

---

Problem Sheet 2 <sup>1</sup>

---

**Exercise 1.** To fulfill the requirements for a certain degree, a student can choose to take any 7 out of a list of 20 courses, with the constraint that at least 1 of the 7 courses must be a statistics course. Suppose that 5 of the 20 courses are statistics courses.

1. How many choices are there for which 7 courses to take?
2. Explain intuitively why the answer to part 1. is *not*  $\binom{5}{1} \cdot \binom{19}{6}$ .

**Solution 1.** 1. The result can be derived by either the “complement” approach or by “partitioning”

- *Complement:* Denote the sample space of a student taking exactly 7 courses with  $S$ , and the event of a student taking 7 courses such that at least 1 of them is a statistics course with  $A \subset S$ . By the definition of the complement  $S = (A \cup A^c)$ . As  $A$  and  $A^c$  are mutually exclusive ( $(A \cap A^c) = \emptyset$ ), it follows that  $|S| = |A| + |A^c|$ , implying  $|A| = |S| - |A^c|$ . A student can choose 7 courses out of the 20,  $\binom{20}{7}$  ways.  $A^c \subset S$  corresponds to the event, that a student takes exactly 7 courses, but none of them is a statistics course, so they have to choose 7 out of 15, meaning  $\binom{15}{7}$ . Substituting back to the formula

$$|A| = \binom{20}{7} - \binom{15}{7} = 71085.$$

- *Partitioning:* Denote with  $B_i$  the event, that a student takes exactly 7 courses, such that  $i$ -many of these 7 are statistics courses. Then the previously defined event  $A$  can be partitioned into the disjoint events  $A = B_1 \cup B_2 \cup B_3 \cup B_4 \cup B_5$ , so  $|A| = |B_1| + |B_2| + |B_3| + |B_4| + |B_5|$  follows. Focus on the event  $B_3$ , and the others follow analogously. First, the student must choose 3 statistics courses out of the 5 offered, that they can do  $\binom{5}{3}$  ways. Then they have to pick the remaining  $7 - 3 = 4$  courses from the 15 non-statistics courses, that they can do  $\binom{15}{4}$  different ways. Following the multiplication rule  $|B_3| = \binom{5}{3} \binom{15}{4}$ . Thus

$$|A| = \binom{5}{1} \binom{15}{6} + \binom{5}{2} \binom{15}{5} + \binom{5}{3} \binom{15}{4} + \binom{5}{4} \binom{15}{3} + \binom{5}{5} \binom{15}{2} = 71085$$

2. By  $\binom{5}{1}$  we count the ways the student can pick a statistics course and by  $\binom{19}{6}$  we count how many ways they can pick the remaining 6 courses. However this approach counts picking the course “STAT-A” first and then “STAT-B”, “MATH-1”, “MATH-2”, “MATH-3”, “MATH-4”, “MATH-5” from the remaining  $4 + 15 = 19$  statistics and non-statistics courses, and picking the course “STAT-B” first and then “STAT-A”, “MATH-1”, “MATH-2”, “MATH-3”, “MATH-4”, “MATH-5” as two separate instances, even though they correspond to the same set of 7 courses.

---

<sup>1</sup>Exercises are based on the coursebook Statistics 110: Probability by Joe Blitzstein

**Exercise 2.** A fair die is rolled  $n$  times. What is the probability that at least 1 of the 6 values never appears?

**Solution 2.** Denote the event that the value  $i$  never appears with  $A_i$ . Then the probability of interest is

$$P(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6).$$

We can use the *inclusion-exclusion* formula to rewrite this probability as

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6) &= \sum_{1 \leq i \leq 6} P(A_i) - \sum_{1 \leq i < j \leq 6} P(A_i \cap A_j) \\ &+ \sum_{1 \leq i < j < k \leq 6} P(A_i \cap A_j \cap A_k) - \dots - P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6). \end{aligned}$$

