University of Primorska

FAMNIT

Probability

Midterm 2

June 6$^{\text{st}}$, 2017

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					

SOLUTIONS

1. (20) A coin is tossed until either exactly r heads or exactly s tails appear. Denote the number of necessary tosses by $X_{r,s}$. Let $P(H) = 1/2$ and assume that the tosses are independent.

a. (10) Compute $E(X_{1,r})$.

Hint: use the fact that for a nonnegative integer valued random variable X

$$E(X) = \sum_{k=1}^{\infty} P(X \geq k)\,.$$

Another possible approach is to use

$$\sum_{k=1}^{n} ka^k = \frac{(an-n-1)a^{n+1}+a}{(1-a)^2}\,.$$

Solution: the event $\{X_{1,r} \geq k\}$ *for* $k = 1, 2, \ldots, r$ *means that on tosses* $1, 2, \ldots, k-1$ *only tails appear. Therefore* $P(X_{1,r} \geq k) = 2^{-(k-1)}$ *for* $k = 1, \ldots, r$ *and* $P(X_{1,r} \geq k) = 0$ *for* $k \geq r+1$. *It follows*

$$\begin{aligned} E(X_{1,r}) &= \sum_{k=1}^{r} P(X_{1,r} \geq k) \\ &= \sum_{k=1}^{r} 2^{-(k-1)} \\ &= 2(1-2^{-r})\,. \end{aligned}$$

b. (10) Show that

$$E(X_{r,r}) = 2r\left(1 - \binom{2r}{r}2^{-2r}\right)\,.$$

Hint: $\sum_{k=r+1}^{2r+1} P(X_{r+1,r+1} = k) = 1$.

Solution: using the negative binomial distribution we get

$$P(X_{r,r} = k) = 2\binom{k-1}{r-1}2^{-k}$$

for $k = r, r+1, \ldots, 2r-1$. *Therefore*

$$\begin{aligned} E(X_{r,r}) &= 2\sum_{k=r}^{2r-1} k \cdot \binom{k-1}{r-1} 2^{-k} \\ &= 2r \sum_{k=r}^{2r-1} \binom{k}{r} 2^{-k} \\ &= 2r \sum_{k=r+1}^{2r} \binom{k-1}{r} 2^{-(k-1)} \\ &= 2r \left[2 \sum_{k=r+1}^{2r+1} \binom{k-1}{r} 2^{-k} - \binom{2r}{r} 2^{-2r} \right] \\ &= 2r \left[1 - \binom{2r}{r} 2^{-2r} \right] . \end{aligned}$$

The first sum in brackets is equal to 1 *as it is the sum of probabilities in the distribution of* $X_{r+1,r+1}$

2. (20) Each of the two players A and B has an ordinary deck of 52 cards. Both players shuffle their decks well and independently of each other. They both start placing cards on the table one by one face up simultaneously from the top of their respective decks.

a. (10) Let X be the number of times when the two players simultaneously place an Ace on the table. Compute $E(X)$.

Solutions: Define the indicators

$$I_k = \begin{cases} 1 & \text{in the } k\text{-th round both players place an Ace on the table,} \\ 0 & \text{otherwise.} \end{cases}$$

With this definition we have $X = I_1 + \cdots + I_{52}$*. By symmetry, all indicators have equal expectation so*

$$E(X) = 52 \cdot E(I_1) = 52 \cdot P(I_1 = 1) = 52 \cdot \left(\frac{4}{52}\right)^2 .$$

b. (10) Let Y be the number of Aces that are still in the deck of the player B immediately after the first time the player A places an Ace on the table. Let N be the number of cards placed on the table until the player A places an Ace on the table. Assume as known that $E(N) = \frac{53}{5}$. Compute $E(Y)$.

