Homework 1

Solutions

1. (True or False)

   (a) True. $P(B) = P(A^c \cap B) + P(A \cap B) = P(A^c \cap B) + P(A)P(B)$, which implies that $P(A^c \cap B) = P(B)(1 - P(A))$.

   (b) False. For example, let $\mathrm{P}(C) = 0.5$, $\mathrm{P}(A \,|\, C) = 0.5$, $\mathrm{P}(A \,|\, C^c) = 0.5$, $\mathrm{P}(B \,|\, A, C) = \mathrm{P}(B \,|\, A^c, C) = 1$, $\mathrm{P}(B \,|\, A, C^c) = 0$, $\mathrm{P}(B \,|\, A^c, C^c) = 1$. [This can occur if A is the event that heads occurs on the first coin flip, C is the event that heads occurs on the second coin flip, and $B = C \cup A^c$.] We have

$$\mathrm{P}(A, B \,|\, C) = \mathrm{P}(A \,|\, C)\mathrm{P}(B \,|\, A, C) = 0.5, \tag{1}$$

$$\mathrm{P}(B \,|\, C) = \mathrm{P}(A \,|\, C)\mathrm{P}(B \,|\, A, C) + \mathrm{P}(A^c \,|\, C)\mathrm{P}(B \,|\, A^c, C) = 1 \tag{2}$$

   so $\mathrm{P}(A, B \,|\, C) = \mathrm{P}(A \,|\, C)\mathrm{P}(B \,|\, C) = 0.5$, which means that A and B are independent given C. However,

$$\mathrm{P}(A, B \,|\, C^c) = \mathrm{P}(A \,|\, C^c)\mathrm{P}(B \,|\, A, C^c) = 0, \tag{3}$$

$$\mathrm{P}(B \,|\, C^c) = \mathrm{P}(A \,|\, C^c)\mathrm{P}(B \,|\, A, C^c) + \mathrm{P}(A^c \,|\, C^c)\mathrm{P}(B \,|\, A^c, C^c) = 0.5 \tag{4}$$

   so $\mathrm{P}(A, B \,|\, C^c) \neq \mathrm{P}(A \,|\, C^c)\mathrm{P}(B \,|\, C^c) = 0.25$.

   (c) True. Any two events A and B in a partition are disjoint, which means that $P(A \cap B) = 0$. If they are independent then $P(A)\,\mathrm{P}(B) = P(A \cap B) = 0$ so that either $P(A) = 0$ or $P(B) = 0$.

   (d) True. *Alternative 1:* $\mathrm{P}(A|B) = 1$ implies that $P(A \cap B) = P(B)$. This in turn implies

$$\begin{aligned}
\mathrm{P}(B^c|A^c) &= \frac{\mathrm{P}(A^c \cap B^c)}{\mathrm{P}(A^c)} && (5)\\
&= \frac{\mathrm{P}((A \cup B)^c)}{\mathrm{P}(A^c)} \quad \text{by DeMorgan's law} && (6)\\
&= \frac{1 - \mathrm{P}(A \cup B)}{1 - \mathrm{P}(A)} && (7)\\
&= \frac{1 - \mathrm{P}(A) - \mathrm{P}(B) + \mathrm{P}(A \cap B)}{1 - \mathrm{P}(A)} && (8)\\
&= \frac{1 - \mathrm{P}(A)}{1 - \mathrm{P}(A)} \quad \text{because } P(A \cap B) = P(B) && (9)\\
&= 1. && (10)
\end{aligned}$$

   Alternative 2: $P(A^c|B) = 0$ implies that

$$\begin{aligned}
P(A^c \cap B) &= 0 && (11)\\
P(B|A^c)P(A^c) &= 0 && (12)\\
P(B|A^c) &= 0 && (13)\\
P(B^c|A^c) &= 1 && (14)
\end{aligned}$$

(e) True. We condition the left-hand side on A.

$$
\begin{aligned}
P(B|A\cup B) =& P(B|[A\cup B]\cap A)P(A|A\cup B) + P(B|[A\cup B]\cap A^c)P(A^c|A\cup B) && (15)\\
=& P(B|A)P(A|A\cup B) + P(B|B\cap A^c)P(A^c|A\cup B) && (16)\\
=& P(B|A)P(A|A\cup B) + P(A^c|A\cup B) && (17)\\
\geq& P(B|A)[P(A|A\cup B) + P(A^c|A\cup B)] && (18)\\
=& P(B|A) && (19)
\end{aligned}
$$

where we observed that $P(B|B\cap A^c) = 1$

2. (Probability spaces)

(a) We check that $\mathcal{F}_A$ satisfies the conditions:

- If $B \in \mathcal{F}_A$, then $B^c \in \mathcal{F}_A$. If the sample space is A then $B^c = A - B$. If $B \in \mathcal{F}_A$, there is some set $S \in \mathcal{F}$ such that $B = A \cap S$. This implies that $S^c \in \mathcal{F}$ because $\mathcal{F}$ is a σ-algebra. As a result, $S^c \cap A \in \mathcal{F}_A$. We end the proof proving $A - B = S^c \cap A$ by showing that they contain each other. (1) If $\omega \in A - B$, then ω belongs to A and not to B. This means that it cannot belong to S because otherwise it would belong to $B = A \cap S$. This implies $A - B \subseteq S^c \cap A$. (2) If $\omega \in S^c \cap A$, ω belongs to A and not to S. It cannot belong to B because then it would be in S. This implies $S^c \cap A \subseteq A - B$.
- If $B_1, B_2 \in \mathcal{F}_A$, then $B_1 \cup B_2 \in \mathcal{F}$. If $B_1, B_2 \in \mathcal{F}_A$, then there exist $S_1, S_2 \in \mathcal{F}$ such that $B_1 = A \cap S_1$ and $B_2 = A \cap S_2$. $S_1 \cup S_2$ is in $\mathcal{F}$, so $A \cap (S_1 \cup S_2)$ is in $\mathcal{F}_A$. This completes the proof because $A \cap (S_1 \cup S_2) = (A \cap S_1) \cup (A \cap S_2) = B_1 \cup B_2$.
- If $B_1, B_2, \ldots \in \mathcal{F}$ then $\cup_{i=1}^{\infty} B_i \in \mathcal{F}$. By the same argument as the finite case.
- $\mathcal{F}_A$ contains the sample space. $A = A \cap A$, so $A \in \mathcal{F}_A$.

Note that by the definition for any $B \in \mathcal{F}_A$

$$
\mathrm{P}_A(B) := \frac{\mathrm{P}(B)}{\mathrm{P}(A)}. \tag{20}
$$

We check that $\mathcal{P}_A$ satisfies the conditions of a probability measure:

- $\mathrm{P}_A(B) \geq 0$ for any event $B \in \mathcal{F}_A$. This just follows from $\mathrm{P}(B) \geq 0$, and $\mathrm{P}(A) > 0$.
- If $B_1, B_2, \ldots, B_n \in \mathcal{F}_A$ are disjoint then $\mathrm{P}(\cup_{i=1}^{n} B_i) = \sum_{i=1}^{n} \mathrm{P}(B_i)$. Since $B_1, B_2, \ldots, B_n$ are also in $\mathcal{F}$ we have

$$
\begin{aligned}
\mathrm{P}_A(\cup_{i=1}^{n} B_i) &:= \frac{\mathrm{P}(\cup_{i=1}^{n} B_i)}{\mathrm{P}(A)} && (21)\\
&= \frac{\sum_{i=1}^{n} \mathrm{P}(B_i)}{\mathrm{P}(A)} && (22)\\
&= \sum_{i=1}^{n} \mathrm{P}_A(B_i). && (23)
\end{aligned}
$$

- For a countably infinite sequence of disjoint sets $B_1, B_2, \ldots \in \mathcal{F}_{\mathcal{A}}$ $\mathrm{P}\left(\lim_{n\to\infty} \cup_{i=1}^n B_i\right) = \lim_{n\to\infty} \sum_{i=1}^n \mathrm{P}\left(B_i\right)$. By the same argument as the finite case.
- The probability of the sample space equals 1. By the definition

$$\mathrm{P}_A\left(A\right) := \frac{\mathrm{P}\left(A\right)}{\mathrm{P}\left(A\right)} = 1. \tag{24}$$

(b) We check that $\mathcal{P}$ satisfies the conditions of a probability measure:

- $\mathrm{P}\left(B\right) \geq 0$ for any event $B \in \mathcal{F}$. The numerator and denominator are both non-negative by definition.
- If $S_1, S_2, \ldots, S_n \in \mathcal{F}$ are disjoint then $\mathrm{P}\left(\cup_{i=1}^n S_i\right) = \sum_{i=1}^n \mathrm{P}\left(S_i\right)$.

$$\begin{aligned} \mathrm{P}\left(\cup_{i=1}^n S_i\right) &= \frac{\text{number of data points with value in } \cup_{i=1}^n S_i}{N} && (25) \\ &= \frac{\text{number of in } S_1 + \text{number of in } S_2 + \ldots + \text{number of in } S_n}{N} \\ &= \sum_{i=1}^n \mathrm{P}\left(S_i\right). && (26) \end{aligned}$$

- For a countably infinite sequence of disjoint sets $S_1, S_2, \ldots \in \mathcal{F}$ $\mathrm{P}\left(\lim_{n\to\infty} \cup_{i=1}^n S_i\right) = \lim_{n\to\infty} \sum_{i=1}^n \mathrm{P}\left(S_i\right)$. By the same argument as the finite case.
- The probability of the sample space equals 1.

$$\mathrm{P}\left(\Omega\right) := \frac{\text{number of data points with value in } \Omega}{N} = 1. \tag{27}$$

3. (Testing)

(a) Yes, it is reasonable to assume that the test only depends on whether that particular employee is ill, and not the others, and the events *Employee i is ill*, for $1 \leq i \leq 10$, are all independent.