Using the naive definition of probability  $P(A_i) = \frac{5^n}{6^n}$ , as  $n$  many dice throws can have  $6^n$  different outcomes by the multiplication rule, but if we count the throws when the value  $i$  never appeared, a throw can have only 5 favorable outcomes, meaning  $n$  many throws can have  $5^n$  different favorable outcomes. By similar logic  $P(A_i \cap A_j) = \frac{4^n}{6^n}$  for  $i \neq j$ , as the number of favorable outcomes per throw is  $6 - 2 = 4$ , and  $P(A_i \cap A_j \cap A_k) = \left(\frac{3}{6}\right)^n$ , etc. Note that  $P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6) = \left(\frac{6-6}{6}\right)^n = 0$ , but it also follows directly, as  $A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6$  is the event that none of the values appeared at least once, that is clearly impossible. Finally, we have to count the number of terms per sum. For a sum over the intersections involving " $s$ " many events there are  $\binom{6}{s}$  terms, as we have to select the  $s$  many values out of the 6, that we do not want to appear. Hence

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup A_6) &= \binom{6}{1} \left(\frac{5}{6}\right)^n - \binom{6}{2} \left(\frac{4}{6}\right)^n + \binom{6}{3} \left(\frac{3}{6}\right)^n \\ &- \binom{6}{4} \left(\frac{2}{6}\right)^n + \binom{6}{5} \left(\frac{1}{6}\right)^n \\ &= \sum_{i=1}^6 (-1)^{1+i} \binom{6}{i} \left(\frac{6-i}{6}\right)^n. \end{aligned}$$

For example, after 10 throws there is still a 72.8% chance that one of the values never appeared, while after 15 throws, this reduces to 35.6%.

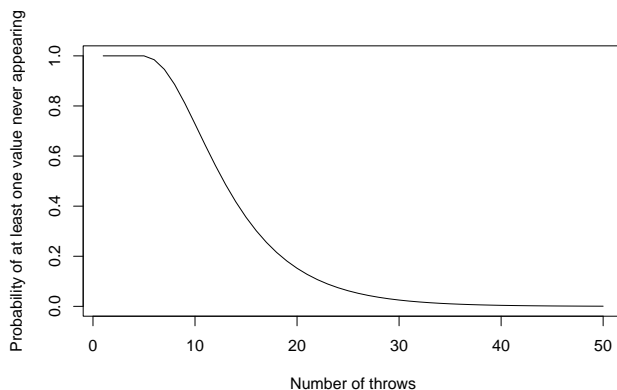


Figure 1: Probability of not seeing at least one of the faces, after  $n$  throws.

**Exercise 3.** A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase “free money” is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention “free money”. What is the probability that it is spam?

**Solution 3.** Let  $S$  be the event that an email is a spam and  $F$  be the event that an email has the “free money” phrase. By Bayes’ rule,

$$P(S|F) = \frac{P(F|S)P(S)}{P(F)} = \frac{0.1 \cdot 0.8}{0.1 \cdot 0.8 + 0.01 \cdot 0.2} = \frac{80}{82}$$

**Exercise 4.** A hat contains 100 coins, where 99 are fair but one is double-headed (always landing Heads). A coin is chosen uniformly at random. The chosen coin is flipped 7 times, and it lands Heads all 7 times. Given this information, what is the probability that the chosen coin is double-headed? (Of course, another approach here would be to *look at both sides of the coin*—but this is a metaphorical coin.)

**Solution 4.** Denote the event of the coin being double-headed with  $DH$ , while the event that the chosen coin lands all 7 times as Heads with  $SH$ . Then the probability of interest is  $P(DH|SH)$  and by Bayes’ rule and the law of total probability,

$$\begin{aligned} P(DH|SH) &= \frac{P(DH, SH)}{P(SH)} \\ &= \frac{P(SH|DH)P(DH)}{P(SH|DH)P(DH) + P(SH|DH^c)P(DH^c)}. \end{aligned}$$

As any of the coins are chosen with equal probability,  $P(DH) = 0.01$ ,  $P(DH^c) = 1 - P(DH) = 0.99$ . If the coin is double-headed, it must always land Heads, so  $P(SH|DH) = 1$ , while if it is a fair coin, then by the multiplication rule  $P(SH|DH^c) = (1/2)^7 = 1/128$ . Therefore

$$\frac{1 \cdot 1/100}{1 \cdot 1/100 + 1/128 \cdot 99/100} = \frac{128}{227} \approx 0.564.$$

**Exercise 5.** A fair coin is flipped 3 times. The toss results are recorded on separate slips of paper (writing “H” if Heads and “T” if Tails), and the 3 slips of paper are thrown into a hat.

