University of Primorska

FAMNIT, Mathematics

Probability

Midterm 2

May 29$^{\text{th}}$, 2019

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) There are 4 aces, 4 kings and 4 queens in a deck of 52 cards. The deck is shuffled well and cards are dealt from the top of the deck one card at a time until the first ace is dealt. Let X be the number of kings and Y the number of queens dealt before the first ace.

a. (10) Determine the distribution of the random vector (X, Y).

Hint: do you have to look at all 52 cards?

Solution: The random vector (X, Y) takes values (k, l) where $0 \leq k, l \leq 4$. We are interested in the relative position of aces, kings and queens. If all other cards are excluded from the deck, we are left with a random permutation of 12 cards. To compute the probabilities $P(X = k, Y = l)$, we have to count all permutations, where k kings and l queens are dealt before the first ace. This happens if we select k kings in $\binom{4}{k}$ ways and l queens in $\binom{4}{l}$ ways. These $k+l$ can be randomly permuted, after which one of four aces is selected. The remaining $12-(k+l+1)$ are randomly permuted. It follows

$$
\begin{aligned}
P(X = k, Y = l) &= \frac{4\binom{4}{k}\binom{4}{l}\,(k+l)!\,(11-k-l)!}{12!} \\
&= \frac{4 \cdot (4!)^2\,(k+l)!\,(11-k-l)!}{12!\,k!\,l!\,(4-k)!\,(4-l)!}\,.
\end{aligned}
$$

b. (5) Are X and Y independent?

Solution: for independence for all pairs $k, l \in \{0, 1, 2, 3, 4\}$ (X, Y) we need to have

$$P(X = k, Y = l) = f(k)\,g(l)$$

for some functions $f, g\colon \{0, 1, 2, 3, 4\} \to \mathbb{R}$. This implies for all pairs $k, l \in \{0, 1, 2, 3, 4\}$ the equality

$$H(k, l) := (k+l)!\,(11-k-l)! = \tilde{f}(k)\,\tilde{g}(l)\,,$$

which for example means that the matrix

$$\begin{bmatrix} H(0,0) & H(0,1) \\ H(1,0) & H(1,1) \end{bmatrix} = \begin{bmatrix} 11! & 10! \\ 10! & 2 \cdot 9! \end{bmatrix} = 9! \begin{bmatrix} 110 & 10 \\ 10 & 2 \end{bmatrix}$$

should be degenerate but this is not true. It follows that X and Y are dependent.

c. (5) Determine the distribution of $X + Y$.

Solution: we use the same procedure as in a. We merge kings and queens. There are again $12!$ *permutation of* 12 *cards, and for* $X+Y=n$ *to happen there are*

$$\binom{8}{n}\cdot n!\cdot 4\cdot(12-n-1)!$$

possibilities. For $n=0,1,\ldots,8$ *we have*

$$P(X+Y=n)=\frac{4\binom{8}{n}\,n!\,(11-n)!}{12!}=\frac{(11-n)(10-n)(9-n)}{2970}\,.$$

2. (20) In the *Poker test* used for testing random number generators, the following problem appears: we have independent, equally distributed random variables $\xi_1, \xi_2, \xi_3, \xi_4, \xi_5$ uniformly distributed on the set $\{0, 1, \ldots, m-1\}$ for a given $m > 0$. Let X be the number of distinct numbers in the set $\{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5\}$. Example: for the set $\{1, 2, 5, 2, 5\}$ we have $X = 3$.

