Lab3 Probability

Sam Reeves

2021-02-16

```
library(tidyverse)
library(openintro)
```
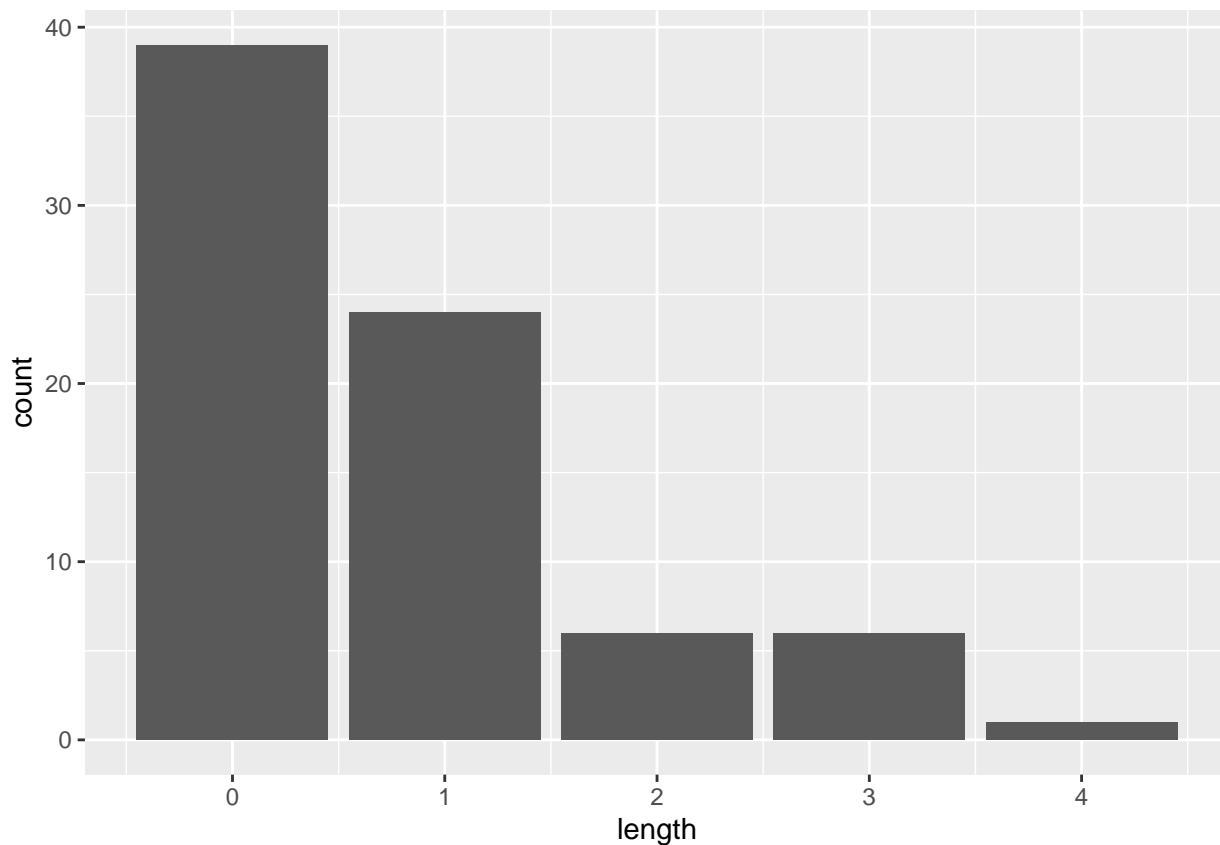
Exercise 1

What does a streak length of 1 mean, i.e. how many hits and misses are in a streak of 1? What about a streak length of 0?

A streak of length 0 is a miss or set of consecutive misses. A streak of length 1 is the first made shot after a miss.

Exercise 2

Describe the distribution of Kobe's streak lengths from the 2009 NBA finals. What was his typical streak length? How long was his longest streak of baskets? Make sure to include the accompanying plot in your answer.

```
kobe_streak <- calc_streak(kobe_basket$shot)
ggplot(data = kobe_streak, aes(x = length)) + geom_bar()
```

It looks like Kobe's most common streak length was 0. He missed a lot of shots and kept trying. His longest streak of baskets was 4 shots.

Exercise 3

In your simulation of flipping the unfair coin 100 times, how many flips came up heads? Include the code for sampling the unfair coin in your response. Since the markdown file will run the code, and generate a new sample each time you Knit it, you should also "set a seed" before you sample. Read more about setting a seed below.

```
outcomes <- c("heads", "tails")
set.seed(1337)

sim_unfair_coin <- sample(outcomes, size = 100,
                          replace = TRUE, prob = c(0.2, 0.8))

heads <- data.frame(sim_unfair_coin) %>%
  filter(sim_unfair_coin == "heads")

length(heads[[1]])
```

```
## [1] 25
```

25 flips came up heads.