(b) We define the events $I_1, \ldots, I_{10}$ to represent each employee being ill, and $T_1, \ldots, T_{10}$ to represent that the corresponding test is positive. The event that at least one test is positive is $\cup_{i=1}^{10} T_i$. By DeMorgan's laws,

$$\begin{aligned} \mathrm{P}(\cup_{i=1}^{10} T_i) &= 1 - \mathrm{P}((\cup_{i=1}^{10} T_i)^c) && (28) \\ &= 1 - \mathrm{P}(\cap_{i=1}^{10} T_i^c). && (29) \end{aligned}$$

We have

$$\begin{aligned} \mathrm{P}(\cap_{i=1}^{10} T_i^c) &= \Pi_{i=1}^{10} \mathrm{P}(T_i^c) \qquad \text{by independence} && (30) \\ &= \Pi_{i=1}^{10} \mathrm{P}(I_i)\mathrm{P}(T_i^c \mid I_i) + \mathrm{P}(I_i^c)\mathrm{P}(T_i^c \mid I_i^c) && (31) \\ &= \left(0.01 \cdot 0.02 + 0.99 \cdot 0.95\right)^{10} && (32) \\ &= 0.543. && (33) \end{aligned}$$

We conclude $\mathrm{P}(\cup_{i=1}^{10} T_i) = 0.457$.

(c) We have

$$\mathrm{P}(\cap_{i=1}^{10} I_i^c \mid \cup_{j=1}^{10} T_j) = \frac{\mathrm{P}(\cap_{i=1}^{10} I_i^c, \cup_{j=1}^{10} T_j)}{\mathrm{P}(\cup_{j=1}^{10} T_j)}, \tag{34}$$

so we only need to compute the numerator; the denominator was computed above.

$$\begin{align} \mathrm{P}(\cap_{i=1}^{10} I_i^c, \cup_{j=1}^{10} T_j) &= \mathrm{P}(\cap_{i=1}^{10} I_i^c) \mathrm{P}(\cup_{j=1}^{10} T_j \mid \cap_{k=1}^{10} I_k^c) \tag{35} \\ &= \Pi_{i=1}^{10} \mathrm{P}(I_i^c) \left(1 - \mathrm{P}(\cap_{j=1}^{10} T_j^c \mid \cap_{k=1}^{10} I_k^c)\right) \tag{36} \end{align}$$

It is reasonable to assume that T_j^c is conditionally independent of T_l^c, $j \neq l$, conditioned on $\cap_k I_k^c$. Even if we fix $\cap_k I_k^c$, no other T_l^c provides any information about T_j^c. Under this assumption

$$\mathrm{P}(\cap_{j=1}^{10} T_j^c \mid \cap_{k=1}^{10} I_k^c) = \Pi_{j=1}^{10} \mathrm{P}(T_j^c \mid \cap_{k=1}^{10} I_k^c). \tag{37}$$

It is reasonable to assume that T_i is conditionally independent of $\cap_{j \neq i} I_j^c$ given I_i because T_i only depends on I_i. Even if we fix I_i, $\cap_{j \neq i} I_j^c$ does not provide any information about T_i. Under this assumption

$$\Pi_{j=1}^{10} \mathrm{P}(T_j^c \mid \cap_{k=1}^{10} I_k^c) = \Pi_{j=1}^{10} \mathrm{P}(T_j^c \mid I_j^c). \tag{38}$$

Putting everything together

$$\begin{align} \mathrm{P}(\cap_{i=1}^{10} I_i^c, \cup_{j=1}^{10} T_j) &= \frac{\Pi_{i=1}^{10} \mathrm{P}(I_i^c)(1 - \Pi_{j=1}^{10} \mathrm{P}(T_j^c \mid I_j^c))}{\mathrm{P}(\cup_{j=1}^{10} T_j)} \tag{39} \\ &= \frac{0.99^{10}(1 - 0.95^{10})}{0.457} \tag{40} \\ &= 0.793. \tag{41} \end{align}$$

4. (Streak of heads)

(a) We represent heads with 1 and tails with 0. To compute the probabilities, we consider the $2^5 = 32$ possible sequences:

