Fund. Math. Topic 2: Probability

Oxbridge Academic Bootcamp 2025

Seppe J. Staelens

August 15, 2025

Introduction

Having become acquainted with the basic concepts of set theory, we will now put additional structure on sets in the form of a probabilistic measure. This forms the basis of probability theory, a mathematical framework to discuss situations that are subject to uncertainty. We will start with some formal definitions and study basic examples, before we discuss Bayes' rule and apply it to the famous Monty Hall problem.

NB: to avoid having to define a σ-algebra, we will only consider finite sample spaces in this course, i.e. sets with a finite number of elements.

Lecture 1: Definitions

1.1 Probability space

Before we can quantify how likely a certain event is to happen, we need to know what all the possibilities are. To this end, we define the **sample space**.

Definition 1. *The* ***sample space*** Ω *is the set of all possible outcomes of a random experiment.*

The prototypical example of a sample space is the set of all possible outcomes of the throw of a die.

Example 1. *The sample space of a die throw is* $\Omega = \{1, 2, 3, 4, 5, 6\}$*, i.e. the set containing the different numbers on a die. An* ***event*** *is a subset of* Ω*, i.e. a set of outcomes that we are interested in. For example, the event of throwing an even number is* $E = \{2, 4, 6\}$.

The collection of all events of interest is referred to as the event space $\mathcal{F}$, a collection of subsets of Ω. Often, the event space will simply be the power set of Ω, i.e. the set of all subsets of Ω. While this does not need to be the case - and sometimes it cannot be the case[1] - we will assume this in the rest of the course unless explicitly stated otherwise.

The question we now want to answer is: what is the probability of a certain event A happening? In the case or our example above, the probability of throwing an even number can be calculated straightforwardly.

[1] Our restriction to finite sample spaces means we need not worry about this. Extending the topics here to infinite spaces requires the introduction of a σ-algebra, a topic beyond the scope of this course.

Under the assumption that all outcomes are equally likely, we can simply count the number of outcomes in the event E and divide it by the total number of outcomes in the sample space Ω. We will denote the probability with $P(E)$, and we have that

$$P(E) = \frac{\#E}{\#\Omega} = \frac{3}{6} = \frac{1}{2}\,.$$

However, this would not have been the case if the die was not fair, i.e. if the outcomes were not equally likely. In that case, we would need more knowledge about the die to determine the probability of an event. The probabilities of different events are determined by the **probability measure**.

Definition 2. *A* ***probability measure*** $P : \mathcal{F} \to [0,1]$ *on a sample space* Ω *is a function from the event space* $\mathcal{F}$ *to the real numbers, such that*

- $P(\Omega) = 1$, *i.e. the probability of the sample space is 1,*
- $P(A) \geq 0$ *for all* $A \in \mathcal{F}$, *i.e. the probability of an event is non-negative,*
- $P(A_1 \cup A_2) = P(A_1) + P(A_2)$ *for all* $A_1, A_2 \in \mathcal{F}$ *such that* $A_1 \cap A_2 = \emptyset$, *i.e. the probability of the union of two disjoint events is the sum of their probabilities.*

Definition 3. *A* ***probability space*** *is a triple* $(\Omega, \mathcal{F}, P)$, *where* Ω *is the sample space,* $\mathcal{F}$ *is the event space and* P *is the probability measure.*

We can now tackle the problem of determining probabilities for a loaded die.

Example 2. *Consider a die with the following probabilities for each outcome, where we understand* $P(i) = P(\{i\})$:

$$\begin{aligned} &P(1) = 1/8, && P(2) = 1/8, && P(3) = 1/8, \\ &P(4) = 1/8, && P(5) = 1/8, && P(6) = 3/8. \end{aligned}$$

We check that this is a valid probability measure, i.e. the probabilities are non-negative and sum to 1:

$$\begin{aligned} P(1) + P(2) + P(3) + P(4) + P(5) + P(6) &= 1/8 + 1/8 + 1/8 + 1/8 + 1/8 + 3/8 \\ \Leftrightarrow P(\{1,2,3,4,5,6\}) &= 1\,. \end{aligned}$$

We can now calculate the probability of throwing an even number:

$$\begin{aligned} P(E) &= P(\{2,4,6\}) \\ &= P(2) + P(4) + P(6) \\ &= 1/8 + 1/8 + 3/8 \\ &= 5/8\,. \end{aligned}$$

Note that this is different from the case of a fair die, where we had $P(E) = 1/2$.

