Probability and Statistics

Exercise sheet 2 - Solutions

MC 2.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Which of the following statements is true? (Exactly one answer is correct.)

(a) $\forall A, B \in \mathcal{F} : \mathbb{P}[A \cap B] = \mathbb{P}[A]\mathbb{P}[B]$.

(b) $\forall A, B \in \mathcal{F} : \mathbb{P}[A \cap B] \geq \mathbb{P}[A]\mathbb{P}[B]$.

(c) $\forall A, B \in \mathcal{F} : \mathbb{P}[A \cap B] \leq \mathbb{P}[A]\mathbb{P}[B]$.

(d) None of these (in)equalities always hold.

Solution: (d) is true. For example, take $A = B$ with $\mathbb{P}[A] = 0.5$. Then, $\mathbb{P}[A \cap B] = \mathbb{P}[A] = 0.5$ and $\mathbb{P}[A]\mathbb{P}[B] = 0.5 \times 0.5 = 0.25$. This disproves (a) and (c). By choosing A, B disjoint with $\mathbb{P}[A] > 0$ and $\mathbb{P}[B] > 0$, we disprove (b).

MC 2.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $A, B \in \mathcal{F}$ with $\mathbb{P}[A] = 0.5$ and $\mathbb{P}[B] = 0.8$. Which of the following statements is always true? (The number of correct answers is between 0 and 4.)

(a) $\mathbb{P}[A \cap B] = 0.5$.

(b) $\mathbb{P}[A \cap B] \leq 0.5$.

(c) $\mathbb{P}[A \cap B] \geq 0.3$.

(d) $\mathbb{P}[A \cap B] \geq 0.5$.

Solution: (b) and (c) are correct. We have $A \cap B \subseteq A$, which gives $\mathbb{P}[A \cap B] \leq \mathbb{P}[A] = 0.5$. Furthermore,

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cup B] = 0.5 + 0.8 - \mathbb{P}[A \cup B] \geq 0.5 + 0.8 - 1 = 0.3.$$

(a) and (d) are generally not true, which can be verified by counterexamples.

MC 2.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $A, B \in \mathcal{F}$. In which cases does the following hold?

$$\mathbb{P}[A \setminus B] = \mathbb{P}[A] - \mathbb{P}[B].$$

(The number of correct answers is between 0 and 4.)

(a) Always.

(b) If $B \subseteq A$.

(c) If $\mathbb{P}[A] \geq \mathbb{P}[B]$.

(d) If $A = \Omega$.

Solution: (b) and (d) are correct. If $B \subseteq A$, then $A = (A \setminus B) \cup B$ (disjoint union), which gives $\mathbb{P}[A] = \mathbb{P}[A \setminus B] + \mathbb{P}[B]$. This is exactly the equality in (b). (d) is a special case of (b). (a) and (c) are generally not true.

Exercise 2.4. Let A and B be two events with

$$\mathbb{P}[A^c] = \frac{1}{2}, \quad \mathbb{P}[B^c] = \frac{1}{2}, \quad \mathbb{P}[A^c \cap B^c] = p.$$

(a) Find as a function of p the probabilities $\mathbb{P}[A \cap B]$, $\mathbb{P}[A \cap B^c]$ and $\mathbb{P}[A^c \cap B]$. What are the possible values of p for the above to be well-defined?

(b) Find as a function of p the probability that at most i of the two events A and B occur, where $i \in \{0, 1, 2\}$.

Solution:

(a) Since

$$(A \cup B)^c = A^c \cap B^c$$

and using the formula from the lecture notes,

$$\mathbb{P}[A \cap B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A \cup B] = (1 - \mathbb{P}[A^c]) + (1 - \mathbb{P}[B^c]) - (1 - \mathbb{P}[A^c \cap B^c]),$$

we get:

$$\mathbb{P}[A \cap B] = \frac{1}{2} + \frac{1}{2} - (1 - p) = p.$$

Since $A \cap B^c$ and $A \cap B$ are disjoint and $(A \cap B^c) \cup (A \cap B) = A$, we obtain:

$$\mathbb{P}[A \cap B^c] = \mathbb{P}[A] - \mathbb{P}[A \cap B] = \frac{1}{2} - p,$$

and similarly,

$$\mathbb{P}[A^c \cap B] = \mathbb{P}[B] - \mathbb{P}[A \cap B] = \frac{1}{2} - p.$$

As all probabilities must lie in $[0, 1]$, p must lie in the interval $\left[0, \frac{1}{2}\right]$.