1. Find the probability that all 3 tosses landed Heads, given that at least 2 were Heads.
2. Two of the slips of paper are randomly drawn from the hat and both the letter H. Given this information, what is the probability that all 3 tosses landed Heads?

**Solution 5.** 1. Denote with  $C$  the number of tosses that landed Heads. Then the probability of interest is  $P(C = 3|C \geq 2)$ . By the definition of conditional probability

$$P(C = 3|C \geq 2) = \frac{P(C = 3 \cap C \geq 2)}{P(C \geq 2)} = \frac{P(C = 3)}{P(C = 3) + P(C = 2)} = \frac{1/8}{1/8 + 3/8} = \frac{1}{4},$$

where we used the fact that the event  $C = 3$  is a subset of the event  $C \geq 2$ , and that the  $C \geq 2$  can be written as the disjoint union of the events  $C = 3$  and  $C = 2$ . The probabilities  $P(C = c)$  can be calculated as

$$P(C = c) = \binom{3}{c} \left(\frac{1}{2}\right)^c \left(\frac{1}{2}\right)^{3-c},$$

because out of three coins the “c” many that landed as Heads can be picked  $\binom{3}{c}$  ways, and following the multiplication rule, those landed as Heads with probability  $(\frac{1}{2})^c$  while the remaining  $3 - c$  coins landed as Tails with probability  $(\frac{1}{2})^{3-c}$ .

2. In addition to the previous notation, denote the number of slips that have the letter “H” on the two slips drawn with  $S$ . Then the probability of interest is  $P(C = 3|S = 2)$ . By Bayes’ rule and the law of total probability

$$P(C = 3|S = 2) = \frac{P(S = 2|C = 3)P(C = 3)}{\sum_{c \in \{0,1,2,3\}} P(S = 2|C = c)P(C = c)}.$$

If less than two of the tosses landed as Heads ( $C = 0$  and  $C = 1$ ), then clearly the two slips drawn cannot have the letter “H” on them, hence  $P(S = 2|C = 0) = P(S = 2|C = 1) = 0$ . Conversely, if all tosses landed as Heads, then all slips have the letter “H” on them, hence the two slips drawn will certainly be with the letter “H”, thus  $P(S = 2|C = 3) = 1$ . Lastly, if one of the slips has the letter “T” on them, then the probability that we picked the other two is  $\frac{1}{3}$ , as we can pick 2 slips out of the 3,  $\binom{3}{2} = 3$  ways, and we can pick two slips with the letter “H” only one way, therefore  $P(S = 2|C = 2) = \frac{1}{3}$ . Calculating the probabilities  $P(C = c)$  as before,

$$P(C = 3|S = 2) = \frac{1 \cdot 1/8}{1 \cdot 1/8 + 1/3 \cdot 3/8} = \frac{1}{2}.$$

By having access to the information that we drew 2 slips with the letter “H” we not only know that at least two of the flips landed as Heads, but we also know that drawing 2 “H” out of 3 “H”-s is three times as likely than drawing 2 ‘H” out of 2 “H” and 1 “T”, hence the conditional probability is “upweighted”.

**Exercise 6.** A bag contains one marble which is either green or blue, with equal probabilities. A green marble is put in the bag (so there are 2 marbles now), and then a random marble is taken out. The marble taken out is green. What is the probability that the remaining marble is also green?

**Solution 6.** Let  $A$  be the event that the initial marble is green,  $B$  be the event that the removed marble is green, and  $C$  be the event that the remaining marble is green. We need to find  $P(C|B)$ . There are several ways to find this; one natural way is to condition on whether the initial marble is green:

$$P(C|B) = P(C|B, A)P(A|B) + P(C|B, A^c)P(A^c|B) = 1 \cdot P(A|B) + 0 \cdot P(A^c|B)$$

To find  $P(A|B)$ , use Bayes’ Rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} = \frac{1 \cdot 0.5}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{0.5}{1 \cdot 0.5 + 0.5 \cdot 0.5} = \frac{2}{3}.$$

So  $P(C|B) = \frac{2}{3}$ .