Advanced information, computation, communication I
EPFL - Fall semester 2021-2022

Week 13 — solutions

December 18, 2021

1 Open Questions

Exercise 1. (*) *The covariance of two random variables X and Y on a sample space S, denoted by $Cov(X,Y)$, is defined as*

$$\mathrm{Cov}(X,Y) = \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))).$$

1. *Show that* $\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ *and use this result to conclude that* $\mathrm{Cov}(X,Y) = 0$ *if X and Y are independent variables.*
   We notice that $\mathbb{E}(X)$ and $\mathbb{E}(Y)$ are real numbers, thus it follows from the linearity of the expected value that $\mathbb{E}(Y\mathbb{E}(X)) = \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X\mathbb{E}(Y))$ and that $\mathbb{E}(\mathbb{E}(X)\mathbb{E}(Y)) = \mathbb{E}(X)\mathbb{E}(Y)$. Then we get

$$\begin{aligned}
\mathrm{Cov}(X,Y) &= \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) \\
&= \mathbb{E}(XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y)) \\
&= \mathbb{E}(XY) - \mathbb{E}(X\mathbb{E}(Y)) - \mathbb{E}(Y\mathbb{E}(X)) + \mathbb{E}(\mathbb{E}(X)\mathbb{E}(Y)) \\
&= \mathbb{E}(XY) - 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y) \\
&= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y).
\end{aligned}$$

   If X and Y are independent, then we know that $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. This implies that

$$\mathrm{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = 0$$

2. *Show that* $V(X+Y) = V(X) + V(Y) + 2\mathrm{Cov}(X,Y)$.
   With the linearity of the expected value we get

$$\begin{aligned}
V(X+Y) = \mathbb{E}((X+Y)^2) - \mathbb{E}(X+Y)^2 &= \mathbb{E}(X^2 + 2XY + Y^2) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\
&= \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - (\mathbb{E}(X)^2 + 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(Y)^2) \\
&= \mathbb{E}(X^2) - \mathbb{E}(X)^2 + \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 + 2(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)) \\
&= V(X) + V(Y) + 2\mathrm{Cov}(X,Y).
\end{aligned}$$

Exercise 2. (*) *Let X be a random variable that outputs a uniformly chosen number in* $\{-10, -9, \ldots, 9, 10\}$, *let Y be a random variable that outputs a uniformly chosen number in* $\{10, 11, \ldots, 29, 30\}$, *and let Z be the random variable that outputs $X + 20$.*

1. *Compute* $\mathbb{E}[X], \mathbb{E}[Y]$ *and* $\mathbb{E}[Z]$.
   It is straightforward to see that $\mathbb{E}[X] = 0$ since i and $-i$ appear with equal probability for all $0 \le i \le 10$. This means that $\mathbb{E}[Z] = \mathbb{E}[X + 20] = \mathbb{E}[X] + E[20] = 0 + 20 = 20$, and since Y has the same distribution as Z, $\mathbb{E}[Y] = 20$.

2. *Compute* $V(X), V(Y)$ *and* $V(Z)$.
   The variance of X is
$$V(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] = \sum_{i=-10}^{10} i^2 \frac{1}{21} = \frac{2}{21} \sum_{i=1}^{10} i^2 = \frac{2 \times 385}{21} = \frac{110}{3}.$$
   Since $X = Z - \mathbb{E}[Z]$ we have $V(Z) = V(X) = \frac{110}{3}$, and the same is true for Y as $Y - \mathbb{E}[Y]$ is a uniform random variable in $\{-10, \dots, 10\}$ so it has the same distribution as X and thus $V(Y) = \frac{110}{3}$.

3. *Compute* $\mathrm{Cov}(X, Y), \mathrm{Cov}(Y, Z)$ *and* $\mathrm{Cov}(X, Z)$.
   Note that X and Y are independent, so $\mathrm{Cov}(X, Y) = 0$ as we have seen in exercise 2. Since the same is true for Z and Y we have $\mathrm{Cov}(Y, Z) = 0$. On the other hand $\mathrm{Cov}(X, Z) \neq 0$ because X and $Z = X + 20$ are very much dependent of each other.
$$\begin{aligned}\mathrm{Cov}(X, Z) &= \mathbb{E}[XZ] - \mathbb{E}[X]\mathbb{E}[Z] = \mathbb{E}[X(X + 20)] - 0 \cdot 20 = \\ &= \mathbb{E}[X^2 + 20X] = \mathbb{E}[X^2] + 20\mathbb{E}[X] = \mathbb{E}[X^2] = V(X) = \frac{110}{3}.\end{aligned}$$

Exercise 3. (**) *Use Chebyshev's inequality to find an upper bound on the probability that the number of tails that come up when a fair coin is tossed n times deviates from the mean by more than $5\sqrt{n}$.*

