Intro DS - Group Assignment 1

Albin Åberg Dahlberg, Emil Wormbs,
Folke Hilding, Oscar Boman, Sofia Alfsson

September 2023

1 Lemma 1.14

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

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

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

The probability measure is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ satisfying the following conditions:

1. The "Something Happens" axiom holds, i.e. $\mathbb{P}(\Omega) = 1$
2. The "Addition Rule" holds, i.e. for $B, C \in \mathcal{F}$

$$B \cap C = \emptyset \qquad \mathbb{P}(B \cup C) = \mathbb{P}(B) + \mathbb{P}(C).$$

For the first property we can write the definition of conditional probabilities as

$$\mathbb{P}(\cdot|A) = \frac{\mathbb{P}(\cdot \cap A)}{\mathbb{P}(A)}.$$

Since $\mathbb{P}(\cdot \cap A) = \mathbb{P}(A)$ we get

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

And such the first property holds. For the second property we use the inclusion-exclusion property to see that

$$\mathbb{P}(B \cup C) = \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(B \cap C)$$

but since $B \cap C = \emptyset$ we get

$$\mathbb{P}(B \cup C) = \mathbb{P}(B) + \mathbb{P}(C) - B \cap C = \mathbb{P}(B) + \mathbb{P}(C) - \emptyset = \mathbb{P}(B) + \mathbb{P}(C)$$

$\square$

2 Lemma 2.8

Property 1.

Given the definition of $\mathbb{1}_A$,

$$\mathbb{1}_A(\omega) = \begin{cases} 0, \omega \notin A \\ 1, \omega \in A, \end{cases}$$

$\mathbb{1}_{A^c}$ can be defined as

$$\mathbb{1}_{A^c}(\omega) = \begin{cases} 1, \omega \notin A \\ 0, \omega \in A. \end{cases}$$

Inserting the definition of $\mathbb{1}_{A^c}$ into the expression $1 - \mathbb{1}_{A^c}$ yields

$$1 - \mathbb{1}_{A^c}(\omega) = \begin{cases} 1-1, \omega \notin A \\ 1-0, \omega \in A \end{cases} = \begin{cases} 0, \omega \notin A \\ 1, \omega \in A \end{cases} = \mathbb{1}_A(\omega).$$

Thus, we have proved that $\mathbb{1}_A = 1 - \mathbb{1}_{A^c}$. $\square$

Property 2.

The indicator function of the intersection of two events can be defined as

$$\mathbb{1}_{A\cap B}(\omega) = \begin{cases} 0, \omega \notin A \cap B \\ 1, \omega \in A \cap B. \end{cases}$$

Thus, if $\omega \in A \cap B$ then $\mathbb{1}_{A\cup B}(\omega) = 1$, and

$$\begin{aligned} &\omega \in A, \omega \in B \\ \Longrightarrow\ &\mathbb{1}_A(\omega) = 1, \mathbb{1}_B(\omega) = 1 \\ \Longrightarrow\ &\mathbb{1}_A(\omega)\mathbb{1}_B(\omega) = 1 = \mathbb{1}_{A\cap B}(\omega). \end{aligned}$$

If $\omega \notin A \cap B$, then either

$$\mathbb{1}_A(\omega) = 1, \mathbb{1}_b(\omega) = 0$$

or

$$\mathbb{1}_A(\omega) = 0, \mathbb{1}_B(\omega) = 1,$$

which implies that

$$\mathbb{1}_A(\omega)\mathbb{1}_B(\omega) = 0 = \mathbb{1}_{A\cap B}(\omega), \omega \notin A \cap B.$$

So the definition of $\mathbb{1}_A(\omega)\mathbb{1}_B(\omega)$ is

$$\mathbb{1}_A(\omega)\mathbb{1}_B(\omega) = \begin{cases} 0, \omega \notin A \cap B \\ 1, \omega \in A \cap B. \end{cases} = \mathbb{1}_{A\cap B}(\omega).$$

Thus, it is proven that $\mathbb{1}_{A\cap B}(\omega) = \mathbb{1}_A \mathbb{1}_B(\omega)$. $\square$

