NAME AND SURNAME: IDENTIFICATION NUMBER:

UNIVERSITY OF PRIMORSKA

FAMNIT, MATHEMATICS

PROBABILITY

MIDTERM 2

JUNE 1st, 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) An urn contains n white balls labeled from 1 to n and b blue balls. You choose balls randomly with replacement until you draw the first blue ball at which point you stop. Let X be the number of different labels that you draw. Let N be the total of balls drawn, and let Y_k be the number of appearances of label k for $k = 1, \ldots, n$.

a. (10) Compute $E(X)$.

Hint: indicators are one possibility.

Solution: define

$$I_k = \begin{cases} 1 & \textit{if we see the ball with label } k; \\ 0 & \textit{else;} \end{cases}$$

We have $X = I_1 + \cdots + I_n$. *By symmetry all indicators have the same expectation. The probability that the ball with label* 1 *comes before the first blue ball is by symmetry equal to* $1/(1+b)$, *and hence*

$$E(X) = \frac{n}{1+b}\,.$$

b. (10) Compute $E(Y_k)$ for $k = 1, 2, \ldots, n$.

Hint: what is the sum $Y_1 + \cdots + Y_n$.

Solution: we have $Y_1 + \cdots + Y_n = N - 1$. *By symmetry all* $E(Y_k)$ *are equal. The random variable* N *is geometric with parameter* $p = b/(b+n)$, *hence*

$$E(N) = \frac{1}{p} = 1 + \frac{b}{n}\,.$$

It follows that

$$nE(Y_k) = E(N) - 1 = \frac{b}{n}\,,$$

and finally

$$E(Y_k) = \frac{b}{n^2}$$

for $k = 1, \ldots, n$.

2. (20) Let the distribution of X and Y be given by

$$P(X=k, Y=l) = \binom{2k}{k} \frac{\left(k+\frac{1}{2}\right)_l}{4^k \cdot 3^{k+l+\frac{1}{2}} \cdot l!} = \frac{[2(k+l)]!}{4^{k+l} \cdot 3^{k+l+\frac{1}{2}} \cdot (k+l)! \cdot k! \cdot l!}$$

for $k, l = 0, 1, 2, \ldots$ where

$$(a)_l = \frac{\Gamma(a+l)}{\Gamma(a)}$$

is the Pochammer symbol.

a. (10) Find the distribution of the sum $Z = X + Y$.

Solution: the possible values for Z are $n = 0, 1, 2, \ldots$. We compute

$$\begin{aligned} P(Z=n) &= \sum_{k=0}^{n} P(X=k, Y=n-k) \\ &= \frac{(2n)!}{4^n \cdot 3^{n+\frac{1}{2}} \cdot n!} \sum_{k=0}^{n} \frac{1}{k! \cdot (n-k)!} \\ &= \frac{(2n)!}{4^n \cdot 3^{n+\frac{1}{2}} \cdot n! \cdot n!} \sum_{k=0}^{n} \frac{n!}{k! \cdot (n-k)!} \\ &= \frac{(2n)!}{4^n \cdot 3^{n+\frac{1}{2}} \cdot n! \cdot n!} \sum_{k=0}^{n} \binom{n}{k} \\ &= \frac{(2n)!}{4^n \cdot 3^{n+\frac{1}{2}} \cdot n! \cdot n!} \cdot 2^n \\ &= \frac{\binom{2n}{n}}{2^n \cdot 3^{n+\frac{1}{2}}} . \end{aligned}$$

b. (10) Find the distribution of X.

Hint: remember that Newton's formula gives

$$(1-x)^{-a} = \sum_{l=0}^{\infty} \frac{(a)_l}{l!} .$$

Solution: using the formula for marginal distributions we compute

$$\begin{aligned}
P(X=k) &= \sum_{l=0}^{\infty} P(X=k, Y=l) \\
&= \sum_{l=0}^{\infty} \frac{\binom{2k}{k}}{4^k \cdot 3^{k+l+\frac{1}{2}}} \cdot \frac{\left(k+\frac{1}{2}\right)_l}{l!} \\
&= \frac{\binom{2k}{k}}{4^k \cdot 3^{k+\frac{1}{2}}} \sum_{l=0}^{\infty} \frac{\left(k+\frac{1}{2}\right)_l}{l! \cdot 3^k} \\
&= \frac{\binom{2k}{k}}{4^k \cdot 3^{k+\frac{1}{2}}} \left(1-\frac{1}{3}\right)^{-k-\frac{1}{2}} \\
&= \frac{\binom{2k}{k}}{\sqrt{2} \cdot 8^k}\,.
\end{aligned}$$

3. (20) Let Y and W be independent with $Y \sim \exp(1)$ and $W \sim \mathrm{N}(0,1)$. Let

$$X = \theta Y + \sigma\sqrt{Y}W\,,$$

where $\theta > 0$ and $\sigma > 0$ are given constants.

a. (10) Find the density of (Y, X).