Chebyshev's inequality states that
$$p(|X(s) - \mathbb{E}(X)| \geq r) \leq \frac{V(X)}{r^2}.$$
In our case X is the number of tails that come up when a coin is tossed n times. The probability of tail is $1/2$ per toss. To compute the expected value and the variance we could simply use the formulas for expected value and variance of Bernoulli trials, but for the sake of demonstration we compute them step by step:

The expected value of X is
$$\begin{aligned}\mathbb{E}(X) &= \sum_{i=0}^{n} i \binom{n}{i} \left(\frac{1}{2}\right)^n = \sum_{i=1}^{n} i \binom{n}{i} \left(\frac{1}{2}\right)^n = \sum_{i=1}^{n} n \binom{n-1}{i-1} \left(\frac{1}{2}\right)^n \\ &= \frac{n}{2^n} \sum_{j=0}^{n-1} \binom{n-1}{j} = n \frac{2^{n-1}}{2^n} = \frac{n}{2}.\end{aligned}$$
Similarly we compute
$$\begin{aligned}\mathbb{E}(X^2) &= \sum_{i=0}^{n} i^2 \binom{n}{i} \left(\frac{1}{2}\right)^n = \sum_{i=1}^{n} i^2 \binom{n}{i} \left(\frac{1}{2}\right)^n = n \sum_{j=0}^{n-1} (j+1) \binom{n-1}{j} \left(\frac{1}{2}\right)^n \\ &= \frac{n}{2^n} \left(\sum_{j=0}^{n-1} j \binom{n-1}{j} + \sum_{j=0}^{n-1} \binom{n-1}{j} \right) = \frac{n}{2^n} \left(\sum_{j=0}^{n-1} (n-1) \binom{n-2}{j-1} + 2^{n-1} \right) \\ &= \frac{n}{2^n} \left((n-1) \sum_{j=1}^{n-1} \binom{n-2}{j-1} + 2^{n-1} \right) = \frac{n}{2^n} \left((n-1) \sum_{i=0}^{n-2} \binom{n-2}{i} + 2^{n-1} \right) \\ &= \frac{n}{2^n} \left((n-1)2^{n-2} + 2^{n-1} \right) = \frac{n}{2^n} (n - 1 + 2)\, 2^{n-2} = \frac{n(n+1)2^{n-2}}{2^n} = \frac{n(n+1)}{4}.\end{aligned}$$
Thus the variance is
$$V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{n(n+1)}{4} - \left(\frac{n}{2}\right)^2 = \frac{n}{4}.$$
Now we can plug this into the formula from above and get
$$p\left(|X(s) - \frac{n}{2}| \geq 5\sqrt{n}\right) \leq \frac{n}{4 \cdot (5\sqrt{n})^2} = \frac{n}{4 \cdot 25 \cdot n} = \frac{1}{100}.$$

Exercise 4. (**) *Let X_n be the random variable that equals the number of tails minus the number of heads when n fair coins are flipped.*

The sample space of possible outcomes for X_n is $S = \{-n, -(n-2), -(n-4), \dots, n-2, n\}$. Define the random variables Y_T, Y_H that count the number of tails and heads, respectively. We know that $\mathbb{E}(Y_T) = \mathbb{E}(Y_H) = \frac{n}{2}$. Furthermore $X_n = Y_T - Y_H$ and $Y_T = n - Y_H$, i.e., Y_H is completely determined by Y_T (and vice versa). Therefore we get

$$p(X_n = -n + 2i) = p\left((Y_T = i) \text{ and } (Y_H = n - i)\right) = \binom{n}{i}\left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{n-i} = \binom{n}{i}\frac{1}{2^n}$$

for $i \in \{0, 1, 2, \dots, n\}$.

1. *What is the expected value of X_n?*

$$\mathbb{E}(X_n) = \mathbb{E}(Y_T) - \mathbb{E}(Y_H) = \frac{n}{2} - \frac{n}{2} = 0.$$

2. *What is the variance of X_n?*

We use the formulas for $\sum_{i=0}^n i\binom{n}{i} = n2^{n-1}$ and $\sum_{i=0}^n i^2\binom{n}{i} = n(n+1)2^{n-2}$, which can be derived using the binomial theorem. We have

$$\mathbb{E}(X_n^2) = \sum_{i=0}^n (-n+2i)^2 p(X_n = -n+2i) = \sum_{i=0}^n (n^2 - 4in + 4i^2)\binom{n}{i}\frac{1}{2^n}$$
$$= \frac{1}{2^n}\left(n^2\sum_{i=0}^n \binom{n}{i} - 4n\sum_{i=0}^n i\binom{n}{i} + 4\sum_{i=0}^n i^2\binom{n}{i}\right)$$
$$= \frac{1}{2^n}\left(n^2 2^n - 4n^2 2^{n-1} + 4n(n+1)2^{n-2}\right) = -n^2 + n(n+1) = n.$$

Thus we get that the variance is

$$V(X_n) = \mathbb{E}(X_n^2) - \mathbb{E}(X_n)^2 = n - 0 = n.$$

Exercise 5. (**) *Consider a Bernoulli trial with success probability $0 \le p \le 1$.*

A run *is a maximal sequence of successes in a sequence of Bernoulli trials. For example, in the sequence S, S, S, F, S, S, F, F, S (where S represents a success and F represents a failure) there are three runs consisting of three successes, two successes, and one success, respectively.*

Let R denote the random variable on the set of sequences of n independent Bernoulli trials that counts the number of runs in this sequence. Find $\mathbb{E}(R)$.

