Homework 3

Autumn 2023

KEY

2023-10-21

Instructions

- This homework is due in Gradescope on Wednesday Oct 25 by midnight PST. There is a 15 minute grace period and submissions made during this time will not be marked as late. Any work submitted past this period is considered late.
- Please answer the following questions in the order in which they are posed.
- Don't forget to (i) make a local copy this document for your work and to (ii) knit the document frequently to make sure there are no compilation errors.
- When you are done, download the PDF file as instructed in section and submit it in Gradescope.

Exercises

1.

a. (Shelley) Below are the last five lines of Shelley's poem *Ozymandias*

```
"My name is Ozymandias, king of kings:
 Look on my works, ye Mighty, and despair!"
 Nothing beside remains. Round the decay
 Of that colossal wreck, boundless and bare
 The lone and level sands stretch far away
```

Let X denote the length of a word which is randomly selected from those lines. What is $X(\text{boundless})$? Write the range of X. Is X a discrete or continuous random variable? Explain briefly.

When we apply the function X on the outcome "boundless", we get a value of 9. This is the number of letters in the word.

The range of X is $\{2, 3, 4, 5, 6, 7, 8, 9, 10\}$. The random variable X is discrete since it takes a finite number of values.

b. (Free throws) As part of the warm up drill, each player on a basketball team is required to shoot free throws until a basket is made. If Amy has a 80% success rate at the foul line, write the PMF of the random variable X that describes the number of misses before she completes the drill. Assume that individual throws constitute independent events.

Write the PMF of X as a formula first, and then write a brief explanation below.

$$f(x) = 0.2^x \, 0.8, \quad x = 0, 1, 2, \ldots$$

The random variable X represents the number of misses before the first basket. It can take values $x = 0, 1, 2, \ldots$. The event $X = x$ corresponds to the outcome with x misses followed by a basket. Since Amy is a 80% shooter, the probability that she misses a basket is 20%. Given that the individual throws are independent, we can simply multiply 0.2 x times and then multiply by 0.8.

2. (Random walk) Suppose a particle moves 4 steps along the x-axis, starting at 0. At each step, it moves one unit to the right or to the left.

a. Each possible outcome of this experiment is an ordered quadruple. For example, the outcome (L, L, L, L) represents the case when the particle moves one unit to the left at each of the four steps to end up at $x = -4$. The outcome (R, L, R, L) represents the case when the particle first moves one unit to the right, then one unit back to the left and so on ending up at $x = 0$.

   What is the size of this sample space? How many elements? Explain your answer very briefly.

   There are 2^4 or 16 possible outcomes in the sample space. We arrive at this answer by a simple application of the **multiplication** rule for counting. Specifically, the first step has 2 possibilities: L, R. For each choice there are two choices for the second and so on.

b. Let E denote the event that the particle ends up at $x = 0$. Write the outcomes in E and calculate $P(E)$ assuming every outcome is equally likely to occur.

   The outcomes in E are as shown below.

$$E = \{(L, R, L, R); (R, L, R, L); (L, L, R, R); (R, R, L, L); (L, R, R, L); (R, L, L, R)\}.$$

   Assuming that each outcome in S is equally likely, we have

$$P(E) = \frac{6}{16}.$$

c. Now suppose the random variable X denotes the position of the particle after 4 steps. Write its PMF in a tabular format. I have created a partial table for you to fill in. Each row should contain a possible value x, then list the outcomes from the original sample space that give that possible value and the probability. (*Hint: in part b you calculated* $P(X = 0)$.)

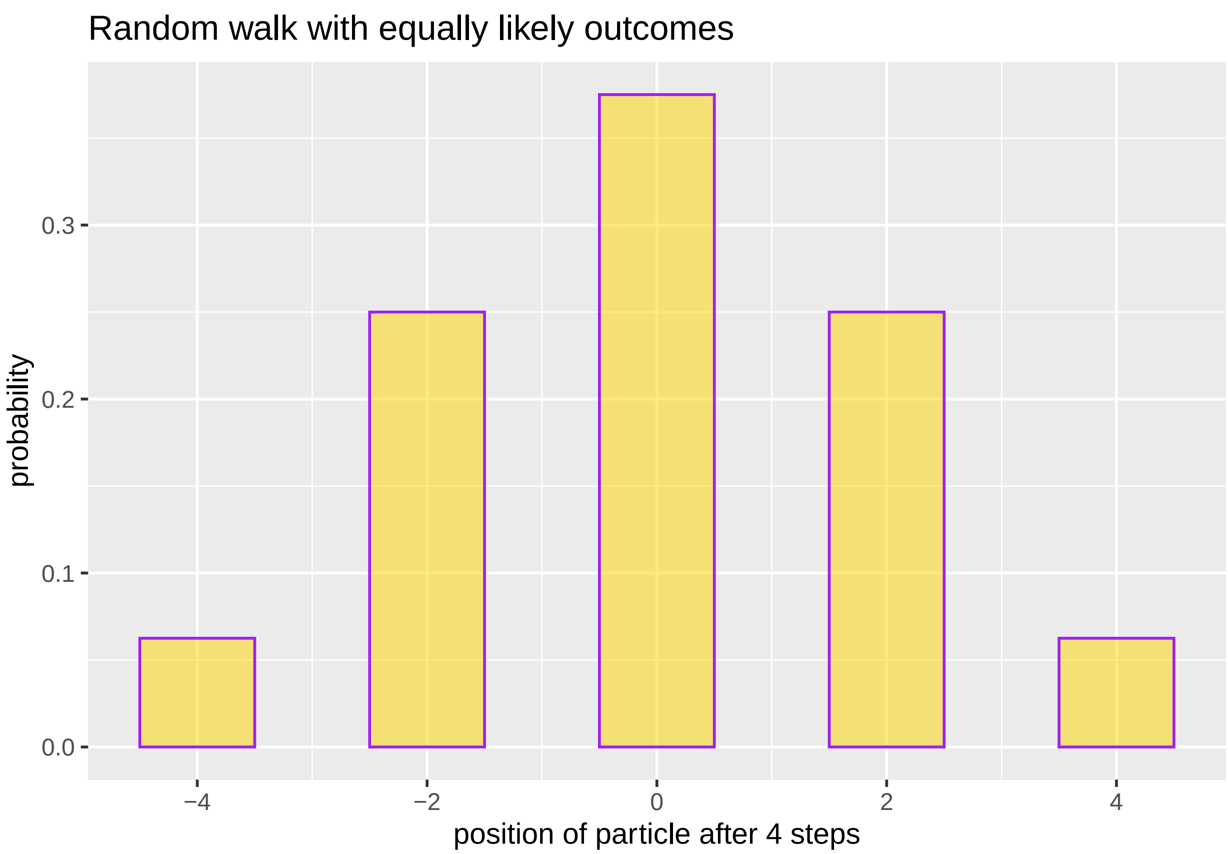
d. Draw the probability histogram of X, the position of the particle after 4 steps. (Don't forget to attach the `tidyverse`package in the setup chunk.)