Hint: for a nonnegative integer valued random variable N *we have* $E(N) = \sum_{n \geq 1} P(N \geq n)$*.*

Solution: Define indicators

$$J_k = \begin{cases} 1 & \text{the player B places an Ace on the table in the } k\text{-th and } N \geq k \\ 0 & \text{otherwise.} \end{cases}$$

The sum $J_1 + \cdots + J_{49}$ *is equal to the number of Aces that player B places on the table, including the moment when the player A places an Ace on the table. It follows that* $Y = 4 - J_1 + \cdots + J_{49}$*. By independence and symmetry it follows that*

$$P(J_k = 1) = \frac{4}{52} P(N \geq k) .$$

It follows

$$E\left(J_1 + \cdots + J_{49}\right) = \frac{4}{52} \sum_{k=1}^{49} P(N \geq k) .$$

For every nonnegative integer valued random variable N *we have*

$$E(N) = \sum_{n \geq 1} P(N \geq n) .$$

The random variable N *can take values* $1, 2, \ldots, 49$. *From the above formula it follows that*

$$\sum_{k=1}^{49} P(N \geq k) = E(N) = \frac{53}{5} .$$

We get

$$E\left(J_1 + \cdots + J_{49}\right) = \frac{4}{52} \cdot \frac{53}{5} .$$

The final result is equal to

$$E(Y) = 4 - E\left(J_1 + \cdots + J_{49}\right) = \frac{207}{65} .$$

3. (20) Let random variables X and Y be independent and equally distributed with

$$P(X=k)=\frac{\binom{2k}{k}\beta^k}{4^k(1+\beta)^{k+\frac{1}{2}}}$$

for $k=0,1,\ldots$ and $\beta>0$.

a. (10) Compute the distribution of $X+Y$.

Hint: check that

$$\binom{2k}{k}=\frac{4^k\left(\frac{1}{2}\right)_k}{k!}\,,$$

where $(a)_k$ *is the Pochhammer symbol defined by:*

$$(a)_k=a(a+1)(a+2)\cdots(a+k-1)\,,\quad (a)_0=1\,.$$

Assume as known that

$$\sum_{k=0}^{n}\binom{n}{k}(a)_k\,(b)_{n-k}=(a+b)_n\ .$$

Solution: for $n=0,1,2,\ldots$ *we compute*

$$\begin{aligned}
P(X+Y=n) &= \sum_{k=0}^{n}P(X=k)P(Y=n-k)\\
&= \frac{\beta^n}{4^n(1+\beta)^{n+1}}\sum_{k=0}^{n}\binom{2k}{k}\binom{2n-2k}{n-k}\\
&= \frac{\beta^n}{4^n(1+\beta)^{n+1}}\,4^n\sum_{k=0}^{n}\frac{1}{k!(n-k)!}\left(\frac{1}{2}\right)_k\left(\frac{1}{2}\right)_{n-k}\\
&= \frac{\beta^n}{n!(1+\beta)^{n+1}}\sum_{k=0}^{n}\binom{n}{k}\left(\frac{1}{2}\right)_k\left(\frac{1}{2}\right)_{n-k}\\
&= \frac{\beta^n}{n!(1+\beta)^{n+1}}\,(1)_n\\
&= \frac{\beta^n}{(1+\beta)^{n+1}}\,.
\end{aligned}$$

b. (10) Justify that $X+Y+1\sim\mathrm{Geom}\left(\frac{1}{1+\beta}\right)$ and use this result to compute $\mathrm{var}(X)$.

Solution: the claim about the distribution can be checked directly. We know that the variance of $Z \sim \mathrm{Geom}(p)$ *is equal to*

$$\mathrm{var}(Z) = \frac{1-p}{p^2}\,.$$

Therefore

$$\mathrm{var}(X+Y) = \mathrm{var}(X+Y+1) = \beta(1+\beta)\,.$$

We have $\mathrm{var}(X) = \mathrm{var}(Y)$ *and*

$$\mathrm{var}(X+Y) = \mathrm{var}(X) + \mathrm{var}(Y)\,,$$

and hence $\mathrm{var}(X) = \dfrac{\beta(1+\beta)}{2}$.