Hint: Show that $R = \sum_{j=1}^n I_j$, where $I_j = 1$ if a run begins at the jth Bernoulli trial and $I_j = 0$ otherwise. Find $\mathbb{E}(I_1)$ and then find $\mathbb{E}(I_j)$, where $1 < j \le n$.

Define $R = \sum_{j=1}^n I_j$, where $I_j = 1$ if a run begins at the jth Bernoulli trial and $I_j = 0$ otherwise. Then R counts how many starting points of a run are in the sequence, i.e., it counts how many runs there are in the sequence. It is easy to see that

$$p(I_1) = p$$

because for I_1 we simply need that the first element of the sequence is a success. For I_j, if $j > 1$, we need that the jth element is a success, but the $(j-1)$th element is a failure, i.e.,

$$p(I_j) = p(1-p).$$

The expected value of I_j is

$$\mathbb{E}(I_1) = p, \quad \mathbb{E}(I_j) = p(1-p).$$

Then we get

$$\mathbb{E}(R) = \mathbb{E}(\sum_{j=1}^{n} I_j) = \sum_{j=1}^{n} \mathbb{E}(I_j) = \mathbb{E}(I_1) + \sum_{j=2}^{n} \mathbb{E}(I_j) = p + (n-1)p(1-p).$$

Exercise 6. (**) *You are given five dice, each in the shape of a different Platonic solid. Every die has numbers written on its faces starting from 1 up to the number of faces. You throw the dice.*

The five Platonic solids are tetrahedron, cube, octahedron, dodecahedron and icosahedron with 4,6,8,12,20 sides respectively. Let's denote with X_i the random variable that represents the value of the i'th die after a throw, respective to the order in the previous sentence.

1. *What is the expected sum of thrown values? What is the expected product of thrown values?*

   Denote the sum of the outcomes as X, and their product as Y, so $X = X_1 + X_2 + X_3 + X_4 + X_5$, and $Y = X_1X_2X_3X_4X_5$. We're asked to compute the expected value of X and Y. Due to the linearity of expected value, we have

   $$\mathbb{E}[X] = \mathbb{E}[X_1 + X_2 + X_3 + X_4 + X_5] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3] + \mathbb{E}[X_4] + \mathbb{E}[X_5].$$

   The X_i's are uncorrelated since the dice throws are independent of one another, so we also have

   $$\mathbb{E}[Y] = \mathbb{E}[X_1X_2X_3X_4X_5] = \mathbb{E}[X_1]\mathbb{E}[X_2]\mathbb{E}[X_3]\mathbb{E}[X_4]\mathbb{E}[X_5].$$

   In order to solve the first part of the exercise we only need $\mathbb{E}[X_i]$ for $i = 1, 2, 3, 4, 5$. This can be computed easily since the dice output a uniform number between 1 and the number of sides of that die. So a die with n sides will give the value $\sum_{i=1}^{n} i\frac{1}{n} = \frac{n(n+1)}{2n} = \frac{n+1}{2}$ on average. This means that

   $\mathbb{E}[X_1] = \frac{4+1}{2} = 2.5,$
   $\mathbb{E}[X_2] = \frac{6+1}{2} = 3.5,$
   $\mathbb{E}[X_3] = \frac{8+1}{2} = 4.5,$
   $\mathbb{E}[X_4] = \frac{12+1}{2} = 6.5,$
   $\mathbb{E}[X_5] = \frac{20+1}{2} = 10.5.$

   Plugging these values in, we obtain $\mathbb{E}[X] = 27.5$ and $\mathbb{E}[Y] = 2687.34375$.

You are now asked to choose four of the dice uniformly at random, throw them and compute their sum.

2. *How would you choose four dice uniformly at random? What is the expected value of this sum?*

