ALY 6050 - Module 1 Project

Analysis of a Betting Strategy in Sports

Jeff Hackmeister

Introduction

The purpose of this study revolves around a better strategy around three different playoff baseball scenarios involving the Boston Red Sox and the New York Yankees. The study will consist of three parts, each part representing a different scenario for the number and location of games played between the two teams. The parameters given for the study are as follows:

- The probability of the Red Sox winning at home is 0.6
- The probability of the Yankees winning at home is 0.57
- For each game of the series, a bet will be placed
  - A Red Sox win results in a $500 profit
  - A Yankees win results in a $520 loss

For the purposes of this study, the results of each game are considered to be independent from previous games

Part One – Best of Three (BOS, NY, BOS)

For the first part of the study, the parameters for the series will be a best of three series, in which a team wins the series by winning two games. The first game will be played in Boston, with game two in New York, and the third game, if necessary, to be played in Boston.

(i) Calculate the probability that the Red Sox will win the series.

There are two different scenarios in which Boston withs the series, either with 2 wins and 0 losses (a sweep) or with a loss in either game 1 or game 2 and then a win in game 3. Starting with the probability of each outcome, winning in either 2 or 3 games.

```
p_win_home <- 0.6  # Red Sox win at home
p_win_away <- 0.43 # Red Sox win at Yankees' home (1 - 0.57)
```

```
# Probability of winning in 2 games
p_win_2 <- p_win_home * p_win_away
p_win_2
```

```
[1] 0.258
```

```r
# Probability of winning in 3 games
p_win_3 <- (p_win_home * (1 - p_win_away) * p_win_home) +
  ((1 - p_win_home) * p_win_away * p_win_home)
p_win_3
```

```
[1] 0.3084
```

With the probabilities of both scenarios calculated, the can be added to get the probability of a Red Sox series win.

```r
# Total probability of Red Sox winning the series
p_total <- p_win_2 + p_win_3
p_total
```

```
[1] 0.5664
```

As shown above, the probability of a Boston series win is 0.5664 or 56.64%.

(ii) Construct the theoretical probability distribution for your net winnings in dollars (X) for the series. From this theoretical calculation you should also compute and record the expected value of the net winnings (the mean of X) and the theoretical standard deviation of X.

To continue the study to incorporate winnings, the probabilities of the Red Sox losing the series need to be calculated as well. Again, this could happen in a sweep or with a single Boston win.

```r
p_lose_2 <- (1 - p_win_home) * (1 - p_win_away)
p_lose_2
```

```
[1] 0.228
```

```r
p_lose_3 <- (p_win_home * (1 - p_win_away) * (1 - p_win_home)) +
  ((1 - p_win_home) * p_win_away * (1 - p_win_home))
p_lose_3
```

```
[1] 0.2056
```

The net winnings are simple to calculate

```r
# Per game winnings
bos_w <- 500
bos_l <- -520

# Net winnings for each outcome
winnings_win_2 <- bos_w * 2
winnings_win_3 <- bos_w * 2 + bos_l
winnings_lose_2 <- bos_l * 2
winnings_lose_3 <- bos_l * 2 + bos_w
```

To calculate the expected mean winnings, each winning amount is multiplied by corresponding scenario and each of those results are added together.

```r
# Expected value (mean) of X
expected_value <- (winnings_win_2 * p_win_2) +
  (winnings_win_3 * p_win_3) +
  (winnings_lose_2 * p_lose_2) +
  (winnings_lose_3 * p_lose_3)
expected_value
```

```
[1] 57.888
```

From this, the expected value of winnings is stated at $57.88 and the standard deviation can be calculated from this amount.

```r
# Variance and standard deviation of X
variance <- ((winnings_win_2 - expected_value)^2 * p_win_2) +
  ((winnings_win_3 - expected_value)^2 * p_win_3) +
  ((winnings_lose_2 - expected_value)^2 * p_lose_2) +
  ((winnings_lose_3 - expected_value)^2 * p_lose_3)
std_dev <- sqrt(variance)

expected_value
```

```
[1] 57.888
```

```r
std_dev
```

```
[1] 795.1491
```

A standard deviation of 1028.629 indicates a wide variability in the possible outcomes for this scenario.

