2 Conditional probability and independence

In this section we study two key probabilistic concepts: *conditional probability* and *independence*. Conditional probability is a useful tool in understanding dependencies between events. Independence describes an absence of dependence, and so the two are closely connected.

2.1 Conditional probability

A deck of 52 cards is shuffled and you are dealt a card from the top. As in Section 1, we can take our sample space $\Omega = \big\{A\heartsuit, A\diamondsuit, A\clubsuit, A\spadesuit, K\heartsuit, \ldots, 2\heartsuit, 2\diamondsuit, 2\clubsuit, 2\spadesuit\big\}$. As each outcome is equally likely to occur, we take $\mathbb{P}$ to be the uniform distribution on Ω, giving each card probability $1/52$ to appear.

Now suppose that before your card is dealt, you glimpse that $A\heartsuit$ is on the bottom of the deck. This card can no longer appear, so it no longer seems reasonable to assign probability $1/52$ to this outcome... but what should the new probabilities be?

Definition 2.1 (Conditional probability)**.** Let $\mathbb{P}$ be a probability distribution on a sample space Ω and let $A, B \subseteq \Omega$ be events where $\mathbb{P}(B) > 0$. The *conditional probability of A given B* is defined by

$$\mathbb{P}\big(A|B\big) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

For $\mathbb{P}(B) = 0$ then we set $\mathbb{P}\big(A|B\big) := 0$.

Remark 2.2.

- It is a good exercise to show that if $\mathbb{P}(B) > 0$ then the map $A \mapsto \mathbb{P}(A|B)$ is a probability distribution [1]. This is often called the *conditional distribution of $\mathbb{P}$ given B*.
- Intuitively, $\mathbb{P}(A|B)$ is the probability of the event A conditional on the occurrence of event B. Noting that $\mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B)$, we can then view the event 'both A and B occur' as equal to the event 'B occurs' <u>and</u> the event 'A occurs conditional on the occurrence of B'.

Example 2.3. Let's return to the card question above. Here we are conditioning on the event $B = \Omega \setminus \{A\heartsuit\}$, since $A\heartsuit$ is now impossible. Definition 2.1 then gives that the new probability distribution should be $\mathbb{P}(\cdot|B)$. In particular, for any outcome $\omega \neq A\heartsuit$ we have $\mathbb{P}\big(\omega|B\big) = (1/52)/(51/52) = 1/51$. Thus we just get the uniform distribution on the remaining 51 cards (which should seem natural) [2].

Example 2.4. We roll a fair dice. What is the probability that the dice shows a prime given the information that the outcome is at most five?

The sample space $\Omega = \{1, \ldots, 6\}$ and $\mathbb{P}$ is the uniform distribution on Ω. Then writing $A = \{2, 3, 5\}$ and $B = \{1, 2, 3, 4, 5\}$, we are asked for $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B) = (3/6)/(5/6) = 3/5$.

[1] That is, check that the properties from Definition 1.6 hold for $A \to \mathbb{P}(A|B)$, provided $\mathbb{P}(B) > 0$.

[2] This example isn't so artificial. Conditional probability often appears in casino games, e.g. see card counting.

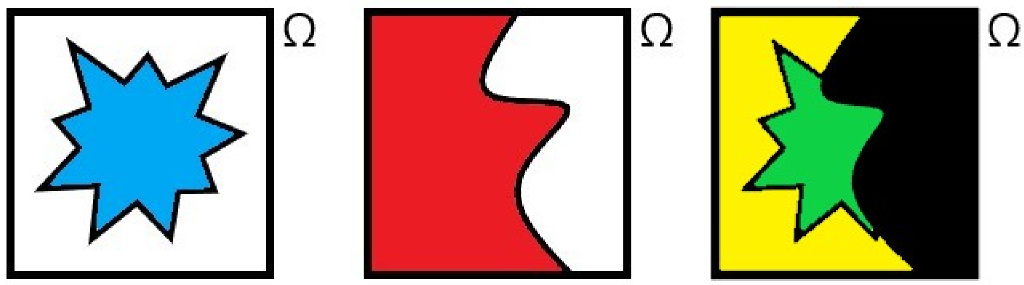

Figure 1: The blue set in the left-hand square denotes an event A, and the red set in the middle square denotes a second event B from the same sample space. Thinking of the areas of these sets as representing their probabilities, the conditional probability $\mathbb{P}(A|B)$ equals to (area of green set)/(area of green and yellow set).