(b) We define the event

$$C_i := \{\text{At most } i \text{ of the two events } A \text{ and } B \text{ occur}\}$$

for all $i \in \{0, 1, 2\}$. We get:

$$\mathbb{P}[C_0] = \mathbb{P}[A^c \cap B^c] = p,$$

$$\mathbb{P}[C_1] = \mathbb{P}[C_0] + \mathbb{P}[A \cup B] - \mathbb{P}[A \cap B] = p + 1 - p - p = 1 - p,$$

and

$$\mathbb{P}[C_2] = \mathbb{P}[C_1] + \mathbb{P}[A \cap B] = 1 - p + p = 1.$$

Exercise 2.5. [Radio signals] Four signals are transmitted sequentially over a communication channel. Each signal is transmitted either correctly or incorrectly. We define the sample space Ω as the set of all 0-1 sequences of length 4 as follows:

$$\Omega = \{\omega = (x_1, x_2, x_3, x_4) \mid x_1, x_2, x_3, x_4 \in \{0, 1\}\},$$

i.e.,

$$\Omega = \{(0,0,0,0), (1,0,0,0), (0,1,0,0), \dots, (1,1,1,1)\},$$

and we interpret $x_i = 1$ as "the i-th signal is transmitted correctly" and $x_i = 0$ as "the i-th signal is transmitted incorrectly" for $i = 1, \dots, 4$. We consider the following events:

- $A \coloneqq \{\text{Exactly one signal is transmitted incorrectly}\}$,
- $B \coloneqq \{\text{At least two signals are transmitted correctly}\}$,
- $C \coloneqq \{\text{At most two signals are transmitted correctly}\}$.

(a) Write the events A, B, and C as subsets of Ω.

(b) Describe in words the events $B \cap C$, $A \cup B$, and $A^c \cap C^c$.

(c) Calculate the probabilities of the events A, B, and C assuming that all elementary events $(x_1, x_2, x_3, x_4) \in \Omega$ are equally likely. What model are we using here?

Solution:

(a) We have

$$\begin{aligned} A &= \{(0,1,1,1),\ (1,0,1,1),\ (1,1,0,1),\ (1,1,1,0)\}, \\ B &= \{(1,1,1,1),\ (0,1,1,1),\ (1,0,1,1),\ (1,1,0,1),\ (1,1,1,0), \\ &\qquad (0,0,1,1),\ (0,1,0,1),\ (0,1,1,0),\ (1,0,0,1),\ (1,0,1,0),\ (1,1,0,0)\}, \\ C &= \{(0,0,0,0),\ (1,0,0,0),\ (0,1,0,0),\ (0,0,1,0),\ (0,0,0,1), \\ &\qquad (1,1,0,0),\ (1,0,1,0),\ (1,0,0,1),\ (0,1,1,0),\ (0,1,0,1),\ (0,0,1,1)\}. \end{aligned}$$

Remark: C can equivalently be formulated as "At least two signals are transmitted incorrectly," making the representation of C as a set completely symmetric to that of B — one only needs to swap 0 and 1 everywhere.

(b)
- $B \cap C = \{\text{Exactly two signals are transmitted correctly}\}$.
- $A \cup B = \{\text{At least two signals are transmitted correctly}\}$ (note that $A \subseteq B$).
- $A^c = \{\text{Either no or at least two signals are transmitted incorrectly}\}$.
  Alternative formulation: "All or at most two signals are transmitted correctly."
  Alternative formulation: "0, 1, 2, or 4 signals are transmitted correctly."
  Alternative formulation: "0, 2, 3, or 4 signals are transmitted incorrectly."
- $C^c = \{\text{At least three signals are transmitted correctly}\}$
  Alternative formulation: "3 or 4 signals are transmitted correctly."
- $A^c \cap C^c = \{\text{All four signals are transmitted correctly}\}$.

(c) Ω has $2^4 = 16$ elements. Since all elementary events (x_1, x_2, x_3, x_4) are equally likely and the probabilities must sum up to one, each elementary event has the probability $\frac{1}{16}$.

To solve this problem rigorously, we need to define a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Since for each $\omega \in \Omega$ the set $\{\omega\}$ must be in the σ-algebra $\mathcal{F}$ (to be able to assign a probability to it) and since Ω is finite, $\mathcal{F}$ must contain all subsets of Ω, i.e., $\mathcal{F} = 2^{\Omega}$. Therefore, we use a Laplace model.

