**Tor Thogersen**

**DA 460 – Fall 2017**

**Lab 1 - Handout 0 R and Handout 0 SAS**

## **Part 2 – R Handout**

1. **Exercise**

> kobe$basket[1:9]

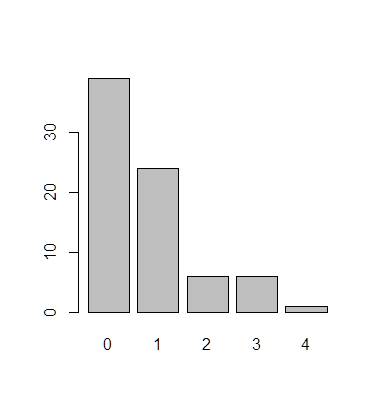
[1] "H" "M" "M" "H" "H" "M" "M" "M" "M

* 1. What does a streak length of 1 mean, i.e.
     1. Kobe made 1 basket and missed the following shot
  2. how many hits and misses are in a streak of 1?
     1. 1 hit, 1 miss
  3. What about a streak length of 0?
     1. 0 hit, 1 mins

1. **Exercise**

> kobe\_streak <- calc\_streak(kobe$basket)

> barplot(table(kobe\_streak))



* 1. Describe the distribution of Kobe’s Streak Lengths from 2009 NBA finals
     1. The distribution starts out at Zero and ends at 4, While the overall shape looks to be normally Distributed with skew to left.
  2. What was his typical Streak Length?
     1. 1 hit, 1 miss
  3. How long was his longest streak of basket?
     1. 4 hit, 1 mins

1. **Exercise**

> sim\_unfair\_coin <-sample(outcomes, size = 100, replace = TRUE, prob = c(0.2,0.08))

> sim\_unfair\_coin

[1] "heads" "heads" "tails" "tails" "heads" "heads" "heads"

[8] "heads" "heads" "heads" "heads" "heads" "heads" "heads"

[15] "tails" "heads" "tails" "heads" "tails" "heads" "heads"

[22] "tails" "heads" "heads" "heads" "tails" "tails" "heads"

[29] "tails" "heads" "tails" "heads" "tails" "heads" "heads"

[36] "tails" "heads" "heads" "tails" "tails" "heads" "heads"

[43] "tails" "tails" "heads" "tails" "tails" "tails" "heads"

[50] "tails" "heads" "heads" "heads" "heads" "heads" "tails"

[57] "heads" "heads" "tails" "tails" "heads" "tails" "heads"

[64] "heads" "heads" "heads" "tails" "tails" "heads" "heads"

[71] "heads" "tails" "heads" "tails" "tails" "heads" "heads"

[78] "heads" "tails" "tails" "heads" "heads" "heads" "tails"

[85] "tails" "heads" "heads" "heads" "heads" "heads" "heads"

[92] "heads" "tails" "heads" "heads" "heads" "heads" "heads"

[99] "heads" "heads"

> table(sim\_unfair\_coin)

sim\_unfair\_coin

heads tails

66 34

* 1. In your simulation of flipping the unfair coin 100 times, how many flips came up heads?
     1. 66

1. **Exercise**

> kobe$basket

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

[16] "M" "M" "H" "H" "H" "M" "M" "H" "M" "H" "H" "H" "M" "M" "M"

[31] "M" "M" "M" "H" "M" "H" "M" "M" "H" "H" "H" "H" "M" "H" "M"

[46] "M" "H" "M" "M" "H" "M" "M" "H" "M" "H" "H" "M" "M" "H" "M"

[61] "H" "H" "M" "H" "M" "M" "M" "H" "M" "M" "M" "M" "H" "M" "H"

[76] "M" "M" "H" "M" "M" "H" "H" "M" "M" "M" "M" "H" "H" "H" "M"

[91] "M" "H" "M" "M" "H" "M" "H" "H" "M" "H" "M" "M" "H" "M" "M"

[106] "M" "H" "M" "H" "H" "H" "M" "H" "H" "H" "M" "H" "M" "H" "M"

[121] "M" "M" "M" "M" "M" "H" "M" "H" "M" "M" "M" "M" "H"

> table(kobe$basket)

H M

58 75

* 1. What changes needs to be made to the “sample” function so that it reflects a shooting percentage of 45%?
     1. Need to change the Prod for Hit and Miss and the sample size
     2. sample(outcomes, size = 133, replace = TRUE, prob = c(0.45,0.55))
  2. Make the adjustment, then run a simulation to sample 133 shots.

> outcomes <-c("H", "M")

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

* 1. Assign the output of this simulation to a new object called sim\_basket.

