Exercise Sheet 6

May 24th 2023

Submission of the homework assignments until June 1st, 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 24 - An ab-normal baby and longlife batteries **[4 pts.]**

a) The biologist Mario tries to analyze the genetics of a local African chimpanzee population. Specifically, he suspects that the brawny male animal "Kong" is the biological father of the baby of the female "Pauline" because the baby has a white spot on its back like its potential father. It is known that the pregnancy in days is approximately normally distributed with parameters $\mu = 235$ and $\sigma^2 = 100$. Unfortunately, Kong was kept apart from the population for a period that began 255 days before the birth of the baby and ended 205 days before the birth of the baby, so exactly around the most probable time of impregnation. If the supposed father was, in fact, the father of the baby, what is the probability that the mother Pauline could have had the very long or very short pregnancy that would have been necessary?

b) Suppose that the number of kilometers an ebike can be used before its battery wears out is exponentially distributed with an average of 1,000 kilometers. If a biker desires to take a 500-km trip, what is the probability that he or she will be able to complete the trip without having to replace the battery? What can be said when the distribution is not exponential?

Solution:

a) Let X denote the length of the pregnancy in days and assume that Kong indeed is the father. Then the probability that the birth could have occurred within the period Kong was in the population is

$$\begin{aligned}
Pr(X > 255 \text{ or } X < 205) &= Pr(X > 255) + Pr(X < 205) \\
&\stackrel{*}{=} Pr(\frac{X-235}{10} > 2) + Pr(\frac{X-235}{10} < -3) \checkmark \\
&= 1 - Pr(\frac{X-235}{10} \leq 2) + Pr(\frac{X-235}{10} > 3) \\
&= 1 - Pr(\frac{X-235}{10} \leq 2) + 1 - Pr(\frac{X-235}{10} \leq 3) \\
&= 1 - \Phi(2) + 1 - \Phi(3) \\
&\approx .0241. \checkmark
\end{aligned}$$

At * X was *standardized* by subtracting the mean and dividing by the standard deviation $\sigma = \sqrt{Var(X)}$ (also known as z-score) to then use tabulated values of the cumulative normal distribution $\Phi(y, 0, 1)$.

b) The total lifetime X (in terms of the driven kilometers) can be modeled by a exponential distribution with parameter $\lambda = \frac{1}{1000}$. Since the exponential distribution is memoryless the probability of having a remaining lifetime Y sufficient for the trip is

$$Pr(Y > 500) = 1 - F_X(500) = e^{-500 \cdot \lambda} = e^{-\frac{1}{2}} \approx .604, \checkmark$$

independently from the kilometers taken before the trip L. If, in contrast, we assume a general distribution that is not exponential (and not memoryless), then the relevant probability is

$$Pr(X > L + 500 | X > L) = \frac{1 - F_X(L + 500)}{1 - F_X(L)}. \checkmark$$

Problem H 25 - Bounds on probability [5 pts.]

Consider the probability generating function of a random variable X defined by

$$G_X(s) = e^{6(s^2-1)}$$

Using this, the expected value and the variance of X were determined to be $\mathbb{E}(X) = 12$ and $Var(X) = 24$, respectively.

a) Show that the probability of X taking a value of 24 or higher is at most equal to $\frac{1}{6}$ by means of Chebyshev's inequality.

b) Denote by M_X the moment generating function of X. Prove the inequality

$$Pr(X \geq t) \leq \frac{M_X(s)}{e^{s \cdot t}}$$

for any $t \geq 0$ and $s \geq 0$. Use this statement to improve the estimate on $Pr(X \geq 24)$ from the first part.

Hint: You should find that the probability is bounded by $\left(\frac{e}{4}\right)^6$.