Exercise 4

What change needs to be made to the sample function so that it reflects a shooting percentage of 45%? Make this adjustment, then run a simulation to sample 133 shots. Assign the output of this simulation to a new object called sim_basket.

```
set.seed(1337)
outcomes <- c("H", "M")
sim_basket <- sample(outcomes, size = 133,
                     replace = TRUE, prob = c(0.45, 0.55))
```
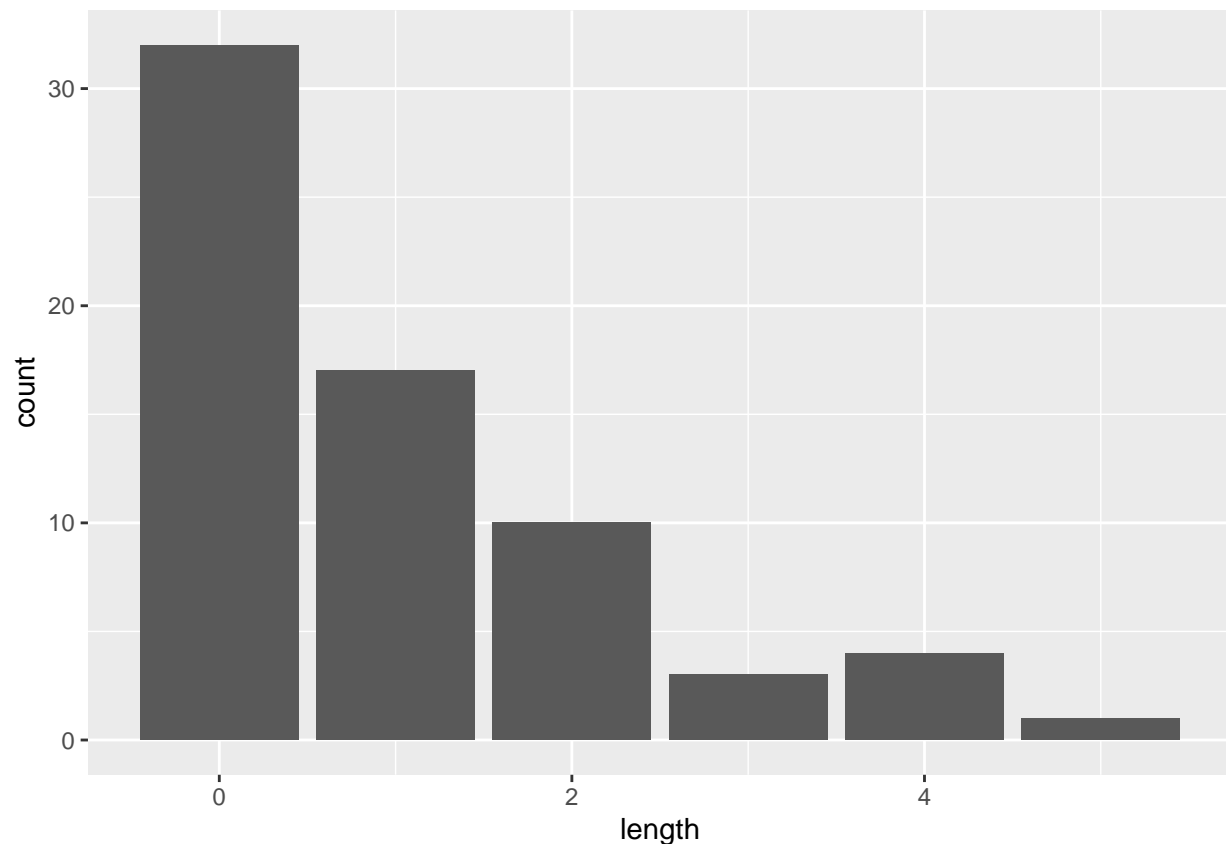
Exercise 5

Using calc_streak, compute the streak lengths of sim_basket, and save the results in a data frame called sim_streak.

```
sim_streak <- data.frame(calc_streak(sim_basket))
```

Exercise 6

Describe the distribution of streak lengths. What is the typical streak length for this simulated independent shooter with a 45% shooting percentage? How long is the player's longest streak of baskets in 133 shots? Make sure to include a plot in your answer.

```
ggplot(data = sim_streak, aes(x = length)) + geom_bar()
```


The typical streak length is 0, and the distribution is skewed right. It looks very similar to Kobe's distribution. The longest streak in the simulation is 5, this only happened one time.

Exercise 7

If you were to run the simulation of the independent shooter a second time, how would you expect its streak distribution to compare to the distribution from the question above? Exactly the same? Somewhat similar? Totally different? Explain your reasoning.

I would set a different seed, and I'd expect the distribution to be roughly similar, with some variation. I expect the results to be mostly the same, but perhaps in a different order. Because each shot is independent, I don't imagine any kindling effects similar to the "hot hand".

Exercise 8

How does Kobe Bryant's distribution of streak lengths compare to the distribution of streak lengths for the simulated shooter? Using this comparison, do you have evidence that the hot hand model fits Kobe's shooting patterns? Explain.

Kobe has significantly more missed shots and streaks of length 1 than our simulation with seed 1337. He also never reached a streak beyond length 4. Considering these facts, and that the probabilities of the outcomes in our simulation are based off of Kobe's career-average percentage, I might even suggest he had a bit of a cold hand.