Probability and Statistics

Exercise sheet 3 - Solutions

MC 3.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let A, B, and C be events with $\mathbb{P}[A \cap B] > 0$ and $\mathbb{P}[C] > 0$. We assume that $\mathbb{P}[A|B] > \mathbb{P}[A]$ and $\mathbb{P}[A|C] > \mathbb{P}[A]$. Which of the following holds? (Exactly one answer is correct.)

(a) $\mathbb{P}[A|B \cap C] > \mathbb{P}[A]$.

(b) $\mathbb{P}[B] = \mathbb{P}[C]$.

(c) $\mathbb{P}[B|A] > \mathbb{P}[B]$.

(d) None of the above.

Solution: Answer (c) is correct. Since $\mathbb{P}[A \cap B] > 0$, we have $\mathbb{P}[A] > 0$ and $\mathbb{P}[B] > 0$, so

$$\mathbb{P}[A \cap B] = \mathbb{P}[A|B]\mathbb{P}[B] = \mathbb{P}[B|A]\mathbb{P}[A].$$

Then

$$\mathbb{P}[A|B] > \mathbb{P}[A] \implies \mathbb{P}[B|A] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[A]} = \frac{\mathbb{P}[A|B]\mathbb{P}[B]}{\mathbb{P}[A]} > \frac{\mathbb{P}[A]\mathbb{P}[B]}{\mathbb{P}[A]} = \mathbb{P}[B].$$

None of the other options hold in general.

MC 3.2. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mathcal{F} = \{\emptyset, \Omega, \{\omega_1, \omega_2\}, \{\omega_3\}\}$. Which of the following define random variables on $(\Omega, \mathcal{F})$? (The number of correct answers is between 0 and 4.)

(a) $X_1(\omega_1) = 1,\ X_1(\omega_2) = 2,\ X_1(\omega_3) = 3$.

(b) $X_2(\omega_1) = 1,\ X_2(\omega_2) = 1,\ X_2(\omega_3) = 2$.

(c) $X_3(\omega_1) = 1,\ X_3(\omega_2) = 2,\ X_3(\omega_3) = 2$.

(d) $X_4(\omega_1) = 1,\ X_4(\omega_2) = 1,\ X_4(\omega_3) = 1$.

Solution: First, we note that in this specific situation, where X_i, $i \in \{1, \dots, 4\}$, takes values in $\{1, 2, 3\}$, we have

$$X_i^{-1}((-\infty, a]) = \begin{cases} \emptyset, & a < 1, \\ X_i^{-1}(\{1\}), & a \in [1, 2), \\ X_i^{-1}(\{1\}) \cup X_i^{-1}(\{2\}), & a \in [2, 3), \\ X_i^{-1}(\{1\}) \cup X_i^{-1}(\{2\}) \cup X_i^{-1}(\{3\}), & a \geq 3. \end{cases}$$

Thus, it suffices to verify measurability of the sets $X_i^{-1}(\{1\})$, $X_i^{-1}(\{2\})$, and $X_i^{-1}(\{3\})$.

(a) We have $X_1^{-1}(\{1\}) = \{\omega_1\} \notin \mathcal{F}$, and so X_1 is **not** a random variable.

(b) We have $X_2^{-1}(\{1\}) = \{\omega_1, \omega_2\} \in \mathcal{F}$, $X_2^{-1}(\{2\}) = \{\omega_3\} \in \mathcal{F}$ and $X_2^{-1}(\{3\}) = \emptyset \in \mathcal{F}$, and so X_2 is a random variable.

(c) We have $X_3^{-1}(\{1\}) = \{\omega_1\} \notin \mathcal{F}$, and so X_3 is **not** a random variable.

(d) We have $X_4^{-1}(\{1\}) = \Omega \in \mathcal{F}$ and $X_4^{-1}(\{2\}) = X_4^{-1}(\{3\}) = \emptyset \in \mathcal{F}$. and so X_4 is a random variable. In fact, constant functions are always random variables.

