Lab 8 - Basketball and Hot Hands

Minsu Kang and Bomin Lyoo

2022-07-17

Lab report

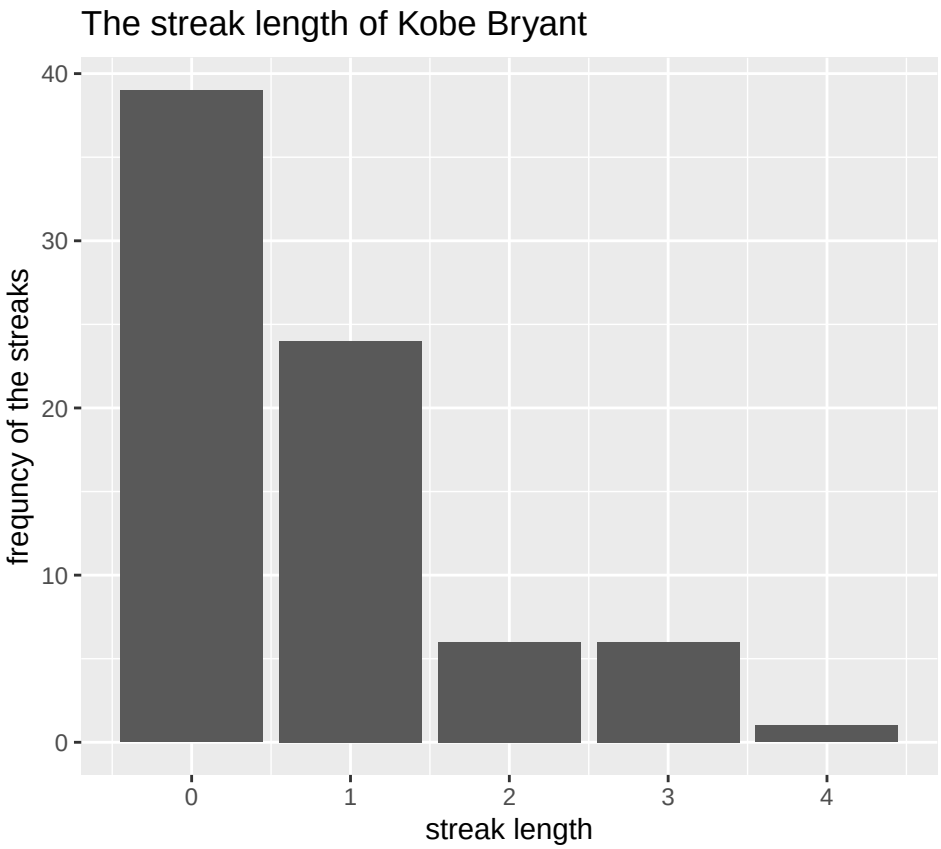
Exercises

Exercise 1

```
i) The streak length means how many shots constantly made a basket.
Streak one means there was only one hit and missed second shot.
ii) The streak zero means missed first shot.(just 'M')
This indicates no hit in a streak.
```

Exercise 2

```
kobe_streak <- calc_streak(kobe_basket$shot)
ggplot(data = kobe_streak) +
  geom_bar(mapping = aes(x = length)) +
  labs(title="The streak length of Kobe Bryant",
       x= "streak length",
       y= "frequncy of the streaks")
```



This bar graph shows Uni-modal, asymmetrical,
right-skewed distribution of streak lengths.
The zero streak is most frequent, as higher streak is hard,
they get lower volumes compare to 0 streaks.
The longest streak length is 4. Kobe constantly shoot 4 hits and missed next shot
and that is the best streak in first quarter of the game.

Exercise 3

```
set.seed(35797)
coin_outcomes <- c("heads", "tails")
sample(coin_outcomes, size = 1, replace = TRUE)
```

```
## [1] "tails"
```

```
sim_fair_coin <- sample(coin_outcomes, size = 100, replace = TRUE)
table(sim_fair_coin)
```

```
## sim_fair_coin
## heads tails
##    56    44
```

```
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
##    17    83
```

i) In the fair coin simulation, we got 56 heads cases in 100 times flipping. In unfair coin simulation, there was 17 head cases, because we adjusted 0.8 possibility to the tail side in weighted coin(unfair coin).