Solution:

a) The probability for X taking at least the value 24 can be estimated by

$$Pr(X \geq 24) = Pr(X - \mathbb{E}(X) \geq 12) \leq Pr(|X - \mathbb{E}(X)| \leq 12). \checkmark$$

The inequality above is valid because $X - \mathbb{E}(X) \geq 12$ implies the event $|X - \mathbb{E}(X)| \geq 12$. We use Chebyshev's inequality to further estimate

$$Pr(|X - \mathbb{E}(X)| \geq 12) \leq \frac{Var(X)}{12^2} = \frac{24}{144} = \frac{1}{6},$$

so in total $Pr(X \geq 24) \leq \frac{1}{6}$. $\checkmark$

b) For $s \geq 0$ the function $f(x) = e^{x \cdot s}$ is monotonously increasing. Thus, if the random variable X is greater than or equal to t then also $e^{s \cdot X}$ is greater than or equal to $e^{s \cdot t}$. Hence, we have $Pr(X \geq t) \leq Pr(e^{s \cdot X} \geq e^{s \cdot t})$. Furthermore, $e^{s \cdot X}$ is a random variable taking non-negative values only. By Markov's inequality we therefore obtain

$$Pr(X \geq t) \leq Pr(e^{s \cdot X} \geq e^{s \cdot t}) \leq \frac{\mathbb{E}(e^{s \cdot X})}{e^{s \cdot t}} = \frac{M_X(s)}{e^{s \cdot t}}. \checkmark$$

Note that the last equations follows from the definition of the moment generating function M_X. In the specific example this is given by

$$M_X(s) = G_X(e^s) = \exp(6 \cdot (e^{2s} - 1)).$$

For $s \geq 0$ the inequality from above yields

$$Pr(X \geq 24) \leq \frac{M_X(s)}{e^{24s}} = \exp(6 \cdot (e^{2s} - 1) - 24s). \checkmark$$

In order to get an estimate as tough as possible s has to be chosen such that the expression $\exp(6 \cdot (e^{2 \cdot s} - 1) - 24s)$ gets as small as possible. Since the exponential function increases monotonously, it is sufficient to achieve that by minimizing the exponent. For this, consider the function $f(s) = 6 \cdot (e^{2s} - 1) - 24s$, its first derivative $f'(s) = 2 \cdot 6 \cdot e^{2s} - 24$ and its second derivative $f''(s) = 4 \cdot 6 \cdot e^{2s}$. At $s = \ln(\frac{24}{2 \cdot 6})/2 = \ln(2)/2$ we have $f'(s) = 0$ and $f''(s) > 0$, so it is the desired extreme value. In addition, $s = \ln(2)/2 > 0$ is a valid parameter for our estimate. Note that for this choice of s it holds that $e^{2s} = \exp(2 \cdot \ln(2)/2) = 2$ as well as $24s = 12\ln(2)$. In conclusion we obtain the estimate that is asked for

$$Pr(X \geq 48) \leq \exp(6 \cdot (2 - 1) - 12\ln(2)) = \exp(6 - 12\ln(2)) = \frac{e^6}{2^{12}} = \left(\frac{e}{4}\right)^6. \checkmark$$

Problem H 26 - Basics of measure theory **[5 pts.]**

Consider the experiment that a fair coin is tossed for countably infinite times. Define by Ω the set of all infinite sequences of heads and tails resulting from the experiment.

a) Provide a proof for Ω being *incountable*.

b) Let $\mathcal{A}$ be the set consisting of all subsets $A \subseteq \Omega$ such that either A or $\Omega \setminus A$ is countable. Prove that $\mathcal{A}$ is a σ-algebra.

c) Define a probability measure Pr that is suitable for the events from $\mathcal{A}$. In particular demonstrate that your definition satisfies the *Kolmogorov axioms*.