4. (20) Let the random variables X and Y have joint density

$$f_{X,Y}(x,y) = \begin{cases} 2 & \text{for } x, y > 0 \text{ and } x + y < 1 \\ 0 & \text{else.} \end{cases}$$

Assume as known that the mapping Φ given by

$$\Phi(x,y) = \left(\frac{x}{x+y}, 1 - x - y \right)$$

is a bijection of $G = \{(x,y)\colon x, y > 0 \text{ in } x + y < 1\}$ to the square $H = (0,1)^2$.

a. (10) Compute the density of the vector

$$(U,V) = \left(\frac{X}{X+Y}, 1 - X - Y \right).$$

Justify that U and V are independent.

Solution: first we need the inverse $\Phi^{-1}(u,v)$, which means that we need to solve the system of equations

$$\frac{x}{x+y} = u, \qquad 1 - x - y = v$$

for $(u,v) \in (0,1)^2$. We get

$$x = x = u(1-v) \quad \textit{and} \quad y = (1-u)(1-v).$$

It follows

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

By the transformation formula we get

$$f_{U,V}(u,v) = \begin{cases} 2(1-v) & \textit{for } (u,v) \in (0,1)^2 \\ 0 & \textit{otherwise.} \end{cases}$$

The density is product of terms that are dependent of u (this is 1) and v respectively, therefore U and V are independent random variables.

b. (10) Show that the random variables X, Y and $1-X-Y$ have the same density.

Solution: for X and Y we compute the marginal density to get

$$f_X(x) = 2(1-x) \quad \textit{and} \quad f_Y(y) = 2(1-y).$$

From the first part it follows that the density of $V = 1 - X - Y$ is equal to

$$f_V(v) = 2(1-v).$$

5. (20) Suppose $m > 1$ begonias and $n > 1$ fuchsias are randomly arranged on a window sill. Every arrangement is equally likely. Denote by X the number of begonias that have a begonia immediately to their right.

a. (10) Express X as a sum of indicators and compute their covariances.

Solution: denote the positions of flowers on the sill from left to right by $k = 1, 2, \ldots, m+n$. For $k = 1, 2, m+n-1$ define

$$I_k = \begin{cases} 1 & \textit{if positions } k \textit{ and } k+1 \textit{ are are occupied by begonias} \\ 0 & \textit{otherwise.} \end{cases}$$

We have $X = I + 1 + \cdots + I_{m+n-1}$.

We will need the probabilities $P(I_k = 1)$, that are the probabilities that two begonias are in the positions k and $k+1$. All the arrangements are equally likely, in the case with no restrictions two randomly chosen flowers are placed in the two positions. Therefore,

$$P(I_k = 1) = \frac{\binom{m}{2}}{\binom{m+n}{2}}$$

Assume $k \leq m+n-2$ and $l = k+1$. The probability $P(I_k = 1, I_l = 1)$ is equal to the probability that there are begonias in positions $k, k+1, k+2$. Since this is a randomly chosen triple, we get

$$P(I_k = 1, I_l = 1) = \frac{\binom{m}{3}}{\binom{m+n}{3}}\,.$$

In this case

$$\operatorname{cov}(I_k, I_l) = \frac{\binom{m}{3}}{\binom{m+n}{3}} - \frac{\binom{m}{2}^2}{\binom{m+n}{2}^2}\,.$$

If $l \geq k+2$, following a similar approach as above, we get

$$\operatorname{cov}(I_k, I_l) = \frac{\binom{m}{4}}{\binom{m+n}{4}} - \frac{\binom{m}{2}^2}{\binom{m+n}{2}^2}\,.$$

b. (10) Express $\operatorname{var}(X)$ by $\operatorname{var}(I_1)$, $\operatorname{cov}(I_1, I_2)$ and $\operatorname{cov}(I_1, I_3)$, where $I_1, \ldots, I_{m+n-1}$ are the indicators defined above. You do not have to simplify the results.

