University of Primorska

FAMNIT, Mathematics

Probability

Midterm 1

April 16^{th}, 2019

Name and surname: ______________________ Identification number: [| | | | | | | |]

Instructions

Read carefully the text of the problems before attempting to solve them. Five problems out of six count for 100%. You are allowed one A4 sheet with formulae and theorems. You have two hours.

Problem	a.	b.	c.	d.	
1.			•	•	
2.					
3.			•	•	
4.			•	•	
5.			•	•	
6.				•	
Total					

1. (20) In the game *Points* two players A and B are tossing a fair coin. Assume the tosses are independent. If head comes up player A gets a point;, if tail comes up player B gets a point. Player A wins if he gets a points before player B gets b points, where a and b are two given positive integers.

a. (10) Express the probability that player A wins with a sum. You do not need to compute the sum.

Solution: assume the players continue to toss the coin forever. Denote by X the number of tosses until heads comes up a times. Player A wins if $X \leq a+b-1$. We know that $X \sim \mathrm{NegBin}(a, \frac{1}{2})$. It follows that

$$P\,(A\ \mathit{wins}) = \sum_{k=a}^{a+b-1} \binom{k-1}{a-1} \left(\frac{1}{2}\right)^k .$$

b. (10) Let us generalize the game of *Points* to players A, B and C. The players are selecting tickets with labels A, B and C from a box independently and with equal probability with replacement. If the ticket with label x is selected the player X gets a point. Player A wins if he collects a points before B collects b points or C collects c points for given numbers a, b and c.

Hint: $H_k = \{A$ collects a points when slecting ticket for the k-th time$\}$.

Solution: similarly as in the previous case, let us think that the players are selecting tickets forever. Denote by X the number of selections until they select a tickets A. We have $X \sim \mathrm{NegBin}\big(a, \frac{1}{3}\big)$. The player A has a chance to win, if the event $H_k = \{X = k\}$ happens for $k = a, a+1, \ldots, a+b+c-2$. If event H_k happens, we know that players B and C together gtt $k-a$ points. Conditionally n that, the points among the players A and B are distributed binomially with probability $\frac{1}{2}$. The player A wins if the number of points l ob the player B will satisfies $l < b$ and $k-a-l < c$ or $k-a-c < l < b$. It follows

$$P(A\ \mathit{wins}) = \sum_{k=a}^{\min\{a+b,a+c\}-1} P(H_k) + \sum_{k=\min\{a+b,a+c\}}^{a+b+c-2} P(H_k)\, P(A\ \mathit{wins} | H_k)\,.$$

The conditional probabilities in the second sum are equal to

$$P(A\ \mathit{wins}|H_k) = \left(\frac{1}{2}\right)^{k-a} \sum_{l=k-a-c+1}^{b-1} \binom{k-a}{l} .$$

We insert in the probabilities for negative binomial to get

$$P(A\ wins) = \left(\frac{1}{3}\right)^a \sum_{k=a}^{\min\{a+b,a+c\}-1} \binom{k-1}{a-1} \left(\frac{2}{3}\right)^{k-a} + \sum_{k=\min\{a+b,a+c\}}^{a+b+c-2} \binom{k-1}{a-1} \left(\frac{1}{3}\right)^k \sum_{l=k-a-c+1}^{b-1} \binom{k-a}{l}$$

with the agreement that the sum is zero if the final index is smaller than the starting one (this happens in the second sum if $a = b = 1$*). If we also agree that* $\binom{m}{r} = 0$*, as soon as* $r < 0$ *or* $r > m$*, the whole expression can be simplified into:*

$$P(A\ wins) = \sum_{k=a}^{a+b+c-2} \binom{k-1}{a-1} \left(\frac{1}{3}\right)^k \sum_{l=k-a-c+1}^{b-1} \binom{k-a}{l}.$$

2. (20) There is one white and two red balls in the urn. We start to pick balls from the urn. Everytime a white ball is picked, we return it to the urn and add another white ball. Everytime a red ball is picked, we do not return it to the urn and add two white balls instead. We assume that individual picks are independent.

a. (10) For $1 \leq m < n$ compute the probability that a red ball is picked for the first time on the m-th step and for the second time on the n-th step. Express the result in closed form.

