da460\_lab2\_grahn

jason grahn

10/2/2018

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?

#head of the basket values  
kobe$basket[1:9]

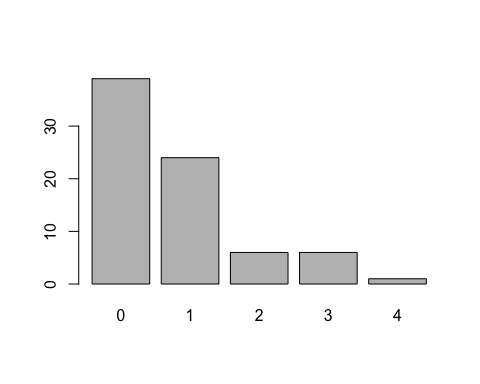
## [1] "H" "M" "M" "H" "H" "M" "M" "M" "M"

A streak length of 1 means that Kobe made one basket and it was immediately followed by a miss; 1 hit and 1 miss. A streak length of zero is a miss that follows another miss.

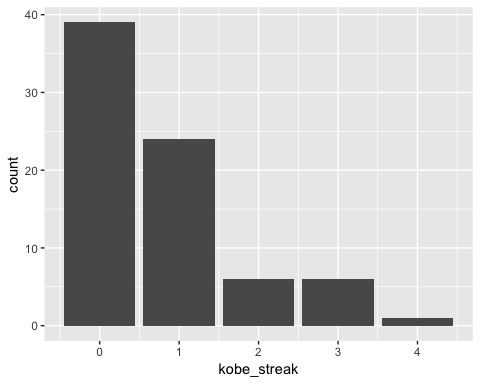
## Exercise 2

### 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?

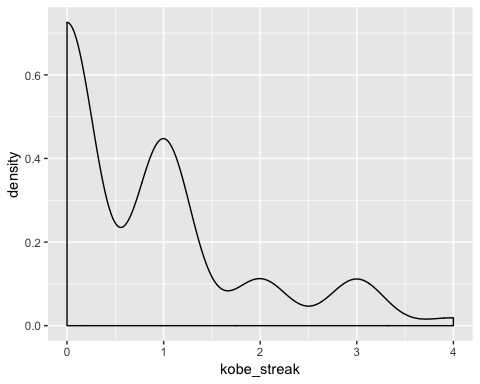
#run the calc\_streak function to generate the kobe\_steak dataset  
kobe\_streak <- calc\_streak(kobe$basket)  
  
#provide a barplot of the kobe\_streak data  
barplot(table(kobe\_streak))



#or do the same tidy style  
tibble(kobe\_streak) %>%   
 ggplot(aes(x = kobe\_streak)) +  
 geom\_bar()



#but really this is better as a density plot  
tibble(kobe\_streak) %>%   
 ggplot(aes(x = kobe\_streak)) +  
 geom\_density()



I dont think a streak of *zero* should count, so his typical streak length is *1*. His longest streak is 4.

## Exercise 3

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

#build the outcomes vector  
outcomes <- c("heads", "tails")  
  
#build a dataset of a fair coin toss  
sim\_fair\_coin <- sample(outcomes, size = 100, replace = TRUE)  
table(sim\_fair\_coin)

## sim\_fair\_coin  
## heads tails   
## 52 48

#build a dataset of an UNFAIR coin toss  
sim\_unfair\_coin <- sample(outcomes, size = 100, replace = TRUE, prob = c(0.2, 0.8))  
table(sim\_unfair\_coin)

## sim\_unfair\_coin  
## heads tails   
## 13 87

In the unfair coin toss, 26 flips came back heads.

## 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.

sim\_basket <- sample(outcomes, size = 133, replace = TRUE, prob = c(0.45, (1-.45)))  
table(sim\_basket)

## sim\_basket  
## heads tails   
## 58 75

table(kobe$basket)

##   
## H M   
## 58 75

table(sim\_basket)

## sim\_basket  
## heads tails   
## 58 75