

Lab3

Avery Davidowitz

2022-09-24

Import

```
library(tidyverse)
library(openintro)
```

Set seed

```
set.seed(40)
```

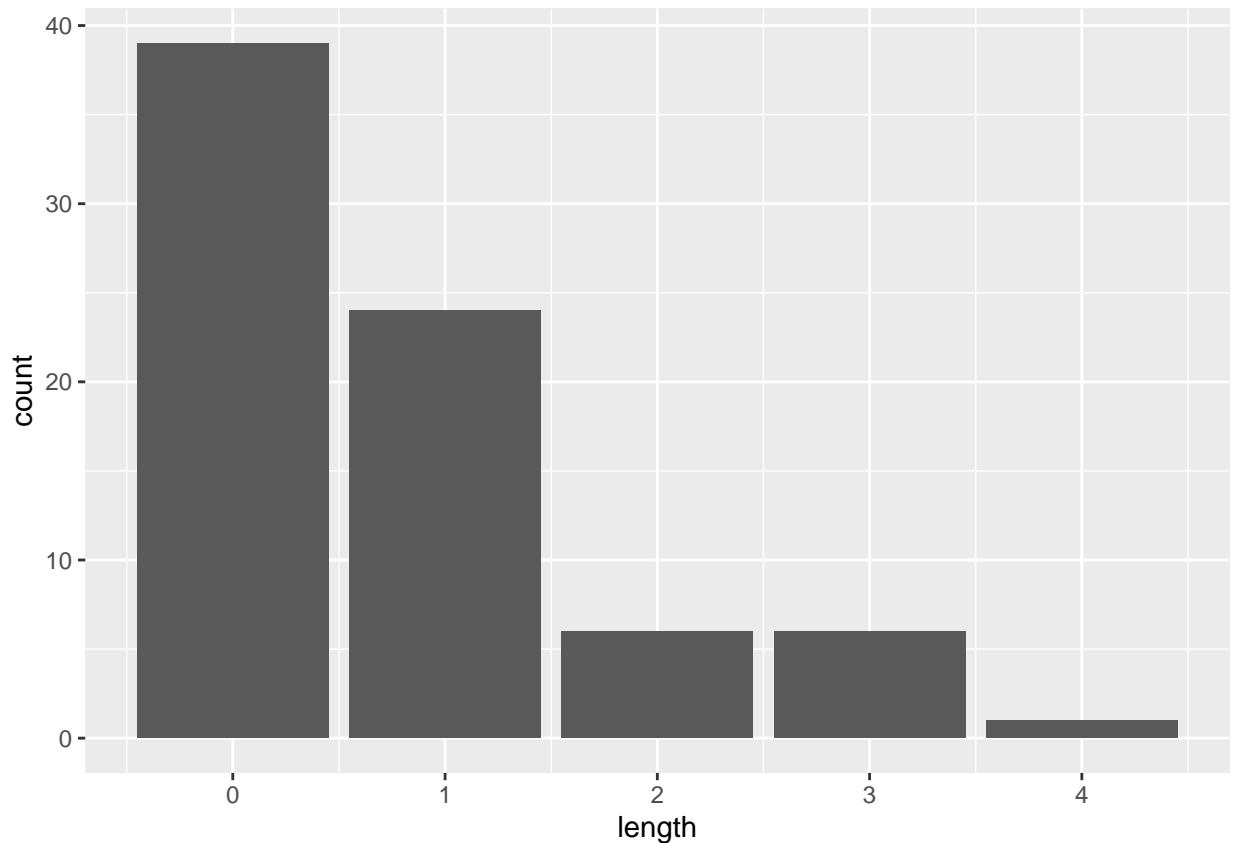
Exercise 1

A streak of length 1 represents making 1 shot and then missing the next shot (HM). A streak of length 0 represents a miss after another miss (M). If the miss was after a hit it wouldn't be part of a new streak.

Exercise 2

As seen in the plot and summary statistics, Kobe typically did not have many long streaks. Mostly Kobe made 1 or missed 1 in the 2009 finals with only a few exceptions. His longest streak was hitting 4 shots in a row.