Property 3.
We can define $\mathbb{1}_{A\cup B}$ as

$$\mathbb{1}_{A\cup B}(\omega) = \begin{cases} 0, \omega \notin A \cup B \\ 1, \omega \in A \cup B. \end{cases}$$

Given the proof of property 2 of the indicator function $\mathbb{1}_{A\cap B}$, we can rewrite the expression of $\mathbb{1}_{A\cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B$ as

$$\mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_{A\cap B}.$$

This gives us 4 cases:
Case 1: $\omega \notin A, \omega \notin B$ yields

$$\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A\cap B}(\omega) = 0 + 0 - 0 = 0$$

Case 2: $\omega \in A, \omega \notin B$ yields

$$\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A\cap B}(\omega) = 1 + 0 - 0 = 1.$$

Case 3: $\omega \notin A, \omega \in B$ yields

$$\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A\cap B}(\omega) = 0 + 1 - 0 = 1$$

Case 4: $\omega \in A, \omega \in B$ yields

$$\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_{A\cap B}(\omega) = 1 + 1 - 1 = 1$$

These four cases show that $\mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B = 1$ if $\omega \in A$ or $\omega \in B$, that is if $\omega \in A \cup B$

$$\mathbb{1}_A(\omega) + \mathbb{1}_B(\omega) - \mathbb{1}_A \mathbb{1}_B(\omega) = \begin{cases} 0, \omega \notin A \cup B \\ 1, \omega \in A \cup B \end{cases} = \mathbb{1}_{A\cup B}(\omega).$$

$\square$

3 Property 4 of Theorem 2.18

We know from the definition of probability triples that $\mathbb{P}(\Omega) = 1$. For a continous random variable. The CDF is defined as a non-decreasing, right-continuous function.

$$F : \mathbb{R} \to [0, 1]$$

The integral

$$\int_{-\infty}^{\infty} f(v) dv$$

can be expressed as the limit

$$\lim_{x\to\infty}\int_{-\infty}^{x} f(v)dv = \lim_{x\to\infty} F(x) = \lim_{x\to\infty} \mathbb{P}(X \le x).$$

Since the domain of F is $\mathbb{R}$, then the limit covers all real numbers, that is, our entire sample space.

$$\lim_{x\to\infty} \mathbb{P}(X \le x) = \mathbb{P}(\Omega) = 1$$

$\square$

4 Solve Exercise 2.59

The PMF of Z given X is represented by the following table:

	X = 0	X = 1
Z = 0	0.5	0
Z = 1	0.5	0.5
Z = 2	0	0.5

The definition of conditional PMF,

$$f_{Z|X}(Z|X) = \frac{f_{Z,X}(z,x)}{f_X(x)},$$

can be tweaked to express the joint PMF

$$f_{Z,X}(z,x) = f_{Z|X}(Z|X)f_X(x).$$

Thus, the joint PMF is given by the conditional PMF multiplied by the PMF of the conditional RV. The joint PMF is represented in the following table:

	X = 0	X = 1
Z = 0, X = 0	0.25	0
Z = 1, X = 0	0.25	0
Z = 1, X = 1	0	0.25
Z = 2, X = 1	0	0.25

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

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

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

where $\mathbb{E}[X|Y]$ is a conditional expectation. The tower property as a whole says that the expected value of the conditional expected value of X given Y is the same as the expected value of X. We denote

$$g(y) = [X|Y = y]$$

and then define the expectation of X given Y as a function of Y:

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

Proof for discrete RVs:

$$\mathbb{E}[\mathbb{E}[X|Y]] = \mathbb{E}[g(Y)] = \sum_y g(y) f_Y(y).$$

Expanding $g(y)$ yields

$$\sum_y \mathbb{E}[X|Y = y] f_Y(y),$$

and by inserting the definition of conditional expectation, we get

$$\sum_y \sum_x x f_{X|Y}(x, y) f_Y(y).$$

Since $f_{X|Y}(x|y) f_Y(y) = f_{X,Y}(x, y)$, we can express it as

$$\sum_y \sum_x x f_{X,Y}(x, y).$$

Since we are summing over all y:s, the y-dimension equals to 1 i.e. "Something Happens in y", meaning that

$$\sum_y \sum_x x f_{X,Y}(x, y) = \sum_x x f_X(x) \cdot 1 = \mathbb{E}[X].$$

$\square$