Probability Homework

Problem 1. Let X and Y be two $\mathbb{N}_0$-valued random variables such that $X = Y + Z$, where Z is a Bernoulli random variable with parameter $p \in (0, 1)$, independent of Y. Only one of the following statements is true. Which one?

(a) $X + Z$ and $Y + Z$ are independent

(b) X has to be $2\mathbb{N}_0 = \{0, 2, 4, 6, \dots\}$-valued

(c) The support of Y is a subset of the support of X

(d) $\mathbb{E}[(X + Y)Z] = \mathbb{E}[(X + Y)]\mathbb{E}[Z]$

(e) none of the above

Solution: The correct answer is (c).

(a) False. Simply take $Y = 0$, so that $Y + Z = Z$ and $X + Z = 2Z$.

(b) False. Take $Y = 0$.

(c) True. For m in the support of Y (so that $\mathbb{P}[Y = m] > 0$), we have

$$\mathbb{P}[X = m] \geq \mathbb{P}[Y = m, Z = 0] = \mathbb{P}[Y = m]\mathbb{P}[Z = 0] = \mathbb{P}[Y = m](1 - p) > 0.$$

Therefore, m is in the support of X.

(d) False. Take $Y = 0$.

(e) False.

Problem 2. A fair die is tossed and its outcome is denoted by X, i.e.,

$$X \sim \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \end{pmatrix}.$$

After that, X independent fair coins are tossed and the number of *heads* obtained is denoted by Y.
Compute:

1. $\mathbb{P}[Y = 4]$.
2. $\mathbb{P}[X = 5|Y = 4]$.
3. $\mathbb{E}[Y]$.
4. $\mathbb{E}[XY]$.

Solution:

1. For $k = 1, \dots, 6$, conditionally on $X = k$, Y has the binomial distribution with parameters k and $\frac{1}{2}$. Therefore,

$$\mathbb{P}[Y = i|X = k] = \begin{cases} \binom{k}{i} 2^{-k}, & 0 \le i \le k \\ 0, & i > k, \end{cases}$$

and so, by the law of total probability.

$$\begin{aligned} \mathbb{P}[Y = 4] &= \sum_{k=1}^{6} \mathbb{P}[Y = 4|X = k]\mathbb{P}[X = k] \\ &= \tfrac{1}{6}(2^{-4} + \binom{5}{4} 2^{-5} + \binom{6}{4} 2^{-6}) \left[= \tfrac{29}{384}\right]. \end{aligned} \tag{1}$$

2. By the (idea behind the) Bayes formula

$$\begin{aligned} \mathbb{P}[X = 5|Y = 4] &= \frac{\mathbb{P}[X = 5, Y = 4]}{\mathbb{P}[Y = 4]} = \frac{\mathbb{P}[Y = 4|X = 5]\mathbb{P}[X = 5]}{\mathbb{P}[Y = 4]} \\ &= \frac{\binom{5}{4} 2^{-5} \times \frac{1}{6}}{\frac{1}{6}\left(2^{-4} + \binom{5}{4} 2^{-5} + \binom{6}{4} 2^{-6}\right)} \left[= \tfrac{10}{29}\right]. \end{aligned}$$

3. Since $\mathbb{E}[Y|X = k] = \frac{k}{2}$ (the expectation of a binomial with $n = k$ and $p = \frac{1}{2}$), the law of total probability implies that

$$\mathbb{E}[Y] = \sum_{k=1}^{6} \mathbb{E}[Y|X = k]\mathbb{P}[X = k] = \tfrac{1}{6} \sum_{k=1}^{6} \tfrac{k}{2} \left[= \tfrac{7}{4}\right].$$

4. By the same reasoning,

$$\begin{aligned} \mathbb{E}[XY] &= \sum_{k=1}^{6} \mathbb{E}[XY|X = k]\mathbb{P}[X = k] = \sum_{k=1}^{6} \mathbb{E}[kY|X = k]\mathbb{P}[X = k] \\ &= \sum_{k=1}^{6} k\mathbb{E}[Y|X = k]\mathbb{P}[X = k] = \tfrac{1}{6} \sum_{k=1}^{6} \tfrac{1}{2} k^2 \left[= \tfrac{91}{12}\right]. \end{aligned}$$

Problem 3.