Table 1: PMF of X: position of particle after 4 steps

x	outcomes from S	$f(x)$
-4	(L, L, L, L)	$\frac{1}{16}$
-2	(L, L, R, L); (L, L, L, R); (L, R, L, L); (R, L, L, L)	$\frac{4}{16}$
0	(L,R,L,R); (R, L, R, L); (L, L, R, R); (R, R, L, L); (L, R, R, L); (R, L, L, R)	$\frac{6}{16}$
2	(R, R, R, L); (R, R, L, R); (R, L, R, R); (L, R, R, R)	$\frac{4}{16}$
4	(R, R, R, R)	$\frac{1}{16}$

```r
walk_df <- tibble(
  x = c(-4,-2,0,2,4),
  f = c(1/16,4/16,6/16,4/16,1/16)
)

ggplot(data = walk_df,
       mapping = aes(x = x, y = f))+
  geom_col(width = 1, fill = "gold", color = "purple",
           alpha = 0.5)+
  scale_x_continuous(breaks = c(-4,-2,0,2,4))+
  labs(title= "Random walk with equally likely outcomes",
       x = "position of particle after 4 steps",
       y = "probability")
```



3. (Random walk revisited)

a. How would the PMF change in problem 2c change if the particle was twice as likely to move to the right as it is to the left? That is, the outcomes are no longer equally likely. Create a table showing the new PMF. Also clearly state any assumption you now need to make in order to re-calculate the probabilities.

Now we are being told that the probability of moving to the right at a particular step is $\frac{2}{3}$. Naturally the probability of moving left is then $\frac{1}{3}$ since those are the only two options for each step.

To find the probability of an outcome, say (L, L, L, L) we can calculate it so long as we assume **independence** between the steps. This means the direction the particle moves on any step is not affected by the direction on any other step. Hence

$$P(L, L, L, L) = P(L) \times P(L) \times P(L) \times P(L) = \frac{1}{3^4}.$$

x	outcomes from S	$f(x)$
-4	(L, L, L, L)	$\frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \frac{1}{3^4}$
-2	(L, L, R, L); (L, L, L, R); (L, R, L, L); (R, L, L, L)	$4 \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = \frac{8}{3^4}$
0	(L,R,L,R); (R, L, R, L); (L, L, R, R); (R, R, L, L); (L, R, R, L); (R, L, L, R)	$6 \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{24}{3^4}$
2	(R, R, R, L); (R, R, L, R); (R, L, R, R); (L, R, R, R)	$4 \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{32}{3^4}$
4	(R, R, R, R)	$\frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \frac{16}{3^4}$

b. Draw the probability histogram of X, the position of the particle after 4 steps in this scenario.

```r
walk_df <- tibble(
  x = c(-4,-2,0,2,4),
  f = c(1/3^4,8/3^4,24/3^4,32/3^4,16/3^4)
)

ggplot(data = walk_df,
       mapping = aes(x = x, y = f))+
  geom_col(width = 1, fill = "gold", color = "purple",
           alpha = 0.5)+
  scale_x_continuous(breaks = c(-4,-2,0,2,4))+
  labs(title= "Right Leaning Random walk",
       x = "position of particle after 4 steps",
       y = "probability")
```

4. Suppose X is a discrete uniform random variable on the integers $1, 2, \ldots, n$, that is, it has PMF

$$f_X(x) = P(X = x) = k, \quad x = 1, 2, \ldots, n$$

where k is some unknown number.

a. Find k

   Since the PMF must add to 1, we can say that

$$\sum_{x=1}^{n} f(x) = 1,$$
$$nk = 1$$
$$k = \frac{1}{n}$$

b. Define the random variable

$$Y = -X$$

   . What is the range of Y? What are the probabilities associated with these values?

   The range of Y is $\{-n, -n+1, \ldots, -3, -2, -1\}$.

   The probabilities of each of these values is just k since Y takes a particular value y only when X takes

$-y$. For example

$$P(Y = -n) = P(X = n) = k$$