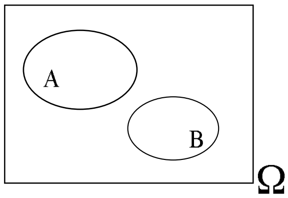
Disjointness:

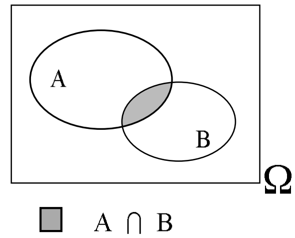


A, B are disjoint

If A and B are disjoint, their intersection is empty, has therefore probability 0:

$$P(A \cap B) = P(\emptyset) = 0.$$

Independence:



If A and B are independent events, the probability of their intersection can be computed as the product of their individual probabilities:

$$P(A \cap B) = P(A) \cdot P(B)$$

If neither of A or B are empty, the probability for the intersection will not be 0 either!

The concept of independence between events can be extended to more than two events:

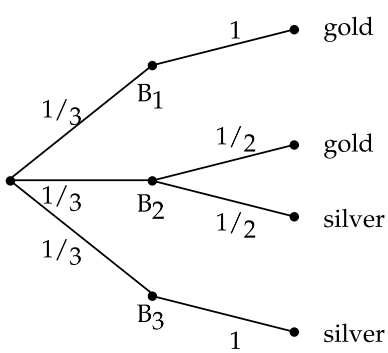
Definition 1.6.2 (Mutual Independence)
A list of events $A_1, \ldots, A_n$ is called *mutually independent*, if for any subset $\{i_1, \ldots, i_k\} \subset \{1, \ldots, n\}$ of indices we have:

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k}).$$

Note: for more than 3 events pairwise independence does not imply mutual independence.

1.7 Bayes' Rule

Example 1.7.1 *Treasure Hunt*
Suppose that there are three closed boxes. The first box contains two gold coins, the second box contains one gold coin and one silver coin, and the third box contains two silver coins. Suppose that you select one of the boxes randomly and then select one of the coins from this box.
What is the probability that the coin you selected is golden?
For a problem like this, that consists of a step-wise procedure, it is often useful to draw a tree (a flow chart) of the choices we can make in each step.
The diagram below shows the tree for the 2 steps of choosing a box first and choosing one of two coins in that box.



The lines are marked by the probabilities, with which each step is done:
Choosing one box (at random) means, that all boxes are equally likely to be chosen: $P(B_i) = \frac{1}{3}$ *for* $i = 1, 2, 3$.
In the first box are two gold coins: A gold coin in this box is therefore chosen with probability 1.
The second box has one golden and one silver coin. A gold coin is therefore chosen with probability 0.5.

How do we piece these information together?
We have two possible paths in the tree, to get a golden coin as a result. Each path corresponds to one event.

$$\begin{aligned} E_1 &= \textit{choose Box 1 and pick one of the two golden coins} \\ E_2 &= \textit{choose Box 2 and pick the golden coin} \end{aligned}$$

We need the probabilities for these two events.
Think: use equation (1.1) to get $P(E_i)$!

$$\begin{aligned} P(E_1) &= P(\textit{ choose Box 1 and pick one of the two golden coins}) = \\ &= P(\textit{ choose Box 1 }) \cdot P(\textit{ pick one of the two golden coins } | B_1) = \\ &= \frac{1}{3} \cdot 1. \end{aligned}$$

and

$$\begin{aligned} P(E_2) &= P(\textit{ choose Box 2 and pick one of the two golden coins}) = \\ &= P(\textit{ choose Box 2 }) \cdot P(\textit{ pick one of the two golden coins } | B_2) = \\ &= \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

The probability to choose a golden coin is the sum of $P(E_1)$ and $P(E_2)$ (since those are the only ways to get a golden coin, as we've seen in the tree diagram).

$$P(\textit{ golden coin }) = \frac{1}{3} + \frac{1}{6} = 0.5.$$

There are several things to learn from this example:

1. Instead of trying to tackle the whole problem, we've divided it into several smaller pieces, that are more manageable (*Divide and conquer* Principle).
2. We identified the smaller parts by looking at the description of the problem with the help of a tree.

   And: if you compare the probabilities on the lines of the tree with the probabilities we used to compute the smaller pieces E_1 and E_2, you'll see that those correspond closely to the branches of the tree.

   The probability of E_1 is computed as the product of all probabilities on the edges from the root to the leaf for E_1.

