MASD, Solutions to Probability Training Exercise Sheet

Exercise 1 (Modelling Probabilities): Provide a suitable model (set of outcomes Ω and probability distribution $\mathbb{P}$) for the following probabilistic experiments. Answer the questions asked in each case.

1. We toss a coin and afterwards we roll a dice. What is the probability that we observe heads and then roll a number smaller or equal to 4?
2. We are given an unfair dice. We roll it 1000 times and record how often we see each number:

1	2	3	4	5	6
107	195	52	492	112	42

   How would you model the experiment of rolling this dice again twice? What are the random variables X_1, X_2 in your model that describe the outcome of the first and second roll, respectively? Compute the probability that the roll sum $X_1 + X_2$ equals 4.
3. We are given a deck of 52 (distinct) cards. They are well organised, starting with hearts $2, 3, 4, 5$. We shuffle them as best we can. What is the probability that after shuffling the first four cards of the deck are again the same, but in possibly different order?
4. We toss a coin $n = 100$ times. A k-run is observing heads k times in a row (without tails in between). Express the probability of observing a 5-run at some point during our experiment in terms of the probability of the events $A_i = \{\text{the first 5-run starts with the } i\text{-th coin toss}\}$. You do not have to compute this probability.

Solutions to Exercise 1: For 1: We encode the outcome of the coin toss by $\{0, 1\}$ (heads is 1 and tails is 0). The dice roll we encode in the number of eyes we observe, i.e. by $\{1, 2, 3, 4, 5, 6\}$. Overall the possible outcomes of the experiment is an element of

$$\Omega = \{0,1\} \times \{1,2,3,4,5,6\} = \{(\omega_1, \omega_2) : \omega_1 \in \{0,1\}\,,\ \omega_2 \in \{1,2,3,4,5,6\}\}\,.$$

As probability distribution $\mathbb{P}$ we choose the uniform distribution since none of these outcomes is more or less likely than any other. Now we compute

$$\begin{aligned}
&\mathbb{P}(\text{coin toss shows heads and dice roll shows a number } \leq 4)\\
=&\mathbb{P}(\{(\omega, \omega_2) \in \Omega : \omega_1 = 1, \omega_2 \leq 4\})\\
=&\mathbb{P}(\{(1,1),(1,2),(1,3),(1,4)\})\\
=&\frac{4}{\#\Omega}\\
=&\frac{4}{12} = \frac{1}{3}\,.
\end{aligned}$$

For 2: Let us first model the experiment of rolling the unfair dice again once. In this case the probability space is $(\Omega_1, \mathbb{P}_1)$, where $\Omega_1 = \{1, 2, 3, 4, 5, 6\}$ is the outcome of one dice roll and it is reasonable to assign the probabilities

$$\mathbb{P}_1(\{i\}) := p_i\,, \quad \text{where } p_1 := \frac{107}{1000}\,,\ p_2 := \frac{195}{1000}\,,\ p_3 := \frac{52}{1000}\,,\ p_4 := \frac{492}{1000}\,,\ p_5 := \frac{112}{1000}\,,\ p_6 := \frac{42}{1000}\,.$$

for $i = 1, 2, 3, 4, 5, 6$. Now we model two such dice rolls. Then the underlying probability space is $(\Omega_2, \mathbb{P}_2)$ where $\Omega_2 = \{1, 2, 3, 4, 5, 6\}^2$ is the outcome of both rolls and we assign the probabilities

$$\mathbb{P}(\{(i, j)\}) := p_i p_j\,, \qquad i, j = 1, 2, 3, 4, 5, 6\,.$$

The two random variables X_1 and X_2 are

$$X_1 : \Omega \to \{1, 2, 3, 4, 5, 6\}, (\omega_1, \omega_2) \mapsto \omega_1\,, \qquad X_2 : \Omega \to \{1, 2, 3, 4, 5, 6\}, (\omega_1, \omega_2) \mapsto \omega_2\,.$$

Put differently, $X_i((\omega_1, \omega_2)) := \omega_i$ for $i = 1, 2$. Then we compute the probability

$$\begin{aligned}\mathbb{P}(X_1 + X_2 = 4) &= \mathbb{P}(\{(1, 3), (3, 1), (2, 2)\}) \\ &\overset{(1)}{=} \mathbb{P}(\{(1, 3)\}) + \mathbb{P}(\{(3, 1)\}) + \mathbb{P}(\{(2, 2)\}) \\ &= 2p_1 p_3 + p_2^2 = 2 \times \frac{107}{1000} \times \frac{52}{1000} + \frac{195^2}{1000^2}\,.\end{aligned}$$

In (1) we used the additivity of probabilities on disjoint sets (in this case each with one element).

For 3: We assign a number to each of the 52 cards starting with hearts 2,3,4,5. Every reshuffled deck of cards is modelled by an assignment of the position in the deck to the card at that position, i.e. by a one to one map of $\{1, 2, \dots, 52\}$ to itself. Mathematically this means that we choose $\Omega = S_{52}$ the set of permutations of the numbers $\{1, 2, \dots, 52\}$. For example the map (upstairs is the argument and below that the value)

$$\omega = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \dots & 52 \\ 3 & 2 & 1 & 4 & 5 & \dots & 52 \end{pmatrix}$$

represents switching the first and third card. Since no shuffling is special we equip Ω with the uniform distribution $\mathbb{P}$. We compute the probability

$$\begin{aligned}&\mathbb{P}(\text{hearts } 2, 3, 4, 5 \text{ in any order are the first four cards}) \\ =&\,\mathbb{P}(\{\omega \in S_{52} : \text{there is a permutation in } \sigma \in S_4 \text{ such that } \omega(i) = \sigma(i) \text{ for all } i = 1, 2, 3, 4\}) \\ =&\sum_{\sigma \in S_4} \mathbb{P}(\{\omega \in S_{52} : \omega(i) = \sigma(i) \text{ for all } i = 1, 2, 3, 4\})\,,\end{aligned}$$

where we used the additivity of probabilities in the last step. Now we compute for some fixed $\sigma \in S_4$ the probability

$$\mathbb{P}(\{\omega \in S_{52} : \omega(i) = \sigma(i) \text{ for all } i = 1, 2, 3, 4\}) = \frac{\#S_{52-4}}{\#S_{52}} = \frac{48!}{52!}\,.$$

Altogether we find

$$\mathbb{P}(\text{hearts } 2, 3, 4, 5 \text{ in any order are the first four cards}) = \#S_4 \times \frac{48!}{52!} = \frac{4! \times 48!}{52!}\,.$$

For 4: Here, a coin toss is modelled by the outcomes $\{0, 1\}$ (1 is heads and 0 is tails). Thus, the $n = 100$ independent coin tosses are modelled by $\Omega = \{0, 1\}^n = \{0, 1\}^{100}$. Since none of the outcomes is distinguished, Ω is equipped with the uniform distribution, i.e.

$$\mathbb{P}(\{\omega\}) = \frac{1}{\#\Omega} = \frac{1}{2^{100}}$$

for all $\omega \in \Omega$. The event of observing a 5-run is the event of observing a 5-run starting at any of the i-th coin tosses, i.e. the event is $\cup_{i=1}^{96} A_i$. Since the A_i are all disjoint we can express its probability as

$$\mathbb{P}\big(\cup_{i=1}^{96} A_i\big) = \sum_{i=1}^{96} \mathbb{P}(A_i)\,.$$

Exercise 2 (Combinatorics):

1. How many subsets of the numbers from 1 to 100 are there that have 10 elements and also contain the numbers 5 and 10?

2. An anagram is a word created from the letters of another word (with exactly the same number of appearances of each letter). How many (possibly nonsensical) anagrams are there of the word "STATISTICS"?

