# Exercise Sheet 5

May 17th 2023

Submission of the homework assignments until May 25th, 11:30 am online in TUM-moodle in groups of two. Please put the *full names* and *student IDs* of you *and* your partner on all parts of your submission. The solution will be discussed in the classes and published on moodle after some time.

## Homework

### Problem H 19 - Indicator Variables [4 pts.]

Let $A_1, \cdots, A_n \subseteq \Omega$ be events and define by $I_{A_1}, \ldots, I_{A_n}$ the corresponding indicator variables. Prove or disprove the following statement:
The events $A_1, \ldots, A_n$ are independent if and only if the random variables $I_{A_1}, \ldots, I_{A_n}$ are independent.

*Solution:* The statement is true - to prove it both implications "$\Leftarrow$" and "$\Rightarrow$" have to be shown. Let us start with "$A_1, \cdots, A_n \subseteq \Omega$ independent" $\Leftarrow$ "$I_{A_1}, \ldots, I_{A_n}$ independent".
We assume that $I_{A_1}, \ldots, I_{A_n}$ are independent. It was shown in the lecture that for independent random variables $Y_1, \ldots, Y_n$ and sets $S_1, \ldots, S_n \subseteq \mathbb{R}$ the events "$Y_1 \in S_1$", ..., "$Y_n \in S_n$" are independent. Consider now $S_i := \{1\}$ or $\{0\} \subseteq \mathbb{R}$ for $i = 1, \ldots, n$, then the events "$I_{A_1} \in S_1$", ..., "$I_{A_n} \in S_n$" are independent by this theorem. By definition of the indicator variable $I_{A_i}$, the event "$I_{A_i} \in S_i = \{1\}$" is equivalent to the event $A_i$ itself, "$I_{A_i} \in S_i = \{0\}$" is equivalent to the complementary. By consequence and since this is for any combination of $\{0\}$ and $\{1\}$, the events $A_1, \ldots, A_n$ are independent. ✓
We continue with "$A_1, \cdots, A_n \subseteq \Omega$ independent" $\Rightarrow$ "$I_{A_1}, \ldots, I_{A_n}$ independent". Assume that the events $A_1, \ldots, A_n$ are independent. To show the independence of the indicator variables we choose an arbitrary element $(s_1, \ldots, s_n) \in W_{I_{A_1}} \times \ldots \times W_{I_{A_n}} \subset \mathbb{R}^n$ and prove that

$$Pr(I_{A_1} = s_1, \ldots, I_{A_n} = s_n) = Pr(I_{A_1} = s_1) \cdot \ldots \cdot Pr(I_{A_n} = s_n)$$

holds for this element. Since $I_{A_i}$ are indicator variables we have $W_{I_{A_i}} = \{0, 1\}$ for all $i$. So, in particular $s_i \in \{0, 1\}$ for all $i$. By this we can write "$I_{A_i} = s_i$" as the event $A_i^{s_i}$ where we define $A_i^1 := A_i$ and $A_i^0 := A_i^c$. ✓
In addition to that we recall from the lecture:
The events $A_1, \ldots, A_n$ are independent if and only if for all $(s_1, \ldots, s_n) \in \{0, 1\}^n$

$$Pr(A_1^{s_1} \cap \ldots \cap A_n^{s_n}) = Pr(A_1^{s_1}) \cdot \ldots \cdot Pr(A_n^{s_n})$$

where $A_i^1 = A_i$ and $A_i^0 = A_i^c$. ✓
By applying this lemma to the independent sets $A_i$ (at *) we get

$$\begin{aligned}
Pr(I_{A_1} = s_1, \ldots, I_{A_n} = s_n) &= Pr(A_1^{s_1} \cap \ldots \cap A_n^{s_n}) \\
&\stackrel{*}{=} Pr(A_1^{s_1}) \cdot \ldots \cdot Pr(A_n^{s_n}) \\
&= Pr(I_{A_1} = s_1) \cdot \ldots \cdot Pr(I_{A_n} = s_n).
\end{aligned}$$

Since this equation holds for all $(s_1, \dots, s_n) \in W_{I_{A_1}} \times \dots \times W_{I_{A_n}}$ the independence of the indicator variables $I_{A_1}, \dots, I_{A_n}$ follows. ✓

**Problem H 20 - Probability Generating Function** **[3 pts.]**

Calculate the probability mass function, the expected value and the variance of a distribution with the probability generating function

$$G_X(s) = \frac{1}{2} \cdot \frac{3+s}{3-s}.$$

---

*Solution:* It helps to recall the geometric series:

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \text{ if } |r| < 1.$$

Having this in mind the form of the given probability generating function can be transformed the following steps:

$$\begin{aligned}
G_X(s) &= \frac{1}{2}\frac{3+s}{3-s} = \frac{1}{2}\left(\frac{3}{3-s} + s \cdot \frac{1}{3-s}\right) \\
&= \frac{1}{2}\left(\frac{1}{1-\frac{s}{3}} + \frac{s}{3} \cdot \frac{1}{1-\frac{s}{3}}\right) \\
&= \frac{1}{2}\left(\sum_{n=0}^{\infty} \frac{s^n}{3^n} + \frac{s}{3} \cdot \sum_{n=0}^{\infty} \frac{s^n}{3^n}\right) \\
&= \frac{1}{2}\left(\sum_{n=0}^{\infty} \frac{s^n}{3^n} + \sum_{n=0}^{\infty} \frac{s^{n+1}}{3^{n+1}}\right) \\
&= \frac{1}{2}\left(\sum_{n=0}^{\infty} \frac{s^n}{3^n} + \sum_{m=1}^{\infty} \frac{s^m}{3^m}\right) \\
&= \frac{1}{2}\left(1 + 2 \cdot \sum_{m=1}^{\infty} \frac{s^m}{3^m}\right) \\
&= \frac{1}{2} + \sum_{m=1}^{\infty} \frac{1}{3^m} s^m.
\end{aligned}$$

By comparison with

$$G_X(s) = \sum_{m=0}^{\infty} Pr(X=m)s^m = Pr(X=0) + \sum_{m=1}^{\infty} Pr(X=m)s^m$$

we see that the probability mass function is

$$f_X(X=m) = \begin{cases} \frac{1}{2} & \text{for} \quad m = 0 \\ (\frac{1}{3})^m & \text{for} \quad m = 1, 2, \dots \end{cases} . ✓$$

The expected value follows from

$$G'_X(s) = \frac{1}{2} \cdot \frac{(3-s)+(3+s)}{(3-s)^2} = \frac{3}{(3-s)^2}$$

which gives

$$\mathbb{E}(X) = G'_X(1) = \frac{3}{4}. \checkmark$$

The second derivative is

$$G''_X(s) = 3 \cdot \frac{2(3-s)}{(3-s)^4} = \frac{6}{(3-s)^3}$$

by which the variance is obtained:

$$Var(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 = \frac{6}{8} + \frac{3}{4} - (\frac{3}{4})^2 = \frac{15}{16}. \checkmark$$

---

**Problem H 21 - The Clever Chimpanzees** **[5 pts.]**

---

Zeynep is a biologist working in an African national park. Her research is on chimpanzees for which she has to catch and examine in total $m$ different animals. Let us assume that the local population she investigates consists of $n$ animals, and further that after catching and examining a certain chimpanzee Zeynep puts it back to freedom before the next catch. Let us also assume that Zeynep catches each of the $n$ chimpanzees by the same probability and independently from all prior catches, and that she does not stop before having $m$ different animals examined.

**a)** Let $X_i$ be a random variable for the number of times the $i$th chimpanzee gets examined. Calculate the expected value of $X_i$.

**b)** The chimpanzees do not like to be caught so they decide for the following strategy: a smaller group of $k$ animals, $k < n$ and $k \geq m$, is sent close to Zeynep to potentially be caught while all others will hide from her at a safe place. Suppose that each formation to a group of size $k$ is equally likely and that Zeynep still tries to examine $m$ different animals. For which value of $k$ should the chimpanzees decide best such that each of them minimizes its expected number of catches?

---

*Solution:*

**a)** The first part is a modification of the coupon collector's problem. Define by $X = \sum_{i=1}^{n} X_i$ the total number of Zeynep's examinations until $m$ different animals have been examined at least once. Further, we denote by $Y_j$ the number of investigations between the points in time at which exactly $j-1$ and $j$ different chimpanzees have been examined at least once first, respectively.
Similarly to the lecture, we may express $X$ alternatively as the sum of the random variables $Y_j$, i.e. $X = \sum_{j=1}^{m} Y_j$ (in other words, $X$ is split to different phases of the random experiment). The random variables $Y_i$ are geometrically distributed with the parameter $p = (n-j+1)/n$ and expected value $n/(n-j+1)$. According to the linearity of the expected value it follows

$$\mathbb{E}(X) = \sum_{j=1}^{m} \mathbb{E}(Y_i) = \sum_{j=1}^{m} \frac{n}{n-j+1} = n \cdot (H_n - H_{n-m}),$$

where $H_n = \sum_{k=1}^{n} \frac{1}{k}$ is the $n$-th harmonic number. ✓
Since Zeynep catches each of the chimpanzees by the same probability, we additionally know that all $X_i$ are identically distributed. So, by the linearity of the expected value of $X$ we conclude

$$\mathbb{E}(X_i) = \frac{1}{n} \cdot \mathbb{E}(\sum_{k=1}^{n} X_k) = \frac{1}{n} \cdot \mathbb{E}(X) = H_n - H_{n-m}. \checkmark$$