Solution: the probability equals

$$\frac{1}{3} \cdot \frac{2}{4} \cdots \frac{m-1}{m+1} \cdot \frac{2}{m+2} \cdot \frac{m+2}{m+3} \cdots \frac{n}{n+1} \cdot \frac{1}{n+2} = \frac{4}{m(m+1)(n+1)(n+2)} .$$

b. (5) Denote by X the number of selections until all red balls are picked, including the last selection when the second red ball is selected. For all $n = 2, 3, 4, \ldots$ compute $P(X = n)$.

Hint: $\frac{1}{m(m+1)} = \frac{1}{m} - \frac{1}{m+1}$.

Solution: we have

$$P(X = n) = \sum_{m=1}^{n-1} \frac{4}{m(m+1)(n+1)(n+2)} = \frac{4(n-1)}{n(n+1)(n+2)} .$$

c. (5) Show that $P(X < \infty) = 1$.

Solution: The first option. *we have*

$$\begin{aligned} P(X < \infty) &= \sum_{n=2}^{\infty} P(X = n) \\ &= \sum_{n=2}^{\infty} \frac{4(n-1)}{n(n+1)(n+2)} \\ &= \sum_{n=2}^{\infty} \left(-\frac{2}{n} + \frac{8}{n+1} - \frac{6}{n+2} \right) \\ &= 6 \sum_{n=2}^{\infty} \left(\frac{1}{n+1} - \frac{1}{n+2} \right) - 2 \sum_{n=2}^{\infty} \left(\frac{1}{n} + \frac{1}{n+1} \right) \\ &= 6 \cdot \frac{1}{3} - 2 \cdot \frac{1}{2} \\ &= 1 . \end{aligned}$$

Second option. *the event* $\{X > n\}$ *is a disjoint union of the events* A_n *and the events* $A_{n,1}, A_{n,2}, \ldots, A_{n,n}$, *where:*

$$A_n = \{\textit{there is no red ball among the first } n \textit{ balls}\},$$
$$A_{n,m} = \{\textit{among first } n \textit{ balls there is only one red ball selected at the } m\textit{-th pick}\}.$$

We have

$$P(A_n) = \frac{1}{3} \cdot \frac{2}{4} \cdots \frac{m}{m+2} = \frac{2}{(m+1)(m+2)},$$
$$P(A_{n,m}) = \frac{1}{3} \cdot \frac{2}{4} \cdots \frac{m-1}{m+1} \cdot \frac{2}{m+2} \cdot \frac{m+2}{m+3} \cdots \frac{n+1}{n+2} = \frac{4}{m(m+1)(n+2)}.$$

Therefore

$$P(X > n) = P(A_n) + \sum_{m=1}^{n} P(A_{n,m}) = \frac{1}{n+2}\left[\frac{2}{n+1} + \sum_{m=1}^{n}\left(\frac{1}{m} - \frac{1}{m+1}\right)\right]$$
$$= \frac{2+4n}{(n+1)(n+2)},$$

which converges to zero, when n *goes to infinity. This means* $P(X < \infty) = 1$.

3. (20) From a set with N elements we repeatedly and independently select subsets with $n < N$ elements. All the subsets have equal probability of being selected. We repeat the selection independently r times.

a. (5) Fix a subset of size k. What is the probability that this subset will be contained in all r subsets selected.

Solution: $\left[\frac{\binom{N-k}{n-k}}{\binom{N}{n}}\right]^r$ *(with agreement* $\binom{m}{s} = 0$ *as soon as* $s < 0$ *or* $s > m$*).*

b. (15) What is the probability that no element is contained in all the r subsets selected? You do not need to simplify the sums.

