Law of Large Numbers

Mini-Assignment - MTH 361 A/B - Spring 2023

Instructions:

- Please provide complete solutions for each problem. If it involves mathematical computations, explanations, or analysis, please provide your reasoning or detailed solutions.
- Note that some problems have multiple solutions or ways to solve it. Make sure that your solutions are clear enough to showcase your work and understanding of the material.
- Creativity and collaborations are encouraged. Use all of the resources you have and what you need to complete the mini-assignment. Each student must take personal responsibility and submit their work individually. Please abide by the University of Portland Academic Honor Principle.
- **Please save your work as one pdf file, don't put your name in any part of the document, and submit it to the Teams Assignments for this course. Your document upload will correspond to your name automatically in Teams.**
- If you have questions or concerns, please feel free to ask the instructor.

R Packages:

```
library(tidyverse)
library(openintro)
```

I. N-Games of Rock-Paper-Scissors

Materials

Purpose: By playing Rock-Paper-Scissors game and evaluating the likelihood of each decision, students will learn the following:

- probability, independent and dependent events,
- the law of large numbers, and
- simulating random processes using R.

Description: Students will compete in two rounds of Rock-Paper-Scissors - one with eyes open and the other with eyes closed. Play each round 40 times with a partner and note the results.

Rules of the Game: The idea behind this game is that each player can use rock, paper, or scissors hand signs to play. The concept is that rock beats scissors, scissors beat paper, and paper beats rock. When both players raise their hands, they can use the same hand motion to lose, win, or tie. See Figure 1.

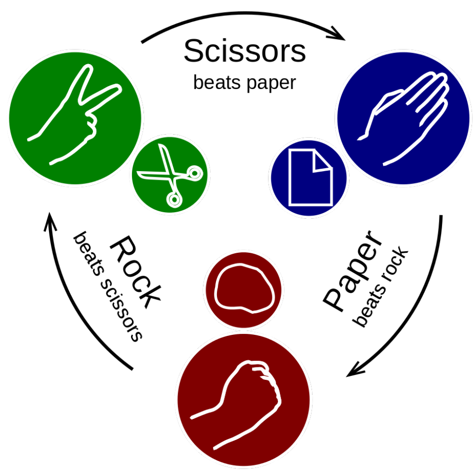


Figure 1: Rock-Paper-Scissors Game Hand Gestures and Rules. This diagram is from

If a player wins, they gain one point, and if a player loses they lose 1 point. Each player starts with 0 points. Below is a table that summarizes the point system of the game.

A/B	Rock (R)	Paper (P)	Scissors (S)
Rock (R)	0/0	-1/1	1/-1
Paper (P)	1/-1	0/0	-1/1
Scissors (S)	-1/1	1/-1	0/0

The theoretical probabilities for each pair of R-P-S are shown in the table below. Notice that all pairs are equally likely.

A/B	Rock (R)	Paper (P)	Scissors (S)
Rock (R)	1/9	1/9	1/9
Paper (P)	1/9	1/9	1/9
Scissors (S)	1/9	1/9	1/9

