Chapter 1

Introduction and basic definitions

1.1 Sample space, events, elementary probability

Exercise 1.1

Prove that $\mathcal{P}(\emptyset) = 0$.

Solution of Exercise 1.1: Events $\mathbf{S}$ (where $\mathbf{S}$ is the sample space) and $\emptyset$ are obviously mutually exclusive. Applying axiom A3 we obtain $\mathcal{P}(\mathbf{S} \cup \emptyset) = \mathcal{P}(\mathbf{S}) + \mathcal{P}(\emptyset)$. Because $\mathcal{P}(\mathbf{S}) = 1$ (axiom A2), we have $1 = 1 + \mathcal{P}(\emptyset)$.

Exercise 1.2

Prove that $\mathcal{P}(\overline{A}) = 1 - \mathcal{P}(A)$ for every event A.

Solution of Exercise 1.2: Because $A \cap \overline{A} = \emptyset$, we have that events A and $\overline{A}$ are mutually exclusive. We can apply axiom A3 and obtain that $\mathcal{P}(A \cup \overline{A}) = \mathcal{P}(A) + \mathcal{P}(\overline{A})$. From A2 we have $\mathcal{P}(A \cup \overline{A}) = \mathcal{P}(\mathbf{S}) = 1$ and thus $1 = \mathcal{P}(A) + \mathcal{P}(\overline{A})$.

Exercise 1.3

For any event $A \subseteq \mathbf{S}$, show that $\mathcal{P}(A) \leq 1$.

Solution of Exercise 1.3: Let $\mathbf{E}$ be a subset of $\mathbf{S}$, and let $\mathbf{E}^c$ denote its complement – being all elements of $\mathbf{S}$ that are not in $\mathbf{E}$. Then $\mathbf{E}$ and $\mathbf{E}^c$ are mutually exclusive, and $\mathbf{E} \cup \mathbf{E}^c = \mathbf{S}$. Then

$$
\begin{aligned}
\mathcal{P}(\mathbf{E}) &\leq \mathcal{P}(\mathbf{E}) + \mathcal{P}(\mathbf{E}^c) & (1.1)\\
&= \mathcal{P}(\mathbf{E} \cup \mathbf{E}^c) & (1.2)\\
&= \mathcal{P}(\mathbf{S}) & (1.3)\\
&= 1 & (1.4)
\end{aligned}
$$

where (1.1) holds by axiom A1, (1.2) holds by axiom A3 and (1.4) holds by axiom A2.

Exercise 1.4

Prove that $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$.

Solution of Exercise 1.4: Observe that $A \cup B = A \cup (B \setminus (A \cap B)) = A \cup \overline{\overline{B} \cup (A \cap B)}$. It follows that $\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B \setminus (A \cap B)) = \mathcal{P}(A) + \mathcal{P}(\overline{\overline{B} \cup (A \cap B)}) = \mathcal{P}(A) + 1 - ((1 - \mathcal{P}(B)) + \mathcal{P}(A \cap B))$, since A and $B \setminus (A \cap B)$ are mutually exclusive as well as $\overline{B}$ and $A \cap B$.

Exercise 1.5

Consider a simplified poker game, where each participant obtains 5 cards drawn randomly out of a deck of 52 cards. What is the probability that you obtain

1. one pair (not as a part of a higher combination)?

2. two pairs (not as a part of a higher combination)?

3. flush (not as a part of a higher combination)?

Solution of Exercise 1.5:

1. In our calculation we do not care about the order in which the cards arrive. Let us calculate how many different combinations of 5 cards represent "one pair". There are 13 different card values (ace, 2, 3, ..., queen, king). We have to choose the face value of the pair and we have 13, or $\binom{13}{1}$ possibilities to choose from. When the face value of the card is fixed, we fully specify the two cards in the pair by choosing their suit. For each value, there are 4 possible suits (hearts, spades, diamonds, clubs). We have to choose exactly two suits out of 4 possibilities. Together, there are $\binom{13}{1}\binom{4}{2}$ possible pair.