Definition 1.7.1 (cover)
A set of k events $B_1, \ldots, B_k$ is called a cover of the sample space Ω, if

(i) the events are pairwise disjoint, i.e.

$$B_i \cap B_j = \emptyset \qquad \text{for all } i, j$$

(ii) the union of the events contains Ω:

$$\bigcup_{i=1}^{k} B_i = \Omega$$

What is a cover, then? – You can think of a cover as several non-overlapping pieces, which in total contain every possible case of the sample space, like pieces of a jig-saw puzzle e.g.
Compare with diagram 1.3.

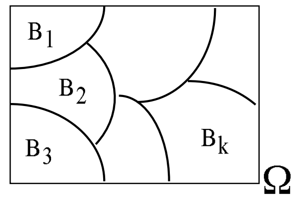


Figure 1.3: $B_1, B_2 \ldots, B_k$ *are a cover of* Ω.

The boxes from the last example, B_1, B_2, and B_3, are a cover of the sample space.

Theorem 1.7.2 **(Total Probability)**
If the set $B_1, \ldots, B_k$ is a cover of the sample space Ω, we can compute the probability for an event A by (cf. fig.1.4):

this is a formal way for "Divide and Conquer"

$$P(A) = \sum_{i=1}^{k} P(B_i) \cdot P(A|B_i).$$

Note: Instead of writing $P(B_i) \cdot P(A|B_i)$ we could have written $P(A \cap B_i)$ - this is the definition of conditional probability cf. def. 1.5.1.

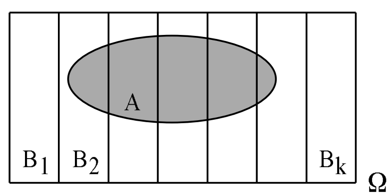
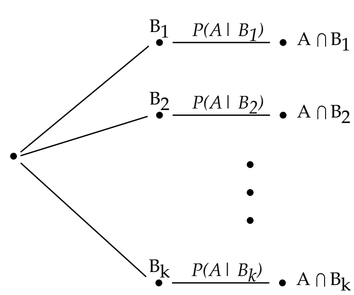


Figure 1.4: *The probability of event A is put together as sum of the probabilities of the smaller pieces (theorem of total probability).*

The challenge in using this Theorem is to identify what set of events to use as cover, i.e. to identify in which parts to dissect the problem.
Very often, the cover $B_1, B_2, \ldots, B_k$ has only two elements, and looks like $E, \bar{E}$.

Tree Diagram:

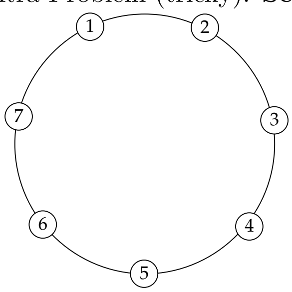


The probability of each node in the tree can be calculated by multiplying all probabilities from the root to the event (1st rule of tree diagrams).
Summing up all the probabilities in the leaves gives $P(A)$ (2nd rule).

Homework: **Powerball - with the Powerball**
Redo the above analysis under the assumption that besides the five numbers chosen from 1 to 49 you choose an additional number, again, between 1 and 49 as the Powerball. The Powerball may be a number you've already chosen or a new one.
You've won, if at least the Powerball is the right number or, if the Powerball is wrong, at least three out of the other five numbers must match.

- Show that the events "Powerball is right", "Powerball is wrong" is a cover of the sample space (for that, you need to define a sample space).
- Draw a tree diagram for all possible ways to win, given that the Powerball is right or wrong.
- What is the probability to win?

Extra Problem (tricky): **Seven Lamps**



A system of seven lamps is given as drawn in the diagram.

Each lamp fails (independently) with probability $p = 0.1$.
The system works as long as not two lamps next to each other fail.
What is the probability that the system works?