00000 00001 00010 00011 00100 00101 00110 00111 01000 01001 01010 01011 01100 01101 01110 01111 10000 10001 10010 10011 10100 10101 10110 10111 11000 11001 11010 11011 11100 11101 11110 11111

We have

$$\begin{align} &\mathrm{P}(\text{sequence equals 00000}) \tag{42} \\ &= \mathrm{P}(\text{1st roll equals } 0, \text{2nd roll equals } 0, \ldots, \text{5th roll equals } 0) \tag{43} \\ &= \mathrm{P}(\text{1st roll equals } 0) \mathrm{P}(\text{2nd roll equals } 0) \cdots \mathrm{P}(\text{5th roll equals } 0) \tag{44} \\ &= \frac{1}{32} \tag{45} \end{align}$$

and by the same argument, all the other sequences also have probability 1/32.

Since the probability of the union of disjoint events is the sum of the individual probabilities,

$$\mathrm{P}(\text{at most 0 heads in a row}) = \frac{1}{32}, \tag{46}$$
$$\mathrm{P}(\text{at most 1 heads in a row}) = \frac{12}{32}, \tag{47}$$
$$\mathrm{P}(\text{at most 2 heads in a row}) = \frac{11}{32}, \tag{48}$$
$$\mathrm{P}(\text{at most 3 heads in a row}) = \frac{5}{32}, \tag{49}$$
$$\mathrm{P}(\text{at most 4 heads in a row}) = \frac{2}{32}, \tag{50}$$
$$\mathrm{P}(\text{at most 5 heads in a row}) = \frac{1}{32}. \tag{51}$$

(b) The code is

```
def p_longest_streak(n, tries):
    p_longest = np.zeros(n+1)
    for i in range(tries):
        current_streak = 0
        longest_streak = 0
        for j in range(n):
            if np.random.rand() > 0.5:
                current_streak = current_streak + 1
            else:
                if current_streak > longest_streak:
                    longest_streak = current_streak
                current_streak = 0
        if current_streak > longest_streak:
            longest_streak = current_streak
        p_longest[longest_streak] = p_longest[longest_streak] + 1./tries
    return p_longest
```

The images are shown in Figure 1.

(c) The probability is 0.319. It is therefore not unlikely to find a streak of 8 or more heads in a sequence of 200 fair coin flips, so it is very possible that the random generator is fine.

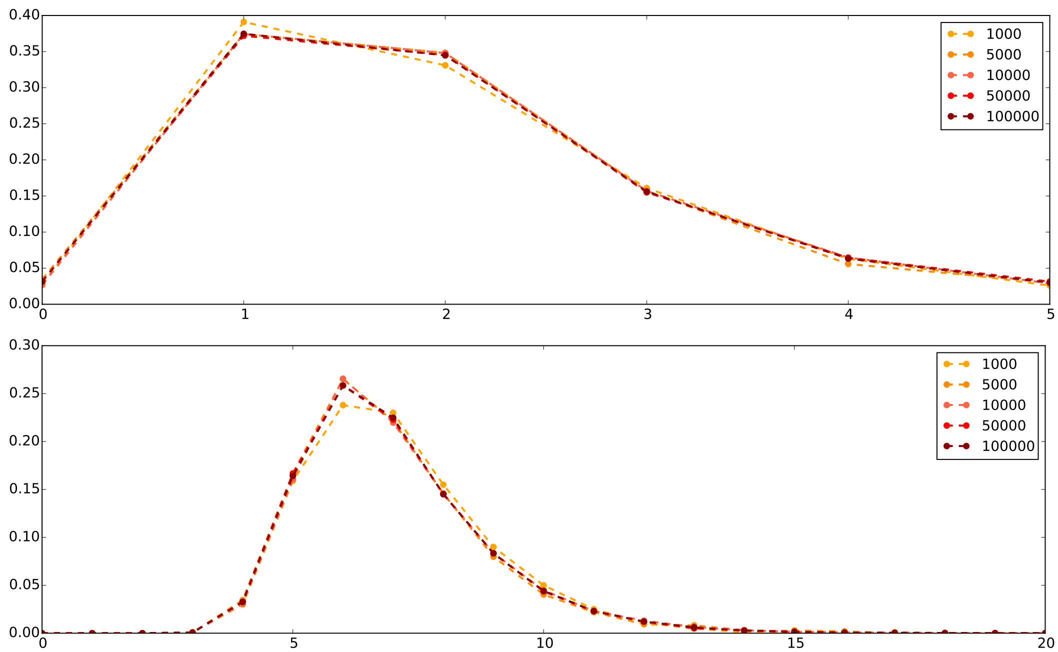

Figure 1: Probability of streaks of heads for sequences of length 5 (above) and 200 (below) estimated using different number of Monte Carlo runs.