Exericse 4

```
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
## 64 69
```

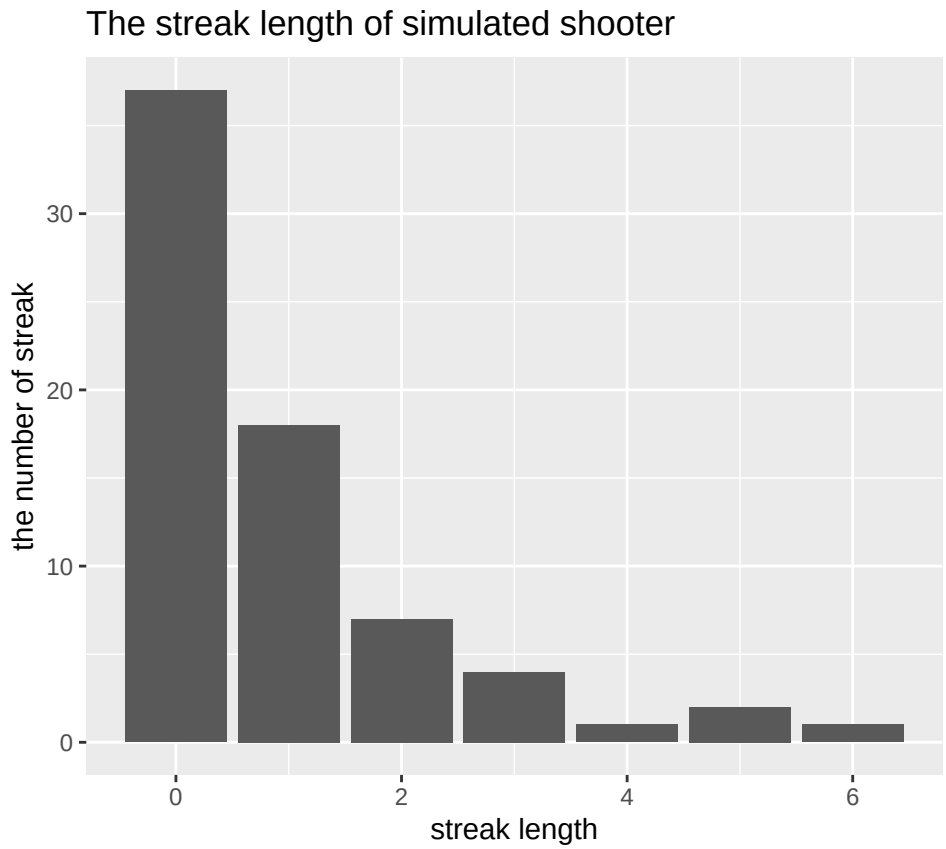
Additional Question 1

```
sim_df <- data.frame(sim_basket)
streak_length_sim_basket <- calc_streak(sim_df)
```

i) As the sim_basket is values, we assigned as a data frame form to the "sim_df", and pass in to cal_streak function that we used for Kobe's shot data.

Question 2

```
ggplot(data = streak_length_sim_basket) +
  geom_bar(mapping = aes(x = length)) +
  labs(title="The streak length of simulated shooter",
       x= "streak length",
       y= "the number of streak")
```




```
i) The simulated shooter streak distribution grabh is Uni-modal, asymmetrical,
right-skewed distribution of streak lengths.
The most frequent streak is 0, and tipically high streak length is lower numbers.
The longest streak is 6.
The longest streak of simulated shooter hits 6times in one streak in 133 times shootings.
```

Question 3

```
i) The distribution is same in every reproduce.
This is because we used set.seed in former code chunk.
seed contains the same results of simulation.
As so, we used set.seed, we can get same result after knitting.
(the result is not changed after the knitting and multiple trying of simulations)
```

Question 4

```
i) Compare simulated shooter and Kobe Bryant,
we can see the distribution is asymmetrical right-skewed,
and distribution is almost same.
We guess Kobe Bryant seems like shooting dependent shot to game audience,
because after game most of them remember the scene he is on fire
```

, when he got constantly made shots.
However, the overall game data stat shows even worse shapes compare to simulation shooter, who shooting independent shots.
So it proves the 'hot hand' is not exist in real life.