   The remaining three cards should be chosen so that they do not establish a higher value combination. In this particular case we have to exclude two pair, three of a kind, full house, and four of a kind. This means that all remaining card face values should be different from each other as well as from the value of the pair. There are $\binom{12}{3}$ possibilities. The suit of the card is irrelevant, giving $\binom{12}{3}\binom{4}{1}$ choices.

   Together, there are $\binom{13}{1}\binom{4}{2}\binom{12}{3}\binom{4}{1}$ 5-tuples containing exactly one pair. The total number of the 5-tuples is $\binom{52}{5}$. Hence, the probability to obtain one pair is

$$\frac{\binom{13}{1}\binom{4}{2}\binom{12}{3}\binom{4}{1}}{\binom{52}{5}}.$$

2. $\dfrac{\binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{11}{1}\binom{4}{1}}{\binom{52}{5}}$

3. $\dfrac{\binom{13}{5}\binom{4}{1}-4\times10}{\binom{52}{5}}$

Exercise 1.6

Is there a probability space $(\mathbf{S}, \mathcal{F}, \mathcal{P})$ such that

1. $\mathbf{S} = \mathbb{N}$ and $\mathcal{P}(n) > 0$ for all $n \in \mathbb{N}$?

2. $\mathbf{S} = \mathbb{N}$ and $\mathcal{P}(n) = 0$ for all $n \in \mathbb{N}$?

3. $\mathbf{S} = \mathbb{R}$ and $\mathcal{P}(r) > 0$ for all $r \in \mathbb{R}$?

We use the convention that $0 \notin \mathbb{N}$.

Solution of Exercise 1.6:

1. Yes. Random experiment attaining this property was define during the lecture: we toss a fair coin until we reach head. Sample space of this experiment is $\mathbf{S} = \{H, TH, TTH, TTTH, \dots\}$.

   Define $\mathcal{F} = 2^{\mathbb{N}}$ and $\mathcal{P}(A) = \sum_{i \in A} \frac{1}{2^i}$, i.e. $\mathcal{P}(i) = 2^{-i}$. It is easy to see that $\mathcal{F}$ is a σ-algebra. Axiom A1 is obviously satisfied, axiom A2 is satisfied because $\sum_{i=1}^{\infty} \frac{1}{2^i} = 1$. Let $A_1, A_2, \dots$ be a sequence of mutually exclusive events. We have

$$\sum_{i=1}^{\infty} \mathcal{P}(A_i) = \sum_{i=1}^{\infty} \sum_{j \in A_i} \frac{1}{2^j} = \sum_{j \in \bigcup_{i=1}^{\infty} A_i} \frac{1}{2^j} = \mathcal{P}\left(\bigcup_{i=1}^{\infty} A_i\right),$$

   where the last two equalities follow from the mutual exclusiveness.

2. No. No matter what σ-algebra and what probability we choose, we have that $\sum_{i=1}^{\infty} \mathcal{P}(i)$ is equal to 0, but it is supposed to be equal to 1 (see axioms A2 and A3').

3. No. Suppose we have a σ-algebra $\mathcal{F}$ and a probability assignment $\mathcal{P}$. Now for each $i \in \mathbb{N}$ let G_i denotes the subset of $\mathbb{R}$ such that for all $r \in G_i$ we have $\mathcal{P}(r) > \frac{1}{2^i}$. Obviously, each G_i is finite and each r is in G_i for some i. In this case we can enumerate $\mathbb{R}$ in the following way: First we enumerate G_1, then G_2 etc. But we can't enumerate $\mathbb{R}$, can we?