Solution: for $i = 1, 2, \ldots, N$ *let* A_i *be the event that the* i*-th is cointained in all chosen subsets. In the previous task we have computed that for any* $i_1, i_2, \ldots, i_k$ *holds*

$$P\big(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}\big) = \left[\frac{\binom{N-k}{n-k}}{\binom{N}{n}}\right]^r .$$

With the help of inclusion exclusion formula we can compute

$$\begin{aligned}
&P\big(A_1^c \cap A_2^c \cap \cdots \cap A_N^c\big) \\
&\quad = 1 - P\big(A_1 \cup A_2 \cup \cdots \cup A_N\big) \\
&\quad = 1 - \sum_{i_1=1}^{N} P(A_{i_1}) + \sum_{1 \le i_1 < i_2 \le N} P(A_{i_1} \cap A_{i_2}) \\
&\qquad\quad - \sum_{1 \le i_1 < i_2 < i_3 \le N} P(A_{i_1} \cap A_{i_2} \cap A_{i_3}) + \cdots \\
&\qquad\quad + (-1)^n \sum_{1 \le i_1 < i_2 < \cdots < i_n \le N} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_n}) \\
&\quad = \sum_{k=0}^{n} (-1)^k \binom{N}{k} \left[\frac{\binom{N-k}{n-k}}{\binom{N}{n}}\right]^r .
\end{aligned}$$

4. (20) There are n white and n black balls in the urn. We are picking the balls randomly without replacement in such a way that the all $\binom{2n}{n}$ arrangements of white and black balls are equally likely.

a. (10) Compute the probability that just after the selection of the $(2k)$-th ball, the numbers of white and black balls among the $2k$ balls selected are equal.

   Solution: because the selections are random, after $2k$ balls we get a random sample of $2k$ balls. The number ob whites among them has hypergeometric distribution with parameters $2k$, n,$2n$. We want the number of white balls to be equal to k. The probability of this happening is

$$\frac{\binom{n}{k}\binom{n}{k}}{\binom{2n}{2k}}\,.$$

b. (10) Let N be the number of selections just after which the numbers of white and black balls among the selected balls are equal. Show that

$$E(N)=\sum_{k=1}^{n}\frac{\binom{n}{k}\binom{n}{k}}{\binom{2n}{2k}}\,.$$

   Solution: write

$$N=I_2+\cdots+I_{2n}\,,$$

   where

$$I_{2k}=\begin{cases} I_{2k}=1\,, & \text{if the number of white balls equals } k \text{ after } 2k \text{ selections} \\ 0 & \text{otherwise.}\end{cases}$$

   Compute

$$P(I_{2k}=1)=\frac{\binom{n}{k}\binom{n}{k}}{\binom{2n}{2k}}\,.$$

5. (20) Let the random variable X have the density

$$f_X(x) = \frac{2}{(e^x + e^{-x})^2} = \frac{1}{2\cosh^2 x}$$

for $x \in \mathbb{R}$.

a. (10) Compute the density of the random variable

$$U = \frac{1}{2}\left(1 + \frac{e^X - e^{-X}}{e^X + e^{-X}}\right).$$

Solution: we can write

$$U = \frac{1}{2}\left(1 + \tanh X\right),$$

therefore we are looking for

$$P(X \leq \operatorname{arctanh}(2u-1)).$$

We have

$$\operatorname{arctanh} x = \frac{1}{2}\ln\left(\frac{1+x}{1-x}\right)$$

or we can compute it directly

$$\begin{aligned} P(U \leq u) &= P\left(\tfrac{1}{2}\left(1 + \tfrac{e^X - e^{-X}}{e^X + e^{-X}}\right) \leq u\right) \\ &= P\left(\tfrac{e^X}{e^X + e^{-X}} \leq u\right) \\ &= P\left(\tfrac{1}{1+e^{-2X}} \leq u\right) \\ &= P\left(X \leq \tfrac{1}{2}\ln\left(\tfrac{u}{1-u}\right)\right). \end{aligned}$$