MC 3.3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let A, B, and C be events in $\mathcal{F}$. Which of the following statements are always true? (The number of correct answers is between 0 and 4.)

(a) If A and B as well as A and C are independent, then A and $B \cap C$ are also independent.

(b) If A and B as well as B and C are independent, then A and C are also independent.

(c) If A, B, and C are independent, then A and $B \cap C$ are also independent.

(d) If A and A are independent, then $\mathbb{P}[A] = 1$ or $\mathbb{P}[A] = 0$.

Solution:

(a) is **not** correct. For instance, consider two independent coin flips, and define

$$\begin{aligned} A &:= \{\text{The first flip results in tails}\}, \\ B &:= \{\text{The second flip results in heads}\} \\ C &:= \{\text{The two flips have the same result}\}. \end{aligned}$$

It is easy to verify that this example disproves (a).

(b) is **not** correct. Take A, B independent and $C = A$. Then it is clear that (b) is generally false.

(c) is corrrect. We have

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

(d) is correct. We have that $\mathbb{P}[A] = \mathbb{P}[A \cap A] = \mathbb{P}[A]\mathbb{P}[A] = (\mathbb{P}[A])^2$ holds only if $\mathbb{P}[A] = 1$ or $\mathbb{P}[A] = 0$.

MC 3.4. Let X and Y be two random variables taking values in $\{1, \ldots, 6\}$ and representing two independent rolls of a die. Which of the following pairs of events are independent? (The number of correct answers is between 0 and 4.)

(a) $\{X \text{ is odd}\}$, $\{X + Y \text{ is even}\}$.

(b) $\big\{X \in \{1, 3\}\big\}$, $\{X + Y = 5\}$.

(c) $\{X = 1\}$, $\{X + Y = 4\}$.

(d) $\{X = 1\}$, $\{X + Y = 13\}$.

Solution: (a) and (d) are correct. We have

$$\mathbb{P}[X \text{ is odd}] = \frac{1}{2},$$

$$\begin{aligned}
\mathbb{P}[X + Y \text{ is even}] &= \mathbb{P}[X \text{ is even}|Y \text{ is even}]\mathbb{P}[Y \text{ is even}] \\
&\quad + \mathbb{P}[X \text{ is odd}|Y \text{ is odd}]\mathbb{P}[Y \text{ is odd}] \\
&= \mathbb{P}[X \text{ is even}]\mathbb{P}[Y \text{ is even}] \\
&\quad + \mathbb{P}[X \text{ is odd}]\mathbb{P}[Y \text{ is odd}] \\
&= \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} \\
&= \frac{1}{2}, \\
\mathbb{P}[X \text{ is odd}, \, X + Y \text{ is even}] &= \mathbb{P}[X \text{ is odd}, \, Y \text{ is odd}] \\
&= \mathbb{P}[X \text{ is odd}]\mathbb{P}[Y \text{ is odd}] \\
&= \frac{1}{4}.
\end{aligned}$$

Thus, we see that

$$\mathbb{P}[X \text{ is odd}, \, X + Y \text{ is even}] = \frac{1}{4} = \frac{1}{2} \times \frac{1}{2} = \mathbb{P}[X \text{ is odd}]\mathbb{P}[X + Y \text{ is even}].$$

The last option is trivially true because

$$\mathbb{P}[X = 1, \, X + Y = 13] = 0 = \mathbb{P}[X = 1]\mathbb{P}[X + Y = 13].$$

Using similar calculations, one can verify that the other options are incorrect.

