

Hao-2

Bruce Hao

September 7, 2016

```
load("more/kobe.RData")
```

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 1 means one and only one H before a M. A streak of zero means no H before the next M.

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?

The distribution of Kobe's streaks are right skewed and concentrated at zero and one. The most typical streak length was zero. The longest streak was 4 baskets.

Exercise 3: In your simulation of flipping the unfair coin 100 times, how many flips came up heads?

```
outcomes <- c('Heads', 'Tails')
table(sample(outcomes, size = 100, replace = TRUE, prob=c(0.2, 0.8)))
```

```
##
## Heads Tails
##    33    67
```

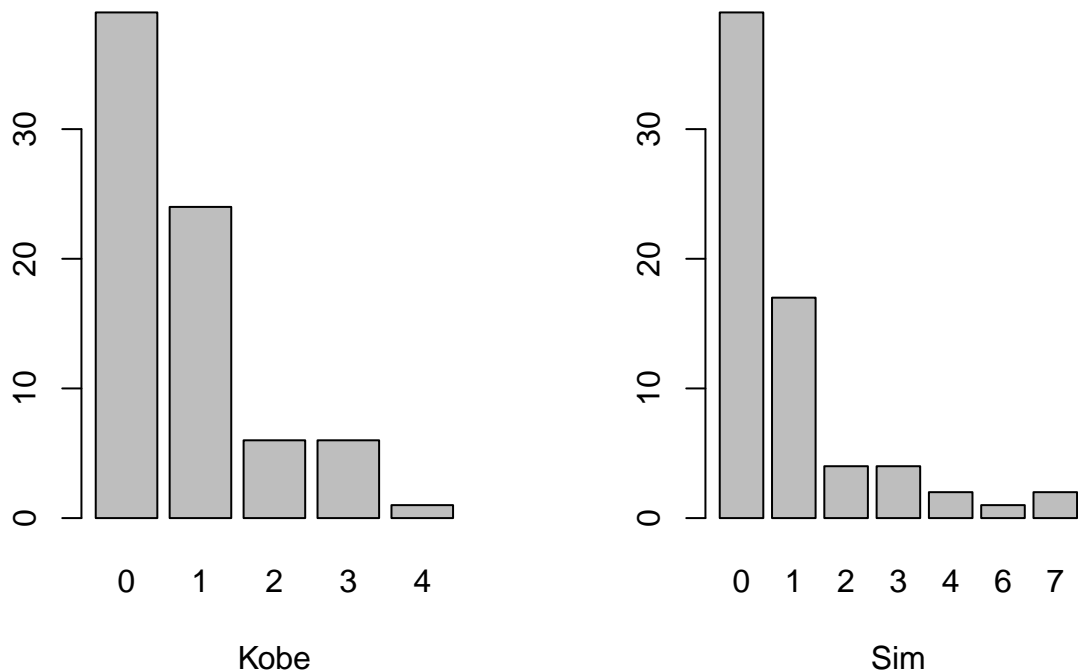
Exercise 4: What change needs to be made to the sample function so that it reflects a shooting percentage of 45%? Make this adjustment, then run a simulation to sample 133 shots. Assign the output of this simulation to a new object called `sim_basket`.

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

```
## sim_basket
##  H  M
## 65 68
```

On your own 1: 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?

```
par(mfrow=c(1,2))
barplot(table(calc_streak(kobe$basket)), xlab = 'Kobe')
barplot(table(calc_streak(sim_basket)), xlab = 'Sim')
```



Based on the charts above, the simulated shooter has more streaks of zero. The typical streak is zero, and the longest streaks is 5 baskets.

On your own 2: 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.

Probably somewhat similar. A sample of 133 observations should be sufficient to ensure that the simulations are do not differ substantially from one another.

On your own 3: 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.

The distribution of Kobe's streaks seem to include more streaks of 2 and 3 than that of the simulation. But the differences are fairly small optically and woud have to be substantiated with more rigorous statistical analysis to determine if there is a real difference or not.