3. You have an even number $n \geq 10$ of friends, $k \leq n$ of them are female and $n-k$ of them male. You invite half of them to a dinner party. How many different dinner party configurations can you organise with exactly 10 female friends present? How many can you organise with at least 10 female friends present?

Solutions to Exercise 2: For 1: Since we are required to have the numbers 5 and 10 in our 10-element subset, we are effectively computing the number of subsets of $\{1,\ldots,100\} \setminus \{5,10\}$ with 8 elements. There are '98 choose 8' of those, or in mathematical terms

$$\#\{I : I \subset \{1,\ldots,100\} \setminus \{5,10\},\ \#I = 8\} = \binom{98}{8}.$$

For 2: To count the number of anagrams we count the number of positions (10 of them at the beginning) where we can place each letter. Then we have to multiply the number of all these options. We start with "A". We can place "A" at any of the 10 positions. Let us keep track of the factors by defining $N_A = 10$. For the letter "C" we are left with 9 positions since "A" already took up one of them. Thus, we set $N_C = 9$. The letter "I" comes up twice. Since we cannot distinguish the two "I"'s from each other the number of choices here is the same as the number of subsets of the remaining 8 positions with 2 elements, i.e. $N_I = \binom{8}{2}$. We are left with 6 positions for the three letters "S", so the number of choices is the same as all 3-element subsets of $\{1,2,3,4,5,6\}$, i.e. $N_S = \binom{6}{3}$. The positions of the remaining "T"'s are fixed, because there are only three positions left, i.e. $N_T = 1$. The answer is thus that there are

$$N_A N_C N_I N_S N_T = 10 \times 9 \times \binom{8}{2} \times \binom{6}{3} = 50400$$

anagrams of the word "STATISTICS".

For 3: First we realise that the answer is zero if $k < 10$ in both cases. Furthermore, if $n < 20$ you also cannot invite 10 female friends since you invite only $n/2$ participants to the party. Therefore, let us consider from now on only the case $k \geq 10$ and $n \geq 20$.

We determine the number $N_{n,k,l}$ of dinner parties you can organise with exactly l female friends present, where $10 \leq l \leq k$. For $l > n/2$ there is no way of organising such a party and therefore $N_{n,k,l} = 0$ when $l > n/2$.

Now let $l \leq n/2$. Since you can choose l female invitees arbitrarily out of k female friends, there are $\binom{k}{l}$ options of choosing the configuration of female participants. For the male participants you can choose $n/2 - l$ out of your $n-k$ male friends., provided $n - k \geq n/2 - l$. Thus, you have $\binom{n-k}{n/2-l}$ possibilities for the male participation in this case. Altogether we have

$$N_{n,k,10} = \binom{k}{10}\binom{n-k}{n/2-10}$$

the number of dinner arrangements with exactly 10 female friends invited, as long as $k \geq 10$, $n \geq 20$ and $n-k \geq n/2 - 10$. In case $k \geq 10$, $n \geq 20$ and $n - k < n/2 - 10$ there is no choice for the male participation and we get

$$N_{n,k,10} = \binom{k}{10}.$$

Finally, to compute the number of dinner configuration with at least 10 female friends invited, we sum up all $N_{n,k,l}$ with $l \geq 10$, i.e. we get

$$\sum_{l=10}^{n/2} N_{n,k,l}\,,$$

possibilities, where

$$N_{n,k,l} = \begin{cases} \binom{k}{l}\binom{n-k}{n/2-l} & \text{if } n-k \geq n/2-l \\ \binom{k}{l} & \text{if } n-k < n/2-l \end{cases}.$$

Exercise 3 (Discrete and Continuous Random Variables): Determine and draw the cdf of the following random variable X:

1. Let X be a roll of the unfair dice from Exercise 1.2.
2. Let X be a continuous random variable with pdf $\rho_X(x) = \frac{1}{2}\mathrm{e}^{-|x|}$ for $x \in \mathbb{R}$.
3. Let $X = \alpha Y + Z$ where Y, Z are independent, $Y \sim \mathrm{Bern}(1/2)$, Z is uniformly distributed on $[0,1]$ and $\alpha \in \{1/2, 1, 2\}$ is a constant. Each value of α is a separate case.
4. Let $X = \min\{Y, 2Z\}$ where Y, Z are as in 3 above.

For the following random variable X, determine the cdf (without drawing) and either the pdf (for continuous random variables) or pmf (for discrete random variables):

5. Let $X = \max\{X_1, X_2, X_3\}$ where X_1, X_2, X_3 are i.i.d. and have the same distribution as the sum of two independent $\mathrm{Bern}(p)$-distributed random variables.
6. Let $X = X_1 + X_2$ with X_1, X_2 i.i.d. random variables that are uniformly distributed on $[-1,1]$.
7. Let $X = \min\{X_1, X_2\}$ with X_1, X_2 i.i.d. random variables that are uniformly distributed on $[-1,1]$.
8. Let $X = X_1 + X_2 + X_3$ with X_1, X_2, X_3 i.i.d. random variables that are uniformly distributed on the interval $[-1,1]$.

Determine the pdf (for continuous random variables) or pmf (for discrete random variables) for the distribution of the following random variable X.

9. Let X have cdf $F_X(x) = 4x$ for $x \in [0, 1/4]$.
10. Let X have cdf $F_X(x) = 1 - \frac{1}{\lfloor x \rfloor^k}$ for any $x \geq 1$, where $\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}$ is rounding down and $k \in \mathbb{N}$.

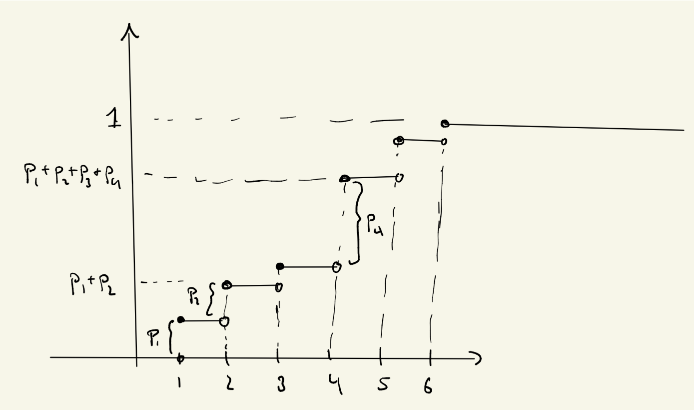
Solutions to Exercise 3: For 1: We use the same notation p_i for $i = 1,2,3,4,5,6$ that we used in Exercise 1.2. The cdf F_X has jumps at positions $x = 1,2,3,4,5,6$, otherwise it is constant. In particular,

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{i=1}^{x} \mathbb{P}(X = i) = \sum_{i=1}^{x} p_i\,,$$

for $x = 1,2,3,4,5,6$. Strictly to the left of $x = 1$ the cdf is zero since dice rolls do not have such values. To the right of $x = 6$ the cdf is equal to 1 since the dice takes a value between $-\infty$ and 6 with probability 1. Altogether we get

$$F_X(x) = \begin{cases} 0 & \text{if } x < 1 \\ \sum_{i=1}^{\lfloor x \rfloor} p_i & \text{if } x \in [1,6] \\ 1 & \text{if } x > 6 \end{cases},$$

where $\lfloor x \rfloor := \max\{n \in \mathbb{Z} : n \leq x\}$ is rounding down.



