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

University of Primorska

FAMNIT, Mathematics

Probability

Written examination

August 23rd, 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) You have two identical decks of cards containing n cards labelled 1 to n. Suppose you merge the two decks into one, shuffle it, and then deal two cards at the time from the top of the deck until there are no cards left. Define

$$A_k = \{\text{the } k\text{-th pair of cards has matching labels}\}$$

for $k = 1, 2, \ldots, n$.

a. (15) Compute $P(A_1 \cap A_2 \cap \cdots \cap A_k)$ for $k = 1, 2, \ldots, n$.

Solution: the cards in positions $(1, 2)$ *in the merged deck are a random pair out of the* $2n$ *cards. Out of all* $\binom{2n}{2}$ *possible pairs of cards, there are* n *pairs with matching labels. The probability of matching labels is* $n/\binom{2n}{2}$. *The cards in positions* $(3, 4)$ *are a random pair out of the remaining* $2n-2$ *cards. The number of labels is reduced by* 1 *each time. Iterating the argument, we get*

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = \prod_{i=0}^{k-1} \frac{n-i}{\binom{2n-2i}{2}}$$

which simplifies to

$$P(A_1 \cap A_2 \cap \cdots \cap A_k) = \prod_{i=0}^{k-1} \frac{1}{2n-2i-1}\,.$$

b. (5) Compute the probability that there are no matching labels among the pairs of cards you deal. You do not need to simplify the sums.

Solution: by symmetry and the inclusion-exclusion principle we have

$$P(\textit{there is at least one mathing pair}) = \sum_{k=1}^{n} (-1)^{k-1} \binom{n}{k} P(A_1 \cap A_2 \cap \cdots \cap A_k)$$

2. (20) An urn contains r white and r black balls. Assume balls are selected one by one at random without replacement.

a. (10) Let M be the number of balls selected until the first white ball is drawn including the white ball. Compute $P(M=k)$ for $k=1,2,\ldots,r+1$.

   Solution: before the first white ball, we need to draw black balls. It follows

$$P(M=k)=\frac{r}{2r}\cdot\frac{r-1}{2r-1}\cdots\frac{r-k+2}{2r-k+2}\cdot\frac{r}{2r-k+1}$$

b. (10) Let N be the number of balls selected until all r white or all r black balls have been drawn. Compute $P(N=k)$ for $k=r,r+1,\ldots,2r-1$.

   Solution: the event $\{N=k\}$ happens, if the first $k-1$ draws produce $r-1$ balls of the same colour, and the last ball is of the same colour. The first $k-1$ draws produce a random sample of all $2r$ balls. The probability that there are $k-1$ white balls among them is by the hypergeometric distribution equal to

$$\frac{\binom{r}{r-1}\binom{r}{k-r}}{\binom{2r}{k-1}}\,.$$

   Conditionally on the above, the probability that the k-th ball will be white is $1/(2r-k+1)$. By symmetry, the overall probability is

$$2\cdot\frac{\binom{r}{r-1}\binom{r}{k-r}}{\binom{2r}{k-1}}\cdot\frac{1}{2r-k+1}\,.$$

3. (20) Let X and Y have the density

$$f_{X,Y}(x,y) = \begin{cases} \frac{2}{\pi} e^{-\frac{x^2+y^2}{2}} & \text{for } x, y > 0; \\ 0 & \text{else;} \end{cases}$$

Define

$$(U,V) = \left(X^2 + Y^2, \frac{Y^2}{X^2+Y^2} \right) .$$

a. (15) Find the density of the vector (U,V).

Solution: the map

$$\Phi(x,y) = \left(x^2 + y^2, \frac{y^2}{x^2+y^2} \right)$$

takes the region $G = \{(x,y) \colon x, y > 0\}$ *bijectively onto* $H = (0,\infty) \times (0,1)$ *and is differentiable on* G. *We have*

$$\Phi^{-1}(u,v) = \left(\sqrt{u(1-v)}, \sqrt{uv} \right)$$

We compute

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

Simplifying we get

$$J_{\Phi^{-1}}(u,v) = \frac{1}{2} \frac{1}{\sqrt{v(1-v)}} .$$

The transformation formula gives

$$f_{U,V}(u,v) = \frac{1}{2\pi} e^{-u/2} \frac{1}{\sqrt{v(1-v)}}$$

on $(0,\infty) \times (0,1)$.

b. (10) Are U and V independent? What are the distributions of these two random variables?

Solution: the density splits into a product on $(0,\infty) \times (0,1)$ *which implies independence. We infer that* $U \sim \exp(1/2)$ *and* $V \sim \mathrm{Beta}(1/2, 1/2)$.

4. (20) Let $X \sim \text{Bin}(n, \frac{1}{2})$. Let Y be integer valued with

$$P\left(Y = l | X = k\right) = \binom{m}{k}\left(\frac{k}{n}\right)^l \left(1 - \frac{k}{n}\right)^{m-l}$$

for $0 \leq l \leq m$, and we interpret $0^0 = 1$.

a. (10) Find $E(Y)$.

Solution: the conditional distribution of Y given $\{X = k\}$ is binomial with parameters m and $\frac{k}{n}$. It follows that

$$E(Y|X = k) = \frac{mk}{n}\,.$$

By the formula for total expectations we have

$$\begin{aligned} E(Y) &= \sum_{k=0}^{n} E(Y|X = k)P(X = k) \\ &= \sum_{k=0}^{n} \frac{mk}{n} P(X = k) \\ &= \frac{m}{n} \sum_{k=0}^{n} kP(X = k) \\ &= \frac{m}{n} E(X) \\ &= \frac{m}{2}\,. \end{aligned}$$

b. (10) Find $\text{var}(Y)$.