Example 2.5. Mrs Smith has two children. If at least one of them is a girl, what is the probability that both children are girls?

We can take our sample space $\Omega = \big\{(B, B), (B, G), (G, B), (G, G)\big\}$, where the first entry in each pair indicates the older child. The uniform distribution on Ω is the natural choice. The probability that both children are girls, given that at least one is a girl, is equal to [3]

$$\mathbb{P}\big(\{(G, G)\}|\{(G, B), (B, G), (G, G)\}\big) = \frac{\mathbb{P}(\{(G, G)\})}{\mathbb{P}(\{(G, G), (G, B), (B, G)\})} = \frac{1/4}{3/4} = \frac{1}{3}.$$

2.2 The law of total probability and Bayes' formula

The following very useful result relates the probability of an event to certain conditional probabilities.

Theorem 2.6 (Law of total probability)**.** *Let $\mathbb{P}$ be a probability distribution on a sample space Ω. Given events $A, B \subseteq \Omega$, we have*

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

Proof. Note that the sets $A \cap B$ and $A \cap B^c$ are disjoint. By Definition 1.6 (iii)(a), the distribution $\mathbb{P}$ is additive for disjoint events, and so

$$\mathbb{P}(A) = \mathbb{P}\big((A \cap B) \cup (A \cap B^c)\big) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}(A|B^c) \cdot \mathbb{P}(B^c),$$

where the final equality uses the definition of $\mathbb{P}(A|B)$ and $\mathbb{P}(A|B^c)$. □

Remark 2.7. The above result easily generalises. Let $B_1, B_2, \ldots \subseteq \Omega$ be a finite or infinite partition of Ω, i.e. $B_1, B_2, \ldots$ are pairwise disjoint and $\bigcup_{i \geq 1} B_i = \Omega$. Then given any event $A \subseteq \Omega$, we have

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

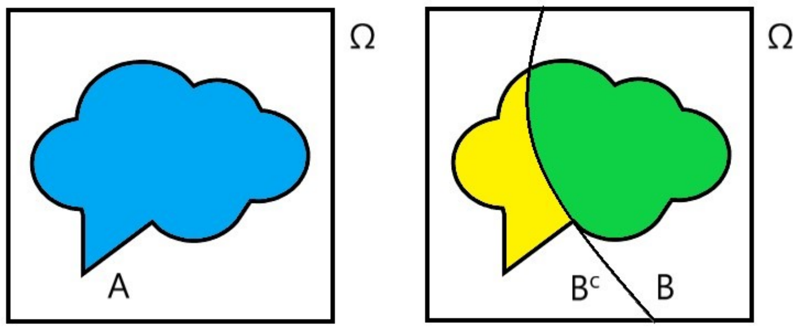

Figure 2: The law of total probability follows almost immediately from decomposing $\mathbb{P}(A)$ into $\mathbb{P}(A \cap B)$ and $\mathbb{P}(A \cap B^c)$ as illustrated.

Example 2.8. If there is ice on a morning then the probability I am late for work is 0.6. Otherwise this probability is just 0.1. Suppose ice is expected on nine days in January. What is the probability I am late to work on a given day in January?

Write L for the event 'I am late at work' and I for the event 'there is ice'. Then we are given $\mathbb{P}(I) = 9/31$, $\mathbb{P}(L|I) = 0.6$ and $\mathbb{P}(L|I^c) = 0.1$. Hence, the desired probability is

$$\mathbb{P}(L) = \mathbb{P}(L|I) \cdot \mathbb{P}(I) + \mathbb{P}(L|I^c) \cdot \mathbb{P}(I^c) = 0.6 \cdot \frac{9}{31} + 0.1 \cdot \frac{22}{31} = 0.245 \quad \text{[3dp]}.$$

Example 2.9 (Phone shop)**.** A shop sells mobile phones from three factories: 60% of the phones are produced at factory $G(enuine)$, 30% are produced at factory $F(ake)$ and 10% at factory $M(ixed)$. Phones from factory G work with probability 0.98. Phones from factory F only work with probability 0.3. At factory M, half of the phones produced are identical to those from G, and the other half are identical to those from factory F. If we buy a phone from the shop, what is the probability it works?