Exercises

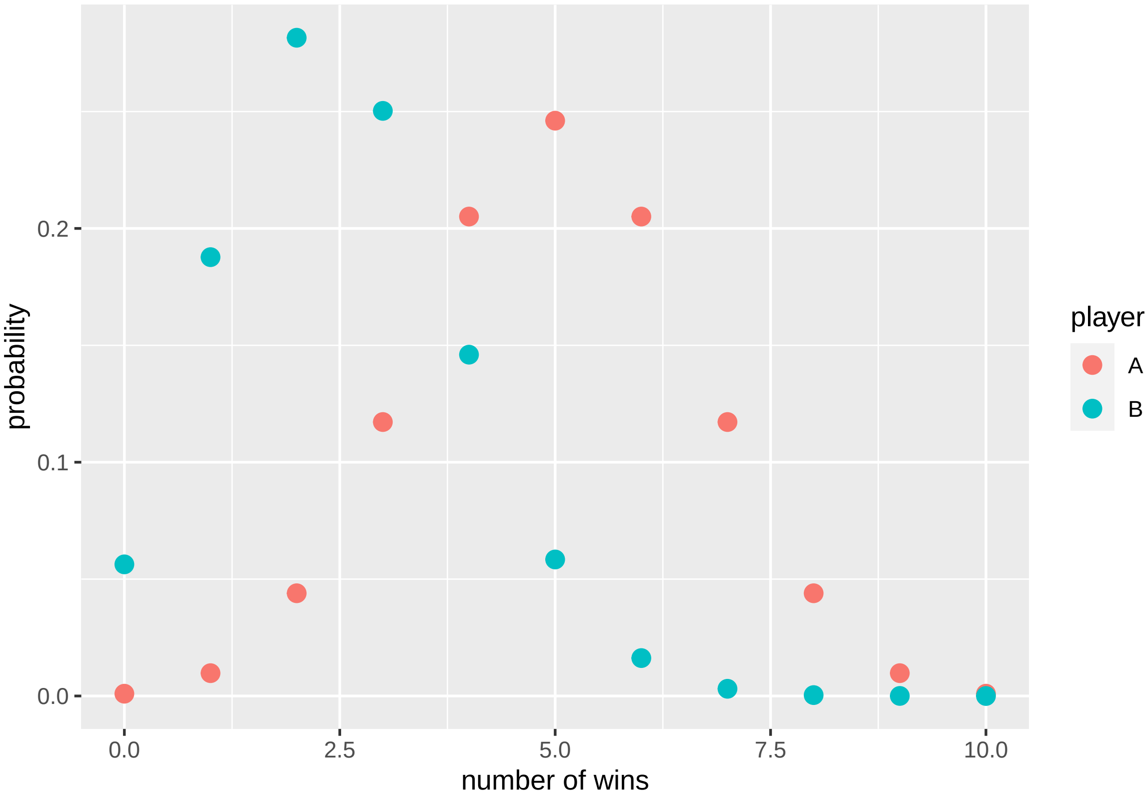
1. **Theoretical Probabilities.** Use the tables above to answer the following questions. Use standard probability notations.

   a. What is the probability that player A wins for one trial? What is the probability that player B wins for one trial?

   b. What is the expected value after 40 trials for player A? What is the expected value after 40 trials for player B? Using these expected values, what is the expected points in each player. Note that a player's win is the other player's loss.

   c. Modify the example R code below. Use the `dbinom` function in R to produce a probability distribution plot of the number of wins in 40 trials for each player. Note that the `dbinom` function is the Binomial Probability Mass Function. You will notice that each player have the same probability distributions. Explain why.

```
## ---Modify these variables---
p_A <- 1/2 # Player A probability of winning
p_B <- 1/4 # Player B probability of winning
n <- 10 # number of trials
## ----------------------------

# all possible outcomes
x <- 0:n

# compute probabilities using Binomial PDF
prob_A <- dbinom(x,n,p_A) # player A
prob_B <- dbinom(x,n,p_B) # player B

# plotting
df <- tibble(n_wins = c(x,x),
             probabilities = c(prob_A,prob_B),
             player = rep(c("A","B"),each=length(x)))
ggplot(data = df, aes(x = n_wins, y = probabilities, color = player)) +
  geom_point(size=3) +
  labs(x = "number of wins", y = "probability")
```



d. Compute the probability that player A win at least the expected value?

e. Compute the probability that player A win at most the expected value?

f. Explain the meaning of the probabilities that you computed in (d) and (e) in terms of the law of large numbers.

2. **Empirical Data.**

   a. *Round 1.* With your partner, play 40 times of Rock-Paper-Scissors with your eyes *open* and record each outcome. Try to strategize in order to win. Use the table below to summarize your results - replace "X" with the number of observations of that particular outcome of R-P-S pair.