- The sample space $\Omega = \{\omega = (x_1, x_2, x_3, x_4) \mid x_1, x_2, x_3, x_4 \in \{0, 1\}\}$ is already given in the problem statement.
- The σ-algebra $\mathcal{F} = 2^\Omega$ is the power set of Ω.
- The probability measure is then defined as the mapping

$$\mathbb{P} : \mathcal{F} \longrightarrow [0, 1], \quad D \longmapsto \mathbb{P}[D] := \frac{|D|}{|\Omega|} = \frac{|D|}{16}.$$

The probabilities of A, B, C are calculated by counting the number of elements in each event:

$$\mathbb{P}[A] = \frac{4}{16} = \frac{1}{4}, \quad \mathbb{P}[B] = \frac{11}{16}, \quad \mathbb{P}[C] = \frac{11}{16}.$$

Exercise 2.6. Let $(\Omega, \mathbb{F}, \mathbb{P})$ be a probability space and let $(B_i)_{i=1}^{\infty}$ be a sequence of almost surely occurring events, i.e., $\mathbb{P}[B_i] = 1$ for all $i \geq 1$. Show that

$$\mathbb{P}\Big[\bigcap_{i=1}^{\infty} B_i\Big] = 1,$$

i.e., almost surely, all (infinitely many) events occur.

Solution: Using De Morgan's rule, we have:

$$\mathbb{P}\Big[\bigcap_{i=1}^{\infty} B_i\Big] = 1 - \mathbb{P}\Big[\bigcup_{i=1}^{\infty} (B_i)^c\Big].$$

Applying the union bound, we obtain:

$$\mathbb{P}\Big[\bigcup_{i=1}^{\infty} (B_i)^c\Big] \leq \sum_{i=1}^{\infty} \mathbb{P}\big[(B_i)^c\big] = \sum_{i=1}^{\infty} \big(1 - \underbrace{\mathbb{P}[B_i]}_{=1}\big) = \sum_{i=1}^{\infty} 0 = 0,$$

since the events B_i occur almost surely. Substituting this result into the first equation gives

$$\mathbb{P}\Big[\bigcap_{i=1}^{\infty} B_i\Big] \geq 1,$$

which implies the desired result, since then

$$1 \leq \mathbb{P}\Big[\bigcap_{i=1}^{\infty} B_i\Big] \leq \mathbb{P}[\Omega] = 1.$$

Exercise 2.7. [Birthdays] We have a class of $n \in \mathbb{N}$ students whose birthdays are randomly distributed throughout the year. To simplify, assume that a year has 365 days and that a birthday is equally likely to fall on any of the 365 days (this is not true in reality, see e.g. statistics in the UK or in Switzerland). Further, assume that the birthdays of the students are independent of each other, i.e., there are for instance no twins. (For now, you can understand "independence" intuitively or as you learned it in high school. More precisely, the assumption is that all distributions of the n students' birthdays over the 365 days are equally likely.) Finally, assume that $1 < n < 365$.

(a) Calculate the probability that there is (at least) one student whose birthday is today.

(b) Alice and Bob are two students from this class. What is the probability that they both have their birthdays today?

(c) What is the probability that Alice and Bob have the same birthday?

(d) What is the probability that (at least) two students have the same birthday?

(e) What is the probability that (at least) two students have their birthdays today?

Solution:

(a) We have:

$$\begin{aligned}&\mathbb{P}[\text{There is at least one student whose birthday is today}]\\&\quad= 1 - \mathbb{P}[\text{No one has a birthday today}]\\&\quad= 1 - \left(\frac{364}{365}\right)^n.\end{aligned}$$

(b) The probability that both Alice and Bob have their birthdays today is:

$$\mathbb{P}[\text{Alice and Bob have their birthdays today}] = \frac{1 \times 1 \times 365^{n-2}}{365^n} = \frac{1}{365} \times \frac{1}{365} \approx 7.506 \times 10^{-6}.$$

(c)

$$\mathbb{P}[\text{Alice and Bob have the same birthday}] = \frac{365}{365} \times \frac{1}{365} = \frac{1}{365} \approx 0.0027.$$

(d)

$$\begin{aligned}&\mathbb{P}[\text{At least two students have the same birthday}]\\&\quad= 1 - \mathbb{P}[\text{All students have different birthdays}]\\&\quad= 1 - \frac{365}{365} \times \frac{364}{365} \times \cdots \times \frac{365-n+1}{365}.\end{aligned}$$

Remark: This is a well-known problem that contradicts common intuition, see Wikipedia. For $n = 23$, the probability is more than 1/2 (50%).

