S&DS 220: Homework 3

Due Friday February 2

Braeden Cullen

Instructions

1. Complete the questions below. Upload your knitted PDF solutions to Gradescope by the due date.
2. Your solutions should be a combination of writing and R code. When writing, use complete sentences!
3. Previous homework assignments already had code chunks created for you. Now it is up to you to insert R code chunks within each problem as needed.
4. You should aim for clear and concise communication (in both words and R code). See the two example solutions below for guidance on how to write your own solutions.

Example exercises

As an example of what solutions should look like, two example exercises with solutions are provided here.

Example exercise 1 and solution: (from Exercise 2.2 in the book)

Consider an experiment where you roll two dice, and subtract the smaller value from the larger value (getting 0 in case of a tie).

(a) What is the probability of getting 0?

Solution.

We obtain a 0 exactly when we roll doubles. There are 6 ways to do this out of a total of 36 ways to roll two dice. And so the probability is 1/6.

(b) What is the probability of getting 4?

Solution.

We get a 4 when we roll $(1, 5)$, $(5, 1)$, $(2, 6)$, or $(6, 2)$. Since there are four ways to do this, the probability of getting 4 is 4/36, or 1/9.

Example exercise 2 and solution: (from Exercise 2.17 in the book)

A standard deck of cards has 52 cards, four each of 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A. In blackjack, a player gets two cards and adds their values. Cards count as their usual numbers, except Aces are 11 (or 1), while K, Q, J are all 10.

(a) "Blackjack" means getting an Ace and a value 10 card. What is the probability of getting a blackjack?

Solution.

First we create the card deck. Only the rank, not the suit, matters. Then we replicate the experiment of drawing a two-card and seeing if we got a blackjack 10,000 times.

```r
# create a 52-card deck
cards <- c(
  rep(2:9, each = 4), # cards 2 through 9
  rep(10, times = 16), # cards with rank 10 (10 through K)
  rep(11, times = 4) # aces
)

blackjack_hands <- replicate(1e4, {
  hand <- sample(x = cards, size = 2, replace = FALSE) # draw two-card hand
  sum(hand) # sum the cards in the hand for the total
})

mean(blackjack_hands == 21) # compute the proportion of blackjacks
```

```
## [1] 0.0474
```

(b) What is the probability of getting 19? (The probability that the sum of your cards is 19, using Ace as 11)

Solution.

The solution to this problem is similar to the solution in part (a), except that we are checking for a total of 19.

```r
mean(blackjack_hands == 19) # compute the proportion of 19s
```

```
## [1] 0.0583
```

Problem set questions !!!!!!!!!!!!!!!!!!!!!!!!!!!

Question 1: Exercise 2.5

Suppose the proportion of M&M's by color is:

Yellow	*Red*	*Orange*	*Brown*	*Green*	*Blue*
0.14	0.13	0.20	0.12	0.20	0.21

(a) What is the probability that a randomly selected M&M is not green?

```
# insert R code here
probabilities <- list(yellow = 0.14, red = 0.13, orange = 0.20, brown = 0.12, green = 0.20, blue = 0.21
1 - probabilities$green
```

```
## [1] 0.8
```

(b) What is the probability that a randomly selected M&M is red, orange, or yellow?

```
# insert R code here
probabilities$yellow + probabilities$red + probabilities$orange
```

```
## [1] 0.47
```

(c) Estimate (using simulation) the probability that a random selection of four M&M's will contain at least one blue M&M.

```
num_sims <- 1e5
num_mms <- 4
mms <- sample(names(probabilities), size = num_sims * num_mms, replace = TRUE, prob = unlist(probabilit
mms <- matrix(mms, nrow = num_sims) # produce a matrix containing the result of every simulation
any_blue <- apply(mms, 1, function(x) "blue" %in% x) # loop over each row (indicated by the 1 for itera
mean(any_blue) # take the mean of this result
```

```
## [1] 0.61035
```

(d) Estimate (using simulation) the probability that a random selection of six M&M's will contain all six colors.

```
num_mandm <- 6
mms <- sample(names(probabilities), size = num_sims * num_mandm, replace = TRUE, prob = unlist(probabil
mms <- matrix(mms, nrow = num_sims)
all_colors <- apply(mms, 1, function(x) length(unique(x)) == 6) # same procedure as before, except chec
mean(all_colors)
```

```
## [1] 0.01309
```

Question 2: Exercise 2.14

In a room of 100 people, estimate (using simulation) the probability that at least two people were not only born on the same day, but also during the same hour of the same day. (For example, both were born between 2 and 3.)

```
set.seed(35)
simulations <- 1000
num_people <- 100
total_matches <- 0

# iterating through the simulations
for(k in 1:simulations) {
  matches_found <- 0 # to count simulations with at least one match
  birthdays <- sample(1:365, 100, replace = TRUE)
  hours <- sample(1:24, 100, replace = TRUE)
  for (i in 1:num_people) {
    for(j in 1:num_people) {
      if(j==i) next # skip if same person is being checked against itself (i = j)
      else {
        if(birthdays[i] == birthdays[j] && hours[i] == hours[j]) {
          matches_found <- matches_found + 1
          break
        }
      }
    }
  }
  if (matches_found > 1) {
    total_matches <- total_matches + 1
  }
}

# calculating the probability
probability <- total_matches / simulations
probability
```

```
## [1] 0.427
```

Question 3: Exercise 2.28

Bob Ross was a painter with a PBS television show, "The Joy of Painting", that ran for 11 years.

