Exercise Sheet 4

May 10th 2023

Due to the holiday submission of the homework assignments until May 19th, 09:00 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 one week after.

Homework

Problem H 14 - Poisson in space [4 pts.]

Every Sunday, Siméon eats his grandmother's famous chocolate cookies for tea. From his lifelong experience he knows that a cookie contains seven chocolate chips on average.

a) Estimate the probability of a cookie containing eight chips by means of the Poisson distribution.

b) Suppose Siméon ate one of grandmother's cookies each Sunday for ten weeks. Assume that the cookies are independent, and estimate the probabilities of the following events:

(i) at least one of the ten cookies had no chocolate chips at all;

(ii) all cookies had at least one chocolate chip;

(iii) exactly one cookie had five or more chocolate chips.

Solution:

a) The parameter λ of a Poisson-distributed random variable can be interpreted as the average number of occurrences of events per unit of time, of space - or per unit of cookies. So, in the example we have $\lambda = 7$, i.e. 7 pieces of chocolate per cookie. If we model the number of chocolate chips in a cookie by the random variable X the distribution is expressed by

$$f_X(k) = \frac{e^{-7} \cdot 7^k}{k!}.$$

By this, we estimate the probability that Siméon finds 8 chips in a cookie by

$$Pr(X = 8) = \frac{e^{-7} \cdot 7^8}{8!} \approx 13\%. \checkmark$$

b) (i) The event is the contrary to having at least one chip in all ten cookies. Having at least one chip in a single cookie again is contrary to the cookie having no chip, and the cookies are assumed to be independent. So,

$$Pr(\exists i \in [1, 10] : X_i = 0) = 1 - Pr(X_i \geq 1 \forall i \in [1, 10])$$
$$= 1 - \Pi_{i=1}^{10}(1 - Pr(X_i = 0)) = 1 - (1 - \frac{e^{-7} \cdot 7^0}{0!})^{10}$$
$$= 1 - (1 - e^{-7})^{10} \approx 0.91\%. \checkmark$$

(ii) Again, one can consider the complementary event: that all cookies do not contain a chocolate chip in this case.

$$Pr(X_i \geq 1 \forall i \in [1, 10]) = (1 - e^{-7})^{10} \approx 99.1\%. \checkmark$$

(iii) For a single cookie the complementary event is to have less than 5 chips. Denote the probability of this by p then

$$p = \sum_{k=0}^{4} Pr(X = k) = \sum_{k=0}^{4} \frac{e^{-7} \cdot 7^k}{k!}.$$

The event of interest then means to choose one of the 10 cookies that has more than 5 chips while all other have less (binomial distribution):

$$Pr(!\,\exists i \in [1, 10] : X_i \geq 5) = \binom{10}{1}(1 - p) \cdot p^9 \approx 1.15 \cdot 10^{-6}. \checkmark$$

Problem H 15 - The wheel of fortune **[6 pts.]**

Daniel and Nicholas play a game where they repeatedly spin a wheel of fortunate that is separated into $n \in \mathbb{N}$ equally sized segments. Each of them spins the wheel once per round and the segments are labeled by the numbers $1, 2, 3, \ldots, n-1, n$.

a) Let us first assume that the wheel has $n = 8$ segments and that the first part of the game is the following: if Daniel gets the wheel to stop at all even numbers at least once first he will receive 10 euros. If, however, Nicholas gets the wheel to stop at all odd numbers at least once first, he will receive this amount (if they need equal numbers of spins both will receive no money). Determine the expected number of spins that Daniel or Nicholas, respectively, needs to have the wheel stopped at all even or odd numbers at least once.

b) For the second part of the game we assume that n is arbitrary. Denote Daniel's and Nicholas's results on the wheel by the random variables X and Y. If $X \geq Y$ Daniel gets $X - Y$ euros of Nicholas's money. If, however, $X < Y$, Daniel has to pay $Y - X$ euros to Nicholas. Let $Z = X - Y$ be Daniel's winning or loss. Use the convolution formula to determine the probability mass function of Z.

Solution:

a) From symmetry reasons it is sufficient to consider Daniel and the even numbers only. In analogy to the coupon collector's problem we may split the experiment to four different phases. In phase 1, stopping at one of the four even numbers is a success which may happen with the probability $p_1 = \frac{4}{8}$. We are interested in the number of spins until the first even number E_1 which is modeled by the geometric distribution, i.e. $X_1 \sim Geo(\frac{4}{8})$. ✓ In phase 2, after the first even number has appeared, there is 3 other even numbers left to again have a success. Thus, $E_2 \sim Geo(\frac{3}{8})$. Similarly, we find for the number of trials needed in phase 3 and 4 $E_3 \sim Geo(\frac{2}{8})$ and $E_4 \sim Geo(\frac{1}{8})$, respectively. The total number of spins required E is the sum of the independent E_i. Its expected value can be easily calculated:

$$\mathbb{E}(E) = \mathbb{E}(E_1) + \mathbb{E}(E_1) + \mathbb{E}(E_3) + \mathbb{E}(E_4) = \frac{8}{4} + \frac{8}{3} + \frac{8}{2} + \frac{8}{1} = \frac{50}{3} \approx 16.7. \checkmark$$

b) From the fact that the segments are equally sized we derive that all numbers have the same probability to appear. So, the two random variables X and Y have the densities

$$f_X(x) \begin{cases} \frac{1}{n} & \text{if } x \in [n] \\ 0 & \text{else} \end{cases}$$

and

$$f_{-Y}(y) \begin{cases} \frac{1}{n} & \text{if } -y \in [n] \\ 0 & \text{else} \end{cases} .$$