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



```
kobe_streak %>% summarize(avg=mean(kobe_streak$length), standard_deviation=sd(kobe_streak$length), longest=max(kobe_streak$length))
```

```
##           avg standard_deviation longest
## 1 0.7631579      0.9915432         4
```

Exercise 3

The simulation with a seed of 40 results in 18 heads.

```
coin_outcomes <- c("heads", "tails")
sim_unfair_coin <- sample(coin_outcomes, size = 100, replace = TRUE,
                          prob = c(0.2, 0.8))
table(sim_unfair_coin)
```

```
## sim_unfair_coin
## heads tails
##    18    82
```

Exercise 4

I needed to add the prob parameter with a .45 success rate for “H”s and had to change the sample size to 133 to match the original Kobe data set.

```
shot_outcomes <- c("H", "M")
sim_basket <- sample(shot_outcomes, size = 133, replace = TRUE,
                    prob = c(0.45, 0.55))
table(sim_basket)
```

```
## sim_basket
## H M
## 57 76
```

Exercise 5

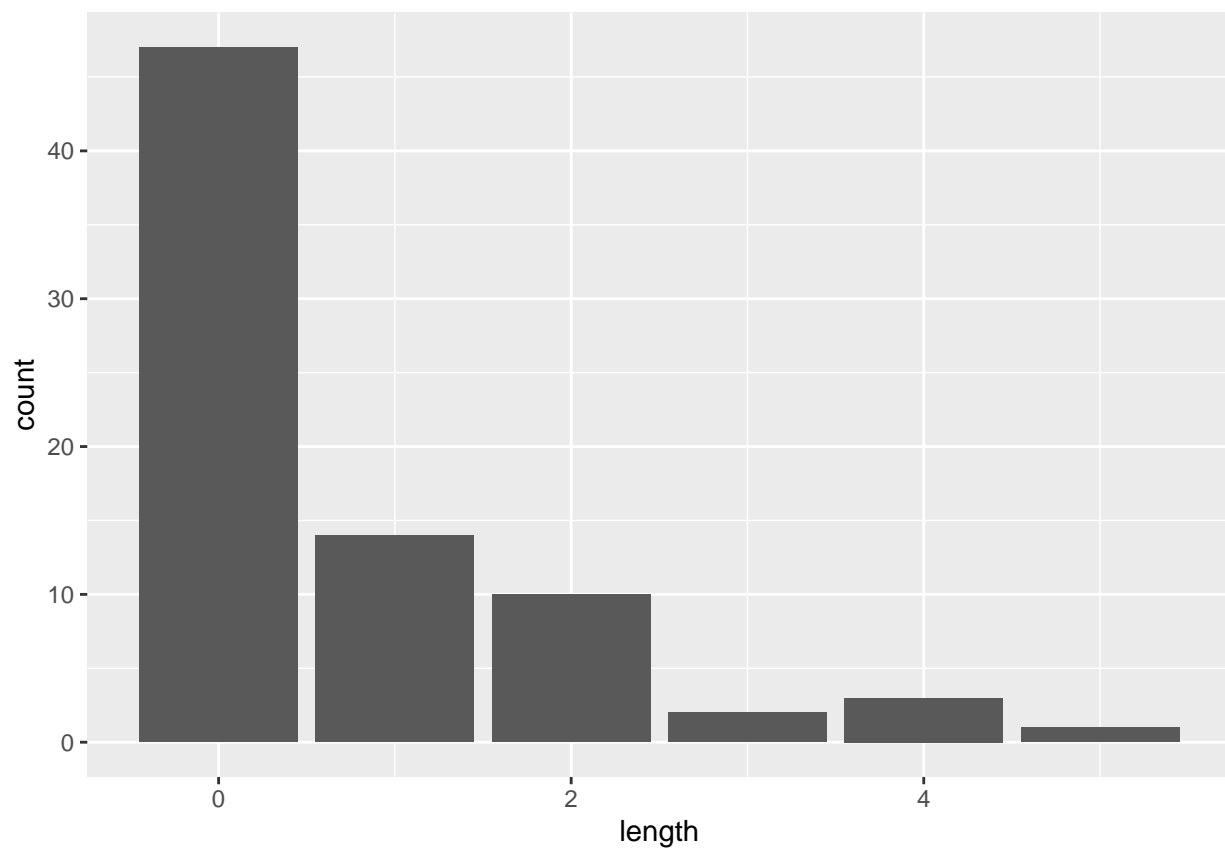
I compute the streak lengths of the simulation.

```
sim_streak <- calc_streak(sim_basket)
```

Exercise 6

The same plot applied to the simulation.

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



```
sim_streak %>% summarize(avg=mean(sim_streak$length), standard_deviation=sd(sim_streak$length), longest=
```

```
##           avg standard_deviation longest
## 1 0.7402597           1.163056         5
```

Exercise 7

Most likely when running the simulation again the results would be somewhat similar. Since we have a sample size of 133 even if there were a few outliers it would generally be the same overall distribution. However, it would be possible to sample 133 times and get drastically different results.

Exercise 8

The distribution between Kobe's finals performance and the simulation are somewhat similar with close summary statistics. The mode remains streaks of length 0 and the streaks are skewed left. However, the quantity of streaks of 1 length are noticeably fewer and we have the maximum streak of 5 vs 4. Ideally, in order to completely eliminate chance from the simulation, I would rerun the simulation many times and normalize the results before comparison to the Kobe data. 1 simulation does not provide enough evidence to determine the truth about "hot hands."