Solution:

a) In order to prove that Ω is incountable we have to find an injective map from a incountable set to Ω. In this case it is convenient to use the Interval $[0, 1)$ that contains incountably many real numbers. From the previous math courses it is known that any real number r can be expressed as a (infinite) decimal fraction. For the interval $[0, 1)$ such an expansion by a decimal fraction is equivalent to an infinite sequence of the integers between 0 and 9 specifying the decimal digits of r. If we now encode each number by a quadruple of the events heads and tails, i.e. a four-digit binary number such as "THTH"$= 0 \cdot 8 + 1 \cdot 4 + 0 \cdot 2 + 1 \cdot 1 = 5$, we obtain an infinite sequence of heads and tails that we can assign to r. So, we have found an injective map from $[0, 1)$ to Ω and hence proven that Ω is incountable. ✓

b) Since the set Ω is incountable there is no consistent way to assign a meaningful probability to every subset of Ω. Instead, we restrict our focus to events $A \subseteq \Omega$ that either are itself countable or that have a countable complement as it is demanded in the task. Let $\mathcal{A}$ be the set containing all A of that kind. We now will prove that $\mathcal{A}$ is a σ-algebra.
By definition, a σ-algebra $\mathcal{A}$ on the set Ω has to satisfy the following three properties:

(1) $\Omega \in \mathcal{A}$.

(2) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$.

(3) $A_i \in \mathcal{A}$ for all $i \in I \subseteq \mathbb{N} \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{A}$.

For (1) we notice that the complement of Ω is the empty set which is countable. Hence, Ω^c is countable and so $\Omega \in \mathcal{A}$.
For (2) consider an arbitrary set $A \in \mathcal{A}$. Now there is two possibilities: either A is countable. In this case $A^c = \Omega \setminus A$ is in $\mathcal{A}$ as well because its complement $\Omega \setminus (\Omega \setminus A) = A$ is countable. Or A is incountable, then $\Omega \setminus A$ has to be countable by the definition of $\mathcal{A}$. This yields that $\Omega \setminus A$ is in $\mathcal{A}$ also in this case. (With above ✓)
For (3) let $\bigcup_{i \in I} A_i$ be an arbitrary but countable union of sets $A_i \subseteq \mathcal{A}$. In the first case it could be that there is a $j \in I$ such that A_j is incountable. By definition of $\mathcal{A}$ we know that then $\Omega \setminus A_j$ is countable, and further know that $\Omega \setminus (\bigcup_{i \in I} A_i) \subseteq \Omega \setminus A_j$. Thus, the complement of the union has to be countable. In the second case assume that each of the sets A_i is countable, then also their union is countable. For instance, this becomes apparent from Cantor's first diagonal argument.
As each set A_i is countable we can assign each of its elements by an index j. By this convention, $a_{i,j}$ is the jth element in the set A_i. By Cantor's first diagonal

argument we are able to enumerate all elements of the union $\bigcup_{i\in I} A_i$ by the scheme illustrated below.

$$\begin{array}{ccccccc}
a_{1,1} & \to & a_{1,2} & & a_{1,3} & \to & a_{1,4} \\
 & \swarrow & & \nearrow & & \swarrow & \\
a_{2,1} & & a_{2,2} & & a_{2,3} & & a_{2,4} \\
\downarrow & \nearrow & & \swarrow & & \nearrow & \\
a_{3,1} & & a_{3,2} & & a_{3,3} & & a_{3,4} \\
 & \swarrow & & \nearrow & & & \\
a_{4,1} & & a_{4,2} & & a_{4,3} & & a_{4,4} \\
\downarrow & \nearrow & & & & & \\
\dots & & & & & &
\end{array}$$

Here we eventually may skip all elements $a_{i,j}$ for which A_i does not exist or for which there is no element in A_i with index j. ✓

c) From the instruction we know that the coin is fair. Hence, all sequences should have the same probability. However, a single sequence cannot have a positive probability because otherwise the union of arbitrary many sequences, still an admissible event, would tend to have infinite probability. Consequentially, the probability of a single sequence has to be 0.
By the same argument we also obtain that every countable set of sequences has to be assigned by probability 0. In reverse, the complement of a countable set has to have the probability 1. Thus, we define Pr by

$$Pr(A) = \begin{cases} 0 & \text{if } A \text{ countable} \\ 1 & \text{if } A \text{ incountable} \end{cases}. \checkmark$$