(iii) Create a simulation of 10,000 different 3 game series by using R to create 10,000 random values for X. Let these random values be denoted Y. Each Y value denotes an outcome of the series as defined by the probability distribution. These 10,000 outcomes represent a statistical sample of possible outcomes. Use this sample of outcomes to estimate the expected net win by using a 95% confidence interval. Does this confidence interval contain the theoretical E(X)

To simulate this series 10,000 times, a function will need to be written.

```
# Step 3: Simulate 10,000 series and compute a 95% confidence interval.
set.seed(123) # For reproducibility
n_simulations <- 10000

# Function to simulate a series
simulate_series <- function() {
  games <- c("Boston", "New York", "Boston")
  red_sox_wins <- 0
  yankees_wins <- 0
  net_winnings <- 0

  for (game in games) {
    if (game == "Boston") {
      red_sox_win <- runif(1) < p_win_home
    } else {
      red_sox_win <- runif(1) < p_win_away
    }

    if (red_sox_win) {
      red_sox_wins <- red_sox_wins + 1
      net_winnings <- net_winnings + 500
    } else {
      yankees_wins <- yankees_wins + 1
      net_winnings <- net_winnings - 520
    }

    if (red_sox_wins == 2 || yankees_wins == 2) {
      break
    }
  }

  return(net_winnings)
}
```

```r
# Simulate 10,000 series
simulated_winnings <- replicate(n_simulations, simulate_series())

# Sample mean and 95% confidence interval
sample_mean <- mean(simulated_winnings)
sample_std_dev <- sd(simulated_winnings)
confidence_interval <- sample_mean + c(-1, 1) * 1.96 *
  (sample_std_dev / sqrt(n_simulations))

sample_mean
```

```
[1] 64.702
```

```r
confidence_interval
```

```
[1] 49.14894 80.25506
```

```r
expected_value
```

```
[1] 57.888
```

After 10,000 simulations, the mean of the sample is 64.702 and the confidence interval is between 49.14894 and 80.25506. The expected value of 57.88 does fall with the confidence interval.

(iv) Construct a frequency distribution for Y. Use the Chi-squared goodness of fit test to verify how closely the distribution of Y has estimated the distribution of X.

For the frequency distribution, the simulated_winnings variable is converted to a table.

```r
# Frequency distribution of simulated winnings
freq_dist <- table(simulated_winnings)
freq_dist
```

```
simulated_winnings
-1040  -540   480  1000
 2231  2077  3083  2609
```

The Chi-squared test is run using the chisq.test function along with the frequency distribution and previously calculated scenario probabilities.

```
# Theoretical probabilities
theoretical_probs <- c(p_win_2, p_win_3, p_lose_2, p_lose_3)

# Chi-squared test
chi_squared_test <- chisq.test(freq_dist, p = theoretical_probs)
chi_squared_test
```

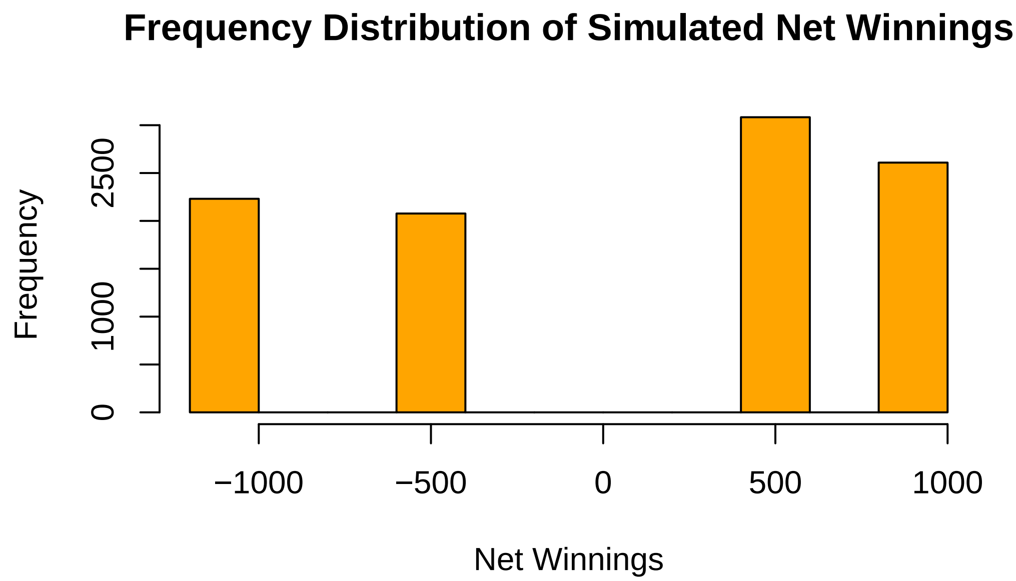
```

	Chi-squared test for given probabilities

data:  freq_dist
X-squared = 807.57, df = 3, p-value < 2.2e-16
```

The chi-square goodness of fit test yielded x-squared = 936.87 with df = 3 and p-value < 2.2e-16, leading to rejection of the null hypothesis at = 0.05. This indicates a statistically significant difference between this distribution and the theoretical distribution.

