Ch 3.1 Practice: Probability Mass Functions

Solutions

1. **Let X be a discrete random variable with probability mass function given by**

X	0	1	2	3
p(x)	1/4	1/2	1/8	1/8

a. **Verify that p is a valid probability mass function.**

The probabilities sum to 1 and they are all positive so it is valid.

b. **Find $P(X \geq 2)$.**

```
1/8+1/8
```

```
## [1] 0.25
```

c. **Find $P(X \geq 2 | X \geq 1)$.**

$P(X \geq 2|X \geq 1) = \frac{(X \geq 2 \cap X \geq 1)}{P(X \geq 1)} = 0.25/0.75 = 1/3$

d. **Find $P(X \geq 2 \cup X \geq 1)$.**

$P(X \geq 2 \cup X \geq 1) = 0.75$

d. **Use simulation to estimate the pmf of X.**

```
x <- 0:3
p_x <- c(0.25,0.5,0.125,0.125)
sample_values <- sample(x,10000,prob=p_x,replace=TRUE)

proportions(table(sample_values))
```

```
## sample_values
##      0      1      2      3
## 0.2555 0.4920 0.1268 0.1257
```

e. **Use your simulated sample from above what is the estimate of $P(X \geq 2)$? Is your result close to the actual value?**

```
mean(sample_values >= 2)
```

```
## [1] 0.2525
```

2. **Let X be a discrete random variable with pmf given by**

X	0	1	2
p(x)	c/4	c/2	c

Find the value of C that makes p a valid probability mass function.

$c(0.25 + .5 + 1) = 1 \rightarrow c = 0.5714$

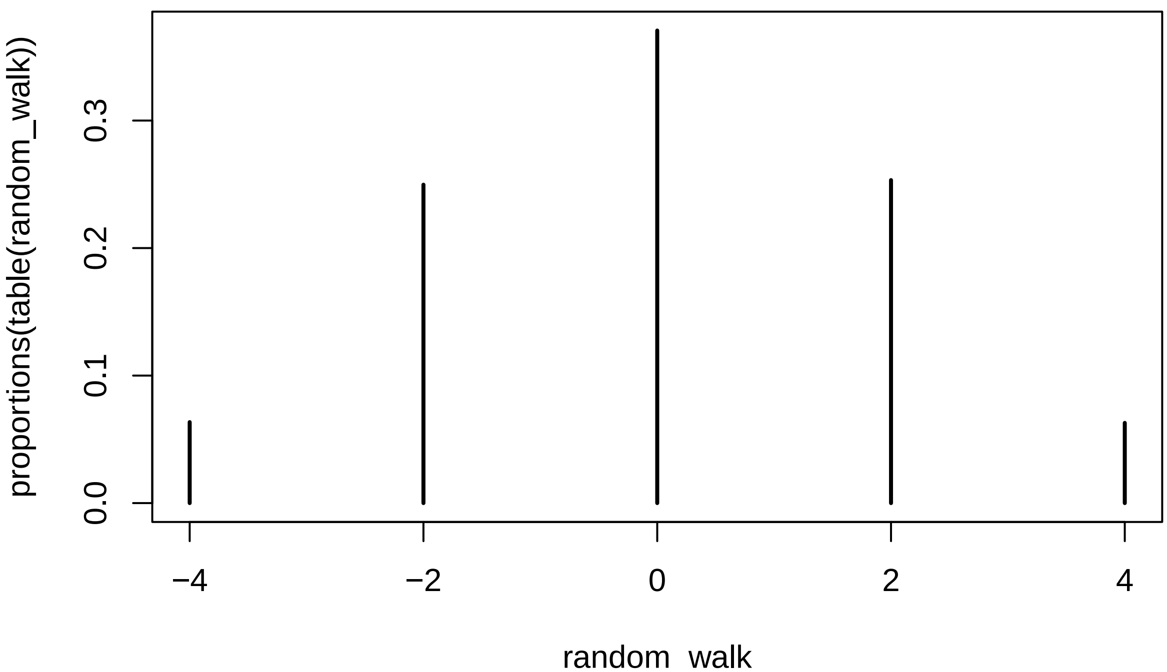
3. Suppose a particle moves along the x-axis beginning at 0. It moves one integer to the left or right with equal probability. Use simulation to estimate the probability distribution of its position after four steps, and plot this distribution.

```
random_walk <- replicate(10000,{
  walks <- sample(c(-1,1),4,replace=TRUE)
  results <- sum(walks)
})

proportions(table(random_walk))
```

```
## random_walk
##     -4     -2      0      2      4
## 0.0635 0.2497 0.3706 0.2533 0.0629
```

```
plot(proportions(table(random_walk)))
```



4. Let's learn a bit about your campus. Go to the Chico State Institutional Research Factbook here: https://www.csuchico.edu/ir/fact-book/index.shtml. Click on the "Enrollment" box. You should now see a dashboard with a bubble chart in the middle. This display shows the percent of students currently enrolled in majors from each college. Click on the tiny little bubble and select "Exclude" because we can't tell what college those students are from. So lets not worry about them for now.

Use the percentages of students enrolled in each college at our university to simulate the probability that in a group of 10 students selected at random, at least one other person is in the same major as you.

```
colleges <- c("ECC", "BUS", "AGR", "CME", "HFA", "BSS", "NSC", "UGRD")
college_prop <- c(.1553, .1422, .0579, .1601, .0669, .2553, .1194, .0431)

same.college <- replicate(10000, {
  students <- sample(colleges, size=30, prob=college_prop, replace=TRUE)
  at.least.1.stats <- sum(students == "NSC") >=1
})

mean(same.college)
```

```
## [1] 0.982
```

5. Let X be the number of siblings of Chico State students. Ask at least 15 other students how many siblings they have and report the pmf below.

I asked 15 fellow faculty and stored their responses in a variable called `siblings`.

```
siblings <- c(0,0, 1, 3, 2, 1, 3, 0, 7, 1, 1, 3, 0, 0, 2)
```

There are 4 unique values; 0, 1, 2, 3 and 7. So the PMF would look like:

X	0	1	2	3	7
p(x)	5/15	4/15	2/15	3/15	1/15

6. Fifty people put their names in a hat. They then all randomly choose one name from the hat. Let X be the number of people who get their own name. Estimate the pmf of X using simulation and create a plot

```
simulate.x <- replicate(10000, {
  hat <- 1:50
  names <- sample(hat, 50, replace = FALSE)
  x <- sum(hat == names)
})

proportions(table(simulate.x))
```

```
## simulate.x
##      0      1      2      3      4      5      6      7
## 0.3703 0.3612 0.1863 0.0625 0.0161 0.0027 0.0007 0.0002
```

```
plot(proportions(table(simulate.x)))
```