University of Primorska

FAMNIT, Mathematics

Probability

Exam

June 17^{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					

1. (20) The Smiths and the Joneses leave for vacation the same week (from Monday to Sunday). There are 10 tourist destinations on the island where both families are staying. Each day both families select a destination at random among the destinations they have not yet visited. The choices of the two families are independent.

a. (5) What is the probability that the Smiths and the Joneses meet on Monday, Tuesday and Wednesday? On other days they may or may not meet.

Solution: Let (i, j, k) be a triple of different destinations. The probability that both families will choose these destinations on Monday, Tuesday and Wednesday is by the assumption of independence

$$\left(\frac{1}{10} \cdot \frac{1}{9} \cdot \frac{1}{8}\right)^2 .$$

To event we are interested in is a disjoint union over all unordered triples of destinations of which there are $10 \cdot 9 \cdot 8$. The final probability is $1/720$.

b. (15) What is the probability that the two families meet *exactly* three times during the entire week?

Hint: first compute the probability that the two families meet on Monday, Tuesday and Wednesday and do not meet on other days. Use the inclusion-exclusion formula.

Solution: Let B_{123} be the event that families meet on days 1, 2 and 3 but not on other days. We have

$$B_{123} := (A_1 \cap A_2 \cap A_3) \setminus (A_4 \cup A_5 \cup A_6 \cup A_7) ,$$

where A_i is the event that families meet on day i, $i = 1, 2, \ldots, 7$. The inclusion-exclusion principle gives

$$\begin{aligned}
P(B_{123}) &= P(A_1 \cap A_2 \cap A_3) - P\left(\bigcup_{k=4}^{7} (A_1 \cap A_2 \cap A_3 \cap A_k)\right) \\
&= P(A_1 \cap A_2 \cap A_3) - 4\,P(A_1 \cap A_2 \cap A_3 \cap A_4) \\
&\qquad + 6\,P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) \\
&\qquad - 4\,P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6) \\
&\qquad + P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5 \cap A_6 \cap A_7) \\
&= \frac{1}{10 \cdot 9 \cdot 8} - \frac{4}{10 \cdot 9 \cdot 8 \cdot 7} + \frac{6}{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6} - \frac{4}{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5} \\
&\qquad + \frac{1}{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5 \cdot 4} \\
&= \frac{31}{40320} .
\end{aligned}$$

The probability we need to find is

$$\binom{7}{3} P(B_{123}) = 35\, P(B_{123}) = \frac{31}{1152} \doteq 0{,}0269\,.$$

2. (20) An urn contains $B \geq 2$ white and R red balls. We randomly select balls from the urn one by one without replacement. Let X be the number of balls until we select the first white ball including the white ball, and let Y be the number of balls until the second white ball including the second white ball.

a. (10) Find the *joint* distribution of the random variables X and Y.

Solution: The possible values for X and Y are integer pairs (k, l) with $1 \leq k < l \leq R+2$. The event $\{X = k, Y = l\}$ happens if we first get $k-1$ red balls,a white ball, then $l-k-1$ red balls and a white ball. Denote by $N = B + R$ the total number of balls. The probability is calculated as

$$
\begin{aligned}
P(X = k, Y = l) &= \frac{R}{N} \cdot \frac{R-1}{N-1} \cdots \frac{R-k+2}{N-k+2} \cdot \frac{B}{N-k+1} \cdot \\
&\qquad \cdot \frac{R-k+1}{N-k} \cdots \frac{R-l+3}{N-l+2} \cdot \frac{B-1}{N-l+1} \\
&= \frac{B(B-1)\, R!\, (N-l)!}{(R-l+2)!\, N!},
\end{aligned}
$$

or

$$
P(X = k, Y = l) = \frac{\binom{N-l}{B-2}}{\binom{N}{B}} = \frac{B(B-1)\, R!\, (N-l)!}{(R-l+2)!\, N!}.
$$

b. (10) Show that for all $l = 2, 3, \ldots, R+2$ and $k = 1, 2, \ldots, l-1$ we have

$$
P(X = k, Y = l) = \frac{1}{l-1} P(Y = l).
$$

Solution: The formula for marginal distributions gives

$$
P(Y = l) = \sum_{k=1}^{l-1} P(X = k, Y = l).
$$

All the terms in the sum are equal hence

$$
P(Y = l) = (l-1) \cdot \frac{B(B-1)\, R!\, (N-l)!}{(R-l+2)!\, N!}.
$$

The assertion follows.