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

Hint: express X with the indicators

$$I_k = \begin{cases} 1 & \textit{if the number } k \textit{ appears in the set } \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5\} \\ 0 & \textit{otherwise.} \end{cases}$$

Solution: for $k = 0, 1, \ldots, m-1$ we define indicators

$$I_k = \begin{cases} 1 & \textit{if the number } k \textit{ appears in the set } \{\xi_1, \xi_2, \xi_3, \xi_4, \xi_5\} \\ 0 & \textit{otherwise.} \end{cases}$$

We can write $X = I_0 + \cdots + I_{m-1}$. All indicators have the same distribution so we can compute

$$P(I_k = 0) = \left(\frac{m-1}{m}\right)^5, \qquad P(I_k = 1) = 1 - \left(\frac{m-1}{m}\right)^5.$$

It follows

$$E(X) = m\left[1 - \left(\frac{m-1}{m}\right)^5\right].$$

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

Hint: use $P(I_k = 1, I_l = 1) = 1 - P\big(\{I_k = 0\} \cup \{I_l = 0\}\big)$.

Solution: First approach: *we use* $\mathrm{var}(X) = E(X^2) - \big(E(X)\big)^2$ *and we compute*

$$\begin{aligned} E(X^2) &= \sum_{k=0}^{m-1} E(I_k^2) + \sum_{\substack{0 \le k,l < m \\ k \ne l}} E(I_k I_l) \\ &= \sum_{k=0}^{m-1} P(I_k = 1) + \sum_{\substack{0 \le k,l < m \\ k \ne l}} P(I_k = 1, I_l = 1). \end{aligned}$$

We use the hint in the second term:

$$\begin{aligned} P(I_k = 1, I_l = 1) &= 1 - P\big(\{I_k = 0\} \cup \{I_l = 0\}\big) \\ &= 1 - P(I_k = 0) - P(I_l = 0) + P(I_k = 0, I_l = 0)\,. \end{aligned}$$

Similarly as in a. for $k \neq l$ we can compute

$$P(I_k = 0, I_l = 0) = \left(\frac{m-2}{m}\right)^5 .$$

It follows

$$P(I_k = 1, I_l = 1) = 1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-2}{m}\right)^5 .$$

We combine the results and we get

$$\begin{aligned} \operatorname{var}(X) &= m\left[1 - \left(\frac{m-1}{m}\right)^5\right] + m(m-1)\left[1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-2}{m}\right)^5\right] \\ &\qquad - m^2\left[1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-1}{m}\right)^{10}\right] \\ &= m\left(\frac{m-1}{m}\right)^5 + m(m-1)\left(\frac{m-2}{m}\right)^5 - m^2\left(\frac{m-1}{m}\right)^{10} . \end{aligned}$$

Second approach: *we can write the variance as a sum of variances and covariances:*

$$\operatorname{var}(X) = \sum_{k=0}^{m-1} \operatorname{var}(I_k) + \sum_{\substack{0 \le k,l < m \\ k \neq l}} \operatorname{cov}(I_k, I_l)\,.$$

By symmetry, all variances are the same and also all covariances are the same. We have

$$\operatorname{var}(I_k) = P(I_k = 0)\,P(I_k = 1) = \left(\frac{m-1}{m}\right)^5 \left[1 - \left(\frac{m-1}{m}\right)^5\right]$$

and

$$\operatorname{cov}(I_k, I_l) = P(I_k = 1, I_l = 1) - P(I_k = 1)\,P(I_l = 1)\,.$$

We compute for $k \neq l$

$$P(I_k = 1, I_l = 1) = 1 - 2\left(\frac{m-1}{m}\right)^5 + \left(\frac{m-2}{m}\right)^5 ,$$

and hence

$$\begin{aligned}\operatorname{cov}(I_k, I_l) &= P(I_k = 1, I_l = 1) - P(I_k = 1)\, P(I_l = 1) \\ &= \left(\frac{m-2}{m}\right)^5 - \left(\frac{m-1}{m}\right)^{10} .\end{aligned}$$

We combine the results to get

$$\begin{aligned}\operatorname{var}(X) = m\left(\frac{m-1}{m}\right)^5 &\left[1 - \left(\frac{m-1}{m}\right)^5\right] \\ &+ m(m-1)\left[\left(\frac{m-2}{m}\right)^5 - \left(\frac{m-1}{m}\right)^{10}\right] \\ = m\left(\frac{m-1}{m}\right)^5 &+ m(m-1)\left(\frac{m-2}{m}\right)^5 - m^2\left(\frac{m-1}{m}\right)^{10} .\end{aligned}$$

3. (20) Let U and V be independent and $U, V \sim \exp(1)$.

a. (10) Let $X = U - V$. Compute the density of X.