Exercise 1.7

Three players a,b,c play rock-paper-scissors in a following way: First, a plays with b. Then the winner plays with c. In third round winner of the second round plays with the one who did not play in second round. Generally, in $(i+1)$-th round the winner of i-th round plays with the one who did not play in i-th round. The game proceeds until one of the players wins two consecutive rounds. If we write the play as a sequence of names of players who won respective rounds, atomic events are words of a language $(acb)^*(aa + acc + acbb) + (bca)^*(bb + bcc + bcaa)$ together with two independent words $(acb)^\omega$, $(bca)^\omega$. This defines the sample space S. Now consider a σ-algebra 2^S. Let us have a function that assigns 2^{-k} to each word of length k and 0 to an infinite word. Does this function induce a probability measure on this space? What is the probability that the player a (b,c) wins a play?

Solution of Exercise 1.7:

measure Yes, because an image of the function is subset of $\mathbb{R}^+$ and probability of S is $2 \cdot (\sum_{k=2}^{\infty} 2^{-k} + 0) = 1$

player a An event "a wins" is a set described by a regular expression $(acb)^*aa + (bca)^*bcaa$. Probability of this set is

$$\sum_{k=0}^{\infty} 2^{-(3k+2)} + \sum_{k=0}^{\infty} 2^{-(3k+4)} = (2^{-2} + 2^{-4}) \cdot \sum_{k=0}^{\infty} 2^{-3k} = (2^{-2} + 2^{-4}) \cdot \frac{8}{7}$$

player b symmetric

player c An event "c wins" is a set described by a regullar expression $(acb)^*acc + (bca)^*bcc$. Probability of this set is

$$2 \cdot \sum_{k=0}^{\infty} 2^{-(3k+3)} = 2^{-2} \cdot \frac{8}{7}$$

Exercise 1.8

Let us consider the problem to decide whether the two polynomials $P(x)$ and $Q(x)$ are equal. Polynomials can be specified in a number of ways, namely e.g. in the canonical form as

$$P(x) = p_d x^d + p_{d-1} x^{d-1} + \cdots + p_1 x + p_0, p_i \in \mathbb{R}.$$

In general any expression can be considered to be a polynomial if it is built up from a variable and constants using only addition, subtraction, multiplication, and raising expressions to constant positive integer, e.g. $(x+2)^3 - 11$. An obvious algorithm to decide equality of two polynomials is to transform both input polynomials to the canonical form and compare respective coefficients. This algorithm takes $\mathcal{O}(d^2)$ multiplications, with d being the degree of the polynomial.

Consider the following randomized algorithm to test the polynomial equality:

Computing the values of $P(r)$ and $Q(r)$ is in $\mathcal{O}(d)$. If uniform and random integer in line 1 is generated in $\mathcal{O}(d)$, than the algorithm is in $\mathcal{O}(d)$.

What is the probability that Algorithm 1 gives the wrong answer?

Solution of Exercise 1.8: There are two possible outcomes of the algorithm:

- The algorithm returns 'NO'. This happens only if $P(r) \neq Q(r)$, what implies that polynomials are not equal and the answer 'NO' is always correct.

Algorithm 1 Randomized Polynomial Equality Testing

Require: Two polynomials $P(x)$ and $Q(x)$ of degree at most d.

1: Select uniformly and randomly an integer $r \in_R \{1, \ldots, 100d\}$
2: **if** $P(r) = Q(r)$ **then**
3:     **return** 'YES'
4: **else**
5:     **return** 'NO'
6: **end if**

- The algorithm returns 'YES'. This happens if $P(r) = Q(r)$. Assuming that $P(x) \neq Q(x)$, this happens if $P(r) - Q(r) = 0$, i.e. r is a root of the polynomial $P(x) - Q(x)$. Since $P(x)$ and $Q(x)$ are of degree at most d, the polynomial $P(x) - Q(x)$ is of degree at most d as well. From the fundamental theorem of algebra we have that polynomial with degree d has no more than d roots. Therefore, there are no more than d integers such that $P(r) = Q(r)$, provided that $P(x) \neq Q(x)$.