We suppose that X and Y are independent so $id(X) = X$ and $g(Y) = -Y$ are independent as well (this follows from a theorem of the lecture). Applying the convolution formula gives

$$f_Z(z) = f_{X+(-Y)}(z) = \sum_{x=1}^{n} f_X(x) \cdot f_{-Y}(z - x). \checkmark$$

z may take only integer values between $1 - n$ and $n - 1$ and, further, the factor $f_{-Y}(z-x)$ is non-zero if and only if $(x-z) \in [n]$. In particular, careful consideration of the possible outcomes of X, Y and Z yields that x has to be an integer between $k = \max\{1, z + 1\}$ and $l = \min\{n, z + n\}$. ✓ In case of $z \leq 0$, this is $k = 1$ and $l = n - |z|$. If, however, $z > 0$ we have $k = |z|+1$ and $l = n$. In any case x may take exactly $n - |z|$ integer values such that $k \leq x \leq l$ holds. Hence, the summation can be simplified:

$$f_Z(z) = \sum_{x=k}^{l} \frac{1}{n} \cdot \frac{1}{n} = \frac{n - |z|}{n^2}. \checkmark$$

In the case that z is not an integer number between $1 - n$ and $n - 1$ it is easy to see that $f_Z(z) = 0$ since here $f_{-Y}(z - x) = 0$ holds. Taken together, we obtain the probability mass function

$$f_Z(z) \begin{cases} \frac{n-|z|}{n^2} & \text{if } z \in \{1 - n, 2 - n, \cdots, n - 2, n - 1\} \\ 0 & \text{else} \end{cases} . \checkmark$$

Problem H 16 - A profitable toy factory **[4 pts.]**

Uncle Scrooge owns a toy factory that yields a yearly profit. He considers a year to be an *excellent* year if it had a profit that both exceeds the year before and the year after. In order to find the frequency of such excellent years Scrooge assigns the last $n+2$ years by a unique rank $\pi : [n+2] \to [n+2]$. For instance, $\pi(k) = 3$ would indicate that the year k was the 3rd best, so for $i \neq j$ $\pi(i) < \pi(j)$ means that year i had a higher profit than year j. Find the expected number of excellent years between years 2 and $n+1$ under the assumption that all ranks have the same probability.
Hint: It helps to find an indicator variable, to restrict considerations on possible π to specific ranks j, k, l for three succeeding years and to then use symmetry arguments.

Solution:
Let X_i be the indicator variable indicating whether or not the profit of year i was *excellent*. More specifically, $X_i = 1$ if $\pi(i) < \pi(i+1)$ and $\pi(i) < \pi(i-1)$ and $X_i = 0$ else. Hence, the total number of excellent years in the period of interest is given by $X = \sum_{i=2}^{n+1} X_i$. From the linearity of the expected value we additionally have that

$$\mathbb{E}(X) = \mathbb{E}(\sum_{i=2}^{n+1} X_i) = \sum_{i=2}^{n+1} \mathbb{E}(X_i).$$

This implies that we only need to determine the expected value of X_i to determine the expected number of excellent years X. ✓
The primary assumption is that all rankings π have the same probability. Thus, the probability $Pr(X_i)$ can be found with the help of the relative frequency of all those rankings for which $\pi(i) < \pi(i+1)$ and $\pi(i) < \pi(i-1)$ holds. It is convenient to first consider the subset of rankings for which the years $i-1$, i and $i+1$ take one of the ranks j, k and l while $j < k < l$. ✓
The relative frequency of having an excellent year i within these rankings is equal to $\frac{1}{3}$. This can be seen from the fact that there is 3! ways to assign the ranks j, k and l to the years $i-1$, i and $i+1$ but only for 2! of them the year i is excellent (i.e. if year i has the best rank $\pi(i) = j < k < l$). ✓ Going from specific ranks j, k, l to arbitrary ones, we find that their choice partitions all possible rankings in disjoint subsets of equal probability. By consequence, the relative frequency of rankings for which year i is an excellent one is $\frac{1}{3}$ in the general case as well. We finally get that the expected total number of excellent years is

$$\mathbb{E}(X) = \sum_{i=2}^{n+1} \frac{1}{3} = \frac{n}{3}. \checkmark$$

Problem H 17 - Roulette in the long run **[3 pts.]**

The classical roulette wheel has 37 fields enumerated by $0, 1, 2, \ldots, 35, 36$ and the ball in the game comes to fall on one of the fields with equal probability $\frac{1}{37}$.