   One way of choosing four out of five dice uniformly at random is by choosing one die uniformly at random and removing it. In order to do this we may throw an icosahedron die and consider its value modulo 5. This gives a uniformly chosen number in $\{0, 1, 2, 3, 4\}$, since the number of sides of an icosahedron is a multiple of 5. Then we add 1 and we obtain a uniformly chosen number in the set $\{1, 2, 3, 4, 5\}$. After removing the die with that index we are left with four dice.
   Another way of doing this is by throwing a regular die, if we obtain 1,2,3,4,5 remove the die with that index and throw again if we obtain 6.
   Denote with Z the sum of values of the four thrown dice. A simple way of computing it would be

to note that each die is used with probability $\frac{4}{5}$, so $\mathbb{E}[Z] = \frac{4}{5}\mathbb{E}[X] = 22$. If you're not convinced, simply note that

$$\begin{aligned}Z =& \frac{1}{5}(X_1 + X_2 + X_3 + X_4) + \frac{1}{5}(X_1 + X_2 + X_3 + X_5) + \\ &+ \frac{1}{5}(X_1 + X_2 + X_4 + X_5) + \frac{1}{5}(X_1 + X_3 + X_4 + X_5) + \\ &+ \frac{1}{5}(X_2 + X_3 + X_4 + X_5) = \\ &= \frac{4}{5}(X_1 + X_2 + X_3 + X_4 + X_5) = \frac{4}{5}X.\end{aligned}$$

Exercise 7. (***)

1. *Show how to use a regular fair die to find a random number in the range* $\{1, 2, \ldots, n\}$ *for* $8 \leq n \leq 12$.

   Throw a die twice, call the two outcomes r_1 and r_2, and consider them as values in $\{0, 1, 2, 3, 4, 5\}$ by subtracting 1. The value $r_1 + 6r_2$ is a uniform number in $\{0, 1, 2, \ldots, 35\}$. Now since 12 divides 36, consider your result modulo 12. You obtain again a uniform number mod 12. There is another way to obtain a uniform number mod 12. We throw a die twice, from the first result we only save the parity, and we save the second throw as r. Then we consider r if the first throw is even and $r + 6$ if the first throw is odd - this is a uniform number mod 12.

   Now that we have a way of producing a uniform number mod 12, we simply take that number if it is in $\{1, 2, \ldots, n\}$, and we throw again if it is in $\{n + 1, \ldots, 12\}$. Note that when $n = 12$ we always obtain an outcome after throwing a die twice, while we might have to throw it multiple times for smaller values of n.

2. *Show how to use a fair coin to find a random number in the range* $\{2, 3, \ldots, n\}$ *for* $2 \leq n \leq 12$.

   Flip the coin four times and write down 0 for a head, and 1 for a tail. Concatenate these digits to obtain a binary number, for example head, head, tails, tails becomes 0011. This is a four digit binary number, meaning that it corresponds to a number in the range $\{0, 1, 2, \ldots, 15\}$. The outcomes are uniformly random, so we obtain one of the first 16 numbers at random. In order to find a random number in the range $\{2, 3, \ldots, n\}$ we simply flip a coin four times, and if the obtained binary number is within this range we take it, and otherwise we flip again.

   Notice that, depending on the value of n, we need to flip the coin a different number of times on average. For example if $n = 3$, then on every four coin flips, there is only a $1/8$ probability of falling in the range $\{2, 3\}$, so on average we flip the coin $4 \cdot 8 = 32$ times. This method is not the most effective way to do it for $n = 3$, but it captures the whole range of n's.

Exercise 8. (*) *After summer, the winter tires of a car (with four wheels) are to be put back. However, the owner has forgotten which tire goes to which wheel, and the tires are installed "randomly", each of the* $4! = 24$ *permutations are equally likely.*

1. *What is the probability that tire* 1 *is installed in its original position?*
   The probability that tire 1 is at its original position is given by finding all the permutations for which tire 1 is well placed: Prob(tire1 well placed)$= \frac{3!}{4!} = \frac{6}{24} = \frac{1}{4}$

2. *What is the probability that all the tires are installed in their original position?*
   The probability that all tires are at their right place corresponds to exactly one of the available permutations: $\frac{1}{24}$.

3. *What is the expected number of tires that are installed in their original positions?*
   Let X_i be the random variable that gives 1 if tire i is correctly placed and 0 otherwise. Then the expected number of tires installed at their original position is $\mathbb{E}[X_1 + X_2 + X_3 + X_4] = 4 \cdot \frac{1}{4} = 1$.

4. *Redo the above for a vehicle with n wheels.*
   For a vehicle with n wheels the corresponding values are $\frac{1}{n}$, $\frac{1}{n!}$ and 1.

5. *What is the probability that none of the tires are installed in their original positions?*
   We want to count the probability of derangements in a set of n. Let A_i be the event that tire i is in its original position. We are interested in the complement of $\cup_i A_i$ which has probability $1 - P(\cup_i A_i)$ of happening. The probability $P(\cup_i A_i)$ can be expressed as an alternating sum:

$$P(\cup_i A_i) = \sum_i P(A_i) - \sum_{i_1<i_2} P(A_{i_1} \cap A_{i_2}) + \sum_{i_1<i_2<i_3} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - ...$$

   The j^{th} sum consists of $\binom{n}{j}$ terms each of them having the value $P(A_{i_1} \cap ... \cap A_{i_j}) = \frac{(n-j)!}{n!}$ which corresponds to the probability that tires 1 to j are in their original position. Consequently,

$$P(\cup_i A_i) = \sum_{j=1}^{n} (-1)^{j-1} \binom{n}{j} \frac{(n-j)!}{n!} = \sum_{j=1}^{n} (-1)^{j-1} \frac{1}{j!}.$$

   And the event that no tire is in its original position has probability:
   $1 - P(\cup_i A_i) = \sum_{j=0}^{n} (-1)^j \frac{1}{j!}$. When $n = 4$ this is equal to $\frac{3}{8}$.