Let's analyze the probability that the algorithm gives incorrect answer 'YES'. The probability space is the triple $(\mathbf{S}, \mathcal{F}, \mathcal{P})$, where the sample space $\mathbf{S}$ is the set $\{1, \ldots, 100d\}$, $\mathcal{F}$ is e.g. $2^{\mathbf{S}}$ and $\mathcal{P} : \mathcal{F} \to \mathbb{R}$ is probability function that assigns to each sample point uniform probability, i.e. $\forall s \in \mathbf{S} \; \mathcal{P}(s) = \frac{1}{100d}$.

Let E_{PQ} be the set of all roots of $P(x) - Q(x)$ in the range $1 \ldots 100d$. The set E_{PQ} also coincides with the event that the algorithm gives incorrect answer for the input $P(x)$ and $Q(x)$. It remains to calculate

$$\mathcal{P}(E_{PQ}) = |E_{PQ}| \frac{1}{100d} \leq \frac{d}{100d} = \frac{1}{100},$$

since $|E_{PQ}| \leq d$.

It is possible to further improve the probability, either by increasing the sample space e.g. to $1 \ldots 1000d$, or by repeating the algorithm a number of times. In the latter case the probability drops exponentially in the number of executions, i.e. after n repeated executions the probability is at most 100^{-n}. In case we remember random choices of previous executions and use different random number r each time, the probability drops further (see the next lecture).

1.2 Conditional probability, independence

Exercise 1.9

A family has two children. What is the probability that both are boys, given that at least one is a boy?

Solution of Exercise 1.9: The older and younger child may each be male or female, so there are four possible combinations of genders, which we assume to be equally likely. Hence we can represent the sample space in the obvious way as

$$\Omega = \{GG, GB, BG, BB\}$$

where $\text{Prob}(GG) = \text{Prob}(BB) = \text{Prob}(GB) = \text{Prob}(BG) = 1/4$. From the definition of conditional probability,

$$
\begin{aligned}
\mathcal{P}(\text{BB}|\text{one boy at least}) &= \mathcal{P}(\text{BB}|\text{GB} \cup \text{BG} \cup BB) \\
&= \frac{\text{Prob}(\text{BB} \cap (\text{GB} \cup \text{BG} \cup \text{BB}))}{\mathcal{P}(\text{GB} \cup \text{BG} \cup BB)} \\
&= \frac{\mathcal{P}(\text{BB})}{\mathcal{P}(\text{GB} \cup \text{BG} \cup \text{BB})} = \frac{1}{3}.
\end{aligned}
$$

Exercise 1.10

Consider a random experiment "tossing a fair coin twice". Define a sample space, an algebra of events (σ-field) and a probability for this random experiment. Give an example of two independent events, an

example of two events that are not independent, an example of two mutually exclusive events and an example of two events that are not mutually exclusive. What is the probability that the head occurred in the second toss, given that it occurred in at least one toss.

Solution of Exercise 1.10: The sample space is e.g. the set $\mathbf{S} = \{HH, HT, TH, TT\}$, algebra of events is e.g. $2^{\mathbf{S}}$. Probability of event E is equal to $\mathcal{P}(E) = \frac{|E|}{4}$ for all $E \subseteq \mathbf{S}$. Events $A = \{HH, HT\}$ and $B = \{HT, TT\}$ are independent[1], because $\mathcal{P}(A \cap B) = \mathcal{P}(\{HT\}) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathcal{P}(A)\mathcal{P}(B)$. Events A and $B' = \{TH, TT\}$ are not independent, because $\mathcal{P}(A \cap B') = \mathcal{P}(\emptyset) = 0 \neq \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathcal{P}(A)\mathcal{P}(B')$. Example of a pair of mutually exclusive events is $\{HH\}$ and $\{TT\}$. Events, that are not mutually exclusive, are e.g. $\{HT, TT\}$ and $\{HT\}$.