(a) 91% of Bob's paintings contain a tree and 85% contain two or more trees. What is the probability that he painted a second tree, given that he painted a tree?

```
prob_tree <- 0.91
prob_two_trees <- 0.85
prob_two_trees / prob_tree
```

```
## [1] 0.9340659
```

(b) 18% of Bob's paintings contain a cabin. Given that he painted a cabin, there is a 35% chance the cabin is on a lake. What is the probability that a Bob Ross painting contains both a cabin and a lake?

```
prob_cabin <- 0.18
prob_lake_given_cabin <- 0.35
prob_cabin * prob_lake_given_cabin
```

```
## [1] 0.063
```

Question 4: Exercise 2.29

Ultimate frisbee players are so poor they don't own coins. So, team captains decide which team will play offense first by flipping frisbees before the start of the game. Rather than flip one frisbee and call a side, each team captain flips a frisbee and one captain calls whether the two frisbees will land on the same side, or on different sides. Presumably, they do this instead of just flipping one frisbee because a frisbee is not obviously a fair coin - the probability of one side seems likely to be different from the probability of the other side.

(a) Suppose you flip two fair coins. What is the probability they show different sides?

```
prob_different = 0.5
prob_different
```

```
## [1] 0.5
```

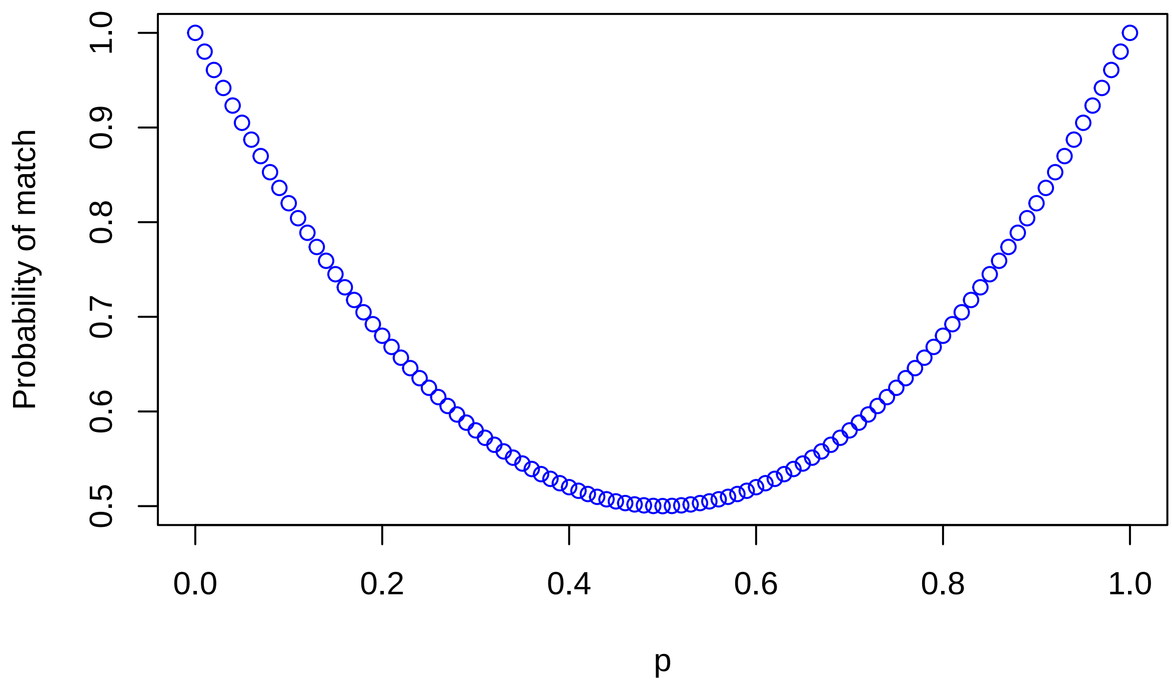
(b) Suppose two captains flip frisbees. Assume the probability that a frisbee lands convex side up is p. Compute the probability (in terms of p) that the two frisbees match.

```
print('prob_match = p^2 + (1-p)^2')
```

```
## [1] "prob_match = p^2 + (1-p)^2"
```

(c) Make a graph of the probability of a match in terms of p. Set the x and y labels inside the `plot` function with `xlab` and `ylab`.

```
prob <- seq(0, 1, by = 0.01)
prob_match <- prob^2 + (1-prob)^2
plot(prob, prob_match, xlab = "p", ylab = "Probability of match", col = "blue")
```



(d) One Reddit user flipped a frisbee 800 times and found that in practice, the convex side lands up 45% of the time. When captains flip, what is the probability of "same?" What is the probability of "different?"

```
prob <- 0.45
prob_match <- prob^2 + (1-prob)^2
prob_different <- 1 - prob_match
prob_match
```

```
## [1] 0.505
```

```
prob_different
```

```
## [1] 0.495
```

(e) What advice would you give to an ultimate frisbee team captain?

```
"Bet on same side up"
```

```
## [1] "Bet on same side up"
```

(f) Is the two-frisbee flip better than a single-frisbee flip for deciding the offense?

```r
"Yes because the probability of a match is higher than the probability of a different side up."
```

```
## [1] "Yes because the probability of a match is higher than the probability of a different side up."
```

Question 5: Exercise 2.33

How many ways are there of getting exactly 4 heads when tossing a coin 10 times? Hint: experiment with the R function `choose`.

```r
choose(10, 4)
```

```
## [1] 210
```

Question 6: Exercise 2.34

How many ways are there of getting exactly 4 heads when tossing a coin 10 times, assuming that the 4th head came on the 10th toss?

```r
choose(9, 3)
```

```
## [1] 84
```