(v) Use your observations of parts (ii) and (iii) above to describe whether your betting strategy is favorable to you.

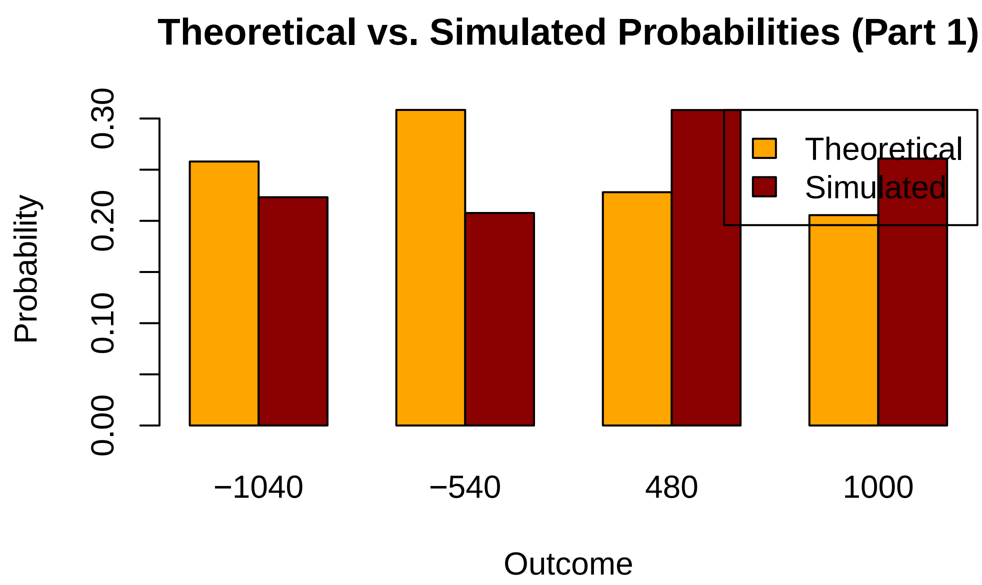
Given the results of the simulations, the better strategy seems mildly favorable. The expected value is positive and falls within the confidence interval. As seen in the winnings frequency chart below, there are more winning scenarios that losing.

```
# Visualization 1: Histogram of Simulated Winnings
hist(simulated_winnings, breaks=10,
     main="Frequency Distribution of Simulated Net Winnings",
     xlab="Net Winnings", ylab="Frequency", col="orange")
```



As seen in this chart, the simulations performed better than the theoretical but they are close.

```r
# Visualization 3: Theoretical vs. Simulated Probabilities
theoretical_probs <- c(p_win_2, p_win_3, p_lose_2, p_lose_3)
simulated_probs <- table(simulated_winnings) / n_simulations

barplot(rbind(theoretical_probs, simulated_probs), beside=TRUE,
        col=c("orange", "darkred"),
        main="Theoretical vs. Simulated Probabilities (Part 1)",
        xlab="Outcome", ylab="Probability",
        legend.text=c("Theoretical", "Simulated"), args.legend=list(x="topright"))
```



However, the large standard deviation makes for a less than certain betting solution. Given all the factors at play in multi-game series, adjusting the wager amount to optimize winning possibilities based on factors such as starting pitchers or the availability of regular starters would go father in optimizing the betting dollars.

Part Two - Best of 3 (NY, BOS, NY)

(i) Calculate the probability that the Red Sox will win the series.

For the next part of the study, the same methodology will be used, only now with the Yankees having home field advantage in the series.

