Lab2

## HotHands Analysis

### Analyzing kobe dataset

load("more/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

kobe$basket[1:9]

## [1] "H" "M" "M" "H" "H" "M" "M" "M" "M"

Comparing data

outcomes <- c("H", "M")  
sim\_basket <- sample(outcomes, size = 1, replace = TRUE)  
kobe$basket

## [1] "H" "M" "M" "H" "H" "M" "M" "M" "M" "H" "H" "H" "M" "H" "H" "M" "M"  
## [18] "H" "H" "H" "M" "M" "H" "M" "H" "H" "H" "M" "M" "M" "M" "M" "M" "H"  
## [35] "M" "H" "M" "M" "H" "H" "H" "H" "M" "H" "M" "M" "H" "M" "M" "H" "M"  
## [52] "M" "H" "M" "H" "H" "M" "M" "H" "M" "H" "H" "M" "H" "M" "M" "M" "H"  
## [69] "M" "M" "M" "M" "H" "M" "H" "M" "M" "H" "M" "M" "H" "H" "M" "M" "M"  
## [86] "M" "H" "H" "H" "M" "M" "H" "M" "M" "H" "M" "H" "H" "M" "H" "M" "M"  
## [103] "H" "M" "M" "M" "H" "M" "H" "H" "H" "M" "H" "H" "H" "M" "H" "M" "H"  
## [120] "M" "M" "M" "M" "M" "M" "H" "M" "H" "M" "M" "M" "M" "H"

sim\_basket

## [1] "M"

Both data sets represent the results of 133 shot attempts, each with the same shooting percentage of 45%. We know that our simulated data is from a shooter that has independent shots. That is, we know the simulated shooter does not have a hot hand.

### Exercises Questions

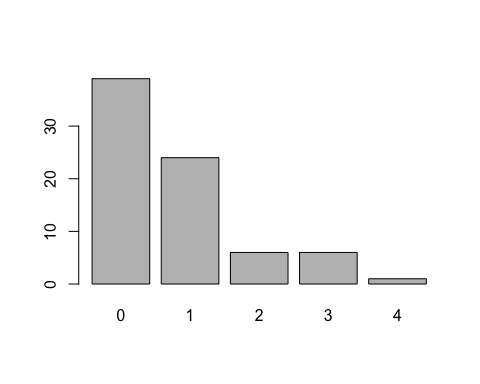
1. What does a streak length of 1 mean, i.e. how many hits and misses are in a streak of 1? What about a streak length of 0?

Answer 1. A streak of length 0: No hit/ zero hit and followed by one miss. A streak of length 1 : One hit and followed by miss.

1. 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?

Answer 2. The distribution is left skewed The typical streak lenght is 0 and the longest streak is 4

kobe\_streak <- calc\_streak(kobe$basket)  
barplot(table(kobe\_streak))



1. In your simulation of flipping the unfair coin 100 times, how many flips came up heads?

sim\_unfair\_coin <- sample(outcomes, size = 100, replace = TRUE, prob = c(0.2, 0.8))  
sim\_unfair\_coin

## [1] "M" "H" "M" "M" "M" "M" "H" "M" "M" "M" "H" "M" "M" "H" "M" "M" "H"  
## [18] "M" "M" "M" "M" "H" "H" "M" "M" "M" "M" "M" "M" "H" "M" "H" "M" "M"  
## [35] "H" "H" "M" "M" "M" "M" "M" "M" "H" "M" "M" "M" "H" "M" "M" "M" "M"  
## [52] "H" "M" "M" "H" "H" "M" "M" "M" "H" "M" "M" "H" "M" "M" "M" "M" "M"  
## [69] "M" "M" "M" "M" "H" "H" "M" "M" "M" "M" "H" "M" "M" "M" "M" "H" "H"  
## [86] "M" "M" "M" "M" "M" "M" "H" "M" "M" "M" "M" "H" "M" "H" "M"

table(sim\_unfair\_coin)

## sim\_unfair\_coin  
## H M   
## 26 74

Answer 3. 17 heads

1. 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)

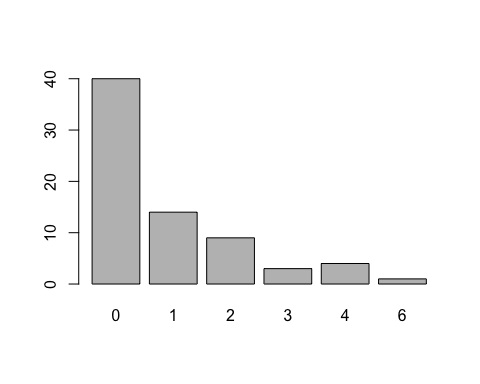
## sim\_basket  
## H M   
## 63 70

### On your own

Comparing Kobe Bryant to the Independent Shooter

Using calc\_streak, compute the streak lengths of sim\_basket.

sim\_basket\_len <- calc\_streak(sim\_basket)  
barplot(table(sim\_basket\_len))



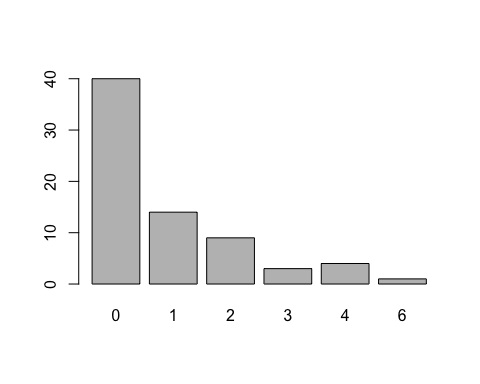
* Describe the distribution of streak lengths. What is the typical streak length for this simulated independent shooter with a 45% shooting percentage? How long is the player’s longest streak of baskets in 133 shots?

Answer) The streak length is left skewed. Typical streak length : 0 Longest streak of baskets in 133 shots: 5

* 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.

Answer)

sim\_basket\_len <- calc\_streak(sim\_basket)  
barplot(table(sim\_basket\_len))



The distribution should be same as the probablilty between two simulations does not change and the shoots are independent of each other.

* 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.

Answer)

Kobe Bryant’s distribution of streak lengths is similar when compared to the distribution of streak lengths for the simulated shooter. Both the data has typical streak of 0. And both data has barplots as left skewed. There is not much evidence that hot hand model fits kobe’s shooting pattern. But it is analyzed by results that the shoots are independent of each other both for Kobe Bryant’s distribution and simulated distribution.