For 2: To determine F_X we have to compute the probability

$$\mathbb{P}(X \leq x) = \int_{-\infty}^{x} \rho_X(y) \mathrm{d}y = \frac{1}{2} \int_{-\infty}^{x} \mathrm{e}^{-|y|} \mathrm{d}y$$

for any $x \in \mathbb{R}$. It is best to distinguish the cases $x < 0$ and $x > 0$. Let us start with $x < 0$. Then

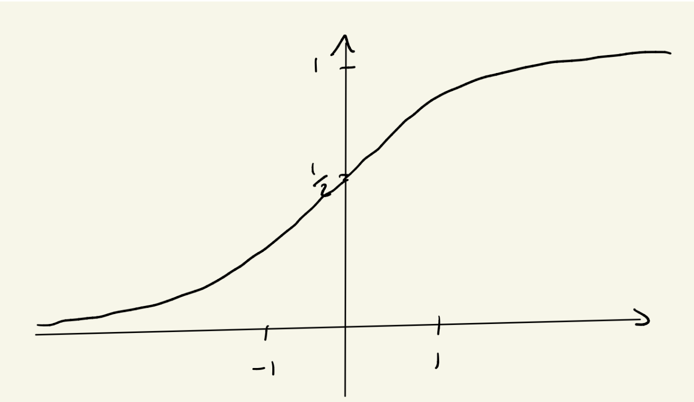
$$\mathbb{P}(X \leq x) = \frac{1}{2} \int_{-\infty}^{x} \mathrm{e}^{-|y|} \mathrm{d}y = \frac{1}{2} \int_{-\infty}^{x} \mathrm{e}^{y} \mathrm{d}y = \frac{1}{2} \mathrm{e}^{y}|_{-\infty}^{x} = \frac{\mathrm{e}^{x}}{2} \,.$$

For $x > 0$ we find

$$\mathbb{P}(X \leq x) = \mathbb{P}(X \leq 0) + \mathbb{P}(X \in (0,x)) = \frac{1}{2} + \frac{1}{2} \int_{0}^{x} \mathrm{e}^{-|y|} \mathrm{d}y = \frac{1}{2} + \frac{1}{2} \int_{0}^{x} \mathrm{e}^{-y} \mathrm{d}y = \frac{1}{2} + \frac{1 - \mathrm{e}^{-x}}{2} \,.$$

Altogether we found

$$F_X(x) = \begin{cases} \frac{\mathrm{e}^{x}}{2} & \text{if } x \leq 0 \\ \frac{1}{2} + \frac{1-\mathrm{e}^{-x}}{2} & \text{if } x > 0 \end{cases} \,.$$



For 3: Here $p = 1/2$, but we present the general case of $Z \sim \mathrm{Bern}(p)$ with $p \in (0,1)$. We compute the probability $\mathbb{P}(X \leq x)$. Here it is good to intersect the event with the events that either $Y = 0$ or $Y = 1$. Together these two events form a partition of the underlying probability space. Thus, we can compute

$$\begin{aligned} \mathbb{P}(X \leq x) &= \mathbb{P}(\alpha Y + Z \leq x) \\ &= \mathbb{P}(\{\alpha \cdot Y + Z \leq x, Y = 0\} \cup \{\alpha \cdot Y + Z \leq x, Y = 1\}) \\ &= \mathbb{P}(\alpha \cdot 0 + Z \leq x, Y = 0) + \mathbb{P}(\alpha \cdot 1 + Z \leq x, Y = 1) \\ &= \mathbb{P}(\mathbb{Z} \leq x, Y = 0) + \mathbb{P}(\alpha + Z \leq x, Y = 1) \,. \end{aligned}$$

Since Y and Z are independent, we can compute further

$$\mathbb{P}(Z \leq x, Y = 0) = \mathbb{P}(Z \leq x)\mathbb{P}(Y = 0) = \mathbb{P}(Z \leq x)(1-p)\,,$$

and

$$\mathbb{P}(\alpha + Z \leq x, Y = 1) = \mathbb{P}(\alpha + Z \leq x)\mathbb{P}(Y = 1) = \mathbb{P}(Z \leq x - \alpha)p\,.$$

Thus, we find

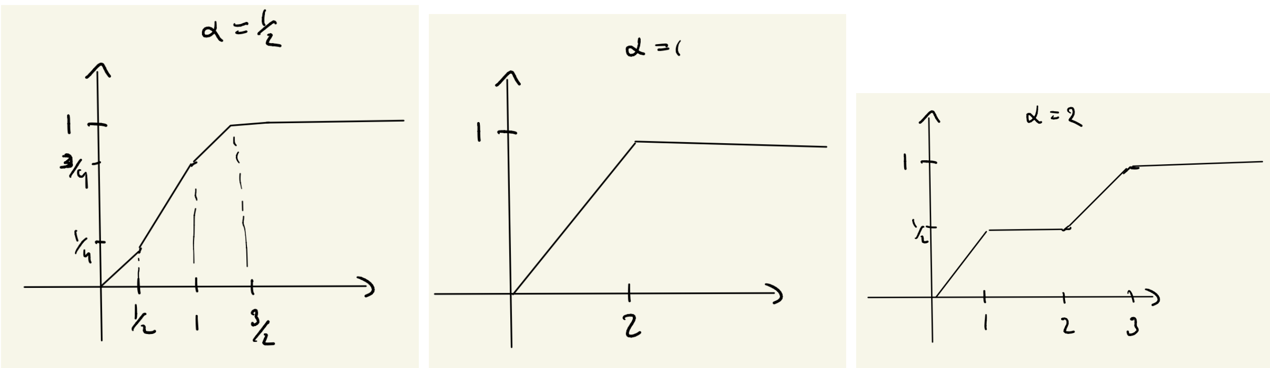
$$F_X(x) = (1-p)F_Z(x) + pF_Z(x-\alpha)\,.$$

Whet remains is to compute F_Z for Z uniformly distributed on $[0,1]$. For $x < 0$ the value $F_Z(x) = 0$ because Z does not take such values with non-vanishing probability. For $x \geq 1$ we have $F_Z(x) = 1$, because all values of X lie between 0 and 1. For $x \in [0,1]$ we compute

$$F_Z(x) = \int_0^x \mathbb{1}_{[0,1]}(y)\,\mathrm{d}y = \int_0^x 1\,\mathrm{d}y = x\,.$$

Thus, we have

$$F_Z(x) = \begin{cases} 0 & \text{if } x < 0 \\ x & \text{if } x \in [0,1] \\ 1 & \text{if } x > 1 \end{cases}\,. \tag{1}$$

