

Lab3

```
download.file("http://www.openintro.org/stat/data/kobe.RData", destfile =
"kobe.RData")
load("kobe.RData")
head(kobe)
```

```
##      vs game quarter time
## 1 ORL      1         1 9:47
## 2 ORL      1         1 9:07
## 3 ORL      1         1 8:11
## 4 ORL      1         1 7:41
## 5 ORL      1         1 7:03
## 6 ORL      1         1 6:01
##                                     description basket
## 1                                Kobe Bryant makes 4-foot two point shot      H
## 2                                Kobe Bryant misses jumper                    M
## 3                                Kobe Bryant misses 7-foot jumper              M
## 4 Kobe Bryant makes 16-foot jumper (Derek Fisher assists)                    H
## 5                                Kobe Bryant makes driving layup              H
## 6                                Kobe Bryant misses jumper                    M
```

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?

```
kobe$basket[1:9] # Kobe's sequence of hits and misses.
## [1] "H" "M" "M" "H" "H" "M" "M" "M" "M"
```

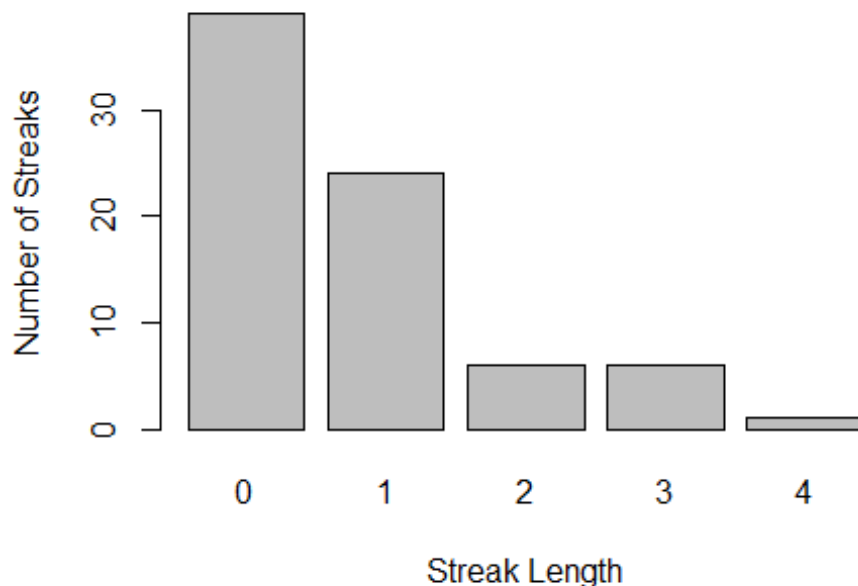
Streak length of 1 means a hit (H) followed by a miss (M). A streak 0 means zero hits and one miss (M).

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?

Generating the Length of Kobe's Shooting Streak

```
kobe_streak <- calc_streak(kobe$basket)
barplot(table(kobe_streak), xlab = "Streak Length", ylab = "Number of
Streaks", main = "Length of All Shooting Streaks (Kobe Bryant)")
```

Length of All Shooting Streaks (Kobe Bryant)



The distribution is skewed right with the streak lengths ranging from 0 to 4 with a median of 0.

```
summary(kobe_streak) # 5 number Summary of Length of ALL Shooting Streaks (Kobe Bryant)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0000  0.0000  0.0000  0.7632  1.0000  4.0000
```

Kobe's typical streak is 0. His longest is 4.

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

Simulating an unfair coin flip 100 times (Heads = 20%, Tails = 80%)

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

## heads
##      17
```

Heads was flipped 17 times.

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

Simulating an independent shooter with 133 shots (Hits = 45%, Misses = 55%).

```
outcomes_bball <- c("H", "M")
sim_basket <- sample(outcomes_bball, size = 133, replace = TRUE, prob =
c(.45, .55))
```

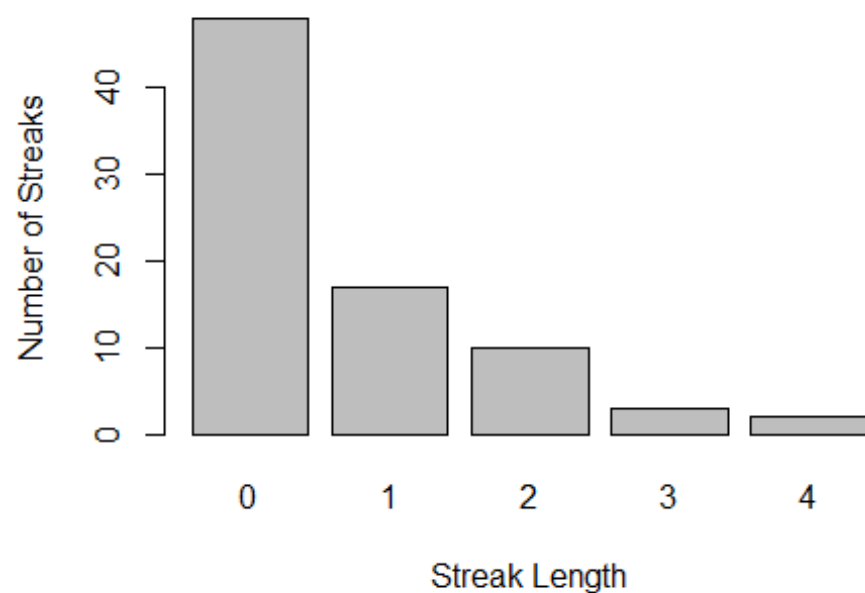
Homework

1. Produce a graph and describe the distribution of streak lengths for the independent shooter. 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?

Generating the Length of Independent Shooter's Shooting Streak

```
sim_streak <- calc_streak(sim_basket)
barplot(table(sim_streak), xlab = "Streak Length", ylab = "Number of Streaks",
main = "Length of All Shooting Streaks (Independent Shooter)")
```

Length of All Shooting Streaks (Independent Shoot



```
summary(sim_streak) # 5 number Summary of Length of ALL Shooting Streaks  
(Independent Shooter)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.   
##  0.000   0.000   0.000   0.675   1.000   4.000
```

The distribution is skewed right with the streak lengths ranging from 0 to 4 with a median of 0. The typical streak length is 0. The longest streak of baskets in 133 shots is 4

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.

If the simulation was a second time, I would expect the results to be somewhat similar. The shooter will still take 133 shots and will still have a hit percentage of 45%. The length of the streak could increase or decrease. However, the median would still be 0.

3. How does Kobe Bryant's distribution of streak lengths from page 2 compare to the distribution of streak lengths for the simulated independent shooter? Based on this comparison, do you have evidence that the hot hand model fits Kobe's shooting patterns? Explain your reasoning.

Both Kobe and the independent shooter's distribution models are right skewed. Both distributions range from 0 to 4 and have a median of 0. The hot hand model does not apply because each shot is independent. The success of any subsequent shot does not increase with the success of the previous event. If the hot hand model applied, the distribution of the bar blot would reflect either a symmetrical or left skewed graph.