**b)** Define the random variable $X_{k,i}$ as the number of times the $i$th chimpanzee gets caught and examined if the size of the group is $k$.
Note here that in case of $k < m$ that is excluded in the instruction, Zeynep will not stop with her research and the chimpanzees have to hide until infinite time. Additionally, each of the chimpanzees that does not hide will get caught infinitely many times.
So, obviously we assume $k \geq m$. Denote by $E_{k,i}$ the event that the $i$th chimpanzee is in the group that is sent to Zeynep. In total, there is $\binom{n}{k}$ ways to form a group of size $k$ and in $\binom{n-1}{k-1}$ of these ways the $i$th chimpanzee is part of this group. As all of these groupings are equally likely, it holds.

$$Pr(E_{k,i}) = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{(n-1)!}{(n-k)! \cdot (k-1)!} \cdot \frac{(n-k)! \cdot k!}{n!} = \frac{k}{n}. \checkmark$$

If the $i$th chimpanzee is not part of this group it also does not get examined. Otherwise, we recall from the first part that it gets $\sum_{j=1}^{m} \frac{1}{k-j+1}$ times caught and examined on expectancy. So, for the expected value of $X_{k,i}$ of $X_{k,i}$ it holds by the partitioning theorem that

$$\mathbb{E}(X_{k,i}) = \mathbb{E}(X_{k,i}|E_{k,i}) \cdot Pr(E_{k,i}) + \mathbb{E}(X_{k,i}|E_{k,i}^c)Pr(E_{k,i}^c) = (\sum_{j=1}^{m} \frac{1}{k-j+1}) \cdot \frac{k}{n} + 0. \checkmark$$

Now define the function $f(k) = 1/n \cdot \sum_{j=1}^{m}(-\frac{k}{k-j+1})$. As we aim at finding the *optimal* parameter $k$, we calculate the derivative with respect to $k$. By the quotient rule we obtain

$$f'(k) = \frac{1}{n} \cdot \sum_{j=1}^{m} \frac{1 \cdot (k-j+1) - k \cdot 1}{(k-j+1)^2} = \frac{1}{n} \cdot \sum_{j=1}^{m}(-\frac{j-1}{(k-j+1)^2}).$$

We notice that none of the summands in the above expression is positive. Furthermore, at least one of the summands is negative because $m > 1$. Hence, $f$ is a strictly monotonously decreasing function. For $k \geq m$ the function is also continuous. Hence, the expected value of $X_{k,i}$ has to decrease as the size of the group $k$ increases. In conclusion, we see that the strategy of the chimpanzees is not working. If they aim at minimizing the personal number of examinations to expect they should not hide. ✓

**Problem H 22 - A date by chance** **[5 pts.]**

**a)** A man and a woman decide to meet for a date at a certain restaurant. However, it is rush hour. If each person independently arrives at a time uniformly distributed between 7 pm and 8 pm, find the probability that the first to arrive has to wait longer than ten minutes.
*Hint: it may help to focus on the two arrival times. Additionally, you may extend your previous knowledge on joint probability distributions to the continuous case without any further proofs.*

**b)** He was not the love of her life. At 8:15 pm the woman orders a taxicab. The taxi company tells that the cab should arrive in $\mu$ minutes. Let $X$ be a geometrically distributed random variable for the woman's waiting time in minutes. After $i$ minutes she becomes troubled and calls the company to ask for her remaining waiting time $X - i$. Show that the expected waiting time still is $\mu$ minutes, i.e. that $\mathbb{E}(X - i | X > i) = \mu$ holds.

---

*Solution:*

**a)** Denote by $X$ and $Y$ the time in minutes past 7 pm the man and the woman arrive, respectively. Then, these random variables are independent and uniformly distributed over the interval $(0, 60)$. The probability of interest is $Pr(Y > X + 10) + Pr(X > Y + 10)$ which by symmetry equals $2Pr(Y > X + 10)$. We obtain it by the following steps where in this continuous case the two-dimensional integral of the joint probability is calculated in analogy to a double sum over the range of interest.

$$\begin{aligned}
2 \cdot Pr(Y > X + 10) &= 2 \cdot \iint_{y > x+10} f(x, y) dx dy \\
&= 2 \cdot \iint_{y > x+10} f_X(x) f_Y(y) dx dy \\
&= 2 \cdot \int_{10}^{60} \int_{0}^{y-10} \left(\frac{1}{60}\right)^2 dx dy \\
&= \frac{2}{(60)^2} \cdot \int_{10}^{60} (y - 10) dy \\
&= \frac{25}{36}. \checkmark\checkmark
\end{aligned}$$

(Note that the joint probability of the two independent variables is just the product.)