Head occurred in the second toss is the event $A = \{HH, TH\}$, head occurred in at least one toss is $B = \{HH, HT, TH\}$. We calculate

$$\mathcal{P}(A|B) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(B)} = \frac{\mathcal{P}(A)}{\mathcal{P}(B)} = \frac{1}{2}\frac{4}{3} = \frac{2}{3}.$$

Exercise 1.11

There are two roads from A to B and two roads from B to C. Each of the four roads has probability p if being blocked by snow, independently of all the others. What is the probability that there is an open road from A to C?

Solution of Exercise 1.11:

$$
\begin{aligned}
\mathcal{P}(\text{open road}) &= \mathcal{P}((\text{open road from A to B}) \cap (\text{open road from B to C})) \\
&= \mathcal{P}(\text{open road from A to B})\mathcal{P}(\text{open road from B to C})
\end{aligned}
$$

using independence. However, p is the same for all roads,

$$
\begin{aligned}
\mathcal{P}(\text{open road}) &= (1 - \mathcal{P}(\text{no road from A to B}))^2 \\
&= (1 - \mathcal{P}((\text{first road blocked}) \cap (\text{second road blocked})))^2 \\
&= (1 - \mathcal{P}((\text{first road blocked})\mathcal{P}(\text{second road blocked})))
\end{aligned}
$$

using independence. Thus

$$\mathcal{P}(\text{open road}) = (1 - p^2)^2. \tag{1.5}$$

Further suppose that there is also a direct road from A to C, which in independently blocked with probability p. Then by Theorem of Total Probability – restricted to just two events!! – and equation (1.5), we have,

$$
\begin{aligned}
\mathcal{P}(\text{open road}) &= \mathcal{P}(\text{open road}|\text{direct road blocked}).p + \mathcal{P}(\text{open road }|\text{direct road open}).(1 - p) \\
&= (1 - p^2)^2.p + 1.(1 - p).
\end{aligned}
$$

Exercise 1.12

The alarm system at a nuclear power plant is not completely reliable. If there is something wrong with the reactor, the probability that the alarm rings is 0.99. On the other hand, the alarm rings on 0.01 of the days when nothing is actually wrong. Suppose that something is wrong with the reactor only one day out of 100. What is the probability that something is actually wrong if the alarm rings?

Solution of Exercise 1.12: Via Bayes' Rule: Let A = something is wrong with reactor and B = alarm rings. We have $\mathcal{P}(B|A) = 0.99$, $\mathcal{P}(B|A^c) = 0.01$, $\mathcal{P}(A) = 0.01$. We desire $\mathcal{P}(A|B)$ – the probability that the alarm indicates an actual fault. The rule states $\mathcal{P}(A|B) = \frac{\mathcal{P}(B|A).\mathcal{P}(A)}{\mathcal{P}(B)}$. Now $\mathcal{P}(B|A)$ is given – its the reliability of the alarm when something is wrong, i.e., 0.99. Also, we have $\mathcal{P}(A)$ – the frequency of accidents, i.e., 0.01. We need to find $\mathcal{P}(B)$! So using total probability,

$$\mathcal{P}(B) = \mathcal{P}(B|A)\mathcal{P}(A) + \mathcal{P}(B|A^c)\mathcal{P}(A^c).$$

---

[1]alternatives are e.g. $\{HH\}$ and $\emptyset$, or $\{HH\}$ and $\mathbf{S}$

Recall, A and A^c are mutually exclusive (and jointly exhaustive too!). We find

$$
\begin{align}
\mathcal{P}(B|A^c) &= 0.01 \tag{1.6}\\
\mathcal{P}(A^c) &= 1 - \mathcal{P}(A) = 0.99 \tag{1.7}\\
\mathcal{P}(B) &= 0.99 * 0.01 + 0.01 * 0.99 = 2 * 0.01 * 0.99 \tag{1.8}\\
\mathcal{P}(B|A)\mathcal{P}(A) &= 0.99 * 0.01 \tag{1.9}
\end{align}
$$

Finally,
$$
Pr(A|B) = 0.99 * 0.01/(2 * 0.01 * 0.99) = 1/2.
$$

In other words, the alarm is right only half the time!