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

$$f_{X,Y}(x,y) = e^{-x} \cdot \frac{1}{\sqrt{2\pi x}} e^{-\frac{(y-x)^2}{2x}}$$

for $x > 0$, $-\infty < y < \infty$.

a. (10) Find the density of the random variable

$$Z = \frac{Y - X}{\sqrt{X}}\,.$$

Solution: We use the transformation formula. Let

$$\Phi(x,y) = \left(X, \frac{Y-X}{\sqrt{X}}\right)\,.$$

We find

$$\Phi^{-1}(x,z) = \left(x, x + \sqrt{x}z\right)\,.$$

It is easily shown that $J_{\Phi^{-1}}(x,z) = \sqrt{x}$. *The transformation formula gives*

$$f_{X,Y}(x,z) = e^{-x} \cdot \frac{1}{\sqrt{2\pi x}} e^{-\frac{z^2}{2}} \cdot \sqrt{x}\,.$$

The term $\sqrt{x}$ *cancels. The density is a product of two terms. One only depends on* x *and the other on* z. *It follows* $Z \sim \mathrm{N}(0,1)$.

b. (10) Show that X and

$$Z = \frac{Y - X}{\sqrt{X}}$$

are independent.

Solution: The claim follows from the form of $f_{X,Z}(x,z)$.

4. (20) Let $U = \{1, 2, \ldots, n\}$. We select subsets $A_1, A_2, \ldots, A_r$ from U in such a way that the choices are independent and all 2^n subsets are equally likely. Let $X = \text{card}(\cup_{j=1}^r A_j)$.

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

Hint:

$$I_i = \begin{cases} 1\,, & \text{if } i \in \cup_{j=1}^r A_j \\ 0\,, & \text{else.} \end{cases}$$

Solution: We have $X = I_1 + \cdots + I_n$. *By independence*

$$P(I_i = 0) = \left(\frac{1}{2}\right)^r$$

since the event $\{I_i = 0\}$ *is the event that the element* i *is not in* A_k *for all* $k = 1, 2, \ldots, r$. *The probability that* i *is not in* A_1 *equals* $\frac{1}{2}$ *and is the same for all other subsets. Since subsets are selected independently the above follows. We have*

$$E(X) = n\,E(I_1) = n\left(1 - \left(\frac{1}{2}\right)^r\right).$$

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

Hint: for covariances start with $P(I_1 = 0, I_2 = 0)$ *and use* $P(I_1 = 1, I_2 = 0) = P(I_2 = 0) - P(I_1 = 0, I_2 = 0)$.

Solution: For the indicators I_1, I_2 *we have*

$$P\left(I_1 = 0, I_2 = 0\right) = \left(\frac{1}{4}\right)^r.$$

This comes from the fact that A_1 *does not contain elements* 1 *and* 2 *with probability* $\frac{1}{4}$ *and choices of subsets are independent. It follows*

$$P(I_1 = 1, I_2 = 0) = P(I_2 = 0) - P(I_1 = 0, I_2 = 0) = \left(\frac{1}{2}\right)^r - \left(\frac{1}{4}\right)^r.$$

By symmetry the same is true for $P(I_1 = 0, I_2 = 0)$. *Finally*

$$P(I_1 = 1, I_2 = 1) = 1 - P(I_1 = 1, I_2 = 0) - P(I_1 = 0, I_2 = 1) - P(I_1 = 0, I_2 = 0).$$

We observe that I_1, I_2 *are independent and hence their covariance is* 0. *It follows*

$$\text{var}(X) = n\,\text{var}(I_1) = n\left(\frac{1}{2}\right)^r\left(1 - \left(\frac{1}{2}\right)^r\right).$$

5. (20) Suppose a colony of bacteria starts with one bacterium at time $n = 0$. From time n to time $n+1$ each existing bacterium splits into two with probability p independently of all other bacteria, or stays as it is with probability $q = 1-p$. Denote by Z_n the number of bacteria at time n. Let G_n be the generating function of Z_n.

a. (10) Express G_{n+1} by G_n.