Solution: we use the usual formula

$$\operatorname{var}(X) = \sum_{k=1}^{m+n-1} \operatorname{var}(I_k) + 2 \sum_{1 \leq k < l \leq m+n-1} \operatorname{cov}(I_k, I_l)\,.$$

All the variances are equal to $P(I_k = 1)(1 - P(I_k = 1))$. *There are* $m + n - 1$ *such variances. We need to count how many different covariances there are. We write*

$$\begin{aligned}
&\sum_{1 \le k < l \le m+n-1} \operatorname{cov}(I_k, I_l) \\
&= \sum_{k=1}^{m+n-2} \sum_{l=k+1}^{m+n-1} \operatorname{cov}(I_k, I_l) \\
&= \sum_{k=1}^{m+n-2} \left[\operatorname{cov}(I_k, I_{k+1}) + (m+n-k-1)\operatorname{cov}(I_k, I_{k+2})\right] \\
&= (m+n-2)\operatorname{cov}(I_1, I_2) + \operatorname{cov}(I_1, I_3) \cdot \sum_{k=1}^{m+n-2} (m+n-k-1) \\
&= (m+n-2)\operatorname{cov}(I_1, I_2) + \operatorname{cov}(I_1, I_3) \cdot (1 + 2 + \cdots + (m+n-2)) \\
&= (m+n-2)\operatorname{cov}(I_1, I_2) + \operatorname{cov}(I_1, I_3) \cdot \frac{(m+n-2)(m+n-1)}{2}\,.
\end{aligned}$$

It follows that

$$\begin{aligned}
\operatorname{var}(X) &= (m+n-1)\operatorname{var}(I_1) + 2(m+n-2)\operatorname{cov}(I_1, I_2) \\
&\quad + (m+n-2)(m+n-1)\operatorname{cov}(I_1, I_3)\,.
\end{aligned}$$

6. (20) A permutation π is selected at random out of the set of all permutations of n elements. Every permutation is equally likely to be chosen. The pair (i, j) with $i < j$ is an *inversion* if $\pi(i) > \pi(j)$. Let X be the number of inversions in the random permutation Π.

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

Solution: let $S = \{(i, j)\colon 1 \leq i < j \leq n\}$ be a set of all ordered pairs of numbers smaller than n. Denote by I_s the indicator of the event where the pair $s \in S$ is an inversion. By symmetry, we get $P(I_s = 1) = 1/2$. Since $X = \sum_{s \in S} I_s$, it follows that

$$E(X) = \sum_{s \in S} E(I_s) = \frac{n(n-1)}{4},$$

as there are $n(n-1)/2$ pairs.

b. (10) Compute $\mathrm{var}(X)$.

Solution: we have

$$\mathrm{var}(X) = \sum_{s \in S} \mathrm{var}(I_s) + \sum_{\substack{s,t \in S \\ s \neq t}} \mathrm{cov}(I_s, I_t).$$

Denote $s = (i, j)$ and $t = (k, l)$. For covariances we get several possibilities: the first is $\{i, j\} \cap \{k, l\} = \emptyset$. In this case by symmetry $P(I_s = 1, I_t = 1) = 1/4$ and consequently $\mathrm{cov}(I_s, I_t) = 0$. The second possibility is $i = k < j < l$. By symmetry in this case we have $P(I_s = 1, I_t = 1) = 1/3$ and consequently $\mathrm{cov}(I_s, I_t) = 1/12$. The third possibility is $i < j = k < l$. By symmetry we have $P(I_s = 1, I_t = 1) = 1/6$ and consequently $\mathrm{cov}(I_s, I_t) = -1/12$. The fourth option is $i < k < j = l$ with the same covariance as in the second case.

We are left with counting the terms in the sum of the covariances. The possible three nonzero covariances we appear $2\binom{n}{3}$ times. It follows

$$\mathrm{var}(X) = \frac{n(n-1)}{8} + 2\binom{n}{3} \cdot \frac{1}{12} = \frac{n(n-1)(2n+5)}{72}.$$