Solution: let

$$\Phi(u, v) = (u - v, v)\,.$$

From

$$\Phi^{-1}(x, v) = (x + v, v)$$

follows that

$$J_{\Phi^{-1}}(u, x) = 1\,.$$

Therefore

$$f_{X,V}(x, v) = f_U(x + v) f_V(v)\,.$$

By symmetry, $f_X(x)$ is an even function. We can assume $x \geq 0$ and compute

$$\begin{aligned} f_X(x) &= \int_0^\infty e^{-x-v} e^{-v} dv \\ &= e^{-x} \int_0^\infty e^{-2v} dv \\ &= \frac{1}{2} e^{-x}\,. \end{aligned}$$

It follows

$$f_X(x) = \frac{1}{2} e^{-|x|}$$

for $x \in \mathbb{R}$.

b. (10) Let X and Y be independent and let both have the same distribution as $U - V$. Compute the density of $Z = X - Y$.

Solution: Let $U, V, \tilde{U}, \tilde{V} \sim \exp(1)$ be equally distributed and independent. Let $X = U - V$ in $Y = \tilde{U} - \tilde{V}$. We can write

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

We know that $U + \tilde{V} \sim \Gamma(2, 1)$ and the same holds for the sum in the second parenthesis. By symmetry, the density $f_Z(z)$ will be an even function. We *compute for $z \geq 0$*

$$\begin{aligned} f_Z(z) &= \int_0^\infty (z+x)e^{-z-x} x e^{-x} dx \\ &= z e^{-z} \int_0^\infty x e^{-2x} dx + e^{-z} \int_0^\infty x^2 e^{-2x} dx \\ &= \frac{1}{4} z e^{-z} + \frac{1}{4} e^{-z} \\ &= \frac{1}{4}(z+1)e^{-z} . \end{aligned}$$

Finally, it follows

$$f_Z(z) = \frac{1}{4}(|z|+1)e^{-|z|}$$

for $z \in \mathbb{R}$.

4. (20) Assume that subsequent coin tosses are independent and that the probability of heads is equal to p. Let $q := 1 - p$. Let W_n be the number of tosses it takes to get n consecutive heads including the last n tosses.

a. (10) Justify that

$$P\left(W_n = k+1 | W_{n-1} = k\right) = p$$

and for $l > k+1$

$$P(W_n = l | W_{n-1} = k) = qP(W_n = l - k - 1)\,.$$

Solution: if on the k-th toss we get $n-1$ consecutive heads for the first time, two things can happen: we get heads on the next toss and therefore $W_n = k+1$; or we get tails on the next toss and the "waiting" for n consecutive heads starts all over. The above equations capture this in mathematical notation.

b. (10) Show that

$$E\left(W_n | W_{n-1} = k\right) = k + 1 + qE(W_n)$$

and compute $E(W_n)$.

Solution: from the conditional probabilities we get

$$\begin{aligned} E\left(W_n | W_{n-1} = k\right) &= p(k+1) + \sum_{l=k+1+n}^{\infty} qlP(W_n = l-k-1) \\ &= p(k+1) + q\sum_{m=n}^{\infty}(m+k+1)P(W_n = m) \\ &= p(k+1) + q(k+1) + qE(W_n) \\ &= k+1+qE(W_n)\,. \end{aligned}$$

