PRACTICE PROBLEMS: PROBABILITY

Statistically, the probability of any one of us being here is so small that you'd think the mere fact of existing would keep us all in a contented dazzlement of surprise. —Lewis Thomas

Course: CS 5002
Fall 2018
Due: No due date

OBJECTIVES

After you go through these problems, you will be comfortable with:

- Probability axioms
- Expected values
- Conditional probability
- Baye's Formula
- Independent Events
- Some discrete probability distributions

RELEVANT READING

- Rosen, chapter 7: Discrete Probability
  - Chapter 7.1. An Introduction to Discrete Probability
  - Chapter 7.2: Probability Theory
  - Chapter 7.3 Bayes' Theorem
  - Chapter 7.4 Expected Value and Variance

EXERCISES

Question 1

Consider the following simple experiment: A fair die is tossed, and its face value is observed. If the number on the face value is even, value 1 is assigned to some random variable X. If, on the other hand, an odd number is observed, value 0 is assigned to X.

(a) What is the range of X?

(b) Find probabilities $P(X = 1)$ and $P(X = 0)$.

(a) The range of X is $R_X = \{0, 1\}$.
(b) Since the coin is fair, we know that the likelihood of observing a head is equal to the likelihood of observing a tail, and it equals 0.5. Therefore, the probabilities $\mathbb{P}(X = 1) = \mathbb{P}(X = 0) = 0.5$.

Consider now the second simple experiment, where some coin is tossed three times. We assume that the tosses are independent, and the probability of a head is p. Let Y be the random variable representing the number of heads observed.

(a) What is the range of Y?

(b) Find the probabilities $P(Y = 0)$, $P(Y = 1)$, $P(Y = 2)$, and $P(Y = 3)$.

The range of Y is $R_Y = \{0, 1, 2, 3\}$.
To find the asked probabilities, let's examine the sample space, and define all possible elementary events. Since coin tosses are independent, and we toss a coin three times, we know that we have $2^3 = 8$ elementary events, given as $\{(TTT), (TTH), (THT), (HTT), (HHT), (HTH), (THH), (HHH)\}$. Since all elementary events are equally likely to occur, we find the asked probabilites as follows: $(Y = 0) = \{TTT\}$, and $P(Y = 0) = (1 - p)^3$.
$(Y = 1) = \{HTT, THT, TTH\}$, and $P(Y = 1) = 3(1 - p)^2 p$.
$(Y = 2) = \{THH, HHT, HTH\}$, and $P(Y = 2) = 3(1 - p)p^2$.
$(Y = 3) = \{HHH\}$, and $P(Y = 3) = p^3$.

 Points: ____________ out of 10

Question 2

Consider the experiment of tossing an honest coin repeatedly. Let the random variable X denote the number of tosses required until the first head appears.

(a) Find and sketch the pmf $p_X(x)$ and the cdf $F_X(x)$ of X.

(b) Find:

(i) $P(1 < X < 4)$

(ii) $P(X > 4)$

(a) The pmf of X is given by:

$$p_X(x) = p_X(k) = P(X = k) = \left(\frac{1}{2}\right)^k, k = 1, 2, \ldots$$

Then we can see that:

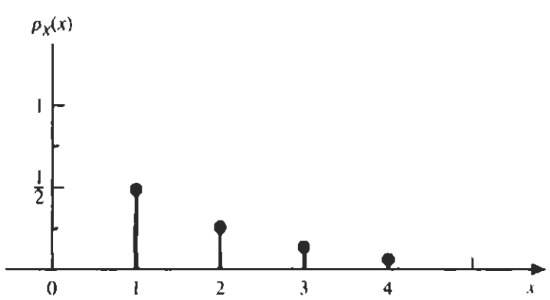
$$F_X(x) = P(X \leq x) = \sum_{k=1}^{m \leq x} p_X(k) = \sum_{k=1}^{m \leq x} \left(\frac{1}{2}\right)^k$$

or we have:

$$F_X(x) = \begin{cases} 0, & x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ \frac{3}{4}, & 2 \leq x < 3 \\ \vdots & \vdots \\ 1 - (\frac{1}{2})^n, & n \leq x < n+1 \\ \vdots & \vdots \end{cases} \tag{1}$$

The pmf and the cdf functions for the given problem are shown in the diagram below.