(e) We have:

$$\begin{aligned}&\mathbb{P}[\text{At least two students have their birthdays today}]\\&\quad= 1 - \mathbb{P}[\text{No one has a birthday today}] - \mathbb{P}[\text{Exactly one student has a birthday today}]\\&\quad= 1 - \left(\frac{364}{365}\right)^n - \binom{n}{1} \times \frac{1}{365} \times \left(\frac{364}{365}\right)^{n-1}\\&\quad= 1 - \left(\frac{364}{365}\right)^n - \frac{n \times 364^{n-1}}{365^n}.\end{aligned}$$

Alternatively,

$$
\begin{aligned}
&\mathbb{P}[\text{At least two students have their birthdays today}] \\
&= \sum_{k=2}^{n} \mathbb{P}[\text{Exactly } k \text{ students have their birthdays today}] \\
&= \sum_{k=2}^{n} \binom{n}{k} \left(\frac{1}{365}\right)^{k} \left(\frac{365}{365}\right)^{n-k}.
\end{aligned}
$$

Remark: For $n = 23$, the probability is less than 0.002 (0.2%).

Exercise 2.8. [Simpson's paradox] We are interested in the probability of a student's success in an entrance exam for two departments of a university. Consider the following events:

$$
\begin{aligned}
A &\coloneqq \{\text{The student is male}\}, \\
B &\coloneqq \{\text{The student applied to Department I}\}, \\
C &\coloneqq \{\text{The student was accepted}\}.
\end{aligned}
$$

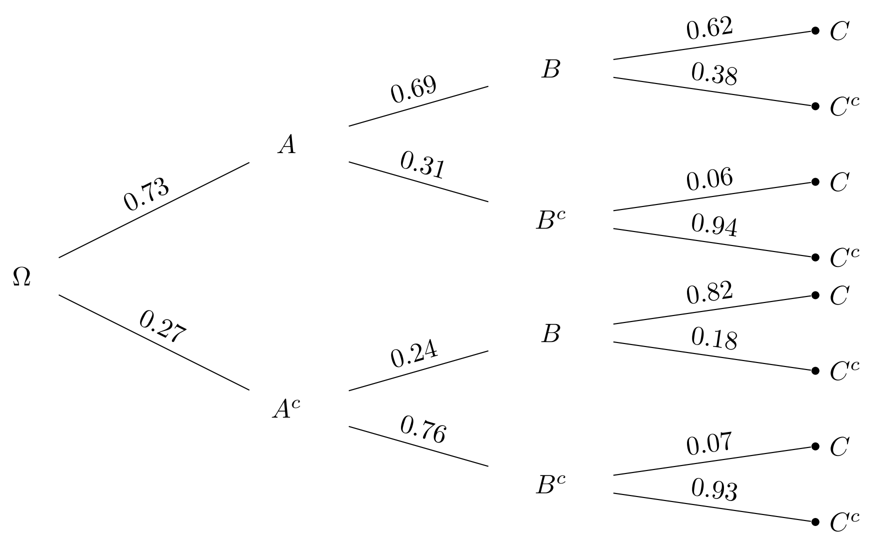
Therefore,

$$
\begin{aligned}
A^c &= \{\text{The student is not male}\}, \\
B^c &= \{\text{The student applied to Department II}\}, \\
C^c &= \{\text{The student was not accepted}\}.
\end{aligned}
$$

We know the following probabilities:

$$
\begin{gathered}
\mathbb{P}[A] = 0.73, \\
\mathbb{P}[B|A] = 0.69, \qquad \mathbb{P}[B|A^c] = 0.24, \\
\mathbb{P}[C|A \cap B] = 0.62, \quad \mathbb{P}[C|A^c \cap B] = 0.82, \\
\mathbb{P}[C|A \cap B^c] = 0.06, \quad \mathbb{P}[C|A^c \cap B^c] = 0.07.
\end{gathered}
$$

Graphically, this is represented as follows:



(a) Explain in words what the above conditional probabilities represent. Do you think that individuals who are not male are disadvantaged in the selection process? Why or why not?

