

lab-3-probability

Pradeep Paladugula

2/4/2020

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

```
load("kobe.RData") # Kobe's shooting stats in few games
head(kobe)
```

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

```
tail(kobe)
```

```
##      vs game quarter    time      description basket
## 128 ORL      3         4    3:57 Bryant Jump Shot: Made (28 PTS)      H
## 129 ORL      3         4    3:33      Bryant Layup Shot: Missed      M
## 130 ORL      3         4    2:02      Bryant 3pt Shot: Missed      M
## 131 ORL      3         4 00:23.9 Bryant 3pt Shot: Missed      M
## 132 ORL      3         4 00:06.9 Bryant 3pt Shot: Missed      M
## 133 ORL      3         4 00:00.5 Bryant Layup Shot: Made (31 PTS)      H
```

```
dim(kobe)
```

```
## [1] 133    6
```

```
kobe$basket[1:9] ## what is another way to get the same data???
```

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

```
kobe[1:9,6]
```

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

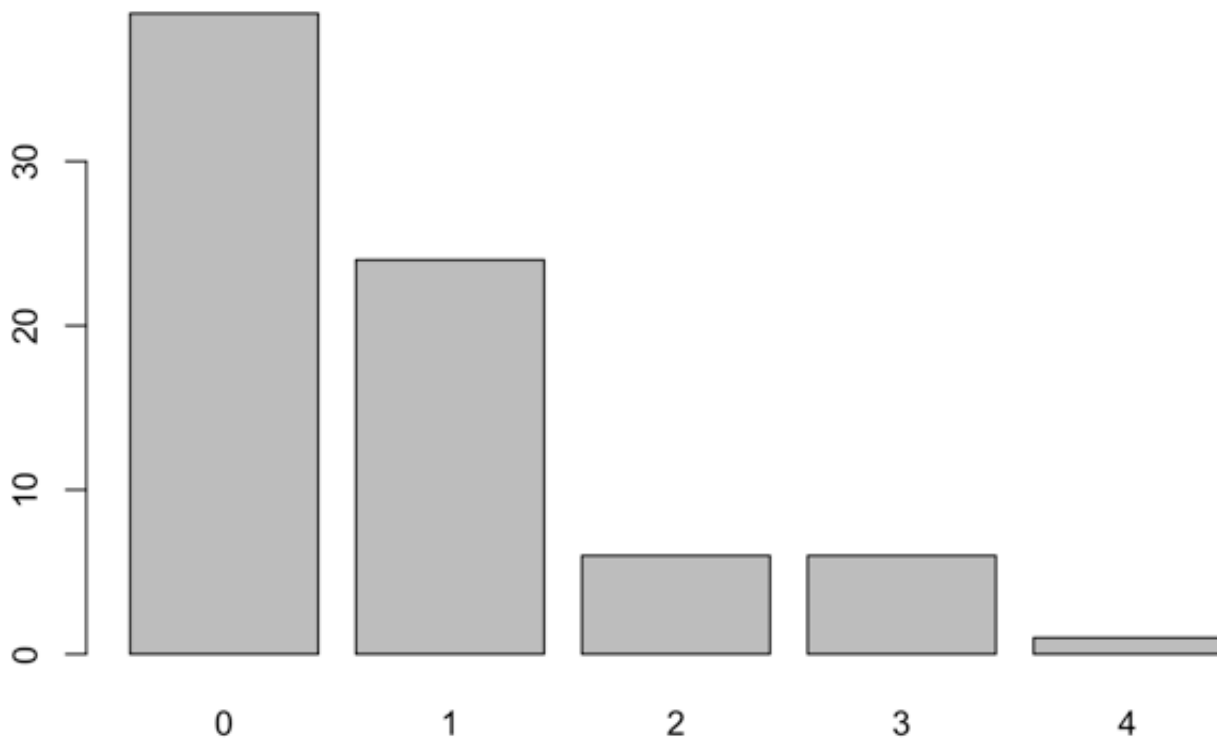
1. What does a streak length of 1 mean, i.e. how many hits and misses are in a streak of 1? What about a streak length of 0?

Solution: 1 H and followed by miss in a streak length of 1 mean. 0 H in a streak length of 0.

```
kobe_streak <- calc_streak(kobe$basket)
kobe_streak
```

```
## [1] 1 0 2 0 0 0 3 2 0 3 0 1 3 0 0 0 0 0 1 1 0 4 1 0 1 0 1 0 1 2 0 1 2 1 0 0 1 0
## [39] 0 0 1 1 0 1 0 2 0 0 0 3 0 1 0 1 2 1 0 1 0 0 1 3 3 1 1 0 0 0 0 0 1 1 0 0 0 1
```

```
barplot(table(kobe_streak))
```



```
table(kobe$basket)/133
```

```
##
##           H           M
## 0.4360902 0.5639098
```

- 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?