A/B	**Rock (R)**	**Paper (P)**	**Scissors (S)**
Rock (R)	X	X	X
Paper (P)	X	X	X
Scissors (S)	X	X	X

   b. *Round 2.* With your partner, play 40 times of Rock-Paper-Scissors with your eyes *closed* and record each outcome. Use the table below to summarize your results - replace "X" with the number of observations of that particular outcome of R-P-S pair.

A/B	**Rock (R)**	**Paper (P)**	**Scissors (S)**
Rock (R)	X	X	X
Paper (P)	X	X	X
Scissors (S)	X	X	X

   c. Now, compare the two rounds (1a) and (1b) by computing the proportion of the number of wins. Use the tables below to record the win proportions and the points. Do you see any evidence of dependence between the variables (rounds and players)? Explain your reasoning.

Win Numbers	**Round 1**	**Round 2**
A	X	X
B	X	X

Win Proporions	**Round 1**	**Round 2**
A	X	X
B	X	X

Points	**Round 1**	**Round 2**
A	X	X
B	X	X

   d. In theory, each player has equally likely to win and has the expected value of winnings to be the same. Based on the results, which round do you think was played more fairly? Explain your reasoning.

3. **N-Games of Rock-Paper-Scissors.** The law of large numbers is a fundamental theorem in probability theory that states that as the number of independent trials or observations increases, the average of the results obtained from these trials will converge to the expected value of the underlying random variable. We can simulate a fair game of Rock-Paper-Scissors using R. Below is an R function that simulates 1 trial of the game. The function can then be used to simulate N trials of Rock-Paper-Scissors.

```
source("rps_simulator.R")

# one trial of R-P-S
N <- 1
rps(k=N)
```

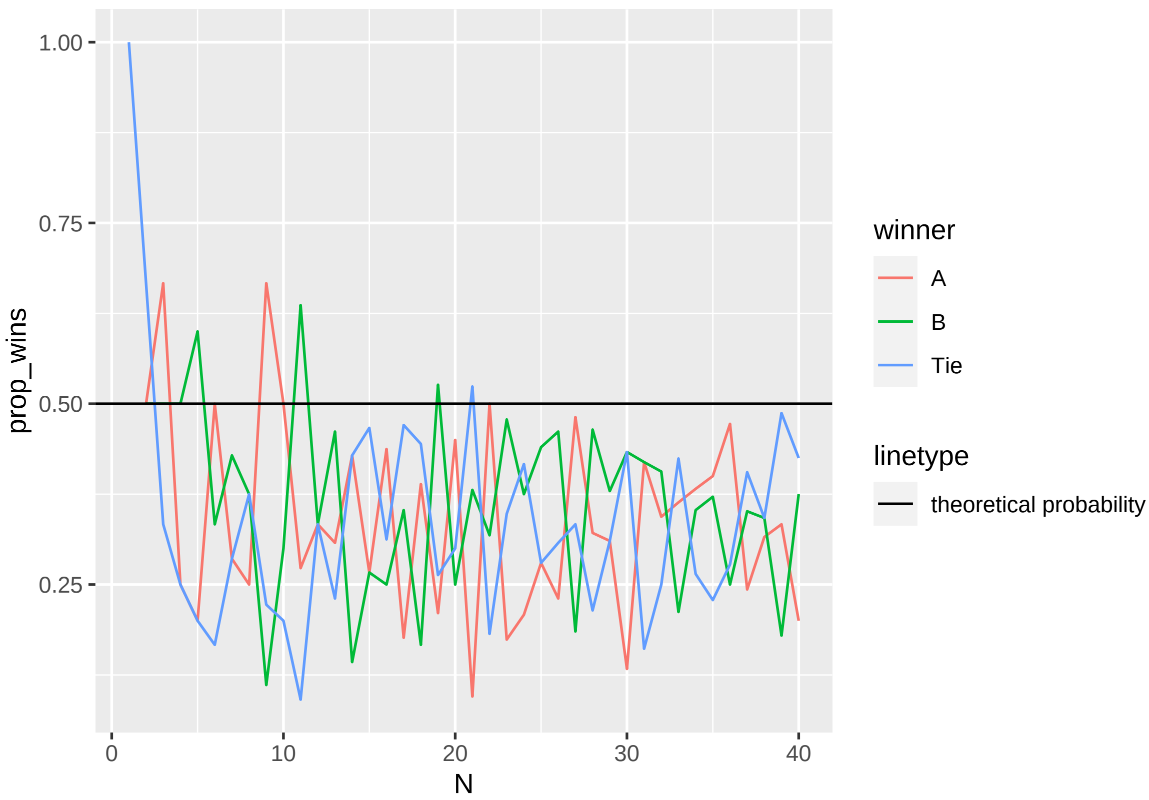
```
## # A tibble: 1 x 1
##   winner
##   <chr>
## 1 Tie
```

