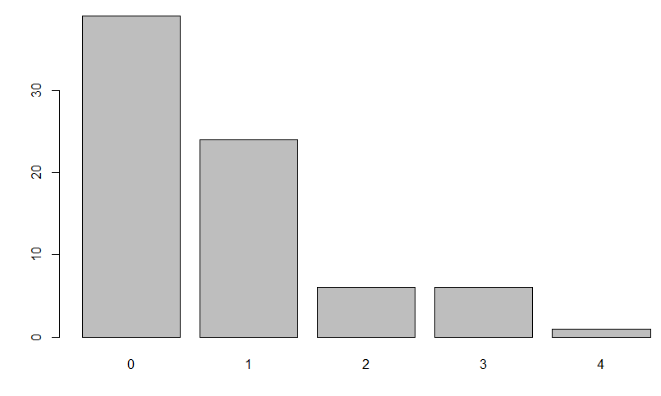
**Lab 2**

**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 length 1 is one hit, or one made basket followed by a miss. A streak of length 0 means a miss and is preceded by a miss.

**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?**

Barplot(table(kobe$basket))

It is a right skew distribution since most data is on the left of the graph, with a unimodal peak streak of 0 hits, which is his typical streak length. His longest streak of baskets is 4 made baskets with no misses.

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

Sim\_unfair\_coin

table(sim\_unfair\_coin)

18 heads and 82 tails.

**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 = 1, replace = TRUE)

……NEW =

sim\_basket <- **sample**(outcomes, size = 133, replace = TRUE, prob = **c**(0.45,

0.55))

Sim\_basket

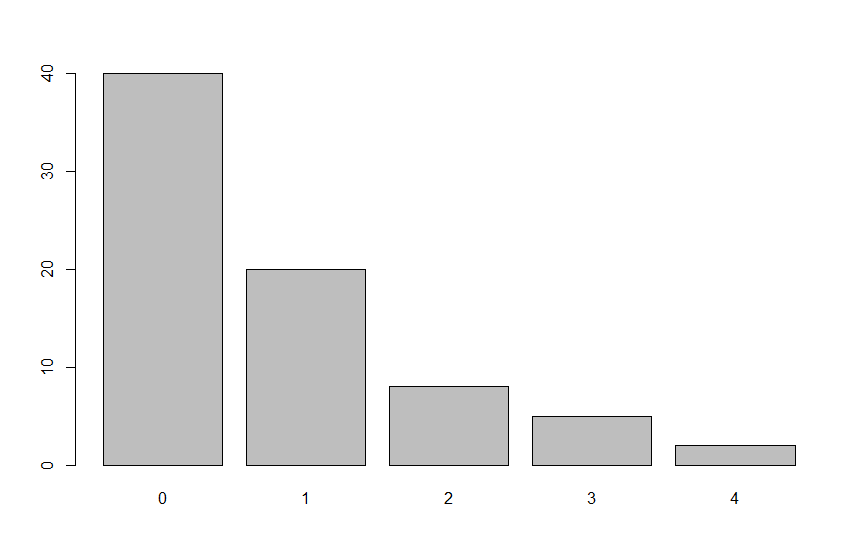
Table(sim\_basket) = 59 hits and 74 misses

**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?**

sim\_streak <- calc\_streak(sim\_basket)

barplot(table(sim\_streak))



Once again, the simulated shooter has a right skewed unimodal distribution, with a typical streak being 0 hits, and highest hit streak being 4 hits with no misses.

1. **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.**

I would expect the streak distribution to be somewhat similar with the distribution above because they have the same independent odds of total hits and misses, so the hits and misses would likely be very close if not the same. I would expect the barplot distribution to follow the same pattern, with the only differences being a chance of having a higher streak than 4, and a slight difference in frequency of each streak length up to 4.