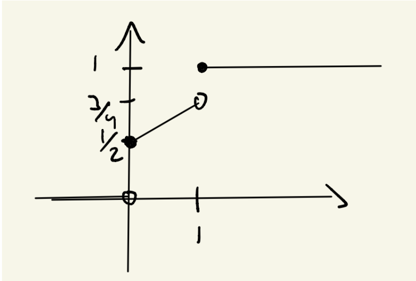


For 4: Here we have to compute a minimum of two independent random variables. This is done most convenient by considering the complement of $\mathbb{P}(X \leq x)$, because this will allow us to use independence directly. We compute

$$\begin{aligned}
\mathbb{P}(X \leq x) &= 1 - \mathbb{P}(X > x) \\
&= 1 - \mathbb{P}(\min\{Y, 2Z\} > x) \\
&\overset{(1)}{=} 1 - \mathbb{P}(Y > x \text{ and } 2Z > x) \\
&\overset{(2)}{=} 1 - \mathbb{P}(Y > x)\mathbb{P}(2Z > x) \\
&= 1 - (1 - \mathbb{P}(Y \leq x))(1 - \mathbb{P}(Z \leq x/2)) \\
&= 1 - (1 - F_Y(x))(1 - F_Z(x/2))\,.
\end{aligned}$$

In (1) we used that $\min\{Y, 2Z\} > x$ is the same as $Y > x$ and $2Z > x$. In (2) we used independence. We already computed F_Z in Exercise 3.3. Let us now compute F_Y for $Y \sim \mathrm{Bern}(p)$. For $x < 0$ we have $F_Y(x) = 0$ and for $x \geq 1$ we have $F_X(x) = 1$ because Y takes values in $\{0,1\}$. For $x \in [0,1$ we have $\mathbb{P}(Y \leq x) = \mathbb{P}(Y = 0) = 1 - p$. Thus,

$$F_Y(x) = (1-p)\mathbb{1}_{[0,1)}(x) + \mathbb{1}_{[1,\infty)}(x)\,. \tag{2}$$



For 5: As with the minimum in Exercise 3.4 it is best to work with the cdf. This time we consider a maximum, so taking complements is not helpful. We compute

$$\begin{aligned}\mathbb{P}(X \le x) &= \mathbb{P}(\max\{X_1, X_2, X_3\} \le x) \\ &\overset{(1)}{=} \mathbb{P}(X_1 \le x,\ X_2 \le x,\ X_3 \le x) \\ &\overset{(2)}{=} \mathbb{P}(X_1 \le x)\mathbb{P}(X_2 \le x)\mathbb{P}(X_3 \le x) \\ &\overset{(3)}{=} \mathbb{P}(X_1 \le x)^3 \\ &= F_{X_1}(x)^3\,.\end{aligned}$$

In (1) we used that the maximum being smaller or equal than x is equivalent all X_i being smaller or equal than x. In (2) we used that X_1, X_2, X_3 are independent. In (3) we used that X_1, X_2, X_3 all have the same distribution and therefore the probability $\mathbb{P}(X_i \le x)$ is the same for all of them.

It remains to compute F_{X_1}, where X_1 is the sum of two independent Bernoulli distributed random variables. For that purpose let Y, Z be i.i.d. Bern(p)-distributed. Then we find

$$\begin{aligned}\mathbb{P}(Y + Z \le x) &\overset{(1)}{=} \mathbb{P}(Y + Z \le x, Y = 0) + \mathbb{P}(Y + Z \le x, Y = 1) \\ &= \mathbb{P}(Z \le x, Y = 0) + \mathbb{P}(1 + Z \le x, Y = 1) \\ &\overset{(2)}{=} \mathbb{P}(Z \le x)\mathbb{P}(Y = 0) + \mathbb{P}(1 + Z \le x)\mathbb{P}(Y = 1) \\ &= F_Z(x)(1-p) + F_Z(x-1)p\,.\end{aligned}$$

In (1) we used that $\{Y = 0\}, \{Y = 1\}$ is a partition of the underlying sample space. In (2) we used independence. We computed the cdf of a Bernoulli random variable Y in (2). Thus, we find

$$F_{X_1}(x) = F_{Y+Z}(x) = \begin{cases} 0 & \text{if } x < 0 \\ (1-p)^2 & \text{if } x \in [0,1) \\ 1 - p^2 & \text{if } x \in [1,2) \\ 1 & \text{if } x \in [2,\infty) \end{cases}\,.$$

We could have gotten this also in different ways. The first alternative is to realise that $X_1 \sim \text{Bin}(2, p)$ and thus

$$\mathbb{P}(X_1 = 0) = (1-p)^2\,, \qquad \mathbb{P}(X_1 = 1) = 2p(1-p)\,, \qquad \mathbb{P}(X_1 = 2) = p^2\,. \tag{3}$$

We can also compute these probabilities by hand. Since X_1 is the sum $Y + Z$ of two Bernoulli random variables, it can only take the values $\{0, 1, 2\}$. For each we can compute the probability

$$\mathbb{P}(X_1 = 0) = \mathbb{P}(Y = 0, Z = 0) = \mathbb{P}(Y = 0)\mathbb{P}(Z = 0) = (1-p)^2\,,$$

and

$$\mathbb{P}(X_1 = 0) = \mathbb{P}(Y = 1, Z = 0) + \mathbb{P}(Y = 0, Z = 1) = \mathbb{P}(Y = 1)\mathbb{P}(Z = 0) + \mathbb{P}(Y = 0)\mathbb{P}(Z = 1) = 2p(1-p)\,,$$

and

$$\mathbb{P}(X_1 = 2) = \mathbb{P}(Y = 1, Z = 1) = \mathbb{P}(Y = 1)\mathbb{P}(Z = 1) = p^2\,.$$

From (3) we can compute the cdf of X_1 via

$$\begin{aligned} F_{X_1}(x) &= \mathbb{P}(X_1 = 0)\mathbb{1}_{[0,1)}(x) + \mathbb{P}(X_1 \le 1)\mathbb{1}_{[1,2)}(x) + \mathbb{1}_{[2,\infty)}(x) \\ &= (1-p)^2\mathbb{1}_{[0,1)}(x) + ((1-p)^2 + 2p(1-p))\mathbb{1}_{[1,2)}(x) + \mathbb{1}_{[2,\infty)}(x) \\ &= (1-p)^2\mathbb{1}_{[0,1)}(x) + (1-p^2)\mathbb{1}_{[1,2)}(x) + \mathbb{1}_{[2,\infty)}(x)\,. \end{aligned}$$

Back to the random variable $X = \max\{X_1, X_2, X_3\}$ with cdf $F_X(x) = F_{X_1}(x)^3$. It is discrete wit value $\{0,1,2\}$. We are asked to compute its pmf $p_X : \{0,1,2\} \to [0,1]$. This means simply that we have to compute the jump heights of F_X, namely

$$p_X(0) = \mathbb{P}(X = 0) = F_X(0) = F_{X_1}(0)^3 = (1-p)^6\,,$$

and

$$\begin{aligned} p_X(1) &= \mathbb{P}(X = 1) \\ &= F_X(1) - \lim_{x\uparrow 1} F_X(x) \\ &= F_X(1) - F_X(0) \\ &= F_{X_1}(1)^3 - (1-p)^6 \\ &= (1-p^2)^3 - (1-p)^6\,, \end{aligned}$$

and

$$\begin{aligned} p_X(2) &= \mathbb{P}(X = 2) \\ &= F_X(2) - \lim_{x\uparrow 2} F_X(x) \\ &= F_X(2) - F_X(1) \\ &= F_{X_1}(2)^3 - (1-p^2)^3 + (1-p)^6 \\ &= 1 - (1-p^2)^3 + (1-p)^6\,. \end{aligned}$$