We are left to demonstrate that Pr satisfies the Kolmogorov axioms. The first is that $Pr(\Omega) = 1$. As Ω is an incountable set this condition trivially is met.
The second Kolmogorov axiom states that for each countable index set I and pairwise disjoint events $A_i \in \mathcal{A}, i \in I$ the identity $Pr(\bigcup_{i\in I} A_i) = \sum_{i\in I} Pr(A_i)$ has to be fulfilled. Similarly as in (b) we distinguish two cases. Either there exists an incountable A_j with $j \in I$. Then also $\bigcup_{i\in I} A_i$ is incountable and we have $Pr(\bigcup_{i\in I} A_i) = 1$. By definition of $\mathcal{A}$ we further know that $\Omega \setminus A_j$ is countable. Since the sets A_i are pairwise disjoint it has to hold that $A_i \subseteq \Omega \setminus A_j$ for all $i \neq j$. Hence we know that all remaining sets A_i are countable and also find

$$\sum_{i\in I} Pr(A_i) = Pr(A_j) + \sum_{i\in I\setminus\{j\}} Pr(A_i) = 1 + \sum_{i\in I\setminus\{j\}} 0 = 1 = Pr(\bigcup_{i\in I} A_i).$$

Or, alternatively, all A_i may be countable, then part (b) tells us that $\bigcup_{i\in I} A_i$ is countable as well. From this we conclude that

$$\sum_{i\in I} Pr(A_i) = \sum_{i\in I} 0 = 0 = Pr(\bigcup_{i\in I} A_i). \checkmark$$

Problem H 27 - Simulating random variables **[8 pts.]**

A software company wants to optimize the customer support in its call center. It was determined by an expert that there is on average 1.5 customers calling per minute. Based on this, the number of calls over a certain period of time shall be simulated.

a) Assume that the number of calls per minute X can be approximated by the Poisson distribution. Write down the probability density function $f_X(x)$ to this specific problem.

b) The simulation will be based on a random number generator that draws numbers $r \in \mathbb{R}$ from the uniform distribution $U([0,1])$. Explain, how, in theory this can be used to generate samples following $f_X(x)$ for an arbitrary X.

c) Define an approximate to the cumulative probability density F_X for the calls per minute $X \in \mathbb{N}_0$. You may subsume all events with a probability less than 10^{-2} to a single one. Use your result to define a map from the uniformly distributed random numbers r to samples of X.

d) Write a piece of code that simulates the incoming calls from minute 1 to minute 20 in the call center based on the previous. Additionally assume that the call center needs (exactly) 2 minutes per customer to solve his or her issue. Use this information to calculate the number of customers waiting in the queue in minutes 1 to 20. Do you observe high numbers of waiting customers? Should there be additional staff members hired to improve the rates of costumers progressed per minute?

For the programming part please include the code and its output for minutes 1 to 20 to your submission. Explain or comment your code. You may prefer to use Python as programming language.

Solution:

a) $f_X(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$ with $\lambda = 1.5$. ✓

b) We need to find an inverse cumulative distribution function to F_X that can be defined as

$$F_X^{-1}(p) := \inf\{x \in \mathbb{R} | F(x) \geq p\}$$

for probabilities $0 \leq p \leq 1$. For continuous X with a strictly increasing F_X we may just take the inverse map, for discrete X the more general definition above applies. In any case, we can redefine $\tilde{X} := F_X^{-1}(U) : r \in U([0,1]) \mapsto x \in W_X$ to obtain samples x of X with the distribution f_X from uniformly distributed random numbers r. ✓

c) Numerically calculating $f_X(X = k)$ up to $k = 6$ and adding up yields

x	$f_X(x)$	$F_X(x)$
0	0.223	0.223
1	0.335	0.558
2	0.251	0.809
3	0.126	0.934
4	0.047	0.981
5	0.014	0.996
≥ 6	$< 10^{-2}$	≈ 1

