Solutions to Exercise Sheet 1

Exercise 1: We have an urn with four balls of colours blue, yellow, green and red, respectively, inside of it. We blindly pick one ball after another without returning the already picked balls to the urn. Create an appropriate probability space for this experiment. How likely is it that we pick green as the third and at the same time red as the fourth ball? Make sure to compute the answer using your specified model.

Solution for Exercise 1: We assign numbers to the colours, identifying (blue, yellow, green, red) with $(1, 2, 3, 4)$. Then we can choose as probability space $\Omega = S_4$, the set of all permutations of the numbers $1, 2, 3, 4$ with $\mathbb{P}$ as the uniform distribution. The interpretation is that for $\sigma \in S_4$ a permutation $\sigma(i)$ encodes the colour chosen in the i-th round. For example, the permutation

$$\sigma(1) = 2\,, \qquad \sigma(2) = 3\,, \qquad \sigma(3) = 1\,, \qquad \sigma(4) = 4\,, \qquad \text{or equivalently} \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 1 & 4 \end{pmatrix}$$

means that we chose the colours yellow, green, blue and red in that order. The uniform distribution is chosen here because none of the sequences of colours we can choose is special. They all have the same probability. Now we compute

$$\mathbb{P}\big(\text{third ball is green and fourth ball is red}\big) = \mathbb{P}\big(\sigma \in S_4 : \sigma_3 = 3 \text{ and } \sigma_4 = 4\big) = \frac{|S_2|}{|S_4|} = \frac{2!}{4!} = \frac{1}{12}\,.$$

Exercise 2: We toss a coin n times. Let A_i be the event that heads is observed in the i-th coin toss. Write the event 'heads is never observed three times in a row' with basic set operations in terms of the events A_i.

Solution for Exercise 2: The intersection $A_k \cap A_{k+1} \cap A_{k+2}$ is the event that heads is the result of the tossed $k, k+1$ and $k+2$. Therefore, the event 'heads is observed three times in a row' is

$$\begin{aligned} &\{\text{heads is observed at tosses } k, k+1 \text{ and } k+2 \text{ for some } k = 1, \ldots, n-2\} \\ &= \cup_{k=1}^{n-2} \left(A_k \cap A_{k+1} \cap A_{k+2}\right). \end{aligned}$$

The event 'heads is never observed three times in a row' is its complement, i.e.

$$\begin{aligned} &\{\text{heads is never observed three times in a row}\} \\ &= \left(\cup_{k=1}^{n-2}(A_k \cap A_{k+1} \cap A_{k+2})\right)^c \\ &= \cap_{k=1}^{n-2}(A_k^c \cup A_{k+1}^c \cup A_{k+2}^c)\,. \end{aligned}$$

Exercise 3: How likely is it that (at least) two students out of $k = 30$ in the tutorial have the same birthday (assume that for each student her/his birthday is uniformly distributed throughout the $n = 365$ days of the year)? On Saturn the year has $n = 10759$ days. To get an estimate on the answer to the sam question on Saturn show that as $k^3/n^2 \to 0$ the probability of a common birthday in the group is approximately $\exp(-\frac{k(k-1)}{2n})$.

Solution for Exercise 3: An appropriate probability space is $\Omega = \{1, \ldots n\}^k$ equipped with the uniform distribution so that for $(\omega_1, \ldots, \omega_k) \in \Omega$ the coordinate ω_i is the birthday of the i-th student.

We consider the complement event that 'All students have different birthdays'. This corresponds to sampling without replacement, i.e.

$$\mathbb{P}\big(\text{All students have different birthdays}\big) = \frac{\#\{\omega \in \Omega : \omega_i \neq \omega_j \ \forall\, i \neq j\}}{|\Omega|} = \frac{n(n-1)\dots(n-k+1)}{n^k}.$$

For the second equality we count the number of all $\omega \in \Omega$ with no entry ω_i appearing twice. As with sampling without replacement this number is $n(n-1)\dots(n-k+1)$ because we have n possibilities for the first entry ω_1, $(n-1)$ possibilities for the second entry ω_2, and so on. Thus, we find for the complement

$$\mathbb{P}\big(\text{Two students have the same birthday}\big) = 1 - \frac{n!}{(n-k)!n^k}.$$

Now we compute the probability of all students having different birthdays in the limit of large n. For that purpose, we write out the product and exponentiate

$$\frac{n(n-1)\dots(n-k+1)}{n^k} = 1\Big(1-\frac{1}{n}\Big)\Big(1-\frac{2}{n}\Big)\dots\Big(1-\frac{k-1}{n}\Big) = \mathrm{e}^{\sum_{l=0}^{k-1}\log\left(1-\frac{l}{n}\right)} = \mathrm{e}^{-\sum_{l=0}^{k-1}\frac{l}{n}+O\left(\frac{k^3}{n^2}\right)}.$$

This is a very useful trick for computing large products because they become sums that are easier to compute. To get

$$\sum_{l=0}^{k-1}\log\big(1-\frac{l}{n}\big) = \sum_{l=0}^{k-1}\frac{l}{n} + O\Big(\frac{k^3}{n^2}\Big)$$

we used the Taylor expansion to first order $\log(1-x) = -x + O(x^2)$. Here $O(y)$ means that this term (coming from the rest term in the Taylor expansion) is bounded by y, i.e. that $|O(y_n)/y_n| < \infty$ for every sequence y_n with $y_n \neq 0$ such that $y_n \to 0$.

Since we have the identity

$$\sum_{l=0}^{k-1} l = \frac{k(k-1)}{2}$$

the estimated probability is

$$\mathbb{P}\big(\text{Two students have the same birthday on Saturn}\big) = 1 - \exp\Big(-\frac{k(k-1)}{2n}\Big) + O\Big(\frac{k^3}{n^2}\Big).$$

Exercise 4: Prove the inclusion-exclusion principle for two events, i.e. show that

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

Solution for Exercise 4: Since $A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$ is a union of three disjoint sets, we conclude

$$\begin{aligned}\mathbb{P}(A \cup B) &\overset{(1)}{=} \mathbb{P}(A \setminus B) + \mathbb{P}(B \setminus A) + \mathbb{P}(A \cap B)\\ &\overset{(2)}{=} \mathbb{P}(A) - \mathbb{P}(A \cap B) + \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B)\\ &= \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).\end{aligned}$$

Here in (1) we used the addition rule for probabilities. In (2) we used $\mathbb{P}(A) = \mathbb{P}(A \setminus B) + \mathbb{P}(A \cap B)$ (as a consequence of the addition rule for probabilities) in the form $\mathbb{P}(A \setminus B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$ for any events A and B.