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

University of Primorska
FAMNIT, Mathematics

Probability

Written examination

August 30$^{\text{th}}$, 2021

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) Eight positions are arranged in a circle. Every position is independently assigned the value 0 or 1 with probability $\frac{1}{2}$ respectively.

a. (15) Find the probability that no 5 contiguous positions are assigned the value 0.

   Solution: let X be the maximum number of contiguous positions that are assigned the value 0. We are looking for

$$P(X < 5) = 1 - P(X = 5) - P(X = 6) - P(X = 7) - P(X = 8)\,.$$

   There are 2^8 equally likely assignments of 0 and 1. Only one assignment corresponds to the event $\{X = 8\}$, 8 correspond to the event $\{X = 7\}$, and 8 to the event $\{X = 6\}$. For $\{X = 5\}$ we need to have 5 contiguous 0 flanked by 1 and the remaining one can be arbitrary. There are 16 assignments corresponding to $\{X = 5\}$. We have

$$P(X < 5) = 1 - \frac{1 + 8 + 8 + 16}{256} = \frac{223}{256} = 0.8710938\,.$$

b. (5) Let B be the event that we do not get 5 contiguous positions with 0s assigned, and let A be the event that we get at least 5 contiguous positions with 1s assigned. Find $P(A|B)$.

   Solution: we need to compute $P(A \cap B)$. Note that

$$P(A \cap B) = P(A) - P(A \cap B^c)\,.$$

   The last intersection is empty so $P(A \cap B) = P(A)$. It follows that

$$P(A|B) = \frac{P(A)}{P(B)} = \frac{33}{223} \doteq 0.148\,.$$

2. (20) Let $X \sim \mathrm{Bin}(n, 1/2)$. For random variables X and Y, suppose that

$$P(X = k, Y = k+1) = P(X = k) \cdot \frac{n-k}{n}\,,$$

and

$$P(X = k, Y = k-1) = P(X = k) \cdot \frac{k}{n}$$

for all $k = 0, 1, 2, \ldots, n$, and $P(X = k, Y = l) = 0$ for $|k - l| > 1$.

a. (10) Find the distribution of Y.

Solution: the distribution of Y is the marginal distribution of (X, Y). We have

$$P(Y = l) = P(X = l+1, Y = l) + P(X = l-1, Y = l)\,.$$

Compute

$$\begin{aligned}
P(Y = l) &= P(X = l+1, Y = l) + P(X = l-1, Y = l) \\
&= P(X = l+1) \cdot \frac{l+1}{n} + P(X = l-1) \cdot \frac{n-l+1}{n} \\
&= \binom{n}{l+1} \left(\frac{1}{2}\right)^n \frac{l+1}{n} + \binom{n}{l-1} \left(\frac{1}{2}\right)^n \frac{n-l+1}{n} \\
&= \binom{n-1}{l} \left(\frac{1}{2}\right)^n + \binom{n-1}{l-1} \left(\frac{1}{2}\right)^n \\
&= \binom{n}{l} \left(\frac{1}{2}\right)^n .
\end{aligned}$$

We used Pascal's identity and interpreted the symbol $\binom{n}{m}$ as 0 when $m > n$ or $m < 0$. Hence $Y \sim \mathrm{Bin}(n, 1/2)$.

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

Solution: we have that $E(X) = E(Y) = n/2$. For the covariance we need $E(XY)$. We compute

$$\begin{aligned}
E(XY) &= \sum_{k=0}^{n} \Big(k(k+1)\, P(X = k, Y = k+1) \\
&\qquad + k(k-1)\, P(X = k, Y = k-1) \Big) \\
&= \sum_{k=0}^{n} \left(k(k+1)\, P(X = k)\, \frac{n-k}{n} + k(k-1)\, P(X = k)\, \frac{k}{n} \right) \\
&= \sum_{k=0}^{n} \frac{k}{n}\, P(X = k) \cdot \big((k+1)(n-k) + (k-1)k \big) \\
&= \sum_{k=0}^{n} \frac{k}{n}\, P(X = k) \cdot \big((n-2)k - n \big) \\
&= \frac{1}{n} \big((n-2)\, E(X^2) - n\, E(X) \big)\,.
\end{aligned}$$

From $\mathrm{var}(X) = n/4$ *we obtain* $E(X^2) = (n^2 + n)/4$*, leading to*

$$\begin{aligned} E(XY) &= \frac{1}{n}\left((n-2)\,\frac{n^2+n}{4} - n\,\frac{n}{2}\right) \\ &= \frac{n^2 - 3n - 2}{4}\,. \end{aligned}$$

Hence

$$\mathrm{cov}(X, Y) = -\frac{3n}{4} - \frac{1}{2}\,.$$

3. (20) Let X, Y and Z be independent with $X, Y \sim \mathrm{N}(0,1)$ and $Z \sim \mathrm{N}\big(0, \frac{1}{2}\big)$.

a. (10) Find the density of $W = \sqrt{(X-Y)^2 + 4Z^2}$.