Solution: we need to compute $E(Y^2)$. We have

$$E(Y^2|X = k) = m \cdot \frac{k}{n} \cdot \left(1 - \frac{k}{n}\right) + \frac{m^2k^2}{n^2}\,.$$

We compute taking into account that

$$E(X^2) = \text{var}(X) + E(X)^2 = \frac{n}{4} + \frac{n^2}{4}\,.$$

We have

$$
\begin{aligned}
E(Y^2) &= \sum_{k=0}^{n} E(Y^2|X=k)P(X=k) \\
&= \sum_{k=0}^{n} \left(m \cdot \frac{k}{n} \cdot \left(1-\frac{k}{n}\right) + \frac{m^2k^2}{n^2}\right) P(X=k) \\
&= \frac{m}{n^2} E\left(X(n-X)\right) + \frac{m^2}{n^2} E(X^2) \\
&= \frac{m}{n} E(X) + \frac{m^2-m}{n^2} E(X^2) \\
&= \frac{m}{2} + \frac{m^2-m}{n^2}\left(\frac{n}{4}+\frac{n^2}{4}\right) .
\end{aligned}
$$

5. (20) Let $Z_0, Z_1, \ldots$ be a branching process with $\mu = E(Z_1) \leq 1$ and $\mathrm{var}(Z_1) > 0$, so that the process dies out with probability 1. Let W be the total number of individuals in the family tree with no offspring. Assume as known that

$$G_W(s) = (s-1)G_{Z_1}(0) + G_{Z_1}\left(G_W(s)\right) .$$

a. (10) Assume $\mu < 1$. Compute $E(W)$.

Solution: we compute

$$G'_W(s) = G_{Z_1}(0) + G'_{Z_1}(G_W(s))G'_W(s) .$$

Letting $s \uparrow 1$ on both sides of the equation, we get the equation

$$E(W) = G_{Z_1}(0) + E(Z_1)E(W) .$$

It follows that

$$E(W) = \frac{G_{Z_1}(0)}{1-\mu} .$$

b. (10) Assume that

$$G_{Z_1}(s) = \frac{1+s^2}{2} .$$

Find the distribution of W.

Solution: the equation for $G_W(s)$ gives that

$$G_W(s) = \frac{(s-1)}{2} + \frac{1+G_W(s)^2}{2} .$$

Solving for $G_W(s)$ we get

$$G_W(s) = 1 \pm \sqrt{1-s} .$$

To have a generating function, we need to take the minus sign. Newton's formula gives

$$G_W(s) = \sum_{k=1}^{n} \binom{1/2}{k} (-1)^{k+1} s^k ,$$

and we have

$$P(W=k) = \binom{1/2}{k} (-1)^{k+1}$$

for $k = 1, 2, \ldots$.

6. (20) The Airbus A380 plane has 525 seats. Typically, carriers sell more tickets than there are seats on the plane, assuming that some passengers with tickets do not present themselves for the flight. Suppose the carrier sells n tickets, and statistics show that 90% of the passengers do present themselves for the flight. If we assume that passengers act independently, the number of passengers who show up is a random variable $S_n = I_1 + I_2 + \cdots + I_n$, where the $I_1, I_2, \ldots$ are independent and $I_k \sim \text{Bernoulli}(0.9)$ for $k = 1, 2, \ldots$.

a. (10) Suppose the carrier sells 575 tickets. What, approximately, is the probability that not all passengers will get on the plane?

Solution: in mathematical terms the question is to approximate the probability $P(S_{575} \geq 526)$. *We have*

$$E(S_{575}) = 575 \cdot 0.9 \doteq 517.5 \quad \textit{and} \quad \mathrm{var}(S_{575}) = 575 \cdot 0.9 \cdot 0.1 \doteq 51.75\,.$$

We estimate using the central limit theorem that

$$\begin{aligned} P(S_{575} \geq 526) &= P\left(\frac{S_{575} - E(S_{575})}{\sqrt{\mathrm{var}(S_{575})}} \geq \frac{526 - E(S_{575})}{\sqrt{\mathrm{var}(S_{575})}}\right) \\ &= P\left(\frac{S_{575} - E(S_{575})}{\sqrt{\mathrm{var}(S_{575})}} \geq 1.18\right) \\ &\approx P(Z \geq 1.18) \\ &\doteq 0.12\,. \end{aligned}$$

b. (10) Suppose the carrier wants that the probability of overbooking is at most 0.05. What, approximately, is the maximum number of tickets the carrier can sell?

Solution: we are looking for the maximum n such that $P(S_n \geq 526) \leq 0.05$. *We approximate*

$$\begin{aligned} P(S_n \geq 526) &= P\left(\frac{S_n - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} \geq \frac{526 - n \cdot 0.9}{\sqrt{n} \cdot 0.3}\right) \\ &\approx P\left(Z \geq \frac{526 - n \cdot 0.9}{\sqrt{n} \cdot 0.3}\right)\,. \end{aligned}$$

If we want this last probability to be 0.05, we need

$$\frac{526 - n \cdot 0.9}{\sqrt{n} \cdot 0.3} = 1.64\,.$$

Solving for n we get $n \approx 571$.