Exercise 1. *Consider a die with the following probabilities:*

$$\begin{aligned} &P(1) = 1/16, && P(2) = 1/16, && P(3) = 1/8, \\ &P(4) = 1/8, && P(5) = 3/8\,. \end{aligned}$$

1. *What does* $P(6)$ *have to be for this to be a valid probability measure?*
2. *What is the probability of throwing an even number?*
3. *What is the probability of throwing a number that is not a multiple of 3?*

1.2 Conditional probability

Sometimes, we will be interested in determining a probability after we already know the outcome of another event. This leads us to the concept of conditional probability.

Example 3. *Consider two people that choose a random number from the set $\{1, 2, 3, 4\}$. Alice chooses a number first, followed by Bob who can not pick the same number as Alice. We will denote the outcome of their choices as (a, b), where a is Alice's choice and b is Bob's choice. The sample space is given by:*

$$\Omega = \begin{pmatrix} (1,2) & (1,3) & (1,4) \\ (2,1) & (2,3) & (2,4) \\ (3,1) & (3,2) & (3,4) \\ (4,1) & (4,2) & (4,3) \end{pmatrix} . \tag{1}$$

Suppose we now want to know the probability that Bob will pick an even number. If Alice would not be picking a number, this probability would be $\frac{2}{4} = \frac{1}{2}$, since there are two even numbers (2 and 4) in the set $\{1, 2, 3, 4\}$.

However, suppose that Alice picks 2. In this case, Bob can only pick from the numbers 1, 3, and 4. Therefore, the probability that Bob picks an even number is now $\frac{1}{3}$. On the other hand, if Alice picks 3, Bob can only pick from the numbers 1, 2, and 4, and the probability that Bob picks an even number is $\frac{2}{3}$.

We will formalize this by defining the **conditional probability** of an event A given another event B.

Definition 4 (Conditional Probability)**.** *The conditional probability of an event A given an event B is defined as:*

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} , \tag{2}$$

provided that $P(B) > 0$.

We can now correctly calculate the conditional probabilities in our example. First, we note that

$$P(\text{Bob picks an even number}|\text{Alice picks 2}) = \frac{P(\text{Bob picks an even number} \cap \text{Alice picks 2})}{P(\text{Alice picks 2})} .$$

For the numerator, we see from Eq. (1) that the only outcome where Bob picks an even number and Alice picks 2 is $(2, 4)$, so we have

$$P(\text{Bob picks an even number} \cap \text{Alice picks 2}) = P((2, 4)) = \frac{1}{12} .$$

Additionally, on the other hand, we have

$$P(\text{Alice picks 2}) = \frac{1}{4} .$$

Therefore, we now find that

$$P(\text{Bob picks an even number}|\text{Alice picks 2}) = \frac{P(\text{Bob picks an even number} \cap \text{Alice picks 2})}{P(\text{Alice picks 2})} = \frac{\frac{1}{12}}{\frac{1}{4}} = \frac{1}{3} .$$

Exercise 2. *Prove now yourself that*

$$P(\text{Bob picks an even number}|\text{Alice picks 3}) = \frac{2}{3} .$$

Lecture 2: Bayes' rule

We start this Lecture by defining the **law of total probability**, which allows us to calculate the probability of an event by considering all possible ways that event can occur. To understand its definition, we first need to define what a **partition** of the sample space is.

Definition 5 (Partition)**.** *A partition of a sample space* Ω *is a collection of disjoint events* $A_1, A_2, \ldots, A_n$ *such that*

- $A_i \cap A_j = \emptyset$ *for all* $i \neq j$ *(the events are disjoint), and*
- $\bigcup_{i=1}^n A_i = \Omega$ *(the events cover the entire sample space).*

This can most easily be visualized as follows:

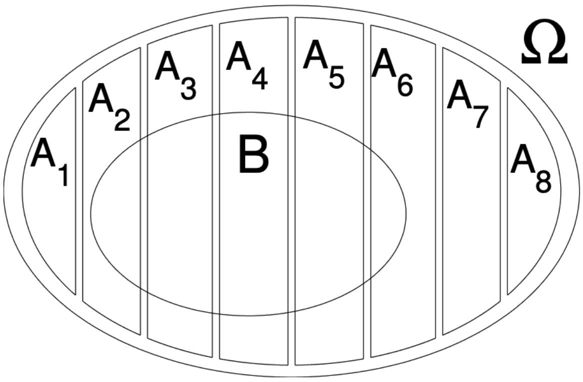

Figure 1: The sets $A_1, A_2, \ldots, A_8$ form a partition of the sample space Ω. The set B is an arbitrary event in the sample space, and overlaps with some of the A_i. *(Figure taken from "Kansrekenen 1" - see References.)*

We are now ready to understand the law of total probability.

