4th problem set for Probability and Statistics — March 11–15

Random Variables

1. We roll two dice – for simplicity, four-sided ones, with numbers 1 through 4. Let X be the maximum of the two rolled numbers. Describe how you will model this situation: what is Ω, what exactly is X as a mathematical object, and what is p_X.

2. 234 people are registered for a lecture. What is the probability that exactly one of them has a birthday today? Ignore leap years, assume that all days are equally likely for birthdays.

(a) Use the binomial distribution.

(b) Use the Poisson approximation: $Bin(n, \lambda/n)$ is approximately $Poi(\lambda)$.

3. Let a random variable X have a Poisson distribution, $X \sim Poi(\lambda)$. Recall the formula for the probability mass function $p_X(k)$. Show that $p_X(k)$ increases for $k \leq \lfloor \lambda \rfloor$ and then decreases, tending to zero as k increases.

4. We have a keyring with five keys, one of which opens our door, but we don't know which one. We try to open a door.

(a) After each attempt, the keyring slips away, and we randomly select again.

(b) We select keys in random order, but each key only once (we can mark them).

In both cases, we examine the number of attempts needed to open the door. What is the distribution of this random variable? That is, what is the probability of opening the door on the kth attempt.

(c) Same as part (a), but two out of ten keys are correct.

(d) Same as part (b), but two out of ten keys are correct.

5. Consider $m + n$ rolls of a fair die. Let X be the number of sixes in the first m rolls, Y be the number of sixes in the last n rolls. What is the distribution of X, Y, and $X + Y$?

6. There are N candies in a bag, of which K are good. We randomly draw n of them, and let X be the number of good candies drawn.

(a) What is the distribution of the random variable X?

(b) What is $P(X = k)$?

Independent Random Variables

Definition: Random variables X_1, X_2 are independent if the events $\{X_1 = x_1\}$ and $\{X_2 = x_2\}$ are independent for every pair of numbers x_1, x_2.

7. Show that events A, B are independent if and only if their indicator random variables are independent.

8. Show that for discrete independent random variables X, Y,

$$P(X \leq x \,\&\, Y \leq y) = P(X \leq x)P(Y \leq y).$$

For simplicity, you may assume that $Im(X) = Im(Y) = \{1, 2, \ldots, n\}$ for some n.

Bonus

9. (St. Petersburg Casino) We toss a coin repeatedly. If the first occurrence of heads is on the nth toss, we receive a reward of 2^n coins. How much would you be willing to pay to participate in this game?

10. * (Absent-minded mathematician) In each pocket, a mathematician carries a box with n matches. Every time they need a match, they randomly select a pocket. When they find an empty box, let X be the number of matches in the other box. Find the probability mass function of the random variable X.

Practice problems

11. n basketball players independently shoot at a hoop. Each player has a probability p of hitting the target with each shot, independent of others. Let X_i denote the order in which player i first hits the target. Let $X = \min(X_1, \ldots, X_n)$.

(a) What is the distribution of X_1, X_2, ...?

(b) What is the distribution of X?

12. Consider X, the number of meteors you observe during an hour of stargazing. Which distribution would you use to describe X?

4th homework, due next week

In a test, there are 20 multiple-choice questions (options a, b, c, d). A correct answer (there is always exactly one correct answer) earns 1 point, a wrong answer deducts $-\frac{1}{4}$ point, and an unanswered question earns zero points. For each question, with probability p Alice has studied the topic that the question asks about, and therefore knows the correct answer. If she doesn't know the correct answer, she knows that she doesn't know the answer and can decide whether to guess.

(a) What is the expected value of the number of points Alice will earn if she only answers questions she knows the answer to?

(b) What if she decides to guess when she doesn't know the correct answer?

(c) How would the penalty for a wrong answer need to change to make the expected values in parts (a) and (b) the same?