> sim\_basket

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

[16] "M" "M" "H" "H" "M" "H" "H" "M" "H" "H" "H" "M" "M" "M" "M"

[31] "H" "H" "H" "H" "H" "H" "H" "M" "H" "M" "H" "M" "M" "H" "M"

[46] "H" "M" "H" "H" "H" "M" "H" "H" "H" "H" "H" "M" "M" "H" "M"

[61] "H" "M" "H" "H" "M" "H" "H" "M" "M" "M" "M" "H" "H" "M" "H"

[76] "M" "M" "H" "H" "M" "M" "H" "M" "H" "M" "M" "M" "H" "M" "M"

[91] "M" "M" "M" "H" "H" "H" "H" "M" "M" "H" "M" "M" "H" "M" "H"

[106] "M" "H" "M" "M" "M" "M" "M" "M" "M" "M" "M" "M" "M" "H" "H"

[121] "M" "H" "M" "H" "M" "H" "H" "M" "H" "H" "H" "M" "M"

> table(sim\_basket)

sim\_basket

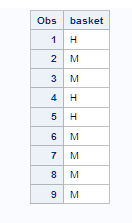
H M

65 68

**Part 2 SAS Handout**

**Exercise 1:**

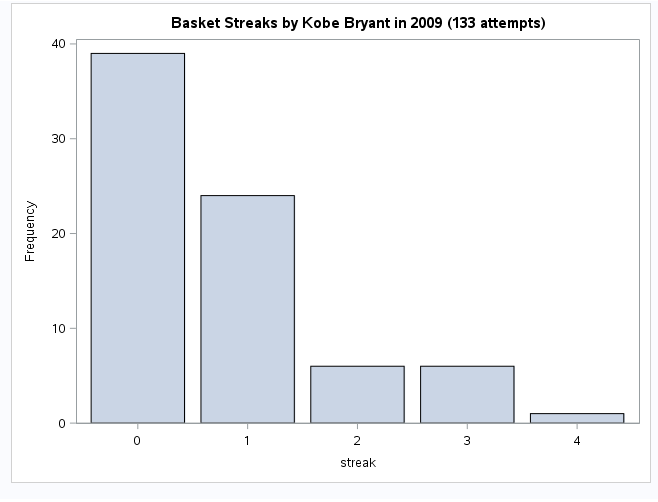
proc print data=kobe;  
 where game=1 and quarter = '1';  
 var basket;  
run;



1. What does a streak length of 1 mean?
   1. 1 hit followed by a miss
2. how many hits and misses are in a streak of 1?
   1. 1 hit, 1 miss
3. What about a streak length of 0?
   1. 0 hits, 1 miss

**Exercise 2:**

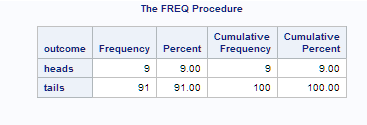
filename cstrk url "http://www.openintro.org/stat/data/calc\_streak.sas";  
%include cstrk;  
%calc\_streak(dset=kobe, streakvar=basket, outset=kobe\_streak);  
proc sgplot data=kobe\_streak;  
title "Basket Streaks by Kobe Bryant in 2009 (133 attempts)"; vbar streak;  
run;



1. Describe the distribution of Bryant’s streak lengths from the 2009 NBA finals.
   1. Normal distribution with left skew, columns starting at 0 ending at 4.
2. What was his typical streak length?
   1. 1 basket made
3. How long was his longest streak of baskets?
   1. 4 consecutive baskets made

**Exercise 3:**

data sim\_unfair\_coin(keep=outcome);  
 do i = 1 to 100;  
 flip = rand("bernoulli", 0.2);  
 if flip = 1 then outcome = "heads";  
 else outcome = "tails";  
 output;  
 end;  
run;  
  
proc freq data = sim\_unfair\_coin;  
 tables outcome;  
run;



1. In your simulation of flipping the unfair coin 100 times, how many flips came up heads?
   1. 9

**Exercise 4:**

1. What change needs to be made to the DATA step so that it reflects a shooting percentage of 45%? Change the do
   1. Change the “do i function from 1 to 100 to 1 to 133”
   2. Change the Bernoulli prob to .45
2. Make this adjustment, and then run a simulation to sample 133 shots.

data sim\_basket(keep=basket);   
 do i = 1 to 133;  
 flip = rand("Bernoulli",0.45);  
 if flip = 1 then basket="H"; else basket="M";   
 output;  
 end;  
run;  
  
proc freq data = sim\_basket;  
 tables basket;  
run;