A second solution! This above way of doing the problem is one many people find somewhat confusing. There is another way and it is the following. Think about ten thousand days of the reactor. On how many days will the reactor have trouble? Here we get $10000 * \mathcal{P}(A) = 100$ days. On how many of those days when theres trouble will the alarm ring? Thats $\mathcal{P}(B|A) * 100$ days $= 99$ days. On the other hand, there are 9900 days when theres no trouble. On how many of them does the alarm ring? Thats $\mathcal{P}(B|A^c) * 9900 = 99$ days. How likely is it that an alarm day is also a problem day? This is $1/2$, because there are as many false alarm days as real alarm days.

Exercise 1.13

Machines A, B, C produce 25%, 35% and 40% of a given commodity. Among the products of machine A, 5% are defective, for B there are 4% and 2% for C. On the condition of taking a defective product, what is the probability that it was produced by machine A, B, C?

Solution of Exercise 1.13: The sample space contains atomic events A_{good}, B_{good}, C_{good}, A_{def}, B_{def} and C_{def}. Algebra of events consists of all subsets of the sample space. We define $D = \{A_{def}, B_{def}, C_{def}\}$ and $A = \{A_{good}, A_{def}\}$ (similarly we define B and C). We are given the probabilities $\mathcal{P}(A) = 0.25$, $\mathcal{P}(B) = 0.35$, $\mathcal{P}(C) = 0.40$, $\mathcal{P}(A_{def}|A) = 0.05$, $\mathcal{P}(B_{def}|B) = 0.04$ and $\mathcal{P}(C_{def}|C) = 0.02$.

First, we are interested in computing the probability $\mathcal{P}(A|D) = \frac{\mathcal{P}(A \cap D)}{\mathcal{P}(D)} = \frac{\mathcal{P}(A_{def})}{\mathcal{P}(D)}$. We have $\mathcal{P}(A_{def}) = \mathcal{P}(A_{def}|A)\mathcal{P}(A) = 0.05 \cdot 0.25 = 0.0125$ and $\mathcal{P}(D) = \mathcal{P}(A_{def}|A)\mathcal{P}(A) + \mathcal{P}(B_{def}|B)\mathcal{P}(B) + \mathcal{P}(C_{def}|C)\mathcal{P}(C) = 0.05 * 0.25 + 0.04 * 0.35 + 0.02 * 0.4 = 0.0345$. Thus, $\frac{\mathcal{P}(A_{def})}{\mathcal{P}(D)} = \frac{0.0125}{0.0345} = \frac{25}{69} \approx 0.36$. The probabilities $\mathcal{P}(B|D)$ and $\mathcal{P}(C|D)$ can be computed similarly.

Exercise 1.14

Give an example of two events A and B that are mutually exclusive, but not independent.

Solution of Exercise 1.14: Let $\mathbf{S} = \{a, b\}$, $\mathcal{F} = 2^{\mathbf{S}}$ and $\mathcal{P}$ is uniquely determined by $\mathcal{P}(a) = \mathcal{P}(b) = \frac{1}{2}$. Events $\{a\}$ and $\{b\}$ are mutually exclusive, but are not independent.

Exercise 1.15

Prove the following statement from the lecture: If A and B are mutually exclusive, then $\mathcal{P}(A \cap B) = 0$. If they are also independent, we obtain either $\mathcal{P}(A) = 0$ or $\mathcal{P}(B) = 0$.

Solution of Exercise 1.15: If A and B are mutually exclusive, then $A \cap B = \emptyset$ and $\mathcal{P}(\emptyset) = 0$. If they are also independent, then $\mathcal{P}(A \cap B) = 0 = \mathcal{P}(A)\mathcal{P}(B)$. But for all real numbers x, y we have $xy = 0$ iff $x = 0$ or $y = 0$.

Exercise 1.16

Give an example of two independent events A and B, such that the conditional probability $\mathcal{P}(A|B)$ is not defined.

