Discrete Probability - Exercises from DataCamp

Exercise 1. The Cavs and the Warriors

Two teams, say the Cavs and the Warriors, are playing a seven game championship series. The first to win four games wins the series. The teams are equally good, so they each have a **50-50** chance of winning each game.

If the Cavs lose the first game, what is the probability that they win the series?

- Assign a variable 'n' as the number of remaining games.

```
n <- 6
```

- Assign a variable `outcomes` as a vector of possible game outcomes, where 0 indicates a loss and 1 indicates a win for the Cavs.

```
outcomes <- c(0, 1)
```

- Assign a variable `l` to a list of all possible outcomes in all remaining games. Use the `rep` function on `list(outcomes)` to create list of length `n`.

```
l <- rep(list(outcomes), n)
```

- Create a data frame named 'possibilities' that contains all combinations of possible outcomes for the remaining games.

```
possibilities <- expand.grid(l)
```

- Create a vector named 'results' that indicates whether each row in the data frame 'possibilities' contains enough wins for the Cavs to win the series.

```
results <- rowSums(possibilities) >= 4
```

- Calculate the proportion of 'results' in which the Cavs win the series. Print the outcome to the console.

```
mean(results)

## [1] 0.34375
```

Exercise 2. The Cavs and the Warriors - Monte Carlo

Confirm the results of the previous question with a Monte Carlo simulation to estimate the probability of the Cavs winning the series after losing the first game.

- The variable `B` specifies the number of times we want the simulation to run. Let's run the Monte Carlo simulation 10^4 times.
- Use the `set.seed` function to make sure your answer matches the expected result after random sampling.
- Create an object called `results` that replicates for `B` iterations a simulated series and determines whether that series contains at least four wins for the Cavs.

```
set.seed(1)
B <- 10^4
results <- replicate(B, {
cavs_wins <- sample(c(0,1),6, replace = TRUE)
sum(cavs_wins) >= 4
})
```

- Calculate the frequency out of `B` iterations that the Cavs won at least four games in the remainder of the series. Print your answer to the console.

```
mean(results)

## [1] 0.3371
```

Exercise 3. A and B play a series - part 1

Two teams, **A** and **B**, are playing a seven series game series. Team **A** is better than team **B** and has a $p > 0.5$ chance of winning each game.

- Let's assign the variable 'p' as the vector of probabilities that team A will win.

```
p <- seq(0.5, 0.95, 0.025)
```

- Given a value '**p**', the probability of winning the series for the underdog team B can be computed with the following function based on a Monte Carlo simulation:

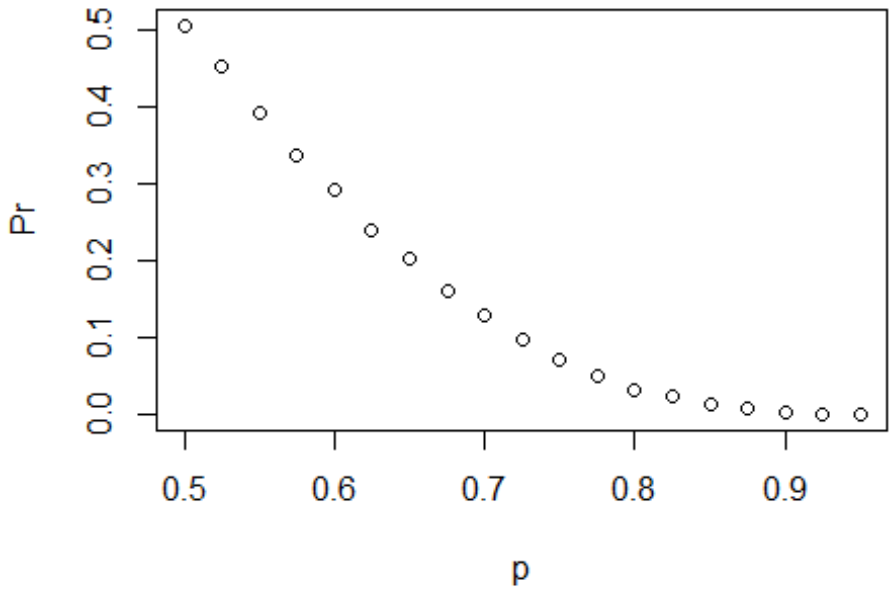
```
prob_win <- function(p){
  B <- 10^4
  result <- replicate(B, {
    b_win <- sample(c(1,0), 7, replace = TRUE, prob = c(1 - p, p))
    sum(b_win)>= 4
    })
  mean(result)
}
```

- Apply the 'prob_win' function across the vector of probabilities that team A will win to determine the probability that team B will win. Call this object 'Pr'.

```
Pr <- sapply(p, prob_win)
```

- Plot the probability 'p' on the x-axis and 'Pr' on the y-axis.

```
plot(p, Pr)
```



Exercise 4. A and B play a series - part 2

Repeat the previous exercise, but now keep the probability that team **A** wins fixed at $p = 0.75$ and compute the probability for different series lengths. For example, wins in best of 1 game, 3 games, 5 games, and so on through a series that lasts 25 games.

*Given a value 'p', the probability of winning the series for the underdog team B can be computed with the following function based on a Monte Carlo simulation:

```
prob_win <- function(N,  p= 0.75){
      B <- 10^4
      result <- replicate(B, {
        b_win <- sample(c(1,0), N, replace = TRUE, prob = c(1 - p, p))
        sum(b_win)>= (N + 1)/2
        })
      mean(result)
    }
```

*Assign the variable 'N' as the vector of series lengths. Use only odd numbers ranging from 1 to 25 games.

```
N <- seq(1,25,2)
```

- Apply the 'prob_win' function across the vector of series lengths to determine the probability that team B will win. Call this object `Pr`.

```
Pr <- sapply(N,prob_win)
```

*Plot the number of games in the series ‘N’ on the x-axis and ‘Pr’ on the y-axis.

```
plot(N, Pr)
```