Write A for the event 'the phone works' and B for the event 'the phone is genuine'. Then, $\mathbb{P}(A|B) = 0.98$ and $\mathbb{P}(A|B^c) = 0.3$. Next we calculate $\mathbb{P}(B)$. Let G_d denote the event that 'the phone comes from factory G', F_d denote the event that 'it comes from F' and M_d the event that 'it comes from M'. As G only makes genuine phones, we have $\mathbb{P}(B|G_d) = 1$. Similarly $\mathbb{P}(B|F_d) = 0$, while $\mathbb{P}(B|M_d) = 1/2$. As $\Omega = G_d \cup F_d \cup M_d$ is a partition, by the law of total probability (with remark following it) we find

$$\mathbb{P}(B) = 1 \cdot \mathbb{P}(G_d) + 0 \cdot \mathbb{P}(F_d) + \mathbb{P}(B|M_d) \cdot \mathbb{P}(M_d) = 0.6 + (0.5) \cdot (0.1) = 0.65.$$

Finally, again using the law of total probability, the desired probability is

$$\mathbb{P}(A) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}(A|B^c) \cdot \mathbb{P}(B^c) = (0.98) \cdot (0.65) + (0.3) \cdot (0.35) = 0.742.$$

Theorem 2.10 (Bayes' formula)**.** *Let $A \subseteq \Omega$ be an event with $\mathbb{P}(A) > 0$. Then, for any event $B \subseteq \Omega$*

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \cdot \mathbb{P}(B)}{\mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}(A|B^c) \cdot \mathbb{P}(B^c)}.$$

[3]This problem is the so called *Boy or Girl paradox* dating back to a 1959 article of Martin Gardner. There are many funny variants. For example, what is the probability that both children are girls given that at least one child is a girl born on a Tuesday? The answer now is neither 1/2 nor 1/3. In case you would like to see another interesting counter-intuitive application of conditional probability, see the Monty Hall problem.

Proof. From Definition 2.1 we have

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B) \cdot \mathbb{P}(B)}{\mathbb{P}(A|B) \cdot \mathbb{P}(B) + \mathbb{P}(A|B^c) \cdot \mathbb{P}(B^c)},$$

where the second equality again uses Definition 2.1 and Theorem 2.6. □

Remark 2.11. Again, as in the law of total probability, there is the following more general variant of Bayes' formula: for a finite or infinite partition $B_1, B_2, \ldots \subseteq \Omega$ of Ω and all $i \geq 1$,

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i) \cdot \mathbb{P}(B_i)}{\sum_{k \geq 1} \mathbb{P}(A|B_k) \cdot \mathbb{P}(B_k)}.$$

Example 2.12 (Medical test)**.** Researchers develop a new medical test for a disease. From surveys, it is known that about 1% of the population carries the disease. It is also known that with probability 0.9 the test is positive when applied to an infected person. Unfortunately, the test also occasionally gives a false positive; the test is positive for a non-infected patient with probability 0.05. If a patient tests positive for the disease, what is the probability that they are infected?

Letting A denote the event that 'the test result is positive' and B denote the event that 'the selected person carries disease', we are given that $\mathbb{P}(B) = 0.01$, $\mathbb{P}(A|B) = 0.9$, and that $\mathbb{P}(A|B^c) = 0.05$. Applying Bayes' formula with these values then gives

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)} = \frac{0.9 \cdot 0.01}{0.9 \cdot 0.01 + 0.05 \cdot 0.99} = 0.153\ldots$$

Note: This test initially appears quite effective in screening for the disease, but our calculation gives that in a series of tests of randomly selected people, one can expect that more than five out of six positive test results are false positives.

Example 2.13 (Twins)**.** About 0.3% of all birth events result in identical twins and 0.7% lead to fraternal twins [4]. Identical twins have the same sex, but the sexes of fraternal twins are independent. If two girls are twins, what is the probability that they are fraternal twins?