Multiply both sides of the equation by $P(W_{n-1} = k)$ and add up. It follows

$$\begin{aligned} E(W_n) &= \sum_{k=n-1}^{\infty} E(W_n | W_{n-1} = k) P(W_{n-1} = k) \\ &= \sum_{k=n-1}^{\infty} \left(k+1+qE(W_n)\right) P(W_{n-1} = k) \\ &= E(W_{n-1}) + 1 + qE(W_n)\,. \end{aligned}$$

Hence

$$E(W_n) = \frac{1}{p} + \frac{1}{p}E(W_{n-1})\,.$$

Since $E(W_1) = p^{-1}$ *we have*

$$E(W_n) = \sum_{k=1}^{n} \frac{1}{p^k} = \frac{1}{p^n} \, \frac{1-p^n}{1-p}\,.$$

5. (20) In a branching process, every individual has two descendants with probability 1/4, and no descendants with probability 3/4.

a. (5) Prove that this process dies out with probability 1.

Solution: the generating function of the number of descendants is equal to $G(s) = \frac{3}{4} + \frac{1}{4}s^2$ *and the equation* $G(s) = s$ *has two solutions,* $s = 1$ *in* $s = 3$*. The value* 1*is the only fixed point of the function* G *in the interval* $[0, 1]$*. Therefore this is the probability that the process dies out.*

b. (5) Let Z_1 be the number of descendants in the first generation. Let N be the number of *all* individuals in the process and let H be the generating function of N. For $k \in \{0, 2\}$ express $E(s^N \mid Z_1 = k)$ by s and $H(s)$.

Solution: if $Z_1 = 0$*, we have* $N = 1$*, therefore* $E(s^N \mid Z_1 = 0) = s$*. Conditionally on* $Z_1 = 2$*, the random variable* N *has the same distribution as the number of all individuals of two independent processes, that are distributed as* N *process, increased by 1. Therefore* $E(s^N \mid Z_1 = 2) = s\big(H(s)\big)^2$*.*

c. (10) Compute the distribution of the number of all individuals in the given process, i. e. for every $n \ge 1$ compute $P(N = n)$.

Hint: $(1 + x)^m = \sum_{k=0}^{\infty} \binom{m}{k} x^k$*.*

Solution: using the total expectation formula we get

$$\begin{aligned} H(s) &= E(s^N) \\ &= P(Z_1 = 0)\, E\big(s^N \mid Z_1 = 0\big) + P(Z_1 = 2)\, E\big(s^N \mid Z_1 = 2\big) \\ &= \frac{s}{4}\left[3 + \big(H(s)\big)^2\right]. \end{aligned}$$

We solve the equation for $H(s)$ *and get*

$$H(s) = \frac{2}{s}\left(1 \pm \sqrt{1 - \tfrac{3}{4}s^2}\right),$$

which is generating function only in the case when we take negative square root. Therefore

$$H(s) = 2\sum_{k=1}^{\infty} (-1)^{k-1} \binom{1/2}{k} \frac{3^k}{4^k}\, s^{2k-1}.$$

The random variable N *takes values* $n = 2k - 1$ *and it follows:*

$$P(N = n) = 2(-1)^{k-1}\binom{1/2}{k}\frac{3^k}{4^k} = -2\,\frac{3^k}{4^k}\,\frac{\left(-\frac{1}{2}\right)_k}{k!} = \frac{3^k}{4^k}\,\frac{\left(\frac{1}{2}\right)_{k-1}}{k!} = 2\,\frac{3^k}{4^k}\,\frac{(n-2)!!}{(n+1)!!},$$

where $(x)_r := x(x+1)(x+2)\cdots(x+r-1)$ *is the Pochhammer simbol.*

6. (20) There are B magical white and R magical red balls in an urn. We randomly select n balls from the urn, where every selection is equally likely. As the balls are magical, just after the selection and before we look at them the balls change their colour to the other with probability $1/4$ independently of one another and independently of the selection procedure. Let M_1 be the number of white balls selected initially and M_2 the number of red balls selected initially. Let N_1 be the number of white balls after the balls change colour and N_2 the number of red balls after the balls change colour.

a. (5) Compute $E(N_1|M_1 = m_1, M_2 = m_2)$.