Exercise 3.5. Let X be a random variable with the distribution function

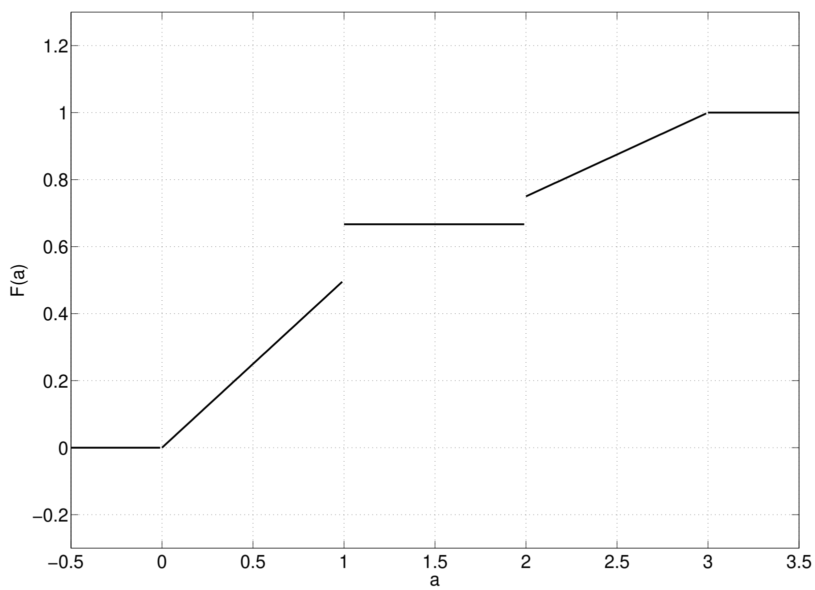
$$F(a) = \begin{cases} 0, & a < 0, \\ \frac{a}{2}, & 0 \le a < 1, \\ \frac{2}{3}, & 1 \le a < 2, \\ \frac{a+1}{4}, & 2 \le a < 3, \\ 1, & 3 \le a. \end{cases}$$

(a) Plot this distribution function.

(b) Determine the following probabilities: $\mathbb{P}[X < 1]$, $\mathbb{P}[X = 2]$, $\mathbb{P}[X = 3]$, $\mathbb{P}[1 < X \le 2]$, $\mathbb{P}[1 \le X < 2]$ and $\mathbb{P}[X \ge 3/2]$.

Solution:

(a) The graph of F is as follows:



(b) We have:

$$\begin{aligned}
\mathbb{P}[X < 1] &= F(1-) = 1/2, \\
\mathbb{P}[X = 2] &= F(2) - F(2-) = 3/4 - 2/3 = 1/12, \\
\mathbb{P}[X = 3] &= 0 \text{ (since } F \text{ is continuous at 3)}, \\
\mathbb{P}[1 < X \leq 2] &= F(2) - F(1) = 3/4 - 2/3 = 1/12, \\
\mathbb{P}[1 \leq X < 2] &= F(2-) - F(1-) = 2/3 - 1/2 = 1/6, \\
\mathbb{P}[X \geq 3/2] &= 1 - F(1.5-) = 1 - 2/3 = 1/3.
\end{aligned}$$

Exercise 3.6. [Riemann zeta function] Let X be a discrete random variable with values in $\mathbb{N} = \{1, 2, 3, \ldots\}$. The distribution of X is given by

$$\mathbb{P}[X = n] = \frac{n^{-s}}{\zeta(s)}, \quad n \in \mathbb{N},$$

where $s > 1$ is a parameter of the distribution, and $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ denotes the Riemann zeta function. For a number $m \in \mathbb{N}$, we define the event E_m as $\{X$ is divisible by m without remainder$\}$, or equivalently, $\{$There exists a $k \in \mathbb{N}$ such that $X = km\}$.

(a) Show that $\mathbb{P}[E_m] = m^{-s}$ for all $m \in \mathbb{N}$.

(b) Let p and q be two distinct prime numbers. Show that E_p and E_q are independent.

Hint: A number n is divisible by two different prime numbers p and q if and only if n is divisible by pq.

(c) Determine $\mathbb{P}\Big[\bigcap_{p \text{ prime}} E_p^c\Big]$.