For 6: Since X is the sum of independent continuous random variables we can use the convolution rule to compute the pdf of X. This is what we do first through

$$\begin{aligned} \rho_{X_1+X_2}(x) &= \int_{\mathbb{R}} \rho_{X_1}(y)\rho_{X_2}(x-y)\mathrm{d}y \\ &= \frac{1}{4}\int_{\mathbb{R}} \mathbb{1}_{[-1,1]}(y)\mathbb{1}_{[-1,1]}(x-y)\mathrm{d}y \\ &= \frac{1}{4}|[-1,1] \cap [x-1, x+1]| \\ &= \frac{1}{4}|[-1,1] \cap [x-1, x+1]|\,. \end{aligned}$$

Here the absolute value around the intervals means its length. At this point we distinguish two cases. If $x \in [-2, 0]$, then

$$[-1,1] \cap [x-1, x+1] = [-1, x+1]$$

and if $x \in [0, 2]$, then

$$[-1,1] \cap [x-1, x+1] = [x-1, 1]\,.$$

For all other cases $x \in \mathbb{R} \setminus [-2, 2]$ the intersection is empty and $\rho_X(x) = 0$. We conclude

$$\rho_X(x) = \begin{cases} \frac{x+2}{4} & \text{if } x \in [-2, 0] \\ \frac{2-x}{4} & \text{if } x \in [0, 2] \end{cases} = \frac{2-|x|}{4}\mathbb{1}_{[-2,2]}(x)\,. \tag{4}$$

For good reason this is called the triangular distribution (draw the function).

Now we compute the cdf for X. Since we already determined ρ_X this is now simple via

$$\mathbb{P}(X \le x) = \int_{-\infty}^{x} \rho_X(y)\mathrm{d}y\,.$$

In particular, $F_X(x) = 0$ for $x < -2$ and $F_X(x) = 1$ for $x > 2$. Now let $x \in [-2, 0]$. Then

$$F_X(x) = \int_{-\infty}^{x} \rho_X(y)\mathrm{d}y = \int_{-2}^{x} \rho_X(y)\mathrm{d}y = \frac{1}{4}\int_{-2}^{x}(y+2)\mathrm{d}y = \frac{x^2-4}{8} + \frac{x+2}{2}\,.$$

For $x \in [0, 2]$ we find

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \rho_X(y)\mathrm{d}y \\ &= \int_{-2}^{0} \rho_X(y)\mathrm{d}y + \int_{0}^{x} \rho_X(y)\mathrm{d}y \\ &\overset{(1)}{=} \frac{1}{2} + \frac{1}{4}\int_0^x (2-y)\mathrm{d}y \\ &= \frac{1}{2} + \frac{4x - x^2}{8}\,. \end{aligned}$$

In (1) we used $F_X(0) = \frac{1}{2}$ which we already computed. As a sanity check we see that $F_X(2) = 1$, which is good because X is a continuous random variable.

For 7: Since we compute the minimum of two independent random variables, we start with the cdf and take the complement. We compute

$$\mathbb{P}(X \le x) = 1 - \mathbb{P}(X > x) = 1 - \mathbb{P}(X_1 > x, X_2 > x) = 1 - \mathbb{P}(X_1 > x)\mathbb{P}(X_2 > x) = 1 - (1 - F_{X_1}(x))^2\,.$$

Now let us compute F_{X_1}. Since X_1 is uniformly distributed on $[-1, 1]$ we have $F_{X_1}(x) = 0$ for $x < -1$ and $F_X(x) = 1$ for $x > 2$. For $x \in [-1, 1]$ we find

$$F_{X_1}(x) = \frac{1}{2}\int_{-1}^{x} \mathrm{d}y = \frac{x+1}{2}\,.$$

We conclude that $F_X(x) = 0$ for $x < -1$ and $F_X(x) = 1$ for $x > 0$. In between we have for $x \in [-1, 1]$ the formula

$$F_X(x) = 1 - \left(1 - \frac{x+1}{2}\right)^2 = 1 - \frac{(x-1)^2}{4}\,.$$

To determine the pdf $\rho_X(x) = F_X'(x)$ we differentiate and find

$$\rho_X(x) = \frac{1-x}{2}\mathbb{1}_{[-1,1]}(x)\,.$$

For 8: Here we use that $X_1 + X_2 + X_3$ is the sum of the two independent random variables $Y = X_1 + X_2$ and $Z = X_3$ (since X_1, X_2, X_3 are independent). We computed the pdf ρ_Y of $X_1 + X_2$ in Exercise 3.6. It is the triangular distribution

$$\rho_Y(x) = \frac{2 - |x|}{4}\mathbb{1}_{[-2,2]}(x)\,.$$

Now we use the convolution rule to compute the pdf of $X = Y + Z$ through

$$\rho_X(x) = \int_{\mathbb{R}} \rho_Y(y)\rho_Z(x-y)\mathrm{d}y = \int \frac{2-|y|}{4}\mathbb{1}_{[-2,2]}(y)\mathbb{1}_{[-1,1]}(x-y)\mathrm{d}y\,.$$

The integrand is non-vanishing only on the interval $[-2, 2] \cap [-1 + x, 1 + x]$. Thus, we distinguish four cases. For $x \in \mathbb{R} \setminus [-3, 3]$ we have $\rho_X(x) = 0$. For $x \in [-3, -1]$ we have $[-2, 2] \cap [-1 + x, 1 + x] = [-2, 1 + x]$ and

$$\rho_X(x) = \int_{-2}^{1+x} \frac{2-|y|}{4}\mathrm{d}y = \int_{-2}^{1+x} \frac{2+y}{4}\mathrm{d}y \overset{(1)}{=} \frac{1}{4}\int_0^{3+x} y\,\mathrm{d}y = \frac{3+x}{4}\,.$$

In (1) we shifted the integration variable $y \to y-2$. For $x \in [-1,1]$ we have $[-2,2] \cap [-1+x, 1+x] = [-1+x, 1+x]$ and

$$\begin{aligned}\rho_X(x) &= \int_{-1+x}^{1+x} \frac{2-|y|}{4} \mathrm{d}y \\ &\overset{(1)}{=} \int_{-1+x}^{0} \frac{2+y}{4} \mathrm{d}y + \int_{0}^{1+x} \frac{2-y}{4} \mathrm{d}y \\ &\overset{(2)}{=} \int_{1+x}^{2} \frac{y}{4} \mathrm{d}y + \int_{1-x}^{2} \frac{y}{4} \mathrm{d}y \\ &= \frac{1}{8}(8-(1+x)^2-(1-x)^2) \\ &= \frac{3-x^2}{4} .\end{aligned}$$

In (1) we split the integral according to whether $y > 0$ or $y < 0$ in order to simplify integration of the absolute value. In (2) we shifted integration variables in both integrals. Instead of computing the last case $x \in [1,3]$ (which is very similar to $x \in [-3,-1]$) we realise that $\rho_X(x)$ must be symmetric, i.e. it must satisfy $\rho_X(-x) = \rho_X(x)$. This is because for each i the random variables X_i and $-X_i$ have the same distribution and therefore X and $-X$ have the same distribution. Thus, we conclude

$$\rho_X(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus [-3,3] \\ \frac{3+x}{4} & \text{if } x \in [-3,-1] \\ \frac{3-x^2}{4} & \text{if } x \in [-1,1] \\ \frac{3-x}{4} & \text{if } x \in [1,3] \end{cases} .$$

Finally we compute the cdf of X by integrating ρ_X in the formula

$$F_X(x) = \int_{-\infty}^{x} \rho_X(y) \mathrm{d}y .$$

The solution again requires the case distinction $x \in (-\infty,-3)$, $x \in [-3,-1)$, $x \in [-1,1)$, $x \in [1,3)$ and $x \in (3,\infty)$. Since it is a longish but simple (and not very enlightening) exercise in integration it is left to the reader.