Exercise 9. (***) *You throw a die until you get a 6. What is the average number of throws conditioned on the event that all throws (before the first 6) gave even numbers?*

Let X be a random variable defined as

$$X = \# \text{ of throws until the first six}$$

and let Y be the event defined as

$$Y = \text{all throws until the first six are even.}$$

First "*technical*" solution - just use definitions:

The goal is to compute $\mathbb{E}[X|Y] = \sum_{i=1}^{\infty} i \cdot \mathbb{P}[X = i|Y]$. We will do this one step at a time - firstly let's compute $\mathbb{P}[X = i|Y]$, i.e. the probability of getting a 6 for the first time on the i'th throw, conditioned on the event that all previous throws were even. We will use Bayes' formula:

$$\mathbb{P}[X = i|Y] = \frac{\mathbb{P}[Y|X = i]\mathbb{P}[X = i]}{\mathbb{P}[Y]}.$$

Now let's compute the three terms of the product separately.

- $\mathbb{P}[X = i] = \left(\frac{5}{6}\right)^{i-1} \frac{1}{6}$ as the first $i-1$ throws must be different from 6, and the i'th throw is a 6
- $\mathbb{P}[Y|X = i] = \left(\frac{2}{5}\right)^{i-1}$ as the first $i-1$ throws must be even, but we condition on them being different from 6, so they can have 2 out of 5 possible values (2 and 4 out of 1,2,3,4,5)
- $\mathbb{P}[Y] = \sum_{j=1}^{\infty} \mathbb{P}[Y|X = j]\mathbb{P}[X = j]$. Here we use conditioning on the first time we get a 6, ranging from the first throw (j=1) to ∞. Plugging in the corresponding values computed above, we obtain $\mathbb{P}[Y] = \sum_{j=1}^{\infty} \left(\frac{2}{5}\right)^{j-1} \left(\frac{5}{6}\right)^{j-1} \frac{1}{6} = \frac{1}{6} \sum_{j=1}^{\infty} \left(\frac{2}{6}\right)^{j-1} = \frac{1}{6} \sum_{j=0}^{\infty} \left(\frac{1}{3}\right)^{j} = \frac{1}{6} \frac{1}{1-1/3} = \frac{1}{4}$.

We plug everything in to obtain

$$\mathbb{P}[X = i|Y] = \frac{\left(\frac{2}{5}\right)^{i-1} \left(\frac{5}{6}\right)^{i-1} \frac{1}{6}}{\frac{1}{4}} = \frac{2}{3} \left(\frac{1}{3}\right)^{i-1}$$

Finally we can compute the expected value formula from its definition

$$\mathbb{E}[X|Y] = \sum_{i=1}^{\infty} i \cdot \mathbb{P}[X = i|Y] = \sum_{i=1}^{\infty} i \cdot \frac{2}{3}\left(\frac{1}{3}\right)^{i-1} = \frac{2}{3}\sum_{i=0}^{\infty}(i+1) \cdot \left(\frac{1}{3}\right)^{i} = \frac{2}{3} \cdot \frac{9}{4} = \frac{3}{2},$$

where $\sum_{i=0}^{\infty}(i+1) \cdot \left(\frac{1}{3}\right)^i = \sum_{i=0}^{\infty} i \cdot \left(\frac{1}{3}\right)^i + \sum_{i=0}^{\infty}\left(\frac{1}{3}\right)^i = \frac{1/3}{(1-1/3)^2} + \frac{1}{1-1/3} = \frac{3}{4} + \frac{3}{2} = \frac{9}{4}$.

Second "*easy way*" solution:
Another, easier way, to compute $\mathbb{E}[X|Y]$ is the following: consider two disjoint cases - obtaining a 6 on the first throw, and not obtaining a 6 on the first throw, both conditioned on Y. Then if we get a 6 on the first throw - the number of throws is exactly 1, but if we don't get a 6, but we get a 2 or a 4 in stead, we have done 1 throw, and we start the game again - i.e. we add 1 to the number of throws and throw again. We formalize this by conditioning $X|Y$ on the events $X = 1|Y$ and $X \neq 1|Y$ in order to obtain

$$\mathbb{E}[X|Y] = \mathbb{E}[(X|Y)|(X = 1|Y)]\mathbb{P}[X = 1|Y] + \mathbb{E}[(X|Y)|(X \neq 1|Y)]\mathbb{P}[X \neq 1|Y]$$