Solution:

(a) Let $m \in \mathbb{N}$. Then we have

$$\mathbb{P}[E_m] = \mathbb{P}\left[\bigcup_{k=1}^{\infty} \{X = km\}\right] \overset{\text{disjoint}}{=} \sum_{k=1}^{\infty} \mathbb{P}[X = km] = \sum_{k=1}^{\infty} \frac{(km)^{-s}}{\zeta(s)}$$
$$= m^{-s} \times \frac{\sum_{k=1}^{\infty} k^{-s}}{\zeta(s)} = m^{-s} \times \frac{\zeta(s)}{\zeta(s)} = m^{-s}.$$

(b) According to the hint, we have $E_p \cap E_q = E_{pq}$, and thus

$$\mathbb{P}[E_p \cap E_q] = \mathbb{P}[E_{pq}] = (pq)^{-s} = p^{-s}q^{-s} = \mathbb{P}[E_p]\mathbb{P}[E_q].$$

Hence, E_p and E_q are independent.

(c) We consider the prime factorization of X. The event $\bigcap_{p \text{ prime}} E_p^c$ corresponds to the situation when no prime number appears in the prime factorization of X, i.e., X must be 1. Thus, we obtain

$$\mathbb{P}\left[\bigcap_{p \text{ prime}} E_p^c\right] = \mathbb{P}[X = 1] = \frac{1}{\zeta(s)}.$$

Exercise 3.7. [First six] Two players, Anja and Beatrice, take turns rolling a (fair) die until a six appears. Anja starts rolling. The player who rolls the first six wins the game. Determine the probabilities of winning for both players.

Solution: Let A be the event that Anja wins, and B be the event that Beatrice wins. First, note that

$$\begin{aligned}
\mathbb{P}[\text{The game never ends}] &= 1 - \sum_{i=1}^{\infty} \mathbb{P}[\text{The game ends after exactly } i \text{ rolls}] \\
&= 1 - \sum_{i=1}^{\infty} \left(\frac{5}{6}\right)^{i-1} \times \frac{1}{6} \\
&= 1 - \frac{1}{6} \times \sum_{i=1}^{\infty} \left(\frac{5}{6}\right)^{i-1} \\
&= 1 - \frac{1}{6} \times 6 \\
&= 0.
\end{aligned}$$

Thus, since one of them must win, but they cannot both win simultaneously, we have

$$\mathbb{P}[A] + \mathbb{P}[B] = 1. \tag{1}$$

Let W be the number of the roll in which the first six appears. W follows a geometric distribution with parameter $1/6$. If W is odd, Anja wins; otherwise, Beatrice wins. Thus, we obtain

$$\begin{aligned}\mathbb{P}[A] = \sum_{n=0}^{\infty} \mathbb{P}[W = 2n+1] &= \sum_{n=0}^{\infty} \left(\frac{5}{6}\right)^{2n} \times \frac{1}{6} \\ &= \frac{1}{6} \sum_{n=0}^{\infty} \left(\left(\frac{5}{6}\right)^2\right)^n = \frac{1}{6} \times \frac{1}{1-(5/6)^2} = \frac{6}{11}\end{aligned}$$

and consequently, $\mathbb{P}[B] = 1 - \mathbb{P}[A] = 5/11$.

Exercise 3.8. [Construction of random variables] The goal of this problem is to construct random variables from a sequence of independent coin flips. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(X_i)_{i\geq 1}$ be an infinite sequence of independent, Bernoulli(1/2)-distributed random variables. We consider the following algorithm:

$$\begin{aligned}&i \coloneqq 1 \\ &\texttt{while} \quad (X_i = X_{i+1} = 1): \\ &\qquad\qquad i \coloneqq i + 2 \\ &Z \coloneqq X_i + 2 \times X_{i+1} \\ &\texttt{return} \quad Z\end{aligned}$$

(a) Show that the algorithm always terminates after a finite number of steps with probability 1.

(b) Show that Z is a uniformly distributed random variable in $\{0, 1, 2\}$.

(c) **[Bonus]** Provide an algorithm that outputs a Bernoulli(1/5)-distributed random variable.