(b)
$P(1 < X < 4) = F_X(3) - F_X(1) = \frac{7}{8} - \frac{1}{2} = \frac{3}{8}$
Since this is a discrete random variable, we can also see the result simply by adding the probabilities of 2 and 3 (since these are the values between 1 and 4):
$P(1 < X < 4) = P(X = 2) + P(X = 3) = \frac{1}{4} + \frac{1}{2} = \frac{3}{8}$
$P(X > 4) = 1 - P(X \leq 4) = 1 - F_X(4) = 1 - \frac{15}{16} = \frac{1}{16}$



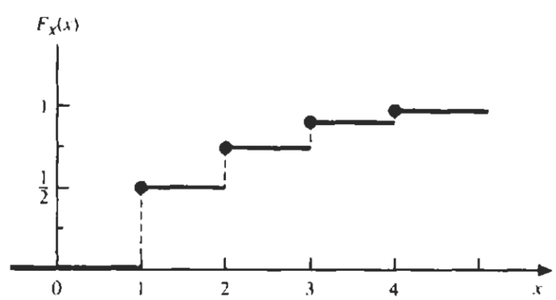


Figure 1: pmf and cdf for the random variable in Problem 2.

Question 3

Consider a discrete random variable X whose pmf is given as:

$$p_X(x) = \begin{cases} \frac{1}{3}, & x = -1, 0, 1 \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

Please find the mean and variance of X.

The pmf $p_X(x)$ is composed of three stems at -1,0, and 1, so the mean can simply be computed as:

$$\mu_X = E(X) = \frac{1}{3}(-1+0+1) = 0$$

Then the variance can be found by the definition:

$$\sigma_X^2 = Var(X) = E[(X-\mu_X)^2] = E(X^2) = \frac{1}{3}[(-1)^2 + 0^2 + 1^2] = \frac{2}{3}$$

Question 4

A lottery game offers \$2 million to the grand prize winner, \$20 to each of 10 000 second prize winner, and \$4 to each of 50 000 third prize winner. The cost of the lottery is \$2 per ticket. Suppose that 1.5 million tickets are sold. What is the expected gain or loss of the ticket?

Sketch of a solution: To solve this problem, we start with the observation that every of the 1.5million tickets is equally as likely to win. With that in mind, we know that:

- The probability of winning the grand prize is equal to $\mathbb{P}[\text{grand prize}]] = \frac{2}{3} \cdot 10^{-6}$
- The probability of winning the prize \$20 is equal to $\mathbb{P}[\$20\text{prize}]] = \frac{2}{3} \cdot 10^{-2}$
- The probability of winning the prize \$4 is equal to $\mathbb{P}[\$4\text{prize}]] = \frac{10}{3} \cdot 10^{-2}$
- The probability of not winning any prize is equal to $\mathbb{P}[\text{no prize}]] = \frac{1.5 \cdot 10^6 - 50001}{1.5 \cdot 10^6}$

For every possible prize, we now find the net gain as follows:

- Grand prize winner: 2million - the price of the ticket
- \$20 winner: = 20 - 5 = \$15
- \$20 winner: = 4 - 5 = -\$1
- Person who didn't win: -\$5

To find the expected gain or loss of the ticket, we now find the expected value of the whole random variable.

Question 5

A company sends millions of people an entry form for a sweepstakes accompanied by an order form for magazine subscriptions. The first, second and third prizes are \$10 000 000, \$1 000 000, and \$50 000, respectively. In order to qualify for a prize, a person is not required to order any magazines, but has to spend 60 cents to mail back the entry form. If 30 million people qualify by sending back their entry forms, what is a person's expected gain or loss?

Please see the previous problem for ideas.

Question 6

An urn contains four balls numbered 2, 2, 5 and 6. If a person selects a set of two balls at random, what is the expected value of the sum of the numbers on the balls?

Question 7

When a pair of balanced dice are rolled and the sum of the numbers showing face up is computed, the result can be any number from 2 to 12, inclusive. What is the expected value of the sum?

Question 8

A person pays $1 to play the following game. The person tosses a fair coin four times. If no heads occur, the person pays an additional $2, if one head occurs, the person pays additional $1, if two heads occur, the person just loses the initial dollar, if three heads occur, the person wins $3, and if four heads occur, the person wins $4. What is the person's expected loss or gain?