Once again, the probabilities for a given game are loaded and used to calculate the probability of each Red Sox winning scenario. Then those are added to produce the final probability of Boston winning the series.

```
# Step 1: Calculate the probability that the Red Sox will win the series.
# Probabilities
p_win_home <- 0.6  # Red Sox win at home
p_win_away <- 0.43 # Red Sox win at Yankees' home (1 - 0.57)

# Probability of winning in 2 games
```

```r
p_win_2 <- p_win_away * p_win_home
p_win_2
```

```
[1] 0.258
```

```r
# Probability of winning in 3 games
p_win_3 <- (p_win_away * (1 - p_win_home) * p_win_away) +
  ((1 - p_win_away) * p_win_home * p_win_away)
p_win_3
```

```
[1] 0.22102
```

```r
# Total probability of Red Sox winning the series
p_total <- p_win_2 + p_win_3
p_total
```

```
[1] 0.47902
```

Which is calculated at 0.47902 or 47.902%.

(ii) Construct the theoretical probability distribution for your net winnings in dollars (X) for the series. From this theoretical calculation you should also compute and record the expected value of the net winnings (the mean of X) and the theoretical standard deviation of X.

Following the same steps as in the first series analysis, the winning amounts are loaded along with the probabilities of the Red Sox losing scenarios before calculating the expected value.

```r
# Step 2: Construct the theoretical probability distribution for net winnings (X).
# Net winnings for each outcome
cat("Winnings for BOS sweep =",winnings_win_2, "\n")
```

```
Winnings for BOS sweep = 1000
```

```r
cat("Winnings for BOS win in 3 =", winnings_win_3, "\n")
```

```
Winnings for BOS win in 3 = 480
```

```r
cat("Winnings for NYY sweep = ", winnings_lose_2, "\n")
```

```
Winnings for NYY sweep =  -1040
```

```r
cat("Winnings for NYY win in 3 =", winnings_lose_3, "\n")
```

```
Winnings for NYY win in 3 = -540
```

```r
# Probabilities for each outcome
cat("Probability for BOS sweep =", p_win_2, "\n")
```

```
Probability for BOS sweep = 0.258
```

```r
cat("Probability for BOS win in 3 =", p_win_3, "\n")
```

```
Probability for BOS win in 3 = 0.22102
```

```r
p_lose_2 <- (1 - p_win_away) * (1 - p_win_home)
cat("Probability for NYY sweep =", p_lose_2, "\n")
```

```
Probability for NYY sweep = 0.228
```

```r
p_lose_3 <- (p_win_away * (1 - p_win_home) * (1 - p_win_away)) +
  ((1 - p_win_away) * p_win_home * (1 - p_win_away))
cat("Probability for NYY win in 3 =", p_lose_3, "\n")
```

```
Probability for NYY win in 3 = 0.29298
```

```r
# Expected value (mean) of X
expected_value <- (winnings_win_2 * p_win_2) +
  (winnings_win_3 * p_win_3) +
  (winnings_lose_2 * p_lose_2) +
  (winnings_lose_3 * p_lose_3)

# Variance and standard deviation of X
variance <- ((winnings_win_2 - expected_value)^2 * p_win_2) +
  ((winnings_win_3 - expected_value)^2 * p_win_3) +
```

```
  ((winnings_lose_2 - expected_value)^2 * p_lose_2) +
  ((winnings_lose_3 - expected_value)^2 * p_lose_3)
std_dev <- sqrt(variance)

cat("Expected Value =", expected_value, "\n")
```

```
Expected Value = -31.2396
```

```
cat("Standard Deviation =", std_dev, "\n")
```

```
Standard Deviation = 799.9905
```

Here the expected value is -31.24 with a standard deviation of 799.9905. So in the scenario the expectation would be to lose money but with a large standard deviation once again.

(iii) Create a simulation of 10,000 different 3 game series by using R to create 10,000 random values for X. Let these random values be denoted Y. Each Y value denotes an outcome of the series as defined by the probability distribution. These 10,000 outcomes represent a statistical sample of possible outcomes. Use this sample of outcomes to estimate the expected net win by using a 95% confidence interval. Does this confidence interval contain the theoretical E(X)