Solution: The Distribution of the Kobe's streak lengths are skewed to the right and also a unimodal distributions. the typical streak length is 0 and the longest streak of baskets is 4.

- In the simulation of flipping the unfair coin 100 times, how many flips came up heads? Read more on the sample function in help. ?sample()

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

```
## [1] "tails" "tails" "tails" "tails" "tails" "heads" "heads" "tails" "heads"
## [10] "heads" "heads" "tails" "tails" "tails" "tails" "tails" "heads" "tails"
## [19] "tails" "tails" "tails" "tails" "tails" "tails" "tails" "tails" "tails"
## [28] "tails" "heads" "heads" "heads" "tails" "tails" "tails" "tails" "tails"
## [37] "tails" "tails" "tails" "tails" "heads" "tails" "tails" "tails" "tails"
## [46] "heads" "heads" "heads" "tails" "heads" "tails" "tails" "tails" "tails"
## [55] "tails" "heads" "tails" "tails" "tails" "heads" "tails" "tails" "tails"
## [64] "tails" "tails" "heads" "tails" "heads" "tails" "tails" "tails" "tails"
## [73] "tails" "heads" "tails" "tails" "heads" "tails" "tails" "tails" "heads"
## [82] "tails" "tails" "tails" "heads" "tails" "tails" "tails" "tails" "tails"
## [91] "tails" "tails" "tails" "tails" "heads" "heads" "tails" "tails" "tails"
## [100] "tails"
```

```
table(sim_unfair_coin)
```

```
## sim_unfair_coin
## heads tails
##      24     76
```

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(.45, 0.55))
sim_basket
```

```
## [1] "H" "H" "M" "H" "M" "H" "M" "M" "M" "M" "M" "M" "H" "H" "H" "H" "H" "M"
## [19] "M" "H" "M" "H" "H" "M" "M" "M" "H" "H" "M" "M" "M" "M" "M" "M" "M" "H"
## [37] "M" "H" "M" "H" "M" "M" "H" "H" "M" "H" "M" "M" "H" "M" "M" "M" "H" "M"
## [55] "M" "M" "H" "H" "H" "M" "H" "M" "H" "M" "H" "H" "M" "M" "M" "M" "M" "M"
## [73] "M" "H" "M" "H" "M" "H" "M" "H" "M" "M" "H" "H" "M" "H" "M" "H" "H" "H"
## [91] "H" "H" "M" "H" "H" "M" "M" "H" "M" "H" "M" "M" "M" "M" "H" "H" "M" "H"
## [109] "M" "M" "M" "M" "H" "M" "H" "H" "H" "M" "H" "H" "H" "H" "M" "M" "M" "H"
## [127] "M" "M" "M" "M" "M" "M" "M"
```

```
table(sim_basket)
```

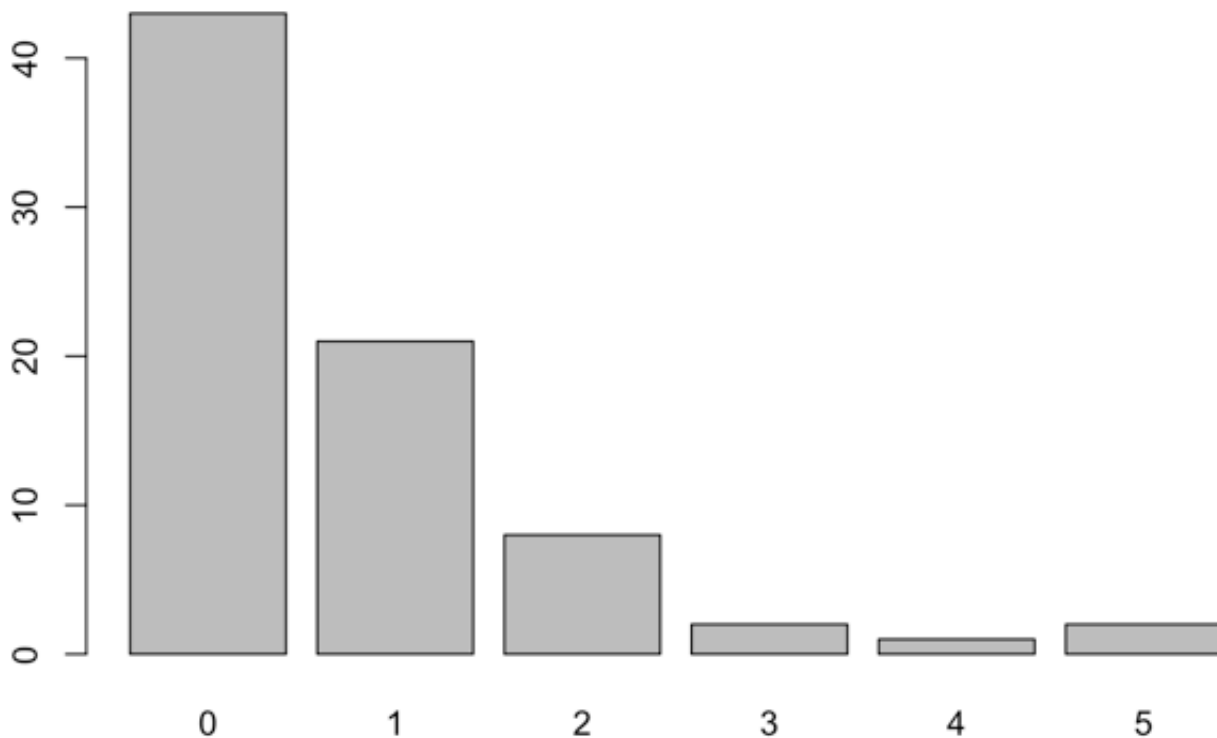
```
## sim_basket
## H M
## 57 76
```

5. 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?