For 9: The random variable X must be continuous. Since $F_X(1/4) = 1$ we conclude from monotonicity of F_X and that $F_X : \mathbb{R} \to [0,1]$ that $F_X(x) = 1$ for all $x > 1/4$. Similarly, we conclude from $F_X(0) = 0$ that $F_X(x) = 0$ for all $x < 0$. In particular, ρ_X must vanish on $\mathbb{R} \setminus [0,1/4]$. Finally for $x \in [0,1/4]$ we compute $\rho_X(x) = F_X'(x) = 4$. Thus

$$\rho_X(x) = 4\mathbb{1}_{[0,1/4]}(x) ,$$

i.e. X is uniformly distributed on $[0,1/4]$.

For 10: Since $F_X(1) = 0$ we see that X can only take values greater or equal to 1. Furthermore, $F_X(x)$ has jumps at every $x \in \mathbb{N}$ and is otherwise constant. We conclude that X must be discrete with values in $\mathbb{N}$. Now we compute the pmf $p_X : \mathbb{N} \to [0,1]$ of X by computing the jump heights

$$p_X(x) = \mathbb{P}(X = x) = F_X(x) - \lim_{y \uparrow x} F_X(y) = F_X(x) - F_X(x-1) = \frac{1}{(x-1)^k} - \frac{1}{x^k} .$$

Exercise 4 (Joint and Marginal Distributions): Determine the joint distribution (by giving either pdf or pmf of (X,Y)) of the following random variables X and Y:

1. Let X, Y be i.i.d. and X have values in $\{-1,0,1\}$ with

$$\mathbb{P}(X = 0) = 2\mathbb{P}(X = -1) = 2\mathbb{P}(X = 1) = 1/2 .$$

2. Let X, Y be i.i.d. $\mathrm{Bern}(p)$ distributed random variables.
3. Let X, Y be jointly Gaussian with $\mathbb{E}X = \mathbb{E}Y = 0$ and $\operatorname{Var} X = \operatorname{Var} Y = 5$ and $\operatorname{Cov}(X, Y) = -3$.
4. Let X, Y be i.i.d. random variables that have the same distribution as the sum of two independent random variables that are uniformly distributed on $[-1, 1]$.

Determine the first marginal distribution (distribution of X) of the random pair (X, Y) with the following joint distribution.

5. Let (X, Y) be distributed as in 3 above.
6. Let (X, Y) have values in $\{g, b, r\} \times \{0, 1\}$ with probabilities summarised in the following table:

$\mathbb{P}(X = x, Y = y)$	$Y = 0$	$Y = 1$
$X = g$	$1/12$	$1/6$
$X = b$	$1/3$	$5/24$
$X = r$	$1/8$	$1/12$

7. Let (X, Y) have values in $\{0, 1, 2, \dots\}^2$ with joint pmf $\mathbb{P}(X = k, Y = l) = \mathrm{e}^{-2\lambda} \frac{\lambda^{k+l}}{k!l!}$.
8. Let (X, Y) be a continuous random variable with values in $\mathbb{R}^2$ and pdf

$$\rho_{(X,Y)}(x, y) = \frac{6}{7}(x + y)^2 \mathbb{1}_{[0,1]}(x) \, \mathbb{1}_{[0,1]}(y) \,.$$

Solutions to Exercise 4: For 1: The joint distribution is a discrete probability distribution on $\{-1, 0, 1\}^2$ and is determined by its pmf $p_{(X,Y)} : \{-1, 0, 1\}^2 \to [0, 1]$ given as

$$p_{(X,Y)}((i, j)) = \mathbb{P}(X = i, Y = j) = \mathbb{P}(X = i)\mathbb{P}(Y = j) = p_i p_j \,,$$

for $i, j = -1, 0, 1$ and where $p_{-1} := \frac{1}{4}$, $p_0 := \frac{1}{2}$, $p_1 := \frac{1}{4}$.

For 2: Here the joint distribution is discrete on $\{0, 1\}^2$ and its pmf is

$$p_{(X,Y)}((0, 0)) = (1 - p)^2 \,, \quad p_{(X,Y)}((1, 0)) = p(1 - p) \,, \quad p_{(X,Y)}((0, 1)) = p(1 - p) \,, \quad p_{(X,Y)}((1, 1)) = p^2 \,.$$

For 3: Since we are given expectation and covariances and since (X, Y) is jointly Gaussian, we know its distribution. It is $(X, Y) \sim N(0, A)$, where

$$A = \begin{pmatrix} \operatorname{Var} X & \operatorname{Cov}(X, Y) \\ \operatorname{Cov}(X, Y) & \operatorname{Var} Y \end{pmatrix} = \begin{pmatrix} 5 & -3 \\ -3 & 5 \end{pmatrix} .$$

The corresponding pdf is

$$\begin{aligned} \rho_{(X,Y)}(x, y) &= \frac{1}{2\pi\sqrt{\det A}} \exp\left(-\frac{1}{2}(x, y) A^{-1} \begin{pmatrix} x \\ y \end{pmatrix} \right) \\ &= \frac{1}{2\pi\sqrt{16}} \exp\left(-\frac{1}{32}(5x^2 + 6xy + 5y^2) \right) . \end{aligned}$$

For 4: Since X, Y are independent, their joint distribution factorises and we have

$$\rho_{(X,Y)}(x, y) = \rho_X(x)\rho_Y(y) = \frac{(2 - |x|)(2 - |y|)}{16} \mathbb{1}_{[-2,2]^2}(x, y) \,,$$

where we used the pdf of the sum of two i.i.d. random variables that are uniformly distributed on $[-1, 1]$ from (4).

For 5: Since (X, Y) from Exercise 4.3 is a 2-dimensional Gaussian distribution, the marginal X is also Gaussian and thus $X \sim N(\mathbb{E}X, \operatorname{Var} X)$. But $\mathbb{E}X = 0$ and $\operatorname{Var} X = 5$ according to Exercise 4.3 and, thus, we have $X \sim N(0, 5)$.

For 6: The random variable X has values in $\{g, b, r\}$. We compute its pmf via

$$p_X(g) = \mathbb{P}(X = g) = \mathbb{P}(X = g, Y = 0) + \mathbb{P}(X = g, Y = 1) = \frac{1}{12} + \frac{1}{6} = \frac{1}{4},$$

and

$$p_X(b) = \mathbb{P}(X = b) = \mathbb{P}(X = b, Y = 0) + \mathbb{P}(X = b, Y = 1) = \frac{1}{3} + \frac{5}{24} = \frac{13}{24},$$

and

$$p_X(r) = \mathbb{P}(X = r) = \mathbb{P}(X = r, Y = 0) + \mathbb{P}(X = r, Y = 1) = \frac{1}{8} + \frac{1}{12} = \frac{5}{24}.$$

For 7: The pmf of (X, Y) has a product structure. Thus, X, Y are independent. We identify that that $\mathbb{P}(X = k) = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!}$ and $\mathbb{P}(Y = l) = \frac{\mathrm{e}^{-\lambda}\lambda^l}{l!}$ are both the pmf of the $\mathrm{Poi}(\lambda)$-distribution. Thus, X, Y are i.i.d. $\mathrm{Poi}(\lambda)$-distributed random variables.

