NAME AND SURNAME: ______________________________

IDENTIFICATION NUMBER: | | | | | | | | |

UNIVERSITY OF PRIMORSKA

FAMNIT, MATHEMATICS

PROBABILITY

WRITTEN EXAMINATION

SEPTEMBER 6$^{\text{th}}$, 2023

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, and a handbook of mathematics. Time allowed: 120 minutes.

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

1. (20) Suppose you toss a fair coin. Tosses are independent. Let A_n be the event that the pattern HH does **not** occur in the first n tosses. Let $p_n = P(A_n)$.

a. (5) Define

$$\begin{aligned} B_1 &= \{\text{first toss is T}\} \\ B_2 &= \{\text{first two tosses are HT}\} \\ B_3 &= \{\text{first two tosses are HH}\} \end{aligned}$$

Argue that for $n \geq 4$

$$\begin{aligned} P(A_n|B_1) &= P(A_{n-1}) \\ P(A_n|B_2) &= P(A_{n-2}) \\ P(A_n|B_3) &= 0\,. \end{aligned}$$

Solution: the statements follow from independence of tosses.

b. (15) Show that for $n \geq 4$ we have

$$p_n = \frac{1}{2}p_{n-1} + \frac{1}{4}p_{n-2}\,.$$

Prove by induction that

$$p_n = \frac{4}{\sqrt{5}}\left(a^{n+2} - b^{n+2}\right)$$

where

$$a = \frac{1+\sqrt{5}}{4} \qquad \text{and} \qquad b = \frac{1-\sqrt{5}}{4}\,.$$

Hint: note that

$$\frac{a}{2} + \frac{1}{4} = a^2$$

and similarly for b.

Solution: for $n \geq 4$ we have by the formula for total probabilities

$$p_n = \frac{1}{2}p_{n-1} + \frac{1}{4}p_{n-2}\,.$$

The formula for p_n is correct for $n = 2, 3$ which we check directly. Assume the formula is correct up to $n \geq 4$. We compute

$$\begin{aligned} p_{n+1} &= \frac{1}{2}p_n + \frac{1}{4}p_{n-1} \\ &= \frac{4}{\sqrt{5}}\left(\frac{1}{2}a^{n+2} + \frac{1}{4}a^{n+1} - \frac{1}{2}b^{n+2} - \frac{1}{4}b^{n+1}\right) \\ &= \frac{4}{\sqrt{5}}\left(a^{n+1}\left(\frac{a}{2} + \frac{1}{4}\right) - b^{n+1}\left(\frac{b}{2} + \frac{1}{4}\right)\right) \\ &= \frac{4}{\sqrt{5}}\left(a^{n+3} - b^{n+3}\right)\,. \end{aligned}$$

This completes the induction step.

2. (20) Two bored statisticians are independently rolling two dice, one die each, until the sum of their outcomes equals 6. Denote the number of rolls by X, including the last roll.

a. (10) Find the distribution of the random variable X.

   Solution: the probability that the sum in a roll is equal to 6 *is* 5/36. *The rolls are independent, which means that we are repeating the experiment until the first "success". It follows* $X \sim \text{Geom}(5/36)$.

b. (10) Let Y be the number of dots on the first die the first time the sum of dots on both dice is 6. Find the distribution of (X, Y).

   Solution: the possible values of the vector (X, Y) *are the pairs* (k, l) *with* $k \geq 1$ *and* $1 \leq l \leq 6$. *The event* $\{X = k, Y = l\}$ *happens if in the first* $k-1$ *rolls the sum* 6 *does not occur, and on the* k*-th roll the first statistician rolls* l *and the other* $6 - l$. *Beacuse of independence we have*

$$P(X = k, Y = l) = \left(\frac{31}{36}\right)^{k-1} \cdot \left(\frac{1}{6}\right)^{2} .$$

3. (20) Let X, Y and Y be independent with $X, Y, Z \sim \exp(1)$. Define

$$(U, V, W) = \left(X + Y, \frac{X}{X+Y}, \frac{Z}{X+Y+Z}\right).$$

a. (10) Find the density of the vector (U, V, W).