a) It is easy to calculate the expected value of the number to be $\mathbb{E}(X) = 18$. You keep track of the game for some limited period of time and find the empirical mean $\overline{X} = 10.943 < 18$. Could it improve the winning chances to now favor numbers greater than 18?

b) By betting on "black" you get the double of the stake paid if the ball stops on one of the 18 black fields, otherwise you loose the bet. Your strategy is to bet 100 dollars on "black" every round. Estimate the probability to have a positive pay-out after 100 and 2000 rounds of the game using Chebyshev's inequality.

Solution:

a) No, the ball and thus the random variable X is memoryless - it does not "know" about which numbers appeared less frequently up to now. In the long run the lack of numbers greater than 18 is going to anneal by the law of large numbers. The contribution of the numbers up to now will be smaller and smaller for increasing number of trials. (Of course, all this assumes the idealizing framework of a perfect roulette wheel etc.) ✓

b) A single round of this betting strategy can be modeled as a Bernoulli experiment with the success probability $p = \frac{18}{37}$. Consequentially, n independent repetitions are equivalent to a binomial experiment with parameters n and p, i.e. $Y \sim B(n, p)$ where Y counts the numbers of successes corresponding to a win of 100 dollars each. The random variable Z of the positive or negative payoff which we are interested in actually is the linear transform $Z = 100 \cdot Y - 100 \cdot (n - Y)$. Clearly, we have $\mathbb{E}(Z) = 100 \cdot n \cdot p - 100 \cdot n \cdot (1 - p) = 100 \cdot n \cdot (2p - 1) = -100 \cdot \frac{n}{37}$ for the expected value and $Var(Z) = Var(200 \cdot Y) = 200^2 \cdot np(1 - p) = 200^2 \cdot \frac{18 \cdot 19}{37^2} \cdot n$. Chebyshev's inequality now yields

$$Pr(Z > 0) \leq Pr(|Z - \mathbb{E}(Z)| \geq |\mathbb{E}(Z)|) \leq \frac{Var(Z)}{\mathbb{E}(Z)^2} = \frac{4 \cdot 18 \cdot 19}{n}.$$

For $n = 100$ this yields $Pr(Z > 0) \leq 13.68$ which is trivial since $Pr(Z > 0) \leq 1$ in any case, for $n = 2000$ $Pr(Z > 0) \leq 0.68$. ✓✓

Problem H 18 - Moment Generating Functions **[4 pts.]**

---

a) Consider a binomial random variable with the parameters p and n. Determine the moment generating function $M_X(s)$ of X. Verify your result by calculating the expected value and the variance of X and comparing these with your previous knowledge on the binomial distribution.

b) Assume the moment generating function of a random variable Y is given by

$$M_Y(s) = e^{3(e^s - 1)}.$$

What is the type of distribution of Y? Determine $Pr(Y = 0)$.

Solution:

a)

$$\begin{aligned} M_X(s) &= \mathbb{E}(e^{sX}) \\ &= \sum_{k=0}^{n} e^{sk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^s)^k (1-p)^{n-k} \\ &= (pe^s + 1 - p)^n \checkmark \end{aligned}$$

where the binomial theorem was applied at the last step.
In order to get the expected value (the first central moment) we need to calculate the first derivative:

$$M_X'(s) = n(pe^s + 1 - p)^{n-1} pe^s.$$

This gives us $\mathbb{E}(X) = M_X'(0) = np$ which indeed matches the expected value of the binomial distribution known from the lecture.
Differentiating a second time yields

$$M_X''(s) = n(n-1)(pe^s + 1 - p)^{n-2}(pe^s)^2 + n(pe^s + 1 - p)^{n-1} pe^s.$$

Thus,

$$\mathbb{E}(X^2) = M_X''(0) = n(n-1)p^2 + np$$

which brings us to the variance by

$$\begin{aligned} Var(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= n(n-1)p^2 + np - n^2p^2 \\ &= np(1-p) \end{aligned}$$

matching the known formula as well. ✓

b) From the lecture we know the relation $M_Y(s) = G_Y(e^s)$ between the moment generating and the probability generating function, and further we now that $G_Y(t) = e^{\lambda(t-1)}$ is the probability generating function of a single Poisson distributed random variable. So, $M_Y(s)$ matches the moment generating function of a random variable from a Poisson distribution with rate $\lambda = 3$, $Y \sim Poi(3)$ ✓(this is guaranteed because of the one-to-one correspondence between the moment generating function and the distribution). By this, we obtain

$$Pr(Y = 0) = \frac{G_X^{(0)}}{0!} = \frac{e^{-3}}{0!} = e^{-3}. \checkmark$$