Theorem 1 (Law of Total Probability)**.** *Let* $A_1, A_2, \ldots, A_n$ *be a partition of the sample space* Ω*. Then, for any event* $B \subseteq \Omega$*, we have*

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B|A_i)P(A_i)\,. \tag{3}$$

The law of total probability will turn out to be particularly useful in combination with Bayes' rule, which follows from the following observation: the definition of conditional probability (Def. 4) tells us that

$$P(A \cap B) = P(A|B)P(B)\,,$$
$$P(A \cap B) = P(B|A)P(A)\,.$$

However, since for any sets A, B we have that $A \cap B = B \cap A$, it must follow that

$$P(A|B)P(B) = P(B|A)P(A)\,.$$

Rewriting the above is widely known as Bayes' theorem.

Theorem 2 (Bayes' Theorem)**.** *Let* A *and* B *be two events with* $P(B) > 0$*. Then, we have*

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}\,. \tag{4}$$

Bayes' theorem thus allows us to rewrite the conditional probability $P(A|B)$ in terms of the reverse conditional probability $P(B|A)$ and the marginal probabilities $P(A)$ and $P(B)$. This simple formula has inspired an entire new way of doing statistics, known as Bayesian statistics.

Exercise 3. *Using the law of total probability, and assuming the existence of a partition $A_1, A_2, \ldots, A_n$ of Ω, prove that Bayes' theorem can be written as*

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{n} P(B|A_j)P(A_j)} \tag{5}$$

for every $i = 1, 2, \ldots, n$.

The best way to understand all of the above definitions is through an example.

Example 4 (Disease testing)**.** *Suppose there exists some kind of tropical disease, which affects 1 in a 1,000 people. It is known that people with this disease have a very high probability of developing an ear infection, say 90%. However, healthy people without the disease can also develop an ear infection, indistinguishable from the one caused by the disease, with a probability of 0.5% - very unlikely.*

The question we now want to answer is: if a person arrives at the doctor with such an ear infection, what is the probability that they actually have the disease? We will calculate this with Bayes' theorem.

Denote with D the event that a person has the disease, and with H the event that a person is healthy. Denote with E the event that a person has an ear infection. The probability we want to calculate is $P(D|E)$, the probability that a person has the disease given that they have an ear infection. Using Bayes' theorem, we can write this as

$$P(D|E) = \frac{P(E|D)P(D)}{P(E)}\,.$$

Using the law of total probability, we can now calculate $P(E)$ as follows:

$$P(E) = P(E|D)P(D) + P(E|H)P(H)\,.$$

Note that this works because a person either has the disease (D) or does not (H). Using the information given in the problem, we can now calculate the different probabilities:

$$\begin{aligned} P(D) &= \frac{1}{1000}\,, & P(H) &= 1 - P(D) = \frac{999}{1000}\,, \\ P(E|D) &= 0.9\,, & P(E|H) &= 0.005\,. \end{aligned}$$

Therefore, we find that

$$\begin{aligned} P(E) &= 0.9 \cdot \frac{1}{1000} + 0.005 \cdot \frac{999}{1000} \\ &= \frac{5.895}{1000}\,. \end{aligned}$$

We can now plug this into Bayes' theorem to find the probability we are looking for:

$$\begin{aligned} P(D|E) &= \frac{P(E|D)P(D)}{P(E)} \\ &= \frac{0.9 \cdot \frac{1}{1000}}{\frac{5.895}{1000}} \\ &= \frac{0.9}{5.895} \\ &\approx 0.15\,. \end{aligned}$$

Therefore, we find that the probability that a person has the disease given that they have an ear infection is approximately 15%.

2.1 The Monty Hall problem

Monty Hall is a famous game show host, who would ask contestants to choose one of three doors. One of the doors would hide a car, while the other two would hide goats. Clearly, the goal of the contestants is to guess the door that has the car behind it.