Hint: for a normal random variable $T \sim \mathrm{N}(0, \sigma^2)$ *we have*

$$T^2 \sim \Gamma\big(\tfrac{1}{2}, \tfrac{1}{2\sigma^2}\big)\,.$$

Solution: the difference $X - Y$ *is independent of* $2Z$*. Both random variables are normal* $\mathrm{N}(0,2)$ *so the squares are independent* $\Gamma\big(\frac{1}{2}, \frac{1}{4}\big)$ *random variables. The sum is* $\Gamma\big(1, \frac{1}{4}\big) = \exp\big(\frac{1}{4}\big)$*. It follows that*

$$P(W \geq w) = P\left(W^2 \geq w^2\right) = e^{-\frac{w^2}{4}}$$

and finally

$$f_W(w) = \frac{w}{2}\, e^{-\frac{w^2}{4}}$$

for all $w > 0$*; elsewhere, the density vanishes.*

b. (5) Show that $X+Y$, $X-Y$ and Z are independent.

Solution: it is enough to show that $X+Y$ *and* $X-Y$ *are independent. The joint density is of the form*

$$c \cdot \exp\left(\frac{-(x+y)^2}{8} - \frac{(x-y)^2}{8}\right)\,.$$

The mixed terms cancel and the density is a product. Independence follows.

c. (5) Let

$$U = \frac{X+Y+\sqrt{(X-Y)^2+4Z^2}}{2} \quad \text{and} \quad V = \frac{X+Y-\sqrt{(X-Y)^2+4Z^2}}{2}\,.$$

Find the density of (U, V). State explicitly where the density if different from 0.

Solution: we use the transformation formula. The support of the random variable $X+Y \sim \mathrm{N}(0,2)$ *is the whole real line, while the support of* W *is* $(0,\infty)$*. The map* $\Phi(t,w) := \big(\frac{t+w}{2}, \frac{t-w}{2}\big)$ *is a bijection from* $\mathbb{R} \times (0,\infty)$ *onto the set* $\{(u,v); u > v\}$*. On the latter set, the density of* (U,V) *will be different from* 0*. Observe that* $\Phi^{-1}(u,v) = (u+v, u-v)$ *and* $J_{\Phi^{-1}} = 2$*. Since* $X+Y$ *is independent of* W*, the transformation formula gives*

$$\begin{aligned} f_{U,V}(u,v) &= f_{X+Y}(u+v) \cdot f_W(u-v) \cdot 2 \\ &= 2 \cdot \frac{1}{2\sqrt{\pi}}\, e^{-\frac{(u+v)^2}{4}} \cdot \frac{(u-v)}{2}\, e^{-\frac{(u-v)^2}{4}} \end{aligned}$$

for all $u > v$*. The density simplifies to*

$$\frac{u-v}{2\sqrt{\pi}}\, e^{-\frac{1}{2}(u^2+v^2)}\,.$$

4. (20) Let $X_1, \dots, X_r$ be independent with $X_k \sim \mathrm{Po}(\lambda_k)$ for $k = 1, 2, \dots, r$. Denote $S_r = X_1 + X_2 + \cdots + X_r$.

a. (10) Compute $E(X_k^2 | S_r = n)$.

Solution: let

$$\lambda = \lambda_1 + \cdots + \lambda_r \quad \textit{and} \quad p_k = \frac{\lambda_k}{\lambda_1 + \cdots + \lambda_r},$$

and observe that $X_k \sim \mathrm{Po}(\lambda_k)$ *and* $S_r - X_k \sim \mathrm{Po}(\lambda - \lambda_k)$ *are independent. For* $0 \le i \le n$ *we have*

$$P(X_k = i | S_r = n) = \frac{P(X_k = i, S_r - X_k = n - i)}{P(S_r = n)} = \binom{n}{i} p_k^i (1 - p_k)^{n-i},$$

so that the conditional distribution of X_k *given* $\{S_r = n\}$ *is binomial* $\mathrm{Bin}(n, p_k)$. *As a result,*

$$E(X_k^2 | S_r = n) = np_k(1 - p_k) + n^2 p_k^2 = np_k + (n^2 - n)p_k^2 .$$

b. (10) Find $E(X_k X_l | S_r = n)$ for $k \neq l$.

Hint: try $E\big((X_k + X_l)^2 | S_r = n\big)$.