Modify the example code below to simulate 100 trials. Make sure you update the theoretical probability of winning that you computed in (1a).

```
# N trials of R-P-S
## ---Modify these variables---
N <- 40 # number of trials
theo_prob <- 1/2 # the theoretical probability of winning
## ----------------------------

# simulations
trials_vect <- 1:N
df_raw <- tibble(N = integer(),
                 winner = character())
for(i in trials_vect){
  # simulate the games
  rps_sims <- rps(k=i)

  # record the results
  df_raw <- df_raw %>% add_row(N = rep(i,i),
                               winner = rps_sims$winner)
}

# compute the proportion of wins
df_props <- df_raw %>%
  group_by(N) %>%
  count(winner) %>%
  mutate(prop_wins = n/N)
```

The code then visualizes the simulations which shows N (the number of trials) as the x-axis and the the proportion of wins as the y-axis. What can you say about the proportions as N approaches infinity? Explain your reasoning in terms of the law of large numbers.

```
# plot the simulations
ggplot(data = df_props, aes(x = N, y = prop_wins, color = winner)) +
  geom_line() +
  geom_hline(aes(yintercept = theo_prob, linetype="theoretical probability"),
             color = "black")
```


1.00
0.75
0.50
0.25
prop_wins
0
10
20
30
40
N
winner
A
B
Tie
linetype
theoretical probability

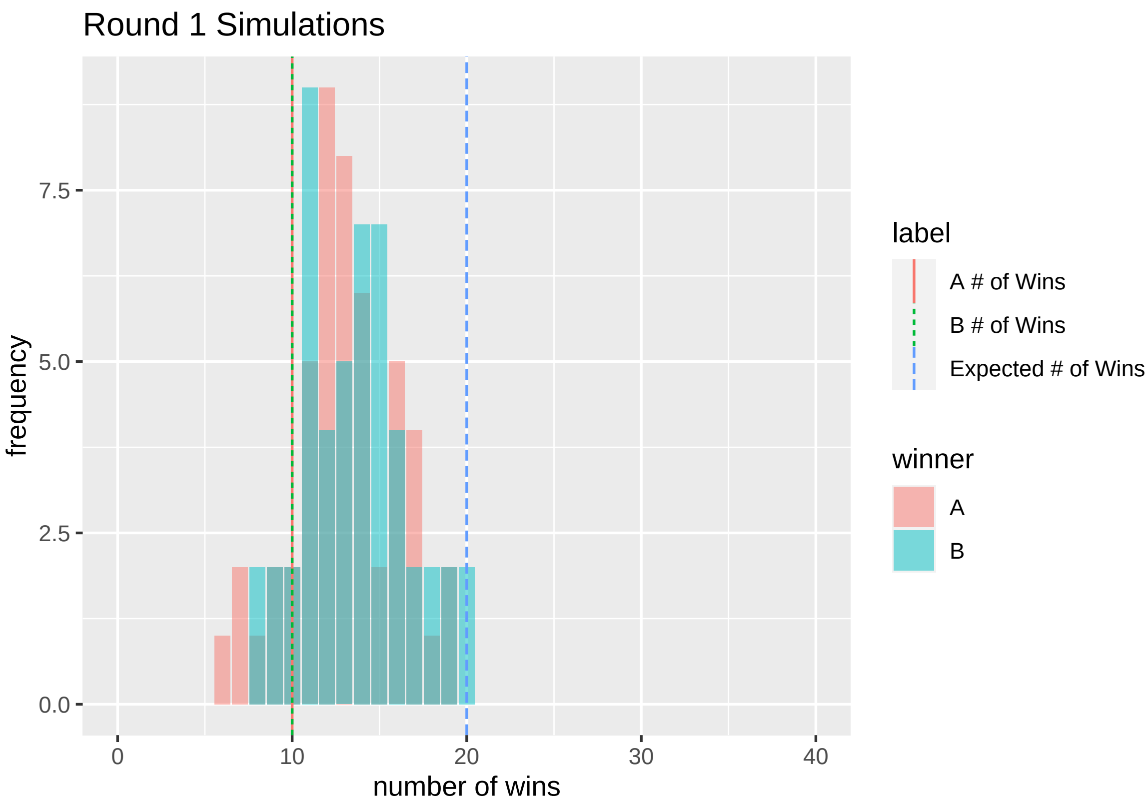
4. (Outstanding Question) **Multiple Rounds.** Suppose that we repeat the Rock-Paper-Scissors game X times with 40 trials each. The law of large numbers tells us that if we repeat an experiment many times, the average result of those experiments will tend to get closer and closer to the true or expected value of the experiment. In this case, we repeat the same number of trials X times and record the number of wins in each player.