For 8: We use that the pdf of the marginal distribution result from integrating over the variable corresponding to Y, i.e.

$$\rho_X(x) = \int_{\mathbb{R}} \rho_{(X,Y)}(x, y)\mathrm{d}y = \frac{6}{7}\mathbb{1}_{[0,1]}(x) \int_0^1 (x + y)^2 \,\mathrm{d}y = \frac{6}{7}\mathbb{1}_{[0,1]}(x) \int_x^{1+x} y^2 \,\mathrm{d}y = \frac{2}{7}(3x^2 + 3x + 1)\mathbb{1}_{[0,1]}(x).$$

Thus, X is a continuous random variable with pdf as above.

Exercise 5 (Conditional Probabilities): We are presented with a hat with $B \in \mathbb{N}$ blue and $R \in \mathbb{N}$ red balls in it. In the first round we blindly pick a ball from the hat. Then we return to the hat the ball we picked and add another one of the same colour (increasing the number of balls in the hat by 1). In the second round we proceed in the same way, but we add two balls of the picked colour (increasing the number of balls in the hat by 2 compared to before the second round). Afterwards we pick a ball blindly from the hat. What is the probability that this last ball is red?

Hint: Use the law of total probability.

Solutions to Exercise 5: Let X_1, X_2 and X_3 random variables with values in $\{b, r\}$, indicating the colours blue and red, that model our picks in each of the three rounds. We are asked the probability $\mathbb{P}(X_3 = r)$. To compute it we use the law of total probability and compute

$$\begin{aligned}
\mathbb{P}(X_3 = r) &= \mathbb{P}(X_3 = r|X_2 = r, X_1 = r)\mathbb{P}(X_2 = r, X_1 = r) \\
&\quad + \mathbb{P}(X_3 = r|X_2 = r, X_1 = b)\mathbb{P}(X_2 = r, X_1 = b) \\
&\quad + \mathbb{P}(X_3 = r|X_2 = b, X_1 = r)\mathbb{P}(X_2 = b, X_1 = r) \\
&\quad + \mathbb{P}(X_3 = r|X_2 = b, X_1 = b)\mathbb{P}(X_2 = b, X_1 = b) \\
&= \mathbb{P}(X_3 = r|X_2 = r, X_1 = r)\mathbb{P}(X_2 = r|X_1 = r)\mathbb{P}(X_1 = r) \\
&\quad + \mathbb{P}(X_3 = r|X_2 = r, X_1 = b)\mathbb{P}(X_2 = r|X_1 = b)\mathbb{P}(X_1 = b) \\
&\quad + \mathbb{P}(X_3 = r|X_2 = b, X_1 = r)\mathbb{P}(X_2 = b|X_1 = r)\mathbb{P}(X_1 = r) \\
&\quad + \mathbb{P}(X_3 = r|X_2 = b, X_1 = b)\mathbb{P}(X_2 = b|X_1 = b)\mathbb{P}(X_1 = b).
\end{aligned}$$

Now we simply compute all of these probabilities. This is simple since we given the past picks we know the number of balls of each colour in the hat. We find

$$\begin{aligned}
\mathbb{P}(X_1 = b) &= \frac{B}{R+B}\,, \\
\mathbb{P}(X_1 = r) &= \frac{R}{R+B}\,, \\
\mathbb{P}(X_2 = b|X_1 = b) &= \frac{B+1}{R+B+1}\,, \\
\mathbb{P}(X_2 = b|X_1 = r) &= \frac{B}{R+B+1}\,, \\
\mathbb{P}(X_2 = r|X_1 = b) &= \frac{R}{R+B+1}\,, \\
\mathbb{P}(X_2 = r|X_1 = r) &= \frac{R+1}{R+B+1}\,, \\
\mathbb{P}(X_3 = r|X_1 = b, X_2 = b) &= \frac{R}{R+B+3}\,, \\
\mathbb{P}(X_3 = r|X_1 = b, X_2 = r) &= \frac{R+2}{R+B+3}\,, \\
\mathbb{P}(X_3 = r|X_1 = r, X_2 = b) &= \frac{R+1}{R+B+3}\,, \\
\mathbb{P}(X_3 = r|X_1 = r, X_2 = r) &= \frac{R+3}{R+B+3}\,.
\end{aligned}$$

Inserting these gives the result.

Exercise 6 (Expectation and Variance): Determine the expectation and variance of the following random variable X:

1. Let X be a roll of the unfair dice from Exercise 1.2.
2. Let $X = 4Y + 3$, where $Y \sim \mathrm{Bern}(p)$
3. Let $X \sim \mathrm{Bin}(n, p)$.
4. Let $X = X_1 + X_2$ with X_1, X_2 i.i.d. random variables that are uniformly distributed on $[-1, 1]$.
5. Let X be a continuous random variable with pdf $\rho_X(x) = \frac{2}{\Gamma(1/4)}\mathrm{e}^{-x^4}$. Evaluating the Γ-function is not required.
6. Let X be the marginal of the 2-dimensional random vector vector (X, Y) with pdf
$$\rho_{(X,Y)}(x, y) = \frac{\sqrt{20}}{2\pi}\mathrm{e}^{-2x^2+2xy-3y^2}\,.$$
7. Let $X = Y + Z + W$ with independent random variables Y, Z, W such that $Y \sim \mathrm{Exp}(1)$, $Z \sim \mathrm{Bin}(3, 1/2)$ and $W \sim \mathrm{Poi}(1)$.
8. Let $X = \log Y$, where Y is uniformly distributed on $[0, 1]$.

Solutions to Exercise 6: For 1: We use the notation from Exercise 1.2. Then

$$\mathbb{E}X = 1 \cdot p_1 + 2 \cdot p_2 + 3 \cdot p_3 + 4 \cdot p_4 + 5 \cdot p_5 + 6 \cdot p_6 = 3.433\,.$$

For the variance we first compute the second moment

$$\mathbb{E}X = 1^2 \cdot p_1 + 2^2 \cdot p_2 + 3^2 \cdot p_3 + 4^2 \cdot p_4 + 5^2 \cdot p_5 + 6^2 \cdot p_6 = 13.539\,.$$

Then

$$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 1.753511\,.$$

For 2: First we compute expectation and variance of Y (or we look it up). We get

$$\mathbb{E}Y = 1 \cdot \mathbb{P}(Y = 1) = p\,,$$

and

$$\operatorname{Var} Y = (0-p)^2\mathbb{P}(Y=0) + (1-p)^2\mathbb{P}(Y=1) = p^2(1-p) + (1-p)^2 p = p(1-p)\,.$$