Solution: the random variables other than X_k *and* X_l *and the sum* $X_k + X_l$ *are independent Poisson random variables with sum* S_r. *From the first part we have*

$$E\big[(X_k + X_l)^2 \big| S_r = n\big] = n(p_k + p_l) + (n^2 - n)(p_k + p_l)^2 .$$

On the other hand, the conditional expectation is linear, so the above expectation equals

$$E(X_k^2 | S_r = n) + 2E(X_k X_l | S_r = n) + E(X_l^2 | S_r = n) .$$

Subtracting the outer two expectations that we know from the first part, we get

$$E(X_k X_l | S_r = n) = (n^2 - n) p_k p_l .$$

5. (20) Let $N, X_1, X_2, \dots$ be independent, non-negative integer valued random variables. Assume that $X_1, X_2, \dots$ are equally distributed. Let $X = X_1 + X_2 + \cdots + X_N$.

a. (10) Is it possible that $X_1 \sim \text{Bernoulli}(p_1)$ and $X \sim \text{Geom}(p_2)$ with $p_1, p_2 \in (0,1)$? Explain.

Solution: we would have

$$G_X(s) = G_N(G_1(s)) \, ,$$

which in the above case means

$$G_N(q_1 + p_1 s) = \frac{p_2 s}{1 - q_2 s}$$

where $q_1 = 1 - p_1$ *and* $q_2 = 1 - p_2$. *This implies*

$$G_N(u) = \frac{p_2\left(\frac{u-q_1}{p_1}\right)}{1 - q_2\left(\frac{u-q_1}{p_1}\right)} \, .$$

But

$$P(N = 0) = G_N(0) = -\frac{p_2 q_1}{p_1 + q_1 q_2} < 0 \, .$$

There is no such N.

b. (10) Find a sufficient and necessary condition on p_1 and p_2 to have $X_1 \sim \text{Bernoulli}(p_1)$ and $X \sim \text{Bin}(m, p_2)$ for some N independent of $X_1, X_2, \dots$ Under that condition find the distribution of N.

Solution: we would have

$$G_N(q_1 + p_1 s) = (q_2 + p_2 s)^m$$

or

$$G_N(u) = \left(q_2 + \frac{p_2(u - q_1)}{p_1}\right)^m \, .$$

Rearrange to get

$$G_N(u) = \left(q_2 - \frac{p_2 q_1}{p_1} + \frac{p_2}{p_1}\, u\right)^m = \left(1 - \frac{p_2}{p_1} + \frac{p_2}{p_1}\, u\right)^m \, .$$

By expanding the right side by the binomial formula we find the G_N *is a generating function if and only if* $p_2 \le p_1$. *In this case* $N \sim \text{Bin}(m, p_2/p_1)$.

6. (20) Consider a game of chance where the player loses 1€ with probability 50%, gains 9€ with probability 5%, and has no loss or gain with probability 45%. Felix plays this game 500 times. Assume that all the games are independent.

a. (10) What, approximately, is the probability that Felix wins 50€ or more in total?

Solution: let X_k be the payout in the k-the game, and let $S_n = X_1 + \cdots + X_n$. We have

$$E(X_1) = -0.05 \quad \textit{and} \quad \mathrm{var}(X_1) = 4.5475\,,$$

and furthermore

$$E(S_{500}) = 500\,E(X_1) = -25 \quad \textit{and} \quad \mathrm{var}(S_{500}) = 500\,\mathrm{var}(X_1) = 2273.75\,.$$

Using the continuity correction in the central limit theorem we get

$$\begin{aligned} P(S_{500} \geq 49.5) &= P\left(\frac{S_{500} - E(S_{500})}{\sqrt{\mathrm{var}(S_{500})}} \geq \frac{49.5 - E(S_{500})}{\sqrt{\mathrm{var}(S_{500})}}\right) \\ &= P\left(\frac{S_{500} - E(S_{500})}{\sqrt{\mathrm{var}(S_{500})}} \geq 1.56\right) \\ &\approx 1 - \Phi(1.56) \\ &\doteq 0.059\,. \end{aligned}$$

Precise result: 0.06354382.

b. (10) Assume that Felix indeed wins 50€ or more in total. Approximate the conditional probability that he wins 9€ in the first game.

Solution: if Felix wins 9€ in the first game he has to win 41€ or more in the remaining 499 games. The probability of the intersection is $0.05 \cdot P(S_{499} \geq 41)$. We compute

$$\begin{aligned} P(S_{499} \geq 40.5) &= P\left(\frac{S_{499} - E(S_{499})}{\sqrt{\mathrm{var}(S_{499})}} \geq \frac{40.5 - E(S_{499})}{\sqrt{\mathrm{var}(S_{499})}}\right) \\ &= P\left(\frac{S_{499} - E(S_{499})}{\sqrt{\mathrm{var}(S_{499})}} \geq 1.37\right) \\ &\approx 1 - \Phi(1.37) \\ &\doteq 0.085\,. \end{aligned}$$

Finally,

$$P(X_1 = 9 | S_{500} \geq 50) = \frac{0.05 \cdot P(S_{499} \geq 41)}{P(S_{500} \geq 50)} \doteq 0.072\,.$$

Precise result: 0.06951523.