Solution of Exercise 1.16: Let $\mathbf{S} = \{a, b\}$, $\mathcal{F} = 2^{\mathbf{S}}$ and $\mathcal{P}(X) = 1$ iff $a \in X$, i.e. $\mathcal{P}(a) = 1$. Let $A = \{a\}$ and $B = \{b\}$. We have $\mathcal{P}(A) = 1$ and $\mathcal{P}(B) = 0$. However, these events are independent, since $\mathcal{P}(A \cap B) = \mathcal{P}(\emptyset) = 0 = 1 \cdot 0 = \mathcal{P}(A) \cdot P(B)$. Finally, using $\mathcal{P}(B) = 0$, we get that the probability $\mathcal{P}(A|B)$ is not defined.

Exercise 1.17

Prove that if $\mathcal{P}(A \cap B) = \mathcal{P}(A)\mathcal{P}(B)$, then $\mathcal{P}(\overline{A} \cap B) = \mathcal{P}(\overline{A})\mathcal{P}(B)$.

Solution of Exercise 1.17: We have $B = (A \cap B) \cup (\overline{A} \cap B)$ and thus $\mathcal{P}(B) = \mathcal{P}(A \cap B) + \mathcal{P}(\overline{A} \cap B)$. From assumption that $\mathcal{P}(A \cap B) = \mathcal{P}(A)\mathcal{P}(B)$ follows $\mathcal{P}(B) = \mathcal{P}(A)\mathcal{P}(B) + \mathcal{P}(\overline{A} \cap B) = (1 - \mathcal{P}(\overline{A}))\mathcal{P}(B) + \mathcal{P}(\overline{A} \cap B) = (\mathcal{P}(B) - \mathcal{P}(\overline{A})\mathcal{P}(B)) + \mathcal{P}(\overline{A} \cap B)$, and, finally, $\mathcal{P}(\overline{A})\mathcal{P}(B)) = \mathcal{P}(\overline{A} \cap B)$.

Exercise 1.18

Let $\{B_1, B_2, \ldots, B_n\}$ be a partition of S such that $\mathcal{P}(B_i) > 0$ for all $1 \le i \le n$. Prove that $\mathcal{P}(A) = \sum_{i=1}^n \mathcal{P}(A|B_i)\mathcal{P}(B_i)$.

Solution of Exercise 1.18: $\sum_{i=1}^n P(A|B_i)P(B_i) = \sum_{i=1}^n \frac{P(A \cap B_i)}{P(B_i)}\mathcal{P}(B_i) = \sum_{i=1}^n P(A \cap B_i)$ and because $B_1, B_2, \ldots, B_n$ are pairwise disjoint, we can apply axiom A3 and obtain $\sum_{i=1}^n \mathcal{P}(A \cap B_i) = \mathcal{P}\left(\bigcup_{i=1}^n (A \cap B_i)\right) = \mathcal{P}(A)$.

Exercise 1.19

Let A, B be two independent events with probabilities $\mathcal{P}(A), \mathcal{P}(B)$. Compute the probability of $A \cup B$.

Solution of Exercise 1.19: Observe that $\overline{A}, \overline{B}$ (the complements of A, B) are independent. Indeed, $\mathcal{P}(\overline{A} \cap \overline{B}) = \mathcal{P}(\overline{(A \cup B)}) = 1 - \mathcal{P}(A \cup B) = 1 - (\mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)) = 1 - (\mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A) \cdot \mathcal{P}(B)) = (1 - \mathcal{P}(A))(1 - \mathcal{P}(B)) = \mathcal{P}(\overline{A}) \cdot \mathcal{P}(\overline{B})$. Now easily $\mathcal{P}(A \cup B) = 1 - \mathcal{P}(\overline{A} \cap \overline{B}) = 1 - \mathcal{P}(\overline{A}) \cdot \mathcal{P}(\overline{B})$. Try to generalize this to n mutually independent events!

Exercise 1.20

Throwing a (six-sided) die until the ace is reached, what is the probability that it will be reached at the 4th or later throw, on the condition that it will not be reached at the first throw?