Now we compute the expectation of X using linearity of the expectation

$$\mathbb{E}X = \mathbb{E}(4Y+3) = 4\mathbb{E}Y + 3 = 4p + 3\,,$$

and the variance using its transformation rules

$$\operatorname{Var}(4Y+3) = 4^2 \operatorname{Var} Y = 16p(1-p)\,.$$

For 3: We determined the expectation in the lecture. It is $\mathbb{E}X = np$. Now we compute the second moment,

$$\begin{aligned}
\mathbb{E}X^2 &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} \\
&= \sum_{k=1}^{n} \frac{k\,n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} \\
&\overset{(1)}{=} np \sum_{k=0}^{n-1} \frac{(k+1)\,(n-1)!}{k!(n-k-1)!} p^k (1-p)^{n-k-1} \\
&\overset{(2)}{=} np \sum_{k=0}^{n-1} \frac{k\,(n-1)!}{k!(n-k-1)!} p^k (1-p)^{n-k-1} + np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} \\
&\overset{(3)}{=} np \sum_{k=1}^{n-1} \frac{(n-1)!}{(k-1)!(n-k-1)!} p^k (1-p)^{n-k-1} + np \\
&\overset{(4)}{=} n(n-1)p^2 \sum_{k=0}^{n-2} \binom{n-2}{k} p^k (1-p)^{n-k-2} + np \\
&\overset{(5)}{=} n(n-1)p^2 + np\,.
\end{aligned}$$

Here, in (1) we changed the summation index $k \to k+1$. In (2) we pulled the sums apart. In (3) we used the binomial theorem to see that the last sum equals 1. In (4) we again changed the summation index $k \to k+1$ and in (5) we used again the binomial theorem. Note the similarity in. strategy to Exercise 5.3 on Exercise Sheet 2. Finally, we compute the variance

$$\operatorname{Var} X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = n(n-1)p^2 + np - n^2p^2 = np(1-p)\,.$$

Another simple way to see this result is to use that X has the same distribution as $\sum_{i=1}^n X_i$ where X_i are i.i.d. Bern(p)-distributed random variables. Therefore by the Bienaymé formula we get

$$\operatorname{Var} X = \operatorname{Var} \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} \operatorname{Var} X_i = \sum_{i=1}^{n} p(1-p) = np(1-p)\,.$$

For 4: We computed the pdf of X in (4), namely

$$\rho_X(x) = \frac{2-|x|}{4}\mathbb{1}_{[-2,2]}(x)\,.$$

Thus, we can compute $\mathbb{E}X$ and $\mathbb{E}X^2$ by simply integrating ρ_X. However, it is quicker to use the computational rules for expectation and variance and that

$$\mathbb{E}X_1 = \mathbb{E}X_2 = 0\,, \qquad \operatorname{Var} X_1 = \operatorname{Var} X_2 = \frac{1}{2}\int_{-1}^{1} x^2 \mathrm{d}x = \frac{1}{3}\,.$$

For the expectation we get $\mathbb{E}X = \mathbb{E}X_1 + \mathbb{E}X_2 = 0$ and for the variance by the Bienaymé formula

$$\operatorname{Var} X = \operatorname{Var} X_1 + \operatorname{Var} X_2 = \frac{2}{3}\,.$$

For 5: By symmetry $\mathbb{E}X = 0$. More precisely, since $\rho_X(x)$ is a symmetric function $f(x) := x\rho_X(x)$ is antisymmetric, i.e. $f(-x) = f(x)$. This implies

$$\mathbb{E}X = \int f(x)\mathrm{d}x \overset{(1)}{=} \int f(-x)\mathrm{d}x \overset{(2)}{=} -\int f(x)\mathrm{d}x = -\mathbb{E}X\,,$$

where we used the change of variables $x \to -x$ in (1) and the antisymmetry in (2). For the variance we compute

$$\frac{\Gamma(1/4)}{2}\operatorname{Var} X = \int x^2 \mathrm{e}^{-x^4}\mathrm{d}x = 2\int_0^\infty x^2 \mathrm{e}^{-x^4}\mathrm{d}x = \frac{1}{2}\int_0^\infty \mathrm{e}^{-x} x^{-1/4}\mathrm{d}x = \frac{\Gamma(3/4)}{2}\,.$$

Thus,

$$\operatorname{Var} X = \frac{\Gamma(3/4)}{\Gamma(1/4)}\,.$$

For 6: There are two ways to approach this problem. The naive approach is to use that

$$\rho_X(x) = \int \rho_{(X,Y)}(x,y)\mathrm{d}y\,,$$

and to perform the integral using the Gaussian integration formula from Exercise 6 on Exercise Sheet 2. However, it is easier to use that $(X,Y) \sim N(0,A)$ is a 2-dimensional Gaussian vector with mean zero and covariance matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ -2 & 6 \end{pmatrix}^{-1} = \frac{1}{10}\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\,.$$

To see this we realise that the exponent can be written in the form

$$-2x^2 + 2xy - 3y^2 = -\frac{1}{2}(x,y)\begin{pmatrix} 4 & -2 \\ -2 & 6 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = -\frac{1}{2}(x,y)A^{-1}\begin{pmatrix} x \\ y \end{pmatrix}\,.$$

We conclude that the marginal X is also Gaussian with $X \sim N(0, a_{11}) = N(0, \frac{3}{10})$ and, thus, $\mathbb{E}X = 0$ and $\operatorname{Var} X = \frac{3}{10}$.

For 7: For the individual random variables we have

$$\mathbb{E}Y = 1\,, \quad \operatorname{Var} Y = 1\,, \quad \mathbb{E}Z = \frac{3}{2}\,, \quad \operatorname{Var} Z = \frac{3}{4}\,, \quad \mathbb{E}W = 1\,, \quad \operatorname{Var} W = 1\,.$$

All of these we computed already during tis course at some point. Now we use the computational rules for expectation and variance (Bienaymé formula) to get

$$\mathbb{E}X = \mathbb{E}Y + \mathbb{E}Z + \mathbb{E}W = \frac{7}{2}\,, \quad \operatorname{Var} X = \operatorname{Var} Y + \operatorname{Var} Z + \operatorname{Var} W = \frac{11}{4}\,.$$

For 8: We use the formula $\mathbb{E}f(Y) = \int f(x)\rho_Y(x)\mathrm{d}x$ with $f(x) = \log x$ to compute

$$\mathbb{E}X = \int_0^1 \log x \,\mathrm{d}x = x\log x|_0^1 - \int_0^1 \mathrm{d}x = -1\,,$$

where integration by parts was used.

Exercise 7 (Law of Large Numbers): Let $f : [0,1] \to \mathbb{R}$ be a continuous function. We determine the integral $\int_0^1 f(x)\mathrm{d}x$ via the following randomised approach (Monte-Carlo-Method). Let $X_1, X_2, \ldots, X_n$ be i.i.d. random variables that are uniformly distributed on $[0,1]$. We set $F_n := \frac{1}{n}\sum_{i=1}^n f(X_i)$. Prove that for any threshold $\varepsilon > 0$ the sequence $F_1, F_2, \ldots$ approximates the integral eventually, i.e. that

$$\mathbb{P}\big(\big|F_n - \textstyle\int_0^1 f(x)\mathrm{d}x\big| \geq \varepsilon\big) \to 0\,, \qquad \text{as} \qquad n \to \infty\,.$$

Solution to Exercise 7: The random variables $f(X_1), f(X_2), \ldots, f(X_n)$ are i.i.d. because $X_1, X_2, \ldots, X_n$ are i.i.d.. Thus, we can use the weak law of large numbers to conclude

$$\mathbb{P}\big(|F_n - \mathbb{E}f(X_1)| \geq \varepsilon\big) \leq \frac{\operatorname{Var} f(X_1)}{\varepsilon^2 n} \to 0\,, \qquad n \to \infty\,.$$

It remains to compute

$$\mathbb{E}f(X_1) = \int f(x)\rho_{X_1}(x)\mathrm{d}x = \int_0^1 f(x)\mathrm{d}x\,.$$