Solution: The wording of the problem is the same as the branching process where a bacterium has one descendant with probability $1-p$ and two descendants with probability p. This means that $G(s) = qs + ps^2$. We have

$$G_{n+1}(s) = G_n(G(s)) = G_n(s(q+ps))\,.$$

b. (5) Show that $E(Z_n) = (1+p)^n$.

Solution: We know that $E(Z_n) = G_n'(1)$. Taking derivatives in the recursion formula gives

$$G_{n+1}'(s) = G_n'\big(s(q+ps)\big)(q+2ps)\,.$$

Plug in $s = 1$ to get

$$E(Z_{n+1}) = E(Z_n)\cdot(1+p)\,.$$

Since $E(Z_0) = 1$ by induction we have

$$E(Z_n) = (1+p)^n\,.$$

c. (5) Compute $P(Z_n = 2^n)$.

Solution: The generating functions G_n are polynomials of degree 2^n. Let a_n be the coefficient of s^{2^n} in G_n. The recursion relation gives that the coefficient of $s^{2^{n+1}}$ in G_{n+1} is $a_n p^{2^n}$. We have $a_0 = 1$. By induction it follows

$$a_n = p^{2^n - 1}\,.$$

6. (25) In his 1953 paper with the title *The Random Character of Stock Market Prices* the eminent statistician M. G. Kendall wrote (the quote is edited slightly):

> It seems that the change in price of a stock from one week to the next is practically independent of the change from that week to the week after. This alone is enough to show that it is impossible to predict the price from week to week from the series of weekly prices itself. And if the series really is wandering, any systematic movements such as trends or cycles which may be "observed" in such series are illusory. The series looks like a "wandering" one, almost as if once a week the *Demon of Chance* drew a random number from a large box with average 0, and variance equal to 1 and added it to the current price to determine next week's price. And this, we may recall, is not the behaviour in some small backwater market. The data derive from the Chicago wheat market over a period of fifty years.

a. (5) Suppose we are interested in the net change of price over the next 52 weeks. Fill in the blanks in the following phrase: the net change in price will be just like the ________ of ___ ________ from a box. We assume that the average of the box is ________ and the variance is ________.

Solution: The change in price will be just like the sum of 52 draws from the box. We assume that the average of the box is 0 and the variance is 1.

b. (5) Find, approximately, the probability that the net change in price over the next 52 weeks will be less than 10. Use $\Phi(1.4) \doteq 0.92$.

Solution: One has $E(S_n) = 0$ *and* $\mathrm{var}(S_n) = 52 \times 1 = 7.2$*. We need to convert* 10 *and* -10 *into standard units. Subtracting the expectation and dividing by the square root of the variance, one finds that the answer is the area between* -1.4 *and* 1.4 *under the standard normal curve. This area is about* 84%*.*

c. (10) Suppose the *Demon of Chance* decides to change the box. He decides to add the same positive amount on each ticket in the box so that the variance of the box does not change but the average of the box goes up by the amount added. What amount should he add so that the sum of 52 draws will be positive with probability approximately 90%? Use $\Phi(1.30) \doteq 0.90$.

Solution: The variance of the sum of draws will not change. The expectation, however, will be increased by $52 \times$ *added number. The table says that the area to the right of* -1.30 *is about* 90%*. When we convert* 0 *to standard units using the new* EV*, we should get* -1.3*. This means*

$$\frac{0 - E(S_n)}{\sqrt{\mathrm{var}(S_n)}} = -\frac{E(S_n)}{7.2} = -1.3\,.$$

It follows that $E(S_n) = 9.4$*. Dividing by* 52 *we get that the new average is* 0.18*. This new average is exactly the amount the Demon of Chance should add.*