We differentiate the above distribution function and we get the density

$$\begin{aligned} f_U(u) &= f_X\left(\tfrac{1}{2}\ln\left(\tfrac{u}{1-u}\right)\right) \cdot \left(\tfrac{1}{2}\ln\left(\tfrac{u}{1-u}\right)\right)' \\ &= \frac{2}{\left(e^{\frac{1}{2}\ln\frac{u}{1-u}} + e^{\frac{1}{2}\cdot\ln\frac{u}{1-u}}\right)^2} \cdot \left(\tfrac{1}{2} \cdot \tfrac{1-u}{u} \cdot \tfrac{1-u+u}{(1-u)^2}\right) \\ &= 1 \end{aligned}$$

on the interval $u \in (0,1)$, which means that the random variable U is distributed uniformly $U(0,1)$.

b. (10) For $p \in (0,1)$ compute

$$P\left(X \leq \frac{1}{2}\log\left(\frac{1+p}{1-p}\right)\right).$$

Solution: we know that $(\tanh x)' = \frac{1}{\cosh^2 x}$. *Let compute*

$$\begin{aligned} P\left(X \leq \frac{1}{2}\log\left(\frac{1+p}{1-p}\right)\right) &= \int_{-\infty}^{\ln\frac{\sqrt{1+p}}{\sqrt{1-p}}} \frac{1}{2\cosh^2 x} dx \\ &= \frac{1}{2}\tanh x\Big|_{-\infty}^{\ln\frac{\sqrt{1+p}}{\sqrt{1-p}}} \\ &= \frac{1}{2}p + \frac{1}{2}. \end{aligned}$$

We have used the definiton $\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$ *and the property* $\lim_{x\to-\infty} \tanh x = -1$.

6. (20) There are n Aces among N cards. The deck of cards is shuffled well and cards are dealt from the top of the deck one by one. Denote the positions of Aces in the deck by $T_1 < T_2 < \cdots < T_n$. Define random variables $W_1 = T_1 - 1$, $W_i = T_i - T_{i-1} - 1$ for $2 \le i \le n$ and $W_{n+1} = N - T_n$. Assume as known that a given number m can be written as the sum of r positive numbers in $\binom{m-1}{r-1}$ ways and as a sum of nonnegative numbers in $\binom{m+r-1}{r-1}$ ways.

a. (10) Compute the joint distribution of the random variables $W_1, W_2, \ldots, W_n$ and the distributions of W_i for $i = 1, 2, \ldots, n+1$.

Solution: let us look at the joint distribution first. The possible values of the random vector $(W_1, \ldots, W_{n+1})$ are $(n+1)$-tuples $(k_1, \ldots, k_{n+1})$, for which holds $k_i \ge 0$ and $\sum_i k_i = N - n$. The event

$$\{W_1 = k_1, \ldots, W_{n+1} = k_{n+1}\}$$

happens if we get an Ace in positions $k_1+1, k_1+k_2+2, \ldots, N-k_n$. The probability for this to happen is

$$\frac{N-n}{N} \cdot \frac{N-n-1}{N-1} \cdots \frac{n}{N-k_1} \cdot \frac{N-n-k_1}{N-k_1-1} \cdots \frac{1}{1} = \frac{1}{\binom{N}{n}} .$$

By symmetry, all of the W_i are equally distributed, therefore is enough to compute the distribution of W_1. The possible values of this random variable are $k = 0, 1, \ldots, N-n$. We compute

$$\begin{aligned} P(W_1 = k) &= P(T_1 = k+1) \\ &= \frac{N-n}{N} \cdot \frac{N-n-1}{N-1} \cdots \frac{N-n-k+1}{N-k+1} \cdot \frac{n}{N-k} \\ &= \frac{\binom{N-k-1}{n-1}}{\binom{N}{n}} . \end{aligned}$$

b. (10) What is the probability that at least once two consecutive Aces are dealt from the deck of cards?

Solution: from the first part it follows that we need to count the $(n+1)$-tuples of nonnegative integers with the sum $N-n$, where at least one of the numbers equals to 0. If there is no restriction, the number of all $(n+1)$-tuples of nonnegative integers with the sum $N-n$ is $\binom{N}{n}$. If we restrict ourselves to $(n+1)$-tuples of positive integers there are $\binom{N-n-1}{n}$ such $(n+1)$-tuples. It follows

$$P(\min_i W_i = 0) = 1 - \frac{\binom{N-n-1}{n}}{\binom{N}{n}} .$$