Let A denote the event that 'the twins are girls' and B the event that 'the twins are fraternal twins'. Then we are given $\mathbb{P}(B) = 0.7, \mathbb{P}(A|B) = 0.25$ and $\mathbb{P}(A|B^c) = 0.5$. Hence, by Bayes' formula,

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)} = \frac{(0.25) \cdot (0.7)}{(0.25) \cdot (0.7) + (0.5) \cdot (0.3)} = \frac{7}{13}.$$

Bayes' formula in court. Bayes' formula and other arguments involving conditional probabilities are often used in court to support arguments. For example, a defendant might argue that the probability that they are guilty given the evidence is very small [5]. A very famous case from the last century, which featured several dubious applications of Bayes' formula, was the O.J. Simpson trial (*California vs Simpson*). You might like to look up *prosecutor's fallacy* on the web to find out more on this topic.

[4]This excludes babies conceived through in-vitro fertilisation where fraternal twin rates are much more likely.

[5]A typical example would be DNA or blood factors recorded at a crime scene, or similarities in paternity tests.

2.3 Independence

Definition 2.14 (Independence)**.** Let $\mathbb{P}$ be a probability distribution on Ω.

(i) Two events $A, B \subseteq \Omega$ are *independent* if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

(ii) More generally, for $k \geq 1$, events $A_1, \ldots, A_k$ are *independent* if for all subsets $\emptyset \neq J \subseteq \{1, 2, \ldots, k\}$, we have

$$\mathbb{P}\Big(\bigcap_{j \in J} A_j\Big) = \prod_{j \in J} \mathbb{P}(A_j).$$

Remark 2.15.

- Independence is a key concept in probability; we will see it reappear throughout the course. Independence simplifies probability calculations, and (happily) this occurs quite often.
- Let $A, B \subseteq \Omega$ with $\mathbb{P}(B) > 0$. Then A, B are independent if and only if the conditional probability $\mathbb{P}(A|B) = \mathbb{P}(A)$. Intuitively, this says that the probability that the event A occurs is unchanged by knowing that B occurs.

Example 2.16. Let A, B be independent events with $\mathbb{P}(A) = 0.4$ and $\mathbb{P}(A \cap B) = 0.1$. Then, $\mathbb{P}(B) = \mathbb{P}(A \cap B)/\mathbb{P}(A) = 0.25$.

Example 2.17 (Archers)**.** Two archers each shoot an arrow at a target. Each archer hits the target with probability 0.3, and the performances are independent. What is the probability that at least one of them hits the target?

Let A_i be the event that the i-th archer hits the target. Then, $\mathbb{P}(A_1) = \mathbb{P}(A_2) = 0.3$. By independence, $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2) = 0.3 \cdot 0.3 = 0.09$. It follows that the desired probability is

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) = 0.51.$$

Example 2.18 (Dice)**.** We roll one dice and consider the events that 'the outcome is at most four' and that 'the outcome is a prime'. Formally, $\Omega = \{1, \ldots, 6\}$, $A = \{1, 2, 3, 4\}$ and $B = \{2, 3, 5\}$.

(a) First suppose the dice are fair. The distribution on Ω is then uniform and we have $\mathbb{P}(A) = |\{1, 2, 3, 4\}|/|\Omega| = 2/3$, $\mathbb{P}(B) = |\{2, 3, 5\}|/|\Omega| = 1/2$ and $\mathbb{P}(A \cap B) = |\{2, 3\}|/|\Omega| = 1/3$. Thus in this case A and B are independent.

(b) Suppose instead that the dice is biased, with $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = 1/4$ while $\mathbb{P}(\{4\}) = \mathbb{P}(\{5\}) = \mathbb{P}(\{6\}) = 1/12$. Then

$$\mathbb{P}(A) = \mathbb{P}(\{2, 3, 5\}) = \frac{7}{12}, \quad \mathbb{P}(B) = \mathbb{P}(\{1, 2, 3, 4\}) = \frac{5}{6}, \quad \text{and} \quad \mathbb{P}(A \cap B) = \mathbb{P}(\{2, 3\}) = \frac{1}{2}.$$

As $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \cdot \mathbb{P}(B)$, A and B are not independent in this case.