Solution: let

$$\Phi(y,w) = (y, \theta y + \sigma\sqrt{y}w)\ .$$

The map is bijective on the open set $U = \{(y,w) \colon y > 0\}$ *and maps* U *onto itself. We compute*

$$\Phi^{-1}(y,x) = \left(y, \frac{x-\theta y}{\sigma\sqrt{y}}\right)\,.$$

All the assumptions of the transformation formula are satisfied and we get

$$J_{\Phi^{-1}}(y,x) = \frac{1}{\sigma\sqrt{y}}\,.$$

It follows that

$$f_{Y,X}(y,x) = f_{Y,W}\left(y, \frac{x-\theta y}{\sigma\sqrt{y}}\right)\cdot J_{\Phi^{-1}}(y,x)\,.$$

By independence $f_{Y,W}(y,w) = f_Y(y)f_W(w)$*, and hence for* $(y,x) \in U$ *we have*

$$f_{Y,X}(y,x) = \frac{1}{\sqrt{2\pi}}\exp\left(-y - \frac{1}{2}\left(\frac{x-\theta y}{\sigma\sqrt{y}}\right)^2\right)\cdot\frac{1}{\sigma\sqrt{y}}\,,$$

and $f_{Y,X}(y,x) = 0$ *else.*

b. (10) Find the density of X. Assume as known that for $a > 0$ and $b \geq 0$ we have

$$\int_0^\infty \frac{e^{-ay-\frac{b}{y}}}{\sqrt{y}}\,dy = \frac{\sqrt{\pi}}{\sqrt{a}}\,e^{-2\sqrt{ab}}\,.$$

Solution: the density of X *is the marginal denisty of* $f_{Y,X}(y,x)$*. We compute*

$$\begin{aligned}
f_X(x) &= \int_0^\infty f_{Y,X}(x,y)\,dy \\
&= \frac{1}{\sigma\sqrt{2\pi}}\int_0^\infty \frac{1}{\sqrt{y}}\exp\left(-y - \frac{x^2}{2\sigma^2 y} + \frac{\theta x}{\sigma^2} - \frac{\theta^2 y}{2\sigma^2}\right)dy \\
&= \frac{1}{\sigma\sqrt{2\pi}}\,e^{\theta x/\sigma^2}\int_0^\infty \frac{1}{\sqrt{y}}\exp\left(-\frac{x^2}{2\sigma^2 y} - \frac{(\theta^2+2\sigma^2)y}{2\sigma^2}\right)dy \\
&= \frac{1}{\sigma\sqrt{2\pi}}\,e^{\theta x/\sigma^2}\,\frac{\sqrt{\pi}}{\sqrt{\frac{\theta^2+2\sigma^2}{2\sigma^2}}}\exp\left(-2\sqrt{\frac{x^2}{2\sigma^2}}\cdot\sqrt{\frac{\theta^2+2\sigma^2}{2\sigma^2}}\right) \\
&= \frac{1}{\sqrt{\theta^2+2\sigma^2}}\,e^{-(|x|\sqrt{\theta^2+2\sigma^2}-\theta x)/\sigma^2}\,.
\end{aligned}$$

4. (20) A deck of cards contains a white and b red cards. We shuffle the deck and place the cards on a table one by one from the top of the deck. Let X be the number of white cards before the first red card, and Y the number of white cards after the last red card.

a. (10) Number the white cards with $k = 1, 2, \ldots, a$ and define

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

and

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

Justify that $X = \sum_{k=1}^{a} I_k$ and $Y = \sum_{l=1}^{a} J_l$. Compute $\mathrm{cov}(X, Y)$.

