = 1)}{\mathbb{P}(X = 1)} = \frac{\mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 0)}{\mathbb{P}(X = 1)} = \mathbb{P}(Y = 0) = 1/2 \tag{12}$$

$$\mathbb{P}(Z = 2|X = 1) = \frac{\mathbb{P}(Y = 1, X = 1)}{\mathbb{P}(X = 1)} = \frac{\mathbb{P}(X = 1) \cdot \mathbb{P}(Y = 1)}{\mathbb{P}(X = 1)} = \mathbb{P}(Y = 1) = 1/2 \tag{13}$$

The PMF of Z given X:

$$PMF_{z|x} = \begin{cases} \mathbb{P}(Z = 0|X = 0) = 1/2 \\ \mathbb{P}(Z = 1|X = 0) = 1/2 \\ \mathbb{P}(Z = 1|X = 1) = 1/2 \\ \mathbb{P}(Z = 2|X = 1) = 1/2 \\ 0 \text{ otherwise} \end{cases} \tag{14}$$

And the joint PMF of (Z, X):

$$PMF_{(z,x)} = \begin{cases} \mathbb{P}(Z = 0, X = 0) = 1/4 \\ \mathbb{P}(Z = 1, X = 0) = 1/4 \\ \mathbb{P}(Z = 1, X = 1) = 1/4 \\ \mathbb{P}(Z = 2, X = 1) = 1/4 \\ 0 \text{ otherwise} \end{cases} \tag{15}$$

5 Proof "tower property" (Theorem 2.60) for a discrete random variable taking a finite number of values

Theorem 1 (The tower property)**.** *Let* (X, Y) *be a* $\mathbb{R}^2$ *valued RV where* $\mathbb{E}[X]$ *is well defined. Then*

$$\mathbb{E}\left[\mathbb{E}[X \mid Y]\right] = \mathbb{E}[X]$$

Proof. We want to proof the tower property for discrete RV's. We denote $g(y) = \mathbb{E}[X \mid Y = y]$ and then define

$$\mathbb{E}[X \mid Y] := g(Y)$$

Lets begin with writing down the LHS:

$$\begin{aligned}
\mathbb{E}\left[\mathbb{E}[X \mid Y]\right] &\overset{\text{g(Y)}}{=} \mathbb{E}[g(Y)] && (16)\\
&\overset{2.53}{=} \sum_y g(y) f_Y(y) && (17)\\
&\overset{\text{g(y)}}{=} \sum_y \mathbb{E}[X \mid Y = y] f_Y(y) && (18)\\
&\overset{\text{def}}{=} \sum_y \left(\sum_x x f_{X|Y}(x \mid y) \right) f_Y(y) && (19)\\
&= \sum_y \sum_x x f_{X|Y}(x \mid y) f_Y(y) && (20)\\
&\overset{2.58}{=} \sum_y \sum_x x f_{X,Y}(x, y) && (21)\\
&= \sum_x \sum_y x f_{X,Y}(x, y) && (22)\\
&= \sum_x x \left(\sum_y f_{X,Y}(x, y) \right) && (23)\\
&= \sum_x x f_X(x) \overset{2.30}{=} \mathbb{E}[X] && (24)
\end{aligned}$$

In step 24 the definition of the marginal probability mass function is used. □