Solution: white balls randomly change colour. If there are m_1 white balls the expected number of white balls after changing colour is $3m_1/4$. Out of m_2 black balls we can expect $m_2/4$ white balls after they change their colour. It follows

$$E(N_1|M_1 = m_1, M_2 = m_2) = \frac{3m_1}{4} + \frac{m_2}{4}\,.$$

b. (5) Compute $E(N_1N_2|M_1 = m_1, M_2 = m_2)$.

Hint: what is $\mathrm{cov}(N_1, N_2|M_1 = m_1, M_2 = m_2)$.

Solution: since magic balls are changing colours independently od one another and independently of the selection procedure, we have

$$\begin{aligned}&\mathrm{cov}(N_1, N_2|M_1 = m_1, M_2 = m_2)\\ &= E(N_1N_2|M_1 = m_1, M_2 = m_2)\\ &\qquad -E(N_1|M_1 = m_1, M_2 = m_2)E(N_2|M_1 = m_1, M_2 = m_2)\\ &= 0\,.\end{aligned}$$

It follows

$$E(N_1N_2|M_1 = m_1, M_2 = m_2) = \left(\frac{3m_1}{4} + \frac{m_2}{4}\right)\left(\frac{m_1}{4} + \frac{3m_2}{4}\right)\,.$$

c. (10) Show that

$$\mathrm{cov}(N_1, N_2) = -\frac{\mathrm{var}(M_1)}{4}\,.$$

Solution: we have $M_1 \sim \text{HiperGeom}(n, B, N)$*.*

$$\begin{aligned} E(N_1) &= \sum_{\substack{k,l \\ k+l=n}} E(N_1 | M_1 = m_1, M_2 = m_2) P(M_1 = m_1, M_2 = m_2) \\ &= \sum_{\substack{k,l \\ k+l=n}} \left(\frac{3m_1}{4} + \frac{m_2}{4} \right) P(M_1 = m_1, M_2 = m_2) \\ &= \frac{3}{4} E(M_1) + \frac{1}{4} E(M_2) \\ &= \frac{3nB}{4N} + \frac{nR}{4N} \\ &= \frac{n(3B+R)}{4N} \end{aligned}$$

On the other hand, if we consider $M_1 + M_2 = n$*, it holds*

$$\text{var}(M_1) = \text{var}(M_2) \qquad \textit{and} \qquad \text{cov}(M_1, M_2) = -\text{var}(M_1)\,.$$

We compute

$$\begin{aligned} &E(N_1 N_2) \\ &\quad= \sum_{\substack{k,l \\ k+l=n}} E(N_1 N_2 | M_1 = m_1, M_2 = m_2) P(M_1 = m_1, M_2 = m_2) \\ &\quad= \sum_{\substack{k,l \\ k+l=n}} \left(\frac{3m_1}{4} + \frac{m_2}{4} \right) \left(\frac{m_1}{4} + \frac{3m_2}{4} \right) P(M_1 = m_1, M_2 = m_2) \\ &\quad= \frac{1}{16} E\left(3M_1^2 + 3M_2^2 + 10 M_1 M_2\right) \\ &\quad= \frac{1}{16} \left(6\text{var}(M_1) + 3E(M_1)^2 + 3E(M_2)^2 - 10\text{var}(M_1) + 10 E(M_1) E(M_2)\right) \\ &\quad= -\frac{1}{4}\text{var}(M_1) + \frac{3n^2B^2}{16N^2} + \frac{3n^2R^2}{16N^2} + \frac{10n^2BR}{16N^2} \\ &\quad= -\frac{1}{4}\text{var}(M_1) + \frac{n^2(3B+R)(B+3R)}{16N^2} \\ &\quad= -\frac{1}{4}\text{var}(M_1) + E(N_1)E(N_2)\,. \end{aligned}$$

The equality follows.