4th problem set for Probability and Statistics — March 18/19

Basic problems about expected value

Recall the formula for expected value:

$$\mathbb{E}(X) = \sum_{x \in Im(X)} x \cdot P(X = x),$$

1. (a) Let $P(X = 100) = p$, $P(X = 0) = 1 - p$. Determine $\mathbb{E}(X)$.
(b) Let $P(Y = 100) = p$, $P(Y = 99) = 1 - p$. Determine $\mathbb{E}(Y)$.

2. Suppose solving one problem takes X minutes, where $X \in \{1, 2, 3, 4, 5\}$. The time required for solving the problem is random (for instance, it may depend on current weather), and the probability function is $p_X(1) = p_X(2) = 0.1$, $p_X(3) = p_X(4) = 0.2$, $p_X(5) = 0.4$. Compute $\mathbb{E}(X)$.

3. (a) We have an unlimited number of black and red socks in a drawer, unpaired. We take out socks without looking, with both colors being equally likely to be drawn. How many socks do we have to take out before we have two of the same color, in the worst case, and on average?
(b) Solve the same problem for three different colors.

4. A new casino offers the following game: we bet x crowns. With probability $1/2$ we lose them, and with a probability of $1/2$ we win $2x$ crowns (in addition to our original x crowns).
(a) We start with k crowns. What is the best strategy if we want to maximize the expected value of our winnings after n rounds? And what is the expected value in this case?
(b) What is the probability of losing all our money with such a strategy?
(c) What strategy would you choose?

Linearity of $\mathbb{E}$

$$\mathbb{E}(a_1 X_1 + \cdots + a_n X_n) = a_1 \mathbb{E}(X_1) + \cdots + a_n \mathbb{E}(X_n)$$

5. We have a biased coin, which lands heads with probability p. We toss it n times.
(a) Let X be the number of positions where the coin landed heads first and then immediately tails. (For example, if $n = 6$ and the sequence is HTHHTH, then $X = 2$.) Determine $\mathbb{E}(X)$.
(b) Does X have binomial distribution? Explain why it doesn't.
(c) Now let Y be the number of occurrences of the sequence HTH (here "occurrence" means that the three symbols occupy three consecutive positions, in this order). What is $\mathbb{E}(Y)$?
[**Hint:** Use linearity. Let A_i be the event that the i-th toss is H and the $(i + 1)$-th toss is T, and let $X_i = I_{A_i}$ be the corresponding indicator variable.]

6. We again have a coin which lands head with probability p, and toss it $\binom{n}{2}$ times. Based on the outcomes, we construct a graph with vertices $V = \{1, 2, \ldots, n\}$. For all pairs $\{i, j\} \in \binom{V}{2}$, we determine if they are connected by an edge – this happens when the corresponding coin toss results in heads. The resulting graph is called a random graph $G(n, p)$.
(a) Show that the expected value of the number of edges in the graph is $p\binom{n}{2}$.
[**Hint:** Use linearity as in the previous case or notice that the number of edges follows a binomial distribution.]
(b) Show that the expected value of the number of triangles in the random graph is $p^3\binom{n}{3}$.

Conditional Expectation

$$\mathbb{E}(X \mid B) = \sum_{x \in Im(X)} x \cdot P(X = x \mid B)$$

$$\mathbb{E}(X) = \sum_i P(B_i) \cdot \mathbb{E}(X \mid B_i) \qquad\qquad B_1, B_2, \dots \text{ is a partition of } \Omega$$

7. My computer occasionally misbehaves: every day, there is a probability p that it freezes. If it freezes two days in a row, I start addressing the issue. What is the expected number of days before I have to address the issue?

8. In a TV quiz show, a participant can choose two questions. For question A, he estimates that he will correctly answer with a probability of 0.8 (and will earn 1,000 CZK for that). For question B, his success probability is only 0.5, but for a correct answer, he receives 2,000 CZK.

(a) What is the expected value of the winnings if he starts with question A?

(b) What if he starts with question B?

(c) Bonus: if the success probabilities are p_A, p_B, and the rewards m_A, m_B, how should the participant decide? * And what if there are more than two questions?

Bonus problems

9. An indicator random variable I_A for an event A is defined as follows:

$$I_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

(a) What is $\mathbb{E}(I_A)$?

(b) Let $A = A_1 \cup A_2 \cup \dots \cup A_n$. Verify the equality

$$1 - I_A = \prod_{i=1}^{n} (1 - I_{A_i}).$$

(c) Expand the above product and use the linearity of expected value. You will obtain the principle of inclusion and exclusion, which you know from discrete mathematics.

10. ***(Coupon collector problem)** There are n different types of Pokémon in a game. Each time you play, you get one random Pokémon. Every Pokémon is equally likely to appear. How many times do you need to play, on average, to collect all the Pokémon? How fast does the expectation grow?

11. * Let M be the number of emails received per day, S be the number of spam emails among them, H be the number of "ham" emails – those that are not spam. Assume that $M \sim Pois(\lambda)$ and that each email, independently of others, has a probability p of being spam.

(a) Express $P(S = k)$ (as an infinite sum) using the joint distribution of M and S.

(b) Derive that $S \sim Pois(p\lambda)$.

(c) Derive that $H \sim Pois((1-p)\lambda)$ and also that H, S are independent random variables.

For Practice

12. King Louis wants to have a male heir so that he can name him Louis again. Each year, his wife gives birth to exactly one child, which is equally likely to be a boy or a girl, independent of previous attempts. All children survive. If a boy is born, Louis will not have any more children. Let S be the number of sons born and D be the number of daughters born.

(a) Determine $\mathbb{E}(S)$.

(b) Determine $\mathbb{E}(D)$.

13. We toss a coin with probability p of landing heads. If two consecutive tosses result in the same outcome, we stop. What is the expected number of tosses we make?