Example 2.19 (Dice)**.** We roll ⚅ ⚁ and set $\Omega = \{1, \dots, 6\}^2$. As the two dice do not influence each other, we expect that any events A, B of the form $A = A' \times \{1, \dots, 6\}$ and $B = \{1, \dots, 6\} \times B'$ for given $A', B' \subseteq \{1, \dots, 6\}$ are independent. That is indeed true for the uniform distribution as $\mathbb{P}(A) = 6|A'|/36 = |A'|/6, \mathbb{P}(B) = 6|B'|/36 = |B'|/6$ and $\mathbb{P}(A \cap B) = |A' \times B'|/36 = |A'| \cdot |B'|/36$.

For a more subtle example, let us study the events A: 'sum is equal to 7' and $B = B' \times \{1, \dots, 6\}$ for a given set $B' \subseteq \{1, \dots, 6\}$. Then,

$$\begin{aligned} A &= \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\}, \\ A \cap B &= \{(i, 7-i) : i \in B'\}. \end{aligned}$$

We have $\mathbb{P}(A) = 1/6, \mathbb{P}(B) = |B'|/6$ and $\mathbb{P}(A \cap B) = |B'|/36$. Thus A and B are independent.

Remark 2.20. To show that events $A_1, \dots, A_k \subseteq \Omega$ with $k \geq 3$ are independent it is **not enough** to only show that $\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$ for all $i \neq j$. [6] To see this, consider the following example.

Example 2.21 (Coins)**.** Toss two fair coins. We can take $\Omega = \{T, H\}^2$ and, as the coins are fair, take the uniform distribution $\mathbb{P}$ on Ω. For $i = 1, 2$, let A_i be the event that 'the i-th coin toss shows heads'. Let A_3 be the event that 'the total number of heads is even'.

Then, $\mathbb{P}(A_1) = \mathbb{P}(A_2) = 1/2$ and $\mathbb{P}(A_3) = \mathbb{P}(\{(H,H)\}) + \mathbb{P}(\{(T,T)\}) = 1/4 + 1/4 = 1/2$. Furthermore $A_1 \cap A_2 = A_1 \cap A_3 = A_2 \cap A_3 = \{(H,H)\}$, and so $\mathbb{P}(A_i \cap A_j) = 1/4$ for all $1 \leq i < j \leq 3$. Thus, all three pairs A_i, A_j with $1 \leq i < j \leq 3$ are independent.

However $\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(\{(H,H)\}) = 1/4$ while $\mathbb{P}(A_1)\mathbb{P}(A_2)\mathbb{P}(A_3) = 1/8$, so A_1, A_2, A_3 are not independent.

The following theorem shows that independence transfers to complementary events.

Theorem 2.22 (Independence of complementary events)**.** *Let $A_1, \dots, A_n \subseteq \Omega$ be independent events. Then, taking any $B_i \in \{A_i, A_i^c\}$ for each $i = 1, \dots, n$, the events $B_1, \dots, B_n$ are independent.*

Proof. We only prove the case $n = 2$; the general case follows by induction. Let A, B be independent. We show that A, B^c are independent. Proposition 1.19(ii) and the independence of A, B give

$$\begin{aligned} \mathbb{P}(A \cap B^c) = \mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B) &= \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) \\ &= \mathbb{P}(A)(1 - \mathbb{P}(B)) \\ &= \mathbb{P}(A)\mathbb{P}(B^c). \end{aligned}$$

Hence, A, B^c are independent. By exchanging the roles of A and B, the same argument shows independence of B and A^c. Then, with $A' = A^c$ and $B' = B$, our findings show that $A' = A^c$ and $(B')^c = B^c$ are independent. This concludes the proof for $n = 2$. □

Example 2.23 (Submarine)**.** A submarine will sink if at least two torpedoes hit it. Suppose three torpedos are fired, that each torpedo hits the submarine with probability p and that these events are all independent. What is the probability the submarine survives?

[6] This is a common mistake, so do look out for it.