We have $\mathbb{E}[(X|Y)|(X = 1|Y)] = 1$ since this is the expected number of throws until the first 6, given that the first throw is 6 (conditioned on Y), and $\mathbb{E}[(X|Y)|(X \neq 1|Y)] = (1 + \mathbb{E}[X|Y])$ since this is the expected number of throws until the first 6 given that the first throw is not a 6, conditioned on Y, as the first throw is even, and all the throws after that one are independent from the previous one. Putting this together we get

$$\mathbb{E}[X|Y] = 1 \cdot \mathbb{P}[X = 1|Y] + (1 + \mathbb{E}[X|Y]) \cdot \mathbb{P}[X \neq 1|Y].$$

The value of $\mathbb{P}[X = 1|Y] = \frac{2}{3}$ was computed above, so $\mathbb{P}[X \neq 1|Y] = \frac{1}{3}$ and by plugging it in we get

$$\begin{aligned}
\mathbb{E}[X|Y] &= \frac{2}{3} + \frac{1}{3} \cdot (1 + \mathbb{E}[X|Y]) \\
\mathbb{E}[X|Y] &= \frac{2}{3} + \frac{1}{3} + \frac{1}{3}\mathbb{E}[X|Y]) \\
\frac{2}{3}\mathbb{E}[X|Y] &= \frac{2}{3} + \frac{1}{3} = 1 \\
\mathbb{E}[X|Y] &= \frac{3}{2}
\end{aligned}$$

Third "*easy way*" solution:
Let's compute $\mathbb{P}[Y]$, i.e. the probability of having all even throws before the first 6. We do this by conditioning Y on the exit of the first throw that we call F. We may write

$$\mathbb{P}[Y] = \sum_{j=1}^{6} \mathbb{P}[Y|F = j]\mathbb{P}[F = j].$$

If the first throw is a 6, then $\mathbb{P}[Y|F = 6] = 1$ since the event Y is satisfied. If the first throw is a $1, 3$ or 5 then $\mathbb{P}[Y|F = 1, 3, 5] = 0$ since we had an odd throw before a 6. If the first throw is a 2 or a 4 then $\mathbb{P}[Y|F = 2, 4] = \mathbb{P}[Y]$ since the throws after the first one are independent and whether this sequence of throws has all evens before the first 6 depends only on the successive throws. So either we hit a 6 on the first throw, or we hit a 2 or 4 and face the same probability again. By summing this together we obtain

$$\begin{aligned}
\mathbb{P}[Y] &= \mathbb{P}[Y|F = 6]\mathbb{P}[F = 6] + \mathbb{P}[Y|F = 4]\mathbb{P}[F = 4] + \mathbb{P}[Y|F = 2]\mathbb{P}[F = 2] \\
&= 1 \cdot \frac{1}{6} + \mathbb{P}[Y]\frac{1}{6} + \mathbb{P}[Y]\frac{1}{6}
\end{aligned}$$

and therefore $P[Y] = \frac{1}{4}$. Now, what is the probability that we hit 6 on the first throw conditioned on Y. Without the conditioning it would be $\frac{1}{6}$, but with the conditioning we get

$$\mathbb{P}[F = 6|Y] = \frac{\mathbb{P}[Y|F = 6]\mathbb{P}[F = 6]}{\mathbb{P}[Y]} = \frac{1 \cdot 1/6}{1/4} = \frac{2}{3}.$$

Now the game becomes the following - we throw a die which has $\frac{2}{3}$ probability of giving a 6, and $1-\frac{2}{3}=\frac{1}{3}$ of giving something else (a 2 or 4). What is the number of throws until the first 6? This is a simple geometric distribution, and it's expected value is $1/(2/3)=\frac{3}{2}$.

<u>Fourth "*combinatorial*" solution</u>:
Consider an infinite sequence of die throws $(a_1, a_2, a_3, a_4, \ldots)$ with $a_i \in \{1, 2, 3, 4, 5, 6\}$. Suppose that in all infinite sequences every value from 1 to 6 occurs (the complement has probability zero). Then define a new sequence $(b_1, b_2, b_3, b_4, b_5, b_6)$ in the following way

$$\begin{aligned} b_1 &= a_1 \\ b_i &= \text{ the first } a_j \text{ different from } b_1, \ldots, b_{i-1}. \end{aligned}$$

For example if the sequence of a's is $(4, 2, 5, 4, 5, 1, 2, 4, 2, 3, 5, 1, 2, 4, 5, 6, 2, 3, 5, 2, 3, \ldots)$ then the sequence of b's is $(4, 2, 5, 1, 3, 6)$. Notice that for two values $x, y \in \{1, 2, 3, 4, 5, 6\}$ the first x comes before the first y in the a list if and only if x comes before y in the b list.