Solution: the map

$$\Phi(x, y, z) = \left(x + y, \frac{x}{x+y}, \frac{z}{x+y+z}\right)$$

takes the set $(0, \infty)^3$ *bijectively onto* $(0, \infty) \times (0, 1)^2$ *and is differentiable. We find*

$$\Phi^{-1}(u, v, w) = \left(uv, u(1-v), \frac{uw}{1-w}\right).$$

We compute

$$J_{\Phi^{-1}}(u, v, w) = \det\begin{pmatrix} v & u & 0 \\ 1-v & -u & 0 \\ \frac{w}{1-w} & 0 & \frac{u}{(1-w)^2} \end{pmatrix} = \frac{u^2}{(1-w)^2}.$$

We have

$$f_{U,V,W}(u, v, w) = e^{-\frac{u}{1-w}} \cdot \frac{u^2}{(1-w)^2}$$

for $u \in (0, \infty)$ *and* $v, w \in (0, 1)$.

b. (10) Show that V and W are independent.

Solution: the density of (V, W) *is the marginal density of the vector* (U, V, W) *Integrating over* u *we get*

$$\begin{aligned} f_{V,W}(v, w) &= \int_0^\infty f_{U,V,W}(u, v, w) du \\ &= \frac{1}{(1-w)^2} \int_0^\infty u^2 e^{-\frac{u}{1-w}} du \\ &= (1-w) \int_0^\infty t^2 e^{-t} dt \\ &= 2(1-w). \end{aligned}$$

The density is a product on $(0, 1)^2$ *which means that* V *and* W *are independent.*

4. (20) A deck of card contains a white cards numbered from 1 to a and b red cards numbered 1 to b. We shuffle the cards so that all permutations of the $a+b$ cards are equally likely, and then deal them one by one from the top until no cards are left. Let X be the number of white cards before the first red card, and Y be the number of white cards after the last red card. Define

$$I_k = \begin{cases} 1 & \text{if the } k\text{-th white card comes before the first red card;} \\ 0 & \text{else;} \end{cases}$$

and

$$J_l = \begin{cases} 1 & \text{if the } l\text{-th white card comes after the last red card;} \\ 0 & \text{else;} \end{cases}$$

a. (10) Compute

$$P(I_k = 1, J_l = 1)$$

for all $1 \le k, l \le a$.

Rešitev: for $k = l$ the event $\{I_k = 1, J_k = 1\}$ is impossible, so $P(I_k = 1, J_k = 1) = 0$. For $k \neq l$ we look at the k-th white and l-th white cards (excluding the other white cards) and the red cards. These $b+2$ cards are randomly permuted by symmetry. The event $\{I_k = 1, J_l = 1\}$ happens if in this permutation of $b+2$ cards, the k-th white card is at the beginning, and the l-th white card is at the end. Out of $(b+2)!$ such permutations there are $b!$ favourable ones. It follows that

$$P(I_k = 1, J_l = 1) = \frac{1}{(b+1)(b+2)} .$$

b. (10) Compute $\text{cov}(X, Y)$.

Solution: note that $X = \sum_{k=1}^{a} I_k$ and $Y = \sum_{l=1}^{b} J_l$, By bilinearity

$$\text{cov}(X, Y) = \sum_{k,l=1}^{a} \text{cov}(I_k, J_l) .$$

We need $E(I_k)$ and $E(J_l)$ to conclude the computation. By a similar symmetry argument as in the first part we get

$$E(I_k) = \frac{1}{b+1} \qquad \textit{and} \qquad E(J_l) = \frac{1}{b+1} .$$

It follows that

$$\text{cov}(I_k, J_k) = -\frac{1}{(1+b)^2} ,$$

and for $k \neq l$

$$\text{cov}(I_k, J_l) = \frac{1}{(b+1)(b+2)} - \frac{1}{(b+1)^2} = -\frac{1}{(b+1)^2(b+2)} .$$

Finally,

$$\text{cov}(X, Y) = -\frac{a}{(b+1)^2} - \frac{a(a-1)}{(b+1)^2(b+2)} = -\frac{a(a+b+1)}{(b+1)^2(b+2)} .$$

5. (20) Suppose you toss a fair coin. Tosses are independent. Denote the number of tosses until the first occurence of the pattern HH by X.

a. (5) Define

$$\begin{aligned} B_1 &= \{\text{first toss is T}\} \\ B_2 &= \{\text{first two tosses are HT}\} \\ B_3 &= \{\text{first two tosses are HH}\} \end{aligned}$$

Write $E\left(s^X|B_i\right)$ for $i = 1,2,3$ using expressions involving constants, s and $G_X(s)$.