**b)** Since $X$ is a geometrically distributed (in this case *discrete*!) random variable, $X$ is memoryless. From the lecture we already know that $Pr(X > y + x | X > x) = Pr(X > y)$ holds for all $x, y \in \mathbb{N}$. We now aim at showing that the *expected value* of $X$ is memoryless which in the specific setup of the task means $\mathbb{E}(X - i | X > i) = \mu = \mathbb{E}(X)$.

We start with the definitions of the conditional expected value as well as the conditional probability and rearrange in the following way:

$$\mathbb{E}(X - i|X > i) = \sum_{j=1}^{\infty}(j - i) \cdot Pr(X = j|X > i) = \sum_{j=1}^{\infty}(j - i) \cdot \frac{Pr(X = j, X > i)}{Pr(X > i)}. \checkmark$$

As $Pr(X = j, X > i) = 0$ for all $j \leq i$ we may neglect the first $i$ summands in the last term of the equation above. Further, $Pr(X = j, X > i) = Pr(X = j)$ for all $j > i$. Thus, we obtain

$$\mathbb{E}(X - i|X > i) = \sum_{j=i+1}^{\infty}(j - i) \cdot \frac{Pr(X = j)}{Pr(X > i)}.$$

Instead of starting the sum at $j = i + 1$, we alternatively may start at $j = 1$ but have to adjust the expressions accordingly such that the offset between $j$ and $i$ is the same. This yields

$$\mathbb{E}(X - i|X > i) = \sum_{j=1}^{\infty}((j + i) - i) \cdot \frac{Pr(X = j + i)}{Pr(X > i)} = \sum_{j=1}^{\infty} j \cdot \frac{Pr(X = j + i)}{Pr(X > i)}. \checkmark$$

Since $X$ is geometrically distributed, the probability of $Pr(X = j + 1)$ is given as $p \cdot (1 - p)^{j+i-1}$ where $p$ is the parameter of the geometric distribution. Additionally, we know from the lecture that $Pr(X > i) = (1 - p)^i$. By inserting the definitions we can conclude the proof:

$$\mathbb{E}(X - i|X > i) = \sum_{j=1}^{\infty} j \cdot \frac{p \cdot (1 - p)^{i+j-1}}{(1 - p)^i} = \sum_{j=1}^{\infty} j \cdot p \cdot (1 - p)^{j-1} = \sum_{j=1}^{\infty} j \cdot Pr(X = j) = \mathbb{E}(X). \checkmark$$

**Problem H 23 - Chebyshev, Markov, Chernoff? [4 pts.]**

Let $X_1, X_2, \ldots, X_n$ be independent random variables from the same distribution with the properties

$$\mathbb{E}(X) = 0 \quad \text{and} \quad \mathbb{E}(e^{tX}) \leq e^{\frac{t^2}{2}} \quad \text{for all } t \in \mathbb{R}.$$

Show that

$$Pr(X_1 + \cdots + X_n \geq s) \leq e^{-\frac{s^2}{2n}}$$

holds for all $n \in \mathbb{N}$ and all $s > 0$.

*Solution:* To prove this statement it helps to consider a more general version of the Markov inequality though not presented in the lecture:
For a random variable $X$, a constant $s > 0$ and a nondecreasing nonnegative function $\phi : W_X \to \mathbb{R}^+$ with $\phi(s) > 0$ it holds

$$Pr(X \geq s) \leq \frac{\mathbb{E}(\phi(X))}{\phi(s)}.$$

Choosing $\phi(x) = \exp(tx)$ with $t \geq 0$ this yields:

$$Pr(\sum_{i=1}^{n} X_i \geq s) \leq e^{-ts}\mathbb{E}(e^{t\sum_{i=1}^{n} X_i}). \checkmark$$

From the functional equation of the exponential function and by independence of the random variables $X_i$ we can express the right hand side as

$$e^{-ts}\mathbb{E}(e^{t\sum_{i=1}^{n} X_i}) = e^{-ts}\prod_{i=1}^{n}\mathbb{E}(e^{tX_i}). \checkmark$$

By our assumption, we have

$$e^{-ts}\prod_{i=1}^{n}\mathbb{E}(e^{tX_i}) \leq e^{-ts}\prod_{i=1}^{n} e^{\frac{t^2}{2}}. \checkmark$$

If we now choose $t = s/n$ it follows the desired inequality

$$Pr(\sum_{i=1}^{n} X_i \geq s) \leq e^{-\frac{s}{n}s}(e^{-\frac{s^2}{2n^2}})^n = e^{-\frac{s^2}{2n^2}}. \checkmark$$

(Note that if the choice of $t$ is not directly apparent one alternatively may determine the optimal $t$. This will lead to the same result.)