1. An urn contains 1 red ball and 10 blue balls. Other than their color, the balls are indistiguishable, so if one is to draw a ball from the urn without peeking - all the balls will be equally likely to be selected. If we draw 5 balls from the urn at once and without peeking, what is the probability that this collection of 5 balls contains the red ball?

2. We roll two fair dice. What is the probability that the sum of the outcomes equals exactly 7?

3. Assume that A and B are *disjoint* events, i.e., assume that $A \cap B = \emptyset$. Moreover, let $\mathbb{P}[A] = a > 0$ and $\mathbb{P}[B] = b > 0$. Calculate $\mathbb{P}[A \cup B]$ and $\mathbb{P}[A \cap B]$, using the values a and b:

Solution:

1. $$\mathbb{P}[\text{"the red ball is selected"}] = \frac{\binom{10}{4}}{\binom{11}{5}} = \frac{5}{11}.$$

2. There are 36 possible outcomes (pairs of numbers) of the above roll. Out of those, the following have the sum equal to $7: (1,6),(2,5),(3,4),(4,3),(5,2),(6,1)$. Since the dice are fair, all outcomes are equally likely. So, the probability is
$$\frac{6}{36} = \frac{1}{6}.$$

3. According to the axioms of probability:
$$\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] = a + b,\ \mathbb{P}[A \cap B] = \mathbb{P}[\emptyset] = 0.$$

Problem 4.

1. Consider an experiment which consists of 2 independent coin-tosses. Let the random variable X denote the number of heads appearing. Write down the probability mass function of X.

2. There are 10 balls in an urn numbered 1 through 10. You randomly select 3 of those balls. Let the random variable Y denote the maximum of the three numbers on the extracted balls. Find the probability mass function of Y. You should simplify your answer to a fraction that does not involve binomial coefficients. Then calculate: $\mathbb{P}[Y \geq 7]$.

3. A fair die is tossed 7 times. We say that a toss is a **success** if a 5 or 6 appears; otherwise it's a **failure**. What is the distribution of the random variable X representing the number of successes out of the 7 tosses? What is the probability that there are exactly 3 successes? What is the probability that there are no successes?

4. The number of misprints per page of text is commonly modeled by a Poisson distribution. It is given that the parameter of this distribution is $\lambda = 0.6$ for a particular book. Find the probability that there are exactly 2 misprints on a given page in the book. How about the probability that there are 2 *or more* misprints?

Solution:

1.

$$\begin{aligned} p_0 &= \mathbb{P}[\{(T,T)\}] = \frac{1}{4}, \\ p_1 &= \mathbb{P}[\{(T,H),(H,T)\}] = \frac{1}{2}, \\ p_2 &= \mathbb{P}[\{(H,H)\}] = \frac{1}{4}, \\ p_k &= 0, \text{ for all other } k. \end{aligned}$$

2. The random variable Y can take the values in the set $\{3, 4, \ldots 10\}$. For any i, the triplet resulting in Y attaining the value i must consist of the ball numbered i and a pair of balls with lower numbers. So,

$$p_i = \mathbb{P}[Y = i] = \frac{\binom{i-1}{2}}{\binom{10}{3}} = \frac{\frac{(i-1)(i-2)}{2}}{\frac{10\cdot 9\cdot 8}{3\cdot 2\cdot 1}} = \frac{(i-1)(i-2)}{240}.$$

Since the balls are numbered 1 through 10, we have

$$\mathbb{P}[Y \geq 7] = \mathbb{P}[Y = 7] + \mathbb{P}[Y = 8] + \mathbb{P}[Y = 9] + \mathbb{P}[Y = 10].$$