. ✓

By this we define the approximate to F_X^{-1} by

$$\tilde{F}_X^{-1}(r) \begin{cases} 0 & \text{for} \quad 0 \le r < 0.223 \\ 1 & \text{for} \quad 0.223 \le r < 0.335 \\ 2 & \text{for} \quad 0.335 \le r < 0.809 \\ 3 & \text{for} \quad 0.809 \le r < 0.934 \\ 4 & \text{for} \quad 0.934 \le r < 0.981 \\ 5 & \text{for} \quad 0.981 \le r < 0.996 \\ 6 & \text{for} \quad 0.996 \le r \le 1 \end{cases} . \checkmark$$

d) This is a sample implementation in Python. In line 25, a random number is generated from $U([0,1])$ to get a sample x by the predefined approximate inverse (line 26). In line 37 every minute half a waiting customer is subtracted because each customer requires 2 minutes. The following lines ensure that there is no negative values for the waiting customers. ✓✓

```
1  #!/usr/bin/python3
2
3  import numpy as np
4  from math import factorial
5
6  # PDF of Poisson distribution
7  def Poi(l,k):
8    return (np.exp(-l)*(l**k)) / factorial(k)
9
10 # Ranges of probability
11 ranges = [np.sum([Poi(1.5,j) for j in range(i)]) for i in np.arange(1,7,1)]
12
13 # Inverse to CDF
14 def InvF(r):
15   x = 0
16   for range in ranges:
17     if r >= range:
18       x += 1
19   return x
20
```

```
21 #Iterate from minute 1 to 20
22 #First minute
23 m = 0
24 in_line = 0
25 r = np.random.rand()
26 in_calls = InvF(r)  # Incoming calls
27 in_line += in_calls # Calls to be handled
28 print("Min \t r \t In \t Inl \t Inl2")
29 print("%i \t %.3f \t %i \t %.1f \t %i" %(m,r,in_calls,in_line,np.round(in_line)))
30 #Following minutes
31 for m in np.arange(2,21,1):
32   r = np.random.rand()
33   in_calls = InvF(r)
34   in_line += in_calls
35   #Progress half a costumer per minute (no negative values for customers in line)
36   if in_line > 0:
37     in_line -= 0.5
38     if in_line < 0:
39       in_line = 0.
40   print("%i \t %.3f \t %i \t %.1f \t %i" %(m,r,in_calls,in_line,np.round(in_line)))
```

Running this piece of code produces the following output, listing the minutes, the drawn random number, the incoming calls, and the number of costumers in line that to be processed (as decimal number and rounded to an integer). As the output depends on the random numbers the numbers will not be reproduced at a second run in general. ✓

Min	r	In	Inl	Inl2
0	0.733	2	2.0	2
2	0.478	1	2.5	2
3	0.174	0	2.0	2
4	0.555	1	2.5	2
5	0.200	0	2.0	2
6	0.174	0	1.5	2
7	0.314	1	2.0	2
8	0.729	2	3.5	4
9	0.921	3	6.0	6
10	0.550	1	6.5	6
11	0.640	2	8.0	8
12	0.450	1	8.5	8
13	0.641	2	10.0	10
14	0.463	1	10.5	10
15	0.418	1	11.0	11
16	0.941	4	14.5	14
17	0.412	1	15.0	15
18	0.388	1	15.5	16
19	0.072	0	15.0	15
20	0.307	1	15.5	16

It can be seen here, though it is only a single run, that there is a large number waiting costumers after some time. To solve this problem the company should decide for hiring additional employees to increase the rate of customers processed per minute. For instance, changing line 37 to `-= 2.0` will lead to a much more efficient call center. ✓