I ran it anyways, and saw that the outcome was what I expected, but 4 was still the highest. There were more 0 streaks and less 2 streaks than the graph above, however.

sim\_basket1 <- sample(outcomes, size = 133, replace = TRUE, prob = c(0.45, 0.55))

table(sim\_basket1)

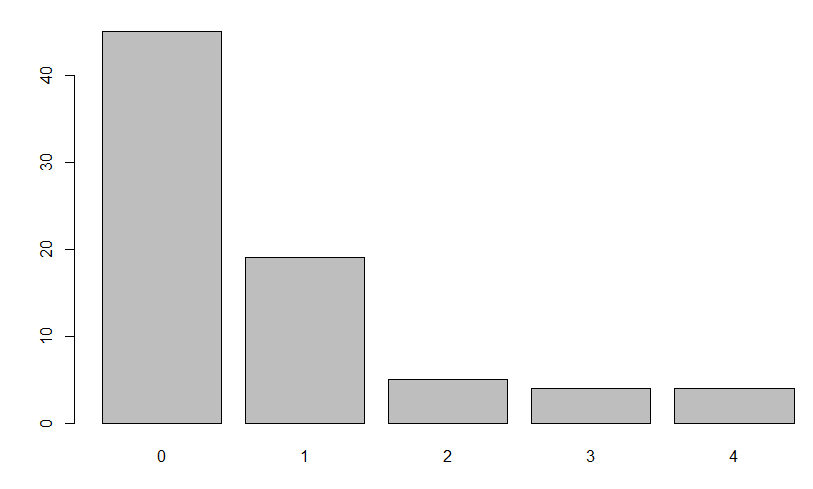
sim\_basket1

**H M**

**57 76**

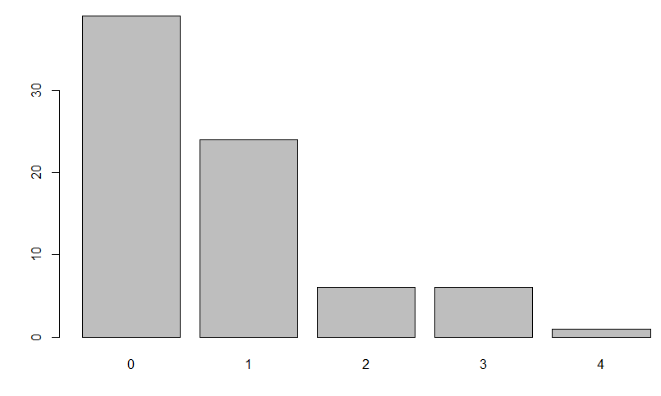
sim\_streak1 <- calc\_streak(sim\_basket1)

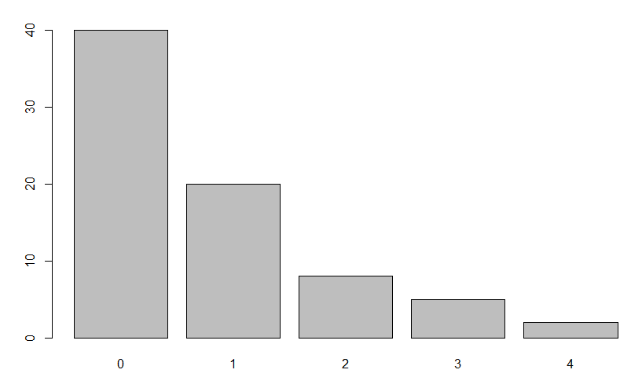
barplot(table(sim\_streak1))



1. **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.**

Kobe Bryant’s distribution of streak lengths is very similar to the distribution for the simulated independent shooters. With this comparison, we have evidence that the hot hand model does not fit Kobe’s shooting patterns on a large sample size basis, as the models suggest that each shot taken is independent of the next.





Here are two graphs that we made, with Kobe’s streaks on the left and the first simulated shooter on the right. They both have right skew, unimodal distribution favoring streaks of 0 and 1, indicating that each shot is more likely than not to not be affected by the previous shot being a hit or miss.