Solution:

(a) To show that the algorithm terminates after a finite number of steps, we need to prove that the while loop runs only a finite number of times. We observe that for $j \geq 0$,

$$\begin{aligned}A_j \coloneqq &\{\text{The while loop runs exactly } j \text{ times}\} \\ = &\left(\bigcap_{i=1}^{2j} \{X_i = 1\}\right) \cap \big(\{X_{2j+1} = 0\} \cup \{X_{2j+2} = 0\}\big).\end{aligned}$$

Due to the independence of the random variables $(X_i)_{i\geq 1}$, we get

$$\begin{aligned}\mathbb{P}[A_j] &= \mathbb{P}\left[\bigcap_{i=1}^{2j}\{X_i=1\}\cap\big(\{X_{2j+1}=0\}\cup\{X_{2j+2}=0\}\big)\right]\\ &= \left(\prod_{i=1}^{2j}\mathbb{P}[X_i=1]\right)\times\underbrace{\mathbb{P}[\{X_{2j+1}=0\}\cup\{X_{2j+2}=0\}]}_{=1-\mathbb{P}[\{X_{2j+1}=1\}\cap\{X_{2j+2}=1\}]=1-\frac{1}{4}=\frac{3}{4}}\\ &= \frac{3}{4}\times\left(\frac{1}{2}\right)^{2j}.\end{aligned}$$

Summing over all cases where the algorithm terminates, i.e., over the disjoint events $(A_j)_{j\geq 1}$, we obtain

$$\begin{aligned}\mathbb{P}[\text{Algorithm terminates}] &= \sum_{j=0}^{\infty}\mathbb{P}[A_j] = \sum_{j=0}^{\infty}\frac{3}{4}\times\left(\frac{1}{2}\right)^{2j}\\ &= \frac{3}{4}\times\sum_{j=0}^{\infty}\left(\frac{1}{4}\right)^{j} = \frac{3}{4}\times\frac{1}{1-\frac{1}{4}} = 1.\end{aligned}$$

(b) From part (a), we know that the event $A := \{\text{Algorithm terminates}\} = \bigcup_{j=0}^{\infty} A_j$ has probability 1 and that the events $(A_j)_{j\geq 1}$ are disjoint. Thus, we obtain

$$\begin{aligned}\mathbb{P}[Z=0] &= \mathbb{P}[\{Z=0\}\cap A] + \underbrace{\mathbb{P}[\{Z=0\}\cap A^c]}_{\leq\mathbb{P}[A^c]=0} = \sum_{j=0}^{\infty}\mathbb{P}[\{Z=0\}\cap A_j]\\ &= \sum_{j=0}^{\infty}\mathbb{P}\left[\bigcap_{i=1}^{2j}\{X_i=1\}\cap\{X_{2j+1}=0\}\cap\{X_{2j+2}=0\}\right]\\ &= \sum_{j=0}^{\infty}\left(\frac{1}{2}\right)^{2j+2} = \frac{1}{4}\times\sum_{j=0}^{\infty}\left(\frac{1}{4}\right)^{j} = \frac{1}{4}\times\frac{1}{1-\frac{1}{4}} = \frac{1}{3},\end{aligned}$$

where we again use the independence of $(X_i)_{i\geq 1}$ and the definition of Z in the algorithm. Similarly, we obtain

$$\mathbb{P}[Z=1] = \mathbb{P}[Z=2] = \frac{1}{3},$$

so Z is uniformly distributed on $\{0,1,2\}$.

(c) Consider the following algorithm:

$$
\begin{aligned}
&i \coloneqq 1 \\
&\texttt{while} \quad (X_i = X_{i+2} = 1) \ \texttt{or} \ (X_{i+1} = X_{i+2} = 1): \\
&\qquad i \coloneqq i + 3 \\
&Z \coloneqq X_i + 2 \times X_{i+1} + 4 \times X_{i+2} \\
&\texttt{if} \quad Z = 4: \\
&\qquad Z' \coloneqq 1 \\
&\texttt{else}: \\
&\qquad Z' \coloneqq 0 \\
&\texttt{return} \quad Z'
\end{aligned}
$$

Following similar reasoning as in parts (a) and (b), we can show that the algorithm terminates with probability 1 and that Z is uniformly distributed on $\{0, 1, 2, 3, 4\}$, implying that Z' is Bernoulli(1/5)-distributed.