To find the person's expected loss or gain, we need to find the expected value of the random variable, X representing the expected end result of this game. The game itself represents a random experiment, where a fair coin is tossed four times. Since the coin is fair, we can assume that the head and the tail occur equally likely, with probability $p = \frac{1}{2}$. We can also assume that all of the coin tosses are mutually indepenent.

With these assumption in mind, we can now write:

$$\begin{aligned} \mathbb{E}[X] &= -2 \cdot (1-p)^4 - 1 \cdot p^2(1-p)^2 + 0 \cdot (1-p) \cdot p^3 + 3p^4 \\ &= 7 - 2 \cdot \frac{1}{16} - \frac{1}{4}\frac{1}{4} + 3\frac{1}{16} = \frac{-2}{16} - \frac{1}{16} + \frac{3}{16} \\ &= -\frac{3}{16} + \frac{3}{16} = 0 \end{aligned}$$

Therefore, the expected value of the game is zero, and this expected value does not take into account that a person already paid $ to play the game. Therefore, the person's expected loss is $1.

Question 9

One urn contains 12 blue balls and 7 white balls, and a second urn contains 8 blue balls and 19 white balls. An urn is selected at random, and a ball is chosen from the urn.

(a) What is the probability that the chosen ball is blue?

(b) If the chosen ball is blue, what is the probability that it came from the first urn?

Question 10

A drug-screening test is used in a large population of people of whom %4 actually use drugs. Suppose that false positive rate is %3, and the false negative rate is %2. This a person who uses drugs tests positive %98 of the time, and a person who does not use drugs tests negative %97 of the time.

(a) What is the probability that a randomly chosen person who tests positive for drugs actually uses drugs?

(b) What is the probability that a randomly chosen person who tests negative for drugs does not use drugs?

Question 11

A student taking a multiple-choice exam does not know the answers to two questions. All have five choices for the answer. For one of the two questions, the student can eliminate two answer choices as incorrect, but has no idea about the other answer choices. For the other question, the student has not clue about the correct answer at all. Assume that whether the student chooses the correct answer on one of the questions does not affect whether the students chooses the correct answer on the other question.

(a) What is the probability that the student will answer both questions correctly?

(b) What is the probability that the student will answer exactly one question correctly?

(c) What is the probability that the student will answer neither question correctly?

Question 12

A company uses two proofreaders X and Y yo check a certain manuscript. X misses %12 of typographical errors, and Y misses %15. Assume that the proofreaders work independently.

(a) What is the probability that a randomly chosen typographical error will be missed by both proofreaders?

(b) If the manuscript contains 1000 typographical errors, what number can be expected to be missed?

Question 13

There are n persons in a room.

(a) What is the probability that at least two persons have the same birthday?

(b) Calculate this probability for $n = 50$.

(c) How large need should n be for this probability to be greater than 0.5?

(a) We assume that every day of the year is equally likely to be a person's birthday. So, every person has a birthday on one of 365 days (ignoring leap years), and there are a total of $(365)^n$ possible outcomes. Let A be the event that no two persons have the same birthday. Then the number of outcomes beloning to A is:

$$n(A) = (365)(364) \cdots (365 - n + 1)$$

Assuming that each outcome is equally likely, then we get:

$$P(A) = \frac{n(A)}{n(S)} = \frac{(365)(364) \cdots (365 - n + 1)}{(365)^n}$$

Let B be the even that at least two persons have the same birthday. Then $B = \bar{A}$ and so $P(B) = 1 - P(A)$.
(b) We now have $n = 50$, and so:

$$P(A) \approx 0.03 \text{ and } P(B) \approx 1 - 0.03 = 0.97 \tag{3}$$

(c) We can try different values of n, and see that for $n = 23$ we get:

$$P(A) \approx 0.493 \text{ and } P(B) = 1 - P(A) \approx 0.507 \tag{4}$$

Meaning that if there are 23 people in the room, the probability that at least two of them have the same birthday exceeds 0.5.

Question	Points	Score
1	10	
2	10	
3	5	
4	10	
5	10	
6	10	
7	10	
8	10	
9	15	
10	20	
11	30	
12	20	
13	30	
Total:	190	