Solution: because of independence of tosses, conditionally on B_1, X has the distribution equal to $1+X$. We have

$$E\left(s^X|B_1\right) = sG_X(s)\,.$$

A similar argument gives

$$E(s^X|B_2) = s^2 G_X(s)\,.$$

Finally,

$$E\left(s^X|B_3\right) = s^2\,.$$

b. (5) Show that

$$G_X(s) = \frac{s}{2}G_X(s) + \frac{s^2}{4}G_X(s) + \frac{s^2}{4}$$

Solution: the formula follows from the law of total expectations noting that $P(B_1) = \frac{1}{2}$ and $P(B_2) = P(B_3) = \frac{1}{4}$.

c. (10) Find $\operatorname{var}(X)$.

Solution: taking derivatives on both sides we get

$$G'_X(s) = \frac{1+s}{2}\cdot G_X(s) + \frac{2s+s^2}{4}\cdot G'_X(s) + \frac{s}{2}\,.$$

Sending $s \uparrow 1$ we get

$$E(X) = 1 + \frac{3}{4}E(X) + \frac{1}{2}\,,$$

or $E(X) = 6$. Taking derivatives again, we get

$$G''_X(s) = \frac{1}{2}G_X(s) + \frac{1+s}{2}\cdot G'_X(s) + \frac{1+s}{2}\cdot G'_X(s) + \frac{2s+s^2}{4}\cdot G''_X(s) + \frac{1}{2}\,.$$

Sending $s \uparrow 1$ we get

$$E\left[X(X-1)\right] = \frac{1}{2} + 6 + 6 + \frac{3}{4}E\left[X(X-1)\right] + \frac{1}{2}\,,$$

or

$$E\left[X(X-1)\right] = 52\,.$$

Finally,

$$\operatorname{var}(X) = E\left[X(X-1)\right] + E(X) - E(X)^2 = 22\,.$$

6. (20) The patrons in the HIT Casino played the game *Colore* 440,000 times in 1999. The probability of winning in the game is $p = 0.00198079$. We assume that subsequent games are independent.

a. (10) The number of winning games in n games is the sum $S_n = X_1 + \cdots + X_n$, where $P(X_i = 1) = p$ and $P(X_i = 0) = 1 - p$, and the random variables $X_1, X_2 \ldots$ are independent. What, approximately, is the probability that there will be 920 winning games or more if $n = 440{,}000$?

Solution: we estimate using the central limit theorem

$$\begin{aligned} P(S_n \geq 920) &= P\left(\frac{S_n - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} \geq \frac{920 - E(S_n)}{\sqrt{\mathrm{var}(S_n)}}\right) \\ &\approx P\left(Z \geq \frac{920 - E(S_n)}{\sqrt{\mathrm{var}(S_n)}}\right), \end{aligned}$$

where $Z \sim N(0,1)$. *We compute* $E(S_n) = np = 871.55$ *and* $\sqrt{\mathrm{var}(S_n)} = \sqrt{npq} = 29.49$, *which yields*

$$\frac{920 - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} = \frac{920 - 871.55}{29.49} = 1.64\,.$$

We have

$$P(S_n \geq 920) \approx 0.050\,.$$

b. (10) Suppose the payout is $x > 0$ when a patron wins in a game. If the patron bets one unit and wins the game, she gets her stake back and an additional x units. The house wants that the probability of losing money in 440,000 games is at most 0.01. Argue that the biggest payout x such that the house loses money with probability 0.01 or less in 440,000 games, is approximately 467.

Solution: the profit of the house in each game is either 1 or $-x$. *Let us denote the profit in game* i *by* Y_i, *so that the total profit after* n *games is* $T_n = Y_1 + \cdots + Y_n$, *where* $P(Y_i = 1) = 1 - p$ *and* $P(Y_i = -x) = p$. *We are looking for the maximum* x *such that*

$$P(T_n < 0) \leq 0.01\,.$$

We compute $E(Y_i) = 1 - p - px$ *and* $\mathrm{var}(Y_i) = pq(x+1)^2$. *We have*

$$\begin{aligned} P(T_n < 0) &= P\left(\frac{T_n - E(T_n)}{\sqrt{\mathrm{var}(T_n)}} < -\frac{E(T_n)}{\sqrt{\mathrm{var}(T_n)}}\right) \\ &\approx P\left(Z < -\frac{n(q - px)}{\sqrt{npq(x+1)^2}}\right) \\ &\approx 0.01\,. \end{aligned}$$

Denote $n = 440{,}000$. *We need to solve the equation*

$$-\frac{n(q - px)}{\sqrt{npq(x+1)^2}} = -2,33\,.$$

Solving for x *gives* $x \approx 467$.