   Modify the example code below, to simulate 500 games with 40 trials in each game. Make sure you update the empirical win numbers you computed in (2c). Note that these simulations are under the assumption of a fair game of R-P-S.

```
# X games and N trials of R-P-S
## ---Modify these variables---
X <- 50 # number of games
N <- 40 # number of trials
empirical_a_round1 <- 10 # player A empirical number of wins of round 1
empirical_b_round1 <- 10 # player B empirical number of wins of round 1
empirical_a_round2 <- 10 # player A empirical number of wins of round 2
empirical_b_round2 <- 10 # player B empirical number of wins of round 2
## ----------------------------

# expected value of wins if N*p for binomial
expected_value <- N*theo_prob

# games
x_vect <- 1:X
df_raw <- tibble(X = integer(),
                 winner = character())
for(x in x_vect){
  # simulate the games
  rps_sims <- rps(k=N)

  # record the results
  df_raw <- df_raw %>% add_row(X = rep(x,N),
                               winner = rps_sims$winner)
}

# compute ties
n_ties <- df_raw %>%
  filter(winner == "Tie") %>%
  group_by(X) %>%
  summarise(n_ties = n())

# compute the proportion of wins and the total scores
df_wins <- df_raw %>%
  filter(winner != "Tie") %>%
  group_by(X) %>%
  count(winner) %>%
  left_join(n_ties, by = c("X" = "X")) %>%
  mutate(prop_wins = n/N,
         total_score = n - (N - n - n_ties))

# data for lines
lines <- tibble(label = c("Expected # of Wins","A # of Wins","B # of Wins"),
                lines_r1 = c(expected_value,empirical_a_round1,empirical_b_round1),
                lines_r2 = c(expected_value,empirical_a_round2,empirical_b_round2))
```

a. Observe the figure below. Does the distribution of wins closely similar to the theoretical distribution shown in (1c)?