Monty would first ask the contestant to choose a door, say door 1. Then, he would open one of the other two doors, say door 3, always revealing a goat. He would then ask the contestant if they want to switch their choice to the other unopened door, door 2 in this case. The contestant can decide to stick with their initial guess, door 1, or switch to door 2. Monty then opens the door corresponding to the final choice, revealing either the car or a goat. While it may seem that the contestant has a 50% chance of winning the car regardless of whether they switch or not, this is not the case. To understand this, we will employ Bayes' theorem.

Without loss of generality, we can assume that the contestant initially chooses door 1. Additionally, we can then also assume that Monty opens door 3, revealing a goat. We want to calculate the probability that door 2 contains the car (D_2), given that Monty opened door 3 (O_3). Using Bayes' theorem (2), we can write this as

$$P(D_2|O_3) = \frac{P(O_3|D_2)P(D_2)}{P(O_3)}\,.$$

In principle, the car could initially be behind any door, such that $P(D_1) = P(D_2) = P(D_3) = \frac{1}{3}$. The probability that Monty opens door 3 given that the car is behind door 2 is $P(O_3|D_2) = 1$: indeed, Monty will not open door 1 (as that is the current choice of the contestant) and he will not open door 2 (as that is where the car is). All that is left to calculate is $P(O_3)$.

To do this, we use the Law of Total Probability (Thm. 1).

$$\begin{aligned} P(O_3) &= P(O_3|D_1)P(D_1) + P(O_3|D_2)P(D_2) + P(O_3|D_3)P(D_3) \\ &= P(O_3|D_1)\cdot\frac{1}{3} + 1\cdot\frac{1}{3} + P(O_3|D_3)\cdot\frac{1}{3}\,, \end{aligned}$$

where we already replaced $P(D_1)$, $P(D_2)$, and $P(D_3)$ with their respective values, as well as $P(O_3|D_2)$ - which we determined earlier. Notice that $P(O_3|D_3) = 0$: if the car is behind door 3, Monty cannot open it. On the other hand, if the car is behind door 1, Monty can open either door 2 or door 3, such that $P(O_3|D_1) = \frac{1}{2}$. Therefore, we find that

$$P(O_3) = \frac{1}{2}\cdot\frac{1}{3} + 1\cdot\frac{1}{3} + 0\cdot\frac{1}{3} = \frac{1}{2}\,.$$

This allows us to calculate that

$$\begin{aligned} P(D_2|O_3) &= \frac{P(O_3|D_2)P(D_2)}{P(O_3)} \\ &= \frac{1\cdot\frac{1}{3}}{\frac{1}{2}} \\ &= \frac{2}{3}\,. \end{aligned}$$

Therefore, we find that the probability that the car is behind door 2, given that Monty opened door 3, is $\frac{2}{3}$.

Exercise 4. *Show that the probability that the car is behind door 1, given that Monty opened door 3, is*

$$P(D_1|O_3) = \frac{P(O_3|D_1)P(D_1)}{P(O_3)} = \frac{1}{3}\,.$$

With the result of the exercise above, we can now conclude that **the contestant should always switch their choice**, as this will increase their chances of winning the car from $\frac{1}{3}$ to $\frac{2}{3}$.

Exercise 5. *Assume the same set-up, but now instead of three doors, there are N doors, with $N \geq 3$.*

1. *Show that the probability that the contestant wins the car by switching their choice is $\frac{1}{N}\frac{N-1}{N-2}$.*
2. *Suppose now additionally that Monty opens p doors with goats, with $1 \leq p \leq N-2$. Show that the probability that the contestant wins the car by switching their choice is $\frac{1}{N}\frac{N-1}{N-p-1}$.*

answer:

We assume the initial pick is door 1, the door $2, \dots, p+1$ are opened, and $p+2, \dots, N$ remain closed. The participant then chooses door $p+2$. Then

$$\begin{aligned} P(D_{p+2}|O_{2\dots p+1}) &= \frac{P(O_{2\dots p+1}|D_{p+2})P(D_{p+2})}{P(O_{2\dots p+1})} \\ &= \frac{\binom{N-2}{p}^{-1}}{(N-p-1)\binom{N-2}{p}^{-1} + \binom{N-1}{p}^{-1}} \end{aligned}$$

Lecture 3: Stochastic variables

In case there is time left on day 5.

References

These notes are based on my own knowledge of these basic mathematical concepts, and the writing has been accelerated by the use of *GitHub copilot* and its implementation in VSCode. Inspiration has been taken from the course notes for *"Kansrekenen I"* (Probabilistic calculus), used in the first year of the Bachelor of Mathematics at the KU Leuven: Prof. Tim Verdonck is the author of the lecture notes.