So,

$$\begin{aligned} \mathbb{P}[Y \geq 7] &= \frac{6\cdot 5}{240} + \frac{7\cdot 6}{240} + \frac{8\cdot 7}{240} + \frac{9\cdot 8}{240} \\ &= \frac{1}{240}(30 + 42 + 56 + 72) \\ &= \frac{200}{240} = \frac{5}{6}. \end{aligned}$$

3. X has a binominal distribution with parameters $n = 7$ and $p = 1/3$, i.e., $X \sim b(7, 1/3)$.

$$\begin{aligned} \mathbb{P}[X = 3] &= \binom{7}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^4 = \frac{560}{2187}, \\ \mathbb{P}[X = 0] &= \left(\frac{2}{3}\right)^7 = \frac{128}{2187}. \end{aligned}$$

4. Let X denote the random variable which stands for the number of misprints on a given page. Then

$$\begin{aligned} \mathbb{P}[X = 2] &= \frac{0.6^2}{2!}e^{-0.6} \approx 0.0988, \\ \mathbb{P}[X \geq 2] &= 1 - \mathbb{P}[X < 2] \\ &= 1 - (\mathbb{P}[X = 0] + \mathbb{P}[X = 1]) \\ &= 1 - \left(\frac{0.6^0}{0!}e^{-0.6} + \frac{0.6^1}{1!}e^{-0.6}\right) \\ &= 1 - \left(e^{-0.6} + 0.6e^{-0.6}\right) \\ &= 1 - 1.6e^{-0.6} \approx 0.122. \end{aligned}$$

Problem 5. Let X and Y be two Bernoulli random variables with the same parameter $p = \frac{1}{2}$. Can the support of their sum be equal to $\{0, 1\}$? How about the case where p is not necesarily equal to $\frac{1}{2}$? Note that no particular dependence structure between X and Y is assumed.

Solution: Let p_{ij}, $i = 0, 1$, $j = 0, 1$ be defined by

$$p_{ij} = \mathbb{P}[X = i, Y = j].$$

These four numbers effectively specify the full dependence structure of X and Y (in other words, they completely determine the distribution of the random vector (X, Y)). Since we are requiring that both X and Y be Bernoulli with parameter p, we must have

$$p = \mathbb{P}[X = 1] = \mathbb{P}[X = 1, Y = 0] + \mathbb{P}[X = 1, Y = 1] = p_{10} + p_{11}. \tag{2}$$

Similarly, we must have

$$\begin{aligned} 1 - p &= p_{00} + p_{01}, & (3) \\ p &= p_{01} + p_{11}, & (4) \\ 1 - p &= p_{00} + p_{10} & (5) \end{aligned}$$

Suppose now that the support of $X + Y$ equals to $\{0, 1\}$. Then $p_{00} > 0$ and $p_{01} + p_{10} > 0$, but $p_{11} = 0$ (why?). Then, the relation (2) implies that $p_{10} = p$. Similarly, $p_{01} = p$ by relation (4). Relations (3) and (5) tell us that $p_{00} = 1 - 2p$. When $p = \frac{1}{2}$, this implies that $p_{00} = 0$ - a contradiction with the fact that $0 \in {}_{X+Y}$.

When $p < \frac{1}{2}$, there is still hope. We construct X and Y as follows: let X be a Bernoulli random variable with parameter p. Then, we define Y depending on the value of X. If $X = 1$, we set $Y = 0$. If $X = 0$, we set $Y = 0$ with probabilty $\frac{1-2p}{1-p}$ and 1 with probability $\frac{p}{1-p}$. How do we know that Y is Bernoulli with probability p? We use the law of total probability:

$$\mathbb{P}[Y = 0] = \mathbb{P}[Y = 0|X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = 0|X = 1]\mathbb{P}[X = 1] = \tfrac{1-2p}{1-p}(1 - p) + p = 1 - p.$$

Similarly,

$$\mathbb{P}[Y = 1] = \mathbb{P}[Y = 1|X = 0]\mathbb{P}[X = 0] + \mathbb{P}[Y = 1|X = 1]\mathbb{P}[X = 1] = (1 - \tfrac{1-2p}{1-p})(1 - p) = p.$$