(b) Calculate $\mathbb{P}[C|A]$ and $\mathbb{P}[C|A^c]$, i.e., the acceptance probabilities for males and for individuals who are not male. Does this match your answer in (a)? Can you explain what is happening here?

Solution:

(a) The interpretation is as follows:

- $\mathbb{P}[B|A]$ and $\mathbb{P}[B|A^c]$ are the probabilities that a male and a non-male student, respectively, applies to Department I.
- $\mathbb{P}[C|A \cap B^c]$ and $\mathbb{P}[C|A^c \cap B^c]$ represent the acceptance rates for males and non-males in Department II.
- Finally, $\mathbb{P}[C|A \cap B^c]$ and $\mathbb{P}[C|A^c \cap B^c]$ represent the acceptance rates for males and non-males in Department II.

We can observe that

$$\mathbb{P}[C \mid B \cap A^c] > \mathbb{P}[C \mid B \cap A]$$

and

$$\mathbb{P}[C|B^c \cap A^c] > \mathbb{P}[C|B^c \cap A].$$

This means that in both departments, the acceptance rate for non-males is higher than for males. Based on this information, we cannot conclude that non-males are disadvantaged.

(b) We have

$$\begin{aligned}
\mathbb{P}[C|A] &= \frac{\mathbb{P}[C \cap A]}{\mathbb{P}[A]} \\
&= \frac{\mathbb{P}[C \cap A \cap B] + \mathbb{P}[C \cap A \cap B^c]}{\mathbb{P}[A]} \\
&= \frac{\mathbb{P}[C|A \cap B]\mathbb{P}[A \cap B] + \mathbb{P}[C|A \cap B^c]\mathbb{P}[A \cap B^c]}{\mathbb{P}[A]} \\
&= \frac{\mathbb{P}[C|A \cap B]\mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[C|A \cap B^c]\mathbb{P}[B^c|A]\mathbb{P}[A]}{\mathbb{P}[A]} \\
&= 0.62 \times 0.69 + 0.06 \times 0.31 \\
&= 0.4464
\end{aligned}$$

and similarly

$$\begin{aligned}
\mathbb{P}[C|A^c] &= \mathbb{P}[C|A^c \cap B]\mathbb{P}[B|A^c] + \mathbb{P}[C|A^c \cap B^c]\mathbb{P}[B^c|A^c] \\
&= 0.82 \times 0.24 + 0.07 \times 0.76 \\
&= 0.25.
\end{aligned}$$

In other words, the (overall) acceptance rates for males and non-males are 0.45 and 0.25, respectively. These numbers suggest a completely different conclusion — non-males seem to be at a disadvantage.

The explanation for this paradox is as follows: The lower overall acceptance rate for non-males is because a large proportion of non-males apply to the department with a low acceptance rate. (Why this is the case is a completely different question and cannot be discussed based on the information given here.)

In fact, we can calculate the acceptance rates for the two departments as follows:

$$\begin{aligned}\mathbb{P}[C|B] &= \frac{\mathbb{P}[C \cap B]}{\mathbb{P}[B]} \\ &= \frac{\mathbb{P}[C \cap B \cap A] + \mathbb{P}[C \cap B \cap A^c]}{\mathbb{P}[B \cap A] + \mathbb{P}[B \cap A^c]} \\ &= \frac{\mathbb{P}[C|B \cap A]\mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[C|B \cap A^c]\mathbb{P}[B|A^c]\mathbb{P}[A^c]}{\mathbb{P}[B|A]\mathbb{P}[A] + \mathbb{P}[B|A^c]\mathbb{P}[A^c]} \\ &= \frac{0.62 \times 0.69 \times 0.73 + 0.82 \times 0.24 \times 0.27}{0.69 \times 0.73 + 0.24 \times 0.27} \\ &= 0.648.\end{aligned}$$

Similar calculations give $\mathbb{P}[C|B^c] \approx 0.065$. Therefore, the above result makes sense when we realize that $\mathbb{P}[B^c|A^c] = 0.76$ (76%) of non-males apply to the highly selective Department II, while $\mathbb{P}[B|A] = 0.69$ (69%) of males apply to the much less selective Department I.

In general, Simpson's Paradox demonstrates that correlation and causality can diverge significantly when an important variable (e.g., the department in this case) is not considered. In practice, it is often unknown whether important variables are missing. More interesting examples of this paradox are well presented at this link.