Let A_i denote the event that the i-th torpedo hits the submarine. From the statement A_1, A_2, A_3 are independent and $\mathbb{P}(A_i) = p$ for all $i \in \{1, 2, 3\}$. If A is the event that the submarine survives then

$$A = \big(A_1^c \cap A_2^c \cap A_3^c\big) \cup \big(A_1 \cap A_2^c \cap A_3^c\big) \cup \big(A_1^c \cap A_2 \cap A_3^c\big) \cup \big(A_1^c \cap A_2^c \cap A_3\big).$$

It can be checked that the four bracketed events are pairwise disjoint, so by Definition 1.19(iii)(a)

$$\mathbb{P}(A) = \mathbb{P}(A_1^c \cap A_2^c \cap A_3^c) + \mathbb{P}(A_1 \cap A_2^c \cap A_3^c) + \mathbb{P}(A_1^c \cap A_2 \cap A_3^c) + \mathbb{P}(A_1^c \cap A_2^c \cap A_3).$$

As A_1, A_2, A_3 are independent, Theorem 2.22 then gives

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3^c) + \mathbb{P}(A_1)\mathbb{P}(A_2^c)\mathbb{P}(A_3^c) + \mathbb{P}(A_1^c)\mathbb{P}(A_2)\mathbb{P}(A_3^c) + \mathbb{P}(A_1^c)\mathbb{P}(A_2^c)\mathbb{P}(A_3)\\ &= (1-p)^3 + 3p(1-p)^2.\end{aligned}$$

Probabilistic ideas have proven very fruitful in many different areas of mathematics. We close this section with an interesting application of independence in number theory.

Example 2.24 (Number theory)**.** Let $n > 1$ be a natural number and $1 < p_1 < p_2 < \ldots < p_\ell \le n$ be all distinct primes dividing n, e.g. if $n = 140 = 2^2 \cdot 5 \cdot 7$ then $\ell = 3$, $p_1 = 2, p_2 = 5, p_3 = 7$. Number theorists are often interested in $\phi(n) := \big|\{1 \le m \le n : \mathrm{hcf}(m,n) = 1\}\big|$. We will prove that

$$\phi(n) = n \cdot \prod_{i=1}^{\ell} \Big(1 - \frac{1}{p_i}\Big). \tag{1}$$

To see this, let $\Omega = \{1, \ldots, n\}$ and $\mathbb{P}$ be the uniform distribution on Ω. For $j = 1, \ldots, \ell$, consider the event

$$A_j = \big\{1 \le m \le n : p_j \text{ divides } m\big\} = \big\{p_j, 2p_j, \ldots, n\big\}.$$

Note that $|A_j| = n/p_j$. However a natural number is divisible by the product of two distinct primes if and only if it is divisible by each prime (this follows from results in 1AC), so

$$A_j \cap A_k = \big\{p_j \cdot p_k, 2p_j \cdot p_k, \ldots, n\big\}, \quad 1 \le j \ne k \le \ell.$$

In particular,

$$\mathbb{P}(A_j \cap A_k) = \frac{n/(p_j \cdot p_k)}{n} = \frac{1}{p_j \cdot p_k} = \frac{1}{p_j} \cdot \frac{1}{p_k} = \mathbb{P}(A_j) \cdot \mathbb{P}(A_k).$$

Hence the events A_j, A_k are independent. You should check that this argument generalises and prove that the events $A_1, \ldots, A_\ell$ are independent.

Let $B = \big\{m \in \Omega : \mathrm{hcf}(m,n) = 1\big\}$ so that $|B| = \phi(n)$. Note that $\mathrm{hcf}(m,n) = 1$ if and only if $p_i \nmid n$ for $i = 1, \ldots, \ell$. Equivalently $B = \bigcap_{i=1}^{\ell} A_i^c$. Hence, using the Theorem 2.22

$$\mathbb{P}(B) = \mathbb{P}(\bigcap_{i=1}^{\ell} A_i^c) = \prod_{i=1}^{\ell} \mathbb{P}(A_i^c) = \prod_{i=1}^{\ell} \left(1 - \frac{1}{p_i}\right).$$

This proves (1), since by definition of the uniform distribution we have $\mathbb{P}(B) = |B|/n = \phi(n)/n$.

Most important takeaways in this chapter. You should

- understand the meaning of conditional probability and know its formal definition,
- be able to compute conditional probabilities in simple scenarios,
- know the law of total probability and Bayes' formula and be able to apply these in standard situations,
- understand the concept of independence and be able to verify independence of events.
- be able to use independence to simplify probability calculations.