Solution: the relative positions of the first white and the red cards are a random permutation of these $1+b$ cards. The probability that the first white card appears before all the red cards is therefore

$$P(I_1 = 1) = \frac{1}{b+1}\,.$$

A similar argument gives

$$P(J_1 = 1) = \frac{1}{b+1}\,.$$

We have

$$P(I_1 = 1, J_1 = 1) = 0 \quad \textit{and} \quad P(I_1 = 1, J_2 = 1) = \frac{1}{(b+1)(b+2)}\,.$$

By bilinearity of covariance we compute

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

By symmetry, all variances and covariances are equal. We have

$$\mathrm{cov}(X, Y) = a\,\mathrm{cov}(I_1, J_1) + a(a-1)\,\mathrm{cov}(I_1, J_2)\,.$$

Since $I_1 J_1 = 0$, we have

$$\mathrm{cov}(I_1, J_1) = -\frac{1}{(1+b)^2}\,.$$

Furthermore,

$$\mathrm{cov}(I_1, J_2) = \frac{1}{(b+1)(b+2)} - \frac{1}{(b+1)^2} = -\frac{1}{(b+1)^2(b+2)}\,.$$

Finally,

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

b. (10) Assume that $b \geq 2$, and that Z is the number of white cards between the first and the second red card. For $l \leq a$ compute

$$\mathrm{cov}(X, Y \mid Z = l)\,.$$

Hint: think about the conditional distribution of (X, Y) before computing.

Solution: if the white cards between the first and the second card and the second red card are removed from the deck, we get a renom permutation of the remaining $a - l$ white and $b - 1$ red cards. The first part gives

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

5. (20) An urn contains balls with labels 0, 1 and 2. We select balls randomly with replacement, where the number of balls selected has the Po(3) distribution. Let Z be the sum of all numbers on the selected balls. We assume the selection process is independent of the random number of selections, and the selections themselves are independent, and we select each ball with probability 1/3.

a. (10) Compute $P(Z = 1)$.

Solution: in mathematical notation we have $Z = X_1 + X_2 + \cdots + X_N$, *where* X_i *is the number on the* i*-the selected ball, and* N *is the random number of selections. The generation function* $G(s)$ *of* X_i *is*

$$G(s) = \frac{1 + s + s^2}{3},$$

which implies that the generating function $K(s)$ *of* Z *is*

$$K(s) = e^{s^2+s-2}.$$

We have

$$K'(s) = (2s + 1)\, e^{s^2+s-2}$$

from which we deduce

$$P(Z = 1) = K'(0) = e^{-2} \doteq 0{,}135.$$

b. (10) Compute $\operatorname{var}(Z)$.

Solution: first we have

$$E(Z) = K'(1) = 3.$$

From

$$K''(s) = (4s^2 + 4s + 3)\, e^{s^2+s-2}$$

we get

$$E\big[Z(Z-1)\big] = K''(1) = 11,$$

and finally

$$\operatorname{var}(Z) = E(Z^2) - \big(E(Z)\big)^2 = E\big[Z(Z-1)\big] + E(Z) - \big(E(Z)\big)^2 = 5.$$

6. (20) You play the following game of chance: you choose 1000 tickets with numbers on them from a box at random with replacement. If the sum of the numbers on tickets is between a and $a + 100$, you win. You can choose the number a. You know that the average of the box is 0.1 and the standard error is 1.5811.

a. (10) Suppose you chose $a = 0$. What, approximatly, is the probability that you win?

Solution: let the sum of the numbers be S_{1000}. The central limit theorem gives

$$\begin{aligned} P(0 \leq S_n \leq 100) &= P(-100 \leq S_{1000} - 100 \leq 0) \\ &= P(-\frac{100}{\sqrt{1000} \cdot 1.5811} \leq \frac{S_{1000} - 100}{\sqrt{1000} \cdot 1.5811} \leq 0) \\ &\approx P(-2 \leq Z \leq 0) \\ &= 0.48\,. \end{aligned}$$

b. (10) Approximately which choice for a gives you the largest probability of winning?

Solution: we compute

$$\begin{aligned} & P(a \leq S_n \leq a + 100) \\ &= P(a - 100 \leq S_{1000} - 100 \leq a) \\ &= P\left(\frac{a - 100}{\sqrt{1000} \cdot 1.5811} \leq \frac{S_{1000} - 100}{\sqrt{1000} \cdot 1.5811} \leq \frac{a}{\sqrt{1000} \cdot 1.5811}\right) \\ &\approx P\left(\frac{a - 100}{\sqrt{1000} \cdot 1.5811} \leq Z \leq \frac{a}{\sqrt{1000} \cdot 1.5811}\right)\,. \end{aligned}$$

The last probability is the integral of the standard normal density over an interval of length $100/(\sqrt{1000} \cdot 1.5811) = 2$. This integral is largest if the interval is symmetric, which means

$$\frac{a}{\sqrt{1000} \cdot 1.5811} = 1\,,$$

in which case the probability of winning is 0.68 and $a \approx 50$.