Now let $p = \mathbb{P}[6 \text{ comes before } 2,4 \mid 6 \text{ comes before } 1,3,5]$. Notice that $p = \mathbb{P}[F = 6|Y] = \mathbb{P}[X = 1|Y]$ from the previous solutions. In order to compute p we only need to consider the sequences the sequences of the b's, and these are simply the permutations of the $\{1, 2, 3, 4, 5, 6\}$. There are in total 6! permutations, and out of these, in $\frac{1}{4}$ of them 6 comes before 1,3,5. This is because for every permutation of the 6 numbers, when looking only at 1,3,5,6, each of them occurs before the other three in the same number of permutations, so we simply divide by four. So the total number of sequences where 6 comes before $1, 3, 5$ is $\frac{6!}{4}$. Out of these sequences, 6 comes before 2,4 if and only if 6 is in the first position. That means that 6 is fixed, and the other 5 values can be permuted freely in the remaining positions, so there are 5! ways to do that. Finally we obtain that

$$\mathbb{P}[6 \text{ comes before } 2,4 \mid 6 \text{ comes before } 1,3,5] = \frac{5!}{6!/4} = \frac{2}{3}.$$

Since we're looking for $\mathbb{E}[\# \text{ of throws until first } 6 \mid 6 \text{ comes before } 1,3,5]$, and the # of throws until first 6 follows a geometric distribution of probability $p = \frac{2}{3}$, the expected number of throws is $\frac{1}{p} = \frac{3}{2}$.

<u>Fifth "*one liner*" solution</u>:
Consider the following game - we throw a die until we get a 1, 3, 5 or 6. This forms a geometric distribution with $p = \frac{4}{6} = \frac{2}{3}$ so the expected number of throws is $\frac{3}{2}$. The expected number of throws doesn't depend of the terminating value, so we expect to throw the die $\frac{3}{2}$ times until the first 6.

2 Exam Questions

Exercise 10. (*) *With* 2.7 *goals on average per soccer game and ten games per weekend, the probability of at least 45 goals per weekend*

- ○ *can be proved to be at most 50%.*
- ○ *can be proved to be 60%.*
- ✓ *can be proved to be at most 60%.*
- ○ *is not sufficiently specified by the data provided to make any type of prediction.*

Due to the additivity of the expected value, the expected number of goals per weekend is $10 \cdot 2.7 = 27$. Because the number of goals per weekend is non-negative, Markov's inequality can be applied, saying that the probability that the number of goals per weekend is at least a is at most the expected number of goals per weekend divided by a. Thus the number of goals is at least 45 with probability at most $\frac{27}{45} = \frac{3}{5} = 60\%$.

Exercise 11. (**) *A die is rolled three times resulting in an ordered triple (r_1, r_2, r_3) of independent random outcomes $r_1, r_2, r_3 \in \{1, 2, 3, 4, 5, 6\}$, and the random variable X is defined by $X((r_1, r_2, r_3)) = r_1 + 2r_2 + 3r_3 \in \mathbf{Z}$.*

The expected value $E(X)$ of X and the probability $p(X \leq r)$ that X takes at most the value r satisfy

- ○ $E(X) = 21$ *and* $p(X \leq 8) = \frac{1}{72}$.
- ✓ $E(X) = 21$ *and* $p(X \leq 8) = \frac{1}{54}$.
- ○ $E(X) = 24.5$ *and* $p(X \leq 8) = \frac{1}{72}$.
- ○ $E(X) = 24.5$ *and* $p(X \leq 8) = \frac{1}{54}$.

With $\mathbb{E}(r_i) = \sum_{j=1}^{6} \frac{j}{6} = \frac{7}{2}$ for $i = 1, 2, 3$ and $\mathbb{E}(X) = \mathbb{E}(r_1) + 2\mathbb{E}(r_2) + 3\mathbb{E}(r_3)$ it follows that $\mathbb{E}(X) = \frac{7}{2}(1 + 2 + 3) = 21$. The outcome $X \leq 8$ is realized only by $r_1 = 1$, $r_2 = 1$, $r_3 = 1$ $(X = 6)$, $r_1 = 2$, $r_2 = 1$, $r_3 = 1$ $(X = 7)$, $r_1 = 3$, $r_2 = 1$, $r_3 = 1$ $(X = 8)$, and $r_1 = 1$, $r_2 = 2$, $r_3 = 1$ $(X = 8)$, thus for 4 of the $6 \cdot 6 \cdot 6$ possibilities (which each have the same probability of $\frac{1}{6 \cdot 6 \cdot 6}$); it follows that $p(X \leq 8) = \frac{4}{6 \cdot 6 \cdot 6} = \frac{1}{3 \cdot 3 \cdot 6} = \frac{1}{54}$.