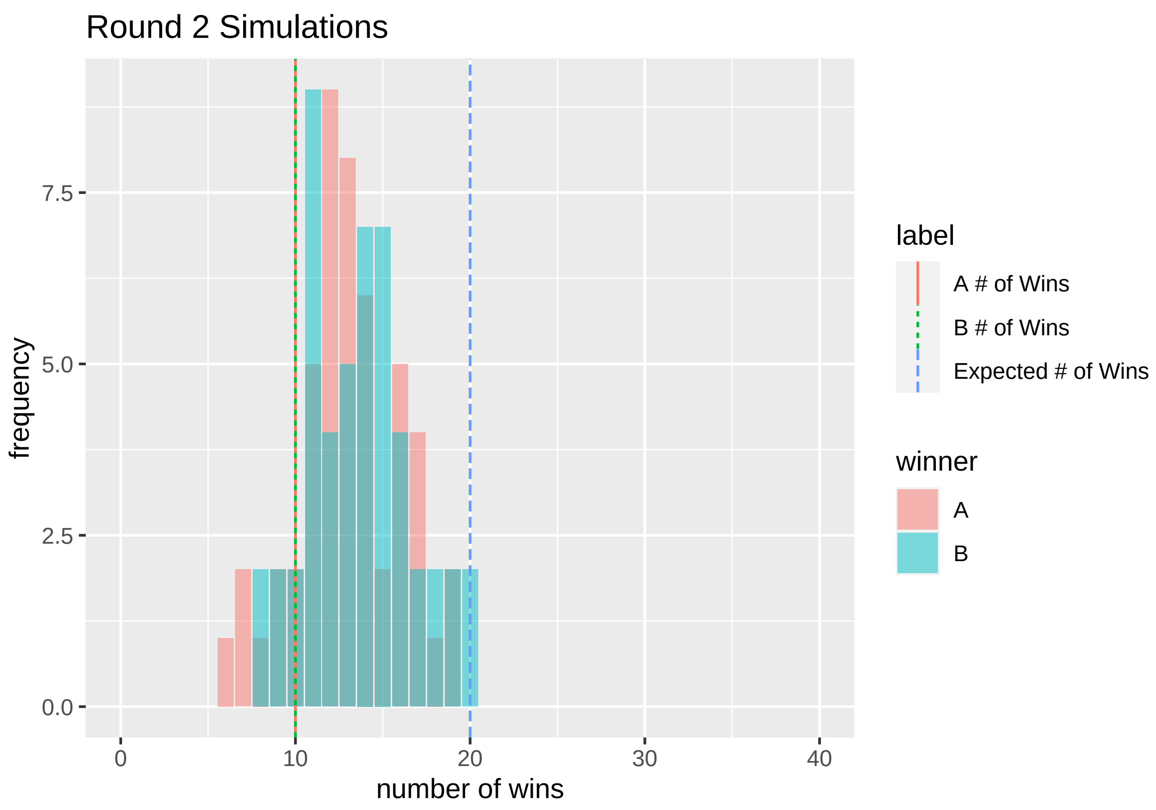
```
# plot results for round 1
ggplot(data = df_wins, aes(x = n, fill = winner)) +
  geom_bar(position = "identity", alpha = 0.50) +
  lims(x = c(0,N)) +
  labs(x = "number of wins", y = "frequency", title = "Round 1 Simulations") +
  geom_vline(data = lines, aes(xintercept = lines_r1, linetype =label, color = label))
```



b. Recall that Round 1 of the game is when your eyes are open and was encourage to strategize. Using the figure in (4a), compare your empirical number of wins of round 1 computed in (2c) to the expected number of wins. How far are these numbers from the theoretical mean? Explain your reasoning.

c. Recall that Round 2 of the game is when your eyes are closed. Use the figure below to compare it to the figure in (1a). How far are the empirical numer of wins of round 2 computed in (2c) to the expected number of wins? Compare these results in (b). Is there evidence of dependence between when eyes are open and eyes are closed? Explain your reasoning.

```
# plot results for round 2
ggplot(data = df_wins, aes(x = n, fill = winner)) +
  geom_bar(position = "identity", alpha = 0.50) +
  lims(x = c(0,N)) +
  labs(x = "number of wins", y = "frequency", title = "Round 2 Simulations") +
  geom_vline(data = lines, aes(xintercept = lines_r2, linetype =label, color = label))
```


Round 2 Simulations
frequency
0.0
2.5
5.0
7.5
number of wins
0
10
20
30
40
label
A # of Wins
B # of Wins
Expected # of Wins
winner
A
B