Example 1.7.2 *Forensic Analysis*
On a crime site the police found traces of DNA (evidence DNA), which could be identified to belong to the perpetrator. Now, the search is done by looking for a DNA match.
The probability for "a man from the street" to have the same DNA as the DNA from the crime site (random match) is approx. 1: 1 Mio.
For the analysis, whether someone is a DNA match or not, a test is used. The test is not totally reliable, but if a person is a true DNA match, the test will be positive with a probability of 1. If the person is not a DNA match, the test will still be positive with a probability of 1:100000.
Assuming, that the police found a man with a positive test result. What is the probability that he actually is a DNA match?
First, we have to formulate the above text into probability statements.
The probability for a random match is

$$P(\text{ match }) = 1 : 1 \text{ Mio } = 10^{-6}.$$

Now, the probabilities for a positive test result:

$$P(\text{ test pos } \mid \text{ match }) = 1$$
$$P(\text{ test pos } \mid \text{ no match }) = 1 : 100000 = 10^{-5}$$

The probability asked for in the question is, again, a conditional probability. We know already, that the man has a positive test result. We look for the probability, that he is a match. This translates to $P(\text{ match } \mid \text{ test pos. })$.
First, we use the definition of conditional probability to re-write this probability:

$$P(\text{ match } \mid \text{ test pos. }) = \frac{P(\text{ match } \cap \text{ test pos. })}{P(\text{ test pos. })}$$

This doesn't seem to help a lot, since we still don't know a single one of those probabilities. But we do the same trick once again for the numerator:

$$P(\text{ match } \cap \text{ test pos. }) = P(\text{ test pos. } \mid \text{ match }) \cdot P(\text{ match })$$

Now, we know both of these probabilities and get

$$P(\text{ match } \cap \text{ test pos. }) = 1 \cdot 10^{-6}.$$

The denominator is a bit more tricky. But remember the theorem of total probabilities - we just need a proper cover to compute this probability.
The way this particular problem is posed, we find a suitable cover in the events match *and* no match*. Using the theorem of total probability gives us:*

$$P(\text{ test pos. }) = P(\text{ match }) \cdot P(\text{ test pos. } \mid \text{ match }) + P(\text{ no match }) \cdot P(\text{ test pos. } \mid \text{ no match })$$

We have got the numbers for all of these probabilities! Plugging them in gives:

$$P(\text{ test pos. }) = 10^{-6} \cdot 1 + (1 - 10^{-6}) \cdot 10^{-5} = 1.1 \cdot 10^{-}5.$$

In total this gives a probability for the man with the positive test result to be a true match of slightly less than 10%!

$$P(\text{ match } \mid \text{ test pos. }) = 10^{-6} \cdot (1.1 \cdot 10^{-}5.) = 1/11.$$

Is that result plausible? - If you look at the probability for a false positive test result and compare it with the overall probability for a true DNA match, you can see, that the test is ten times more likely to give a positive result than there are true matches.This means that, if 10 Mio people are tested, we would expect 10 people to have a true DNA match. On the other hand, the test will yield additional 100 false positive results, which gives us a total of 110 people with positive test results.
This, by the way, is not a property limited to DNA tests - it's a property of every test, where the overall percentage of positives is fairly small, like e.g. tuberculosis tests, HIV tests or - in Europe - tests for mad cow disease.

Theorem 1.7.3 (Bayes' Rule)
If $B_1, B_2, \ldots, B_k$ is a cover of the sample space Ω,

$$P(B_j|A) = \frac{P(B_j \cap A)}{P(A)} = \frac{P(A|B_j) \cdot P(B_j)}{\sum_{i=1}^{k} P(A|B_i) \cdot P(B_i)} \qquad \text{for all } j \text{ and } \emptyset \neq A \subset \Omega.$$

Example 1.7.3
A given lot of chips contains 2% defective chips. Each chip is tested before delivery.
However, the tester is not wholly reliable:

$$\begin{aligned} P(\text{ "tester says chip is good" } \mid \text{ "chip is good" }) &= 0.95 \\ P(\text{ "tester says chip is defective" } \mid \text{ "chip is defective" }) &= 0.94 \end{aligned}$$

If the test device says the chip is defective, what is the probability that the chip actually is defective?

$$\begin{aligned} P(\underbrace{\textit{chip is defective}}_{:=C_d} \mid \underbrace{\textit{tester says it's defective}}_{:=T_d}) = P(C_d|T_d) &\overset{\textit{Bayes' Rule, use } C_d, \bar{C}_d \textit{ as cover}}{=} \\ &= \frac{P(T_d|C_d)P(C_d)}{P(T_d|C_d)P(C_d) + P(T_d|\bar{C}_d)P(\bar{C}_d)} = \\ &= \frac{0.94 \cdot 0.02}{0.94 \cdot 0.02 + \underbrace{(1 - P(\bar{T}_d|\bar{C}_d)}_{0.05} \cdot 0.98} = 0.28. \end{aligned}$$