Exercise 2.9. [Nadal vs. Federer] We analyze a tennis match between Roger Federer and Rafael Nadal. The match is played under the "best of 3" rule, i.e., the winner is the first to win two sets (a maximum of 3 sets are played). We assume that Federer wins each set – independently of the others – with probability $p = \frac{1}{3}$. Let A denote the event that Federer wins the first set and B denote the event that Federer wins the match (i.e., wins two sets).

(a) Express $A \cup B$, $A^c \cap B$, $A \cap B^c$ and $A \setminus B$ in words. Calculate the conditional probabilities $\mathbb{P}[B^c|A]$, $\mathbb{P}[B|A]$ and $\mathbb{P}[B|A^c]$.

(b) Calculate the probability that Federer wins the match.

(c) Calculate the conditional probabilities $\mathbb{P}[A|B]$ and $\mathbb{P}[A|B^c]$.

Solution:

(a) We have

- $A \cup B = \{$Federer wins the first set or wins the match$\}$,
- $A^c \cap B = \{$Federer loses the first set and wins the match$\}$,
- $A \cap B^c = A \setminus B = \{$Federer wins the first set and loses the match$\}$.

To calculate the conditional probabilities, we introduce the events

$$S_i = \{\text{Federer wins the } i\text{-th set}\}, \quad i \in \{1, 2, 3\}.$$

From the problem statement, we know that $\mathbb{P}[S_i] = \frac{1}{3}$, $i \in \{1,2,3\}$, and that S_1, S_2, S_3 are independent. Therefore, we have:

$$\mathbb{P}[B^c|A] = \frac{\mathbb{P}[B^c \cap A]}{\mathbb{P}[A]} = \frac{\mathbb{P}[S_1 \cap S_2^c \cap S_3^c]}{\mathbb{P}[S_1]} = \frac{\mathbb{P}[S_1]\mathbb{P}[S_2^c]\mathbb{P}[S_3^c]}{\mathbb{P}[S_1]} = \mathbb{P}[S_2^c]\mathbb{P}[S_3^c] = \frac{2}{3} \times \frac{2}{3} = \frac{4}{9},$$

$$\mathbb{P}[B|A] = 1 - \mathbb{P}[B^c|A] = \frac{5}{9},$$

$$\mathbb{P}[B|A^c] = \frac{\mathbb{P}[B \cap A^c]}{\mathbb{P}[A^c]} = \mathbb{P}[S_2]\mathbb{P}[S_3] = \frac{1}{3} \times \frac{1}{3} = \frac{1}{9}.$$

(b) Using the law of total probability, we calculate:

$$\mathbb{P}[B] = \mathbb{P}[B|A]\,\mathbb{P}[A] + \mathbb{P}[B|A^c]\,\mathbb{P}[A^c] = \frac{5}{9} \times \frac{1}{3} + \frac{1}{9} \times \frac{2}{3} = \frac{7}{27}.$$

Alternatively, we can express B as the disjoint union

$$B = (S_1 \cap S_2) \cup (S_1 \cap S_2^c \cap S_3) \cup (S_1^c \cap S_2 \cap S_3).$$

Using independence, we get:

$$\mathbb{P}[B] = \mathbb{P}[S_1]\mathbb{P}[S_2] + \mathbb{P}[S_1]\mathbb{P}[S_2^c]\mathbb{P}[S_3] + \mathbb{P}[S_1^c]\mathbb{P}[S_2]\mathbb{P}[S_3] = \frac{1}{3} \times \frac{1}{3} + \frac{1}{3} \times \frac{2}{3} \times \frac{1}{3} + \frac{2}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{7}{27}.$$

(c)

$$A|B = \{\text{Federer won the first set, given that he won the match}\},$$
$$A|B^c = \{\text{Federer won the first set, given that he lost the match}\}.$$

Using Bayes' theorem:

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[B|A]\,\mathbb{P}[A]}{\mathbb{P}[B]} = \frac{\frac{5}{9} \times \frac{1}{3}}{\frac{7}{27}} = \frac{\frac{5}{27}}{\frac{7}{27}} = \frac{5}{7},$$

$$\mathbb{P}[A|B^c] = \frac{\mathbb{P}[B^c|A]\,\mathbb{P}[A]}{\mathbb{P}[B^c]} = \frac{\frac{4}{9} \times \frac{1}{3}}{\frac{20}{27}} = \frac{\frac{4}{27}}{\frac{20}{27}} = \frac{1}{5}.$$