```
calc_streak(sim_basket)
```

```
## [1] 2 1 1 0 0 0 0 0 5 0 1 2 0 0 2 0 0 0 0 0 0 1 1 1 0 2 1 0 1 0 0 1 0 0 3 1 1 2
## [39] 0 0 0 0 0 0 1 1 1 1 0 2 1 5 2 0 1 1 0 0 0 2 1 0 0 0 1 3 4 0 0 1 0 0 0 0 0 0
## [77] 0
```

```
barplot(table(calc_streak(sim_basket)))
```



Solution:

typical streak length for this simulated independent shooter with a 45% shooting percentage: 0

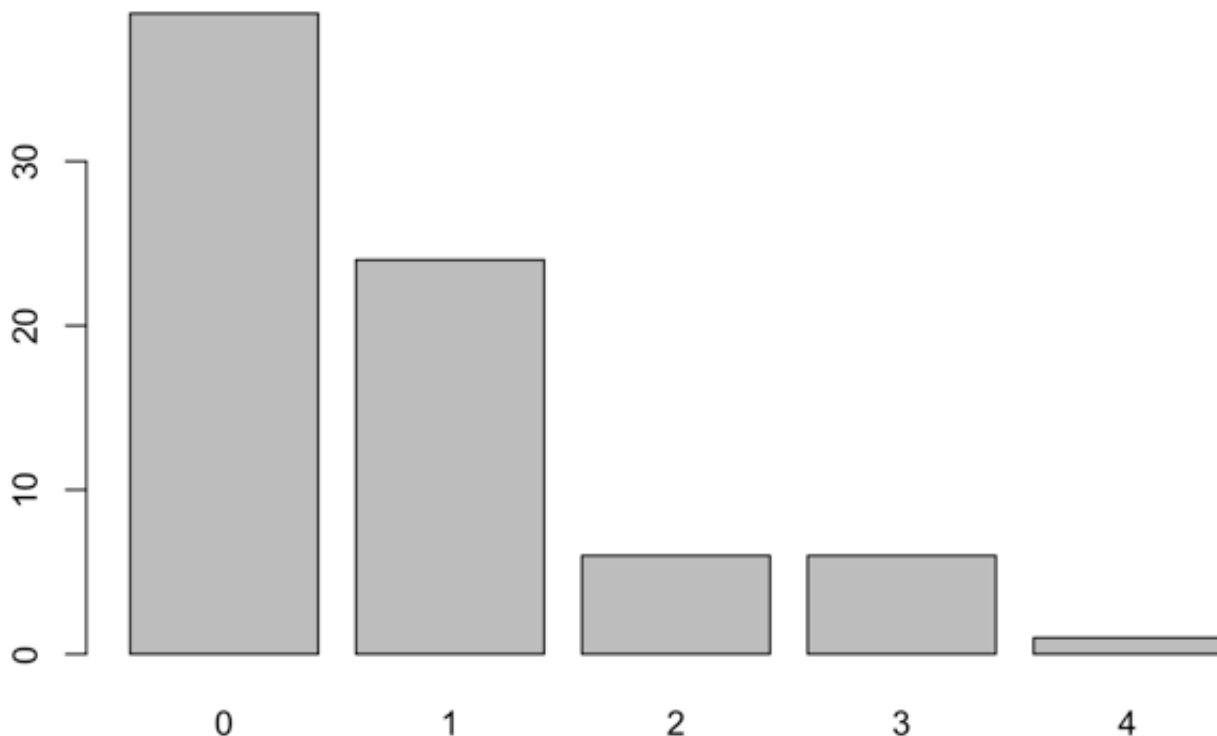
longest streak of baskets in 133 shots: 6

6. 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.(SEEDING THE SIMULATION????)

Solution: It is “Somewhat similar” when the simulation shooter ran for a second time.

7. 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

```
barplot(table(kobe_streak))
```



#Resourse: <https://www.khanacademy.org/math/ap-statistics/quantitative-data-ap/describing-comparing-distributions/v/classifying-distributions>

Solution: Distribution of streak lengths of both Kobe Bryant's and Simulated shooter are somewhat similar, which are skewed right.