Exercise 12. (**) *Let $S = \{1, 2, 3, 4, 5, 6\}$ be the set of outcomes of a fair die throw, and let $x \in S$ be an outcome. For any partition $S = \bigcup_{i=1}^{2} S_{2,i}$ with $|S_{2,i}| = 3$ $(i = 1, 2)$, the random variable X_2 maps x to i if $x \in S_{2,i}$; this results in a random outcome in $\{1, 2\}$. Similarly, for any partition $S = \bigcup_{j=1}^{3} S_{3,j}$ with $|S_{3,j}| = 2$ $(j = 1, 2, 3)$, the random variable X_3 maps x to j if $x \in S_{3,j}$; this results in a random outcome in $\{1, 2, 3\}$. The random variables X_2 and X_3 are independent if*

- ○ $S_{2,1} = \{1, 2, 3\}$, $S_{2,2} = \{4, 5, 6\}$, $S_{3,1} = \{4, 6\}$, $S_{3,2} = \{2, 5\}$, $S_{3,3} = \{1, 3\}$.
- ✓ $S_{2,1} = \{1, 3, 4\}$, $S_{2,2} = \{2, 5, 6\}$, $S_{3,1} = \{1, 5\}$, $S_{3,2} = \{2, 3\}$, $S_{3,3} = \{4, 6\}$.
- ○ $S_{2,1} = \{2, 3, 4\}$, $S_{2,2} = \{1, 5, 6\}$, $S_{3,1} = \{1, 2\}$, $S_{3,2} = \{5, 6\}$, $S_{3,3} = \{3, 4\}$.
- ○ *None of the other answers is correct.*

The random variables X_2 and X_3 are independent if

$$p(X_2 = r_2 \text{ and } X_3 = r_3) = p(X_2 = r_2) \cdot p(X_3 = r_3)$$

for all real numbers r_2 and r_3. The probabilities are zero on both sides if $r_2 \notin \{1, 2\}$ or $r_3 \notin \{1, 2, 3\}$, so we may restrict ourselves to $r_2 \in \{1, 2\}$ and $r_3 \in \{1, 2, 3\}$.

For $r_2 \in \{1, 2\}$ it is the case that $p(X_2 = r_2) = p(x \in S_{2,r_2}) = \frac{|S_{2,r_2}|}{6} = \frac{3}{6} = \frac{1}{2}$ (according to the definition of the random variable X_2) and for $r_3 \in \{1, 2, 3\}$ it is the case that $p(X_3 = r_3) = p(x \in S_{3,r_3}) = \frac{|S_{3,r_3}|}{6} = \frac{2}{6} = \frac{1}{3}$ (according to the definition of the random variable X_3). Therefore, the right-hand side $p(X_2 = r_2) \cdot p(X_3 = r_3)$ of the above equation is always equals to $\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$.

The left-hand side $p(X_2 = r_2 \text{ and } X_3 = r_3)$ is equal to $p(x \in S_{2,r_2} \text{ and } x \in S_{3,r_3})$ which equals $p(x \in S_{2,r_2} \cap S_{3,r_3}) = \frac{|S_{2,r_2} \cap S_{3,r_3}|}{6}$.

With $S_{2,1} = \{1, 2, 3\}$ and $S_{3,3} = \{1, 3\}$ (as in the first answer), we have that $S_{3,3} \subset S_{2,1}$ so that $S_{2,1} \cap S_{3,3} = S_{3,3}$. Thus $p(X_2 = 1 \text{ and } X_3 = 3) = \frac{|S_{2,1} \cap S_{3,3}|}{6} = \frac{|S_{3,3}|}{6} = \frac{2}{6} = \frac{1}{3} \neq \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = p(X_2 = 1) \cdot p(X_3 = 3)$. The first answer is therefore not correct.

Similarly, with $S_{2,1} = \{2, 3, 4\}$ and $S_{3,3} = \{3, 4\}$ (as in the third answer), we again have that $S_{3,3} \subset S_{2,1}$ so that $S_{2,1} \cap S_{3,3} = S_{3,3}$. Thus $p(X_2 = 1 \text{ and } X_3 = 3) = \frac{|S_{2,1} \cap S_{3,3}|}{6} = \frac{|S_{3,3}|}{6} = \frac{2}{6} = \frac{1}{3} \neq \frac{1}{6} = \frac{1}{2} \cdot \frac{1}{3} = p(X_2 = 1) \cdot p(X_3 = 3)$. The third answer is therefore not correct.

For the second answer, however, it is the case that $|S_{2,r_2} \cap S_{3,r_3}| = 1$ for all $r_2 \in \{1, 2\}$ and $r_3 \in \{1, 2, 3\}$. Therefore, for all $r_2 \in \{1, 2\}$ and $r_3 \in \{1, 2, 3\}$ it is the case that $p(X_2 = r_2 \text{ and } X_3 = r_3) = \frac{|S_{2,r_2} \cap S_{3,r_3}|}{6} = \frac{1}{6}$, which equals $\frac{1}{2} \cdot \frac{1}{3} = p(X_2 = r_2) \cdot p(X_3 = r_3)$. The second answer if therefore correct.

$*$ = *easy exercise, everyone should solve it rapidly*
$**$ = *moderately difficult exercise, can be solved with standard approaches*
$***$ = *difficult exercise, requires some idea or intuition or complex reasoning*