Solution of Exercise 1.20: The sample space consists of words $\{1, 2, 3, 4, 5, 6\}^n$ for any n. A word of length n is assigned a probability 6^{-n}. Let us denote A the event that the ace does not occur in the first throw. Clearly $\mathcal{P}(A) = 1 - \mathcal{P}(\overline{A}) = 1 - \mathcal{P}(1) = \frac{5}{6}$. Now let us denote B the event of reaching ace in at least 4 throws. We have

$$
\begin{aligned}
\mathcal{P}(B) &= 1 - \mathcal{P}(\overline{B}) = \\
&= 1 - \mathcal{P}\big(\{1\} \cup (\{2,3,4,5,6\} \cdot \{1\}) \cup (\{2,3,4,5,6\}^2 \cdot \{1\})\big) = \\
&= 1 - \big(\mathcal{P}(1) + \mathcal{P}(\{2,3,4,5,6\} \cdot \{1\}) + \mathcal{P}(\{2,3,4,5,6\}^2 \cdot \{1\})\big) = \\
&= 1 - (\tfrac{1}{6} + \tfrac{5}{6}\tfrac{1}{6} + \tfrac{5}{6}\tfrac{5}{6}\tfrac{1}{6}) = \\
&= 1 - \tfrac{91}{216} = \tfrac{125}{216}
\end{aligned}
$$

An alternative way to calculate this probability is using the fact that the ace is reached with probability one. The probability of B is then the probability that it won't be reached in the first three throws, i.e.

$$
\mathcal{P}(B) = \frac{5}{6}\frac{5}{6}\frac{5}{6}.
$$

The probability we are interested in is $\mathcal{P}(B|A) = \frac{\mathcal{P}(A \cap B)}{\mathcal{P}(A)} = \frac{\mathcal{P}(B)}{\mathcal{P}(A)} = \frac{25}{36}$.

Exercise 1.21

Assume we are throwing a die as in Exercise 1.20. On the condition that the ace is reached in an even number of throws, what is the probability of reaching it in the second throw?

Solution of Exercise 1.21: Note that for $x < 1$ it holds that $\sum_{i=0}^\infty x^i = \frac{1}{1-x}$.

Let us denote C the event that the ace is reached in an even throw. Then

$$
\mathcal{P}(C) = \frac{1}{6} \cdot \sum_{i=0}^\infty \left(\frac{5}{6}\right)^{1+2i} = \frac{5}{36} \cdot \sum_{i=0}^\infty \left(\frac{5}{6}\right)^{2i} = \frac{5}{36} \cdot \sum_{i=0}^\infty \left(\frac{25}{36}\right)^i = \frac{5}{36}\frac{1}{1 - \frac{25}{36}} = \frac{5}{36}\frac{36}{11} = \frac{5}{11}.
$$

Denote D the event that the ace is reached in the second step. We have $\mathcal{P}(D) = \frac{5}{6} \cdot \frac{1}{6} = \frac{5}{36}$. Since $D \cap C = D$, we have $\mathcal{P}(D|C) = \frac{\mathcal{P}(D)}{\mathcal{P}(C)} = \frac{11}{36}$.

Exercise 1.22

Let us consider a binary relation $\sim$ defined on algebra of events by $A \sim B$ iff A is independent of B. Is $\sim$ reflexive, symmetric or transitive?

Solution of Exercise 1.22:

It is **not reflexive**. It suffices to choose A with nonzero probability less than 1.

It is **symmetric**. If $A \sim B$, then $\mathcal{P}(A)\mathcal{P}(B) = \mathcal{P}(A \cap B) = \mathcal{P}(B)\mathcal{P}(A)$ and thus $B \sim A$.

It is **not transitive**. From $A \sim B$ and $B \sim A$ we may not obtain $A \sim A$. Select independent A and B such that to A is assigned a nonzero probability less than 1. We have that $A \sim B$ and $B \sim A$, but not $A \sim A$.