The same simulation function from part 1 is used but changing the location of the games to match the new scenario.

```
# Step 3: Simulate 10,000 series and compute a 95% confidence interval.
set.seed(123) # For reproducibility
n_simulations <- 10000

# Function to simulate a series
simulate_series <- function() {
  games <- c("New York", "Boston", "New York")
  red_sox_wins <- 0
  yankees_wins <- 0
  net_winnings <- 0

  for (game in games) {
    if (game == "Boston") {
      red_sox_win <- runif(1) < p_win_home
    } else {
```

```r
      red_sox_win <- runif(1) < p_win_away
    }

    if (red_sox_win) {
      red_sox_wins <- red_sox_wins + 1
      net_winnings <- net_winnings + 500
    } else {
      yankees_wins <- yankees_wins + 1
      net_winnings <- net_winnings - 520
    }

    if (red_sox_wins == 2 || yankees_wins == 2) {
      break
    }
  }

  return(net_winnings)
}
```

And the simulation is run another 10,000 times.

```r
# Simulate 10,000 series
simulated_winnings <- replicate(n_simulations, simulate_series())

# Sample mean and 95% confidence interval
sample_mean <- mean(simulated_winnings)
sample_std_dev <- sd(simulated_winnings)
confidence_interval <- sample_mean + c(-1, 1) * 1.96 *
  (sample_std_dev / sqrt(n_simulations))

cat("Sample Mean =", sample_mean, "\n")
```

```
Sample Mean = -25.95
```

```r
cat("Confidence Interval =", confidence_interval, "\n")
```

```
Confidence Interval = -41.58931 -10.31069
```

And these results follow the same pattern, the sample mean is a loss of 19.69 with the confidence interval being between -35.41 and -3.97. The expected value does fall within this range.

(iv) Construct a frequency distribution for Y. Use the Chi-squared goodness of fit test to verify how closely the distribution of Y has estimated the distribution of X.

```r
# Frequency distribution of simulated winnings
freq_dist <- table(simulated_winnings)

# Theoretical probabilities
theoretical_probs <- c(p_win_2, p_win_3, p_lose_2, p_lose_3)

# Chi-squared test
chi_squared_test <- chisq.test(freq_dist, p = theoretical_probs)
chi_squared_test
```

```
    Chi-squared test for given probabilities

data:  freq_dist
X-squared = 358.42, df = 3, p-value < 2.2e-16
```

The chi-square goodness of fit test yielded x-squared = 254.55 with df = 3 and p-value < 2.2e-16, leading to rejection of the null hypothesis at   = 0.05. This indicates a statistically significant difference between this distribution and the theoretical distribution.

(v) Use your observations of parts (ii) and (iii) above to describe whether your betting strategy is favorable to you.

This is not a good betting strategy. The same lack of meaningful variables from the first scenario are at play but not without the home field advantage.

Part 3 - Best of 5 (BOS, NY, BOS, NY, BOS)

For this part of the study, the same assumptions are made but now with a best of 5 series. These games will alternate between Boston and NY but start in Boston. This also means that a potentially decisive 5th game would be in Boston.

(i) Calculate the probability that the Red Sox will win the series.

p_total

```r
# Probabilities
p_win_home <- 0.6  # Red Sox win at home
p_win_away <- 0.43 # Red Sox win at Yankees' home (1 - 0.57)

# Probability of winning in 3 games
p_win_3 <- p_win_home * p_win_away * p_win_home
cat("Probability of BOS sweep =", p_win_3, "\n")
```

```
Probability of BOS sweep = 0.1548
```

```r
# Probability of winning in 4 games
p_win_4 <- (p_win_home * p_win_away * (1 - p_win_home) * p_win_away) +
  (p_win_home * (1 - p_win_away) * p_win_home * p_win_away) +
  ((1 - p_win_home) * p_win_away * p_win_home * p_win_away)
cat("Probability of BOS win in 4 =", p_win_4, "\n")
```

```
Probability of BOS win in 4 = 0.176988
```

```r
# Probability of winning in 5 games
p_win_5 <- (p_win_home * p_win_away * (1 - p_win_home) * (1 - p_win_away) * p_win_home) +
  (p_win_home * (1 - p_win_away) * p_win_home * (1 - p_win_away) * p_win_home) +
  ((1 - p_win_home) * p_win_away * p_win_home * (1 - p_win_away) * p_win_home) +
  ((1 - p_win_home) * p_win_away * (1 - p_win_home) * p_win_away * p_win_home)
cat("Probability of BOS win in 5 =", p_win_5, "\n")
```

```
Probability of BOS win in 5 = 0.1585176
```

```r
# Total probability of Red Sox winning the series
p_total <- p_win_3 + p_win_4 + p_win_5
cat("Probability of BOS win series =", p_total, "\n")
```

```
Probability of BOS win series = 0.4903056
```

The probability of Boston winning the series is 0.4903056 or 49.03%

(ii) Construct the theoretical probability distribution for your net winnings in dollars (X) for the series. From this theoretical calculation you should also compute and record the expected value of the net winnings (the mean of X) and the theoretical standard deviation of X.

With more games, there are more potential winning scenarios. Those are loaded along with the probabilities of New York winning in 3, 4, or 5 games.

```
# Net winnings for each outcome
winnings_win_3 <- 1500
winnings_win_4 <- 980
winnings_win_5 <- 460
winnings_lose_3 <- -1560
winnings_lose_4 <- -1060
winnings_lose_5 <- -560

# Probabilities for each outcome
p_lose_3 <- (1 - p_win_home) * (1 - p_win_away) * (1 - p_win_home)
cat("Probability of BOS being swept =", p_lose_3, "\n")
```

```
Probability of BOS being swept = 0.0912
```

```
p_lose_4 <- (p_win_home * (1 - p_win_away) * (1 - p_win_home) * (1 - p_win_away)) +
  ((1 - p_win_home) * p_win_away * (1 - p_win_home) * (1 - p_win_away))
cat("Probability of BOS losing in 4 =", p_lose_4, "\n")
```

```
Probability of BOS losing in 4 = 0.117192
```

```
p_lose_5 <- (p_win_home * p_win_away * (1 - p_win_home) * (1 - p_win_away) * (1 - p_win_home
  (p_win_home * (1 - p_win_away) * p_win_home * (1 - p_win_away) * (1 - p_win_home)) +
  ((1 - p_win_home) * p_win_away * p_win_home * (1 - p_win_away) * (1 - p_win_home)) +
  ((1 - p_win_home) * p_win_away * (1 - p_win_home) * p_win_away * (1 - p_win_home))
cat("Probability of BOS losing in 5 =", p_lose_5, "\n")
```

```
Probability of BOS losing in 5 = 0.1056784
```

```
p_loss_total <- p_lose_3 + p_lose_4 + p_lose_5
cat("Probability of BOS lose series =", p_total, "\n")
```

```
Probability of BOS lose series = 0.4903056
```

With those calculated, the expected mean for winnings can be calucluated along with the variance and standard deviation.

```
# Expected value (mean) of X
expected_value <- (winnings_win_3 * p_win_3) +
  (winnings_win_4 * p_win_4) +
  (winnings_win_5 * p_win_5) +
  (winnings_lose_3 * p_lose_3) +
  (winnings_lose_4 * p_lose_4) +
  (winnings_lose_5 * p_lose_5)

# Variance and standard deviation of X
variance <- ((winnings_win_3 - expected_value)^2 * p_win_3) +
  ((winnings_win_4 - expected_value)^2 * p_win_4) +
  ((winnings_win_5 - expected_value)^2 * p_win_5) +
  ((winnings_lose_3 - expected_value)^2 * p_lose_3) +
  ((winnings_lose_4 - expected_value)^2 * p_lose_4) +
  ((winnings_lose_5 - expected_value)^2 * p_lose_5)
std_dev <- sqrt(variance)

cat("Expected Value =", expected_value, "\n")
```

```
Expected Value = 152.8909
```

```
cat("Variance =", variance, "\n")
```

```
Variance = 910635.1
```

```
cat("Standard Deviation =", std_dev, "\n")
```

```
Standard Deviation = 954.272
```

The expected value is 152.89 with a standard deviation of 954.27. So this also indicates a winning betting strategy but still a large standard deviation.

(iii) Create a simulation of 10,000 different 3 game series by using R to create 10,000 random values for X. Let these random values be denoted Y. Each Y value denotes an outcome of the series as defined by the probability distribution. These 10,000 outcomes represent a statistical sample of possible outcomes. Use this sample of outcomes to

estimate the expected net win by using a 95% confidence interval. Does this confidence interval contain the theoretical E(X)

For this simulation, a similar function can be utilized but expanded for the new series length and format.

```
set.seed(123) # For reproducibility
n_simulations <- 10000

# Function to simulate a series
simulate_series <- function() {
  games <- c("Boston", "New York", "Boston", "New York", "Boston")
  red_sox_wins <- 0
  yankees_wins <- 0
  net_winnings <- 0

  for (game in games) {
    if (game == "Boston") {
      red_sox_win <- runif(1) < p_win_home
    } else {
      red_sox_win <- runif(1) < p_win_away
    }

    if (red_sox_win) {
      red_sox_wins <- red_sox_wins + 1
      net_winnings <- net_winnings + 500
    } else {
      yankees_wins <- yankees_wins + 1
      net_winnings <- net_winnings - 520
    }

    if (red_sox_wins == 3 || yankees_wins == 3) {
      break
    }
  }

  return(net_winnings)
}
```

The same 10,000 simulations are run and then sample mean and confidence interval is calculated.

```r
# Simulate 10,000 series
simulated_winnings <- replicate(n_simulations, simulate_series())

# Sample mean and 95% confidence interval
sample_mean <- mean(simulated_winnings)
sample_std_dev <- sd(simulated_winnings)
confidence_interval <- sample_mean + c(-1, 1) * 1.96 *
  (sample_std_dev / sqrt(n_simulations))

cat("sample_mean =", sample_mean, "\n")
```

```
sample_mean = 101.298
```

```r
cat("Confidence Interval =", confidence_interval, "\n")
```

```
Confidence Interval = 81.38021 121.2158
```

The sample mean was calculated at 101.29, compared to the expected value of 152.89. The expected value is also outside of the confidence interval of 81.38 to 121.22.

(iv) Construct a frequency distribution for Y. Use the Chi-squared goodness of fit test to verify how closely the distribution of Y has estimated the distribution of X.

```r
# Frequency distribution of simulated winnings
freq_dist <- table(simulated_winnings)

# Theoretical probabilities
theoretical_probs <- c(p_win_3, p_win_4, p_win_5, p_lose_3, p_lose_4, p_lose_5)
theoretical_probs <- theoretical_probs / sum(theoretical_probs)
# This normalizes to sum 1

# Chi-squared test
chi_squared_test <- chisq.test(freq_dist, p = theoretical_probs)
chi_squared_test
```

```

	Chi-squared test for given probabilities

data:  freq_dist
X-squared = 2323.2, df = 5, p-value < 2.2e-16
```

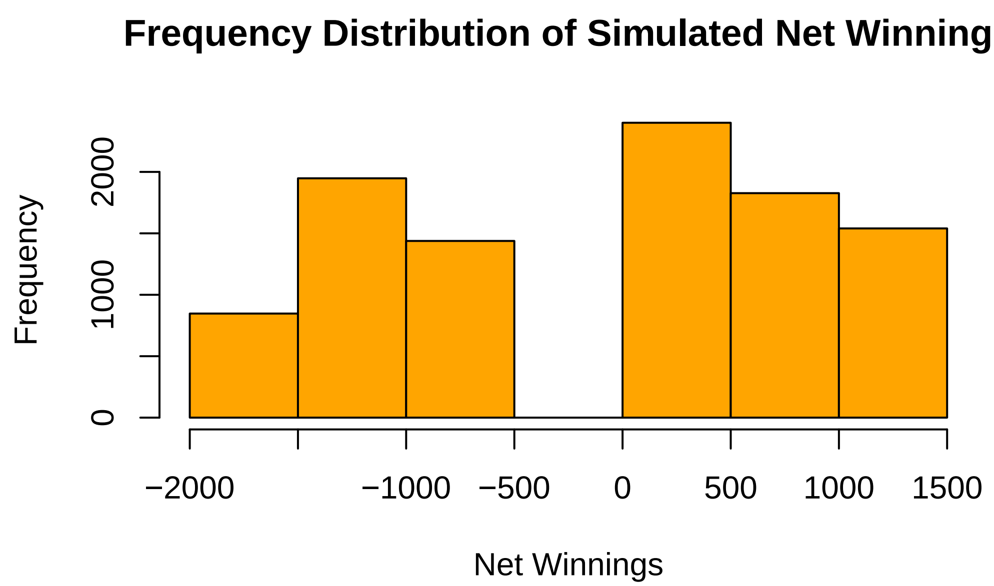
As with the previous Chi-squared tests, the large value indicates a difference between the distribution and the theoretical distribution.

(v) Use your observations of parts (ii) and (iii) above to describe whether your betting strategy is favorable to you.

With the larger number of games involved, and those game being played at home, the strategy of betting on Boston to win each game involves a higher potential profit. The statistical analysis bears this out with the higher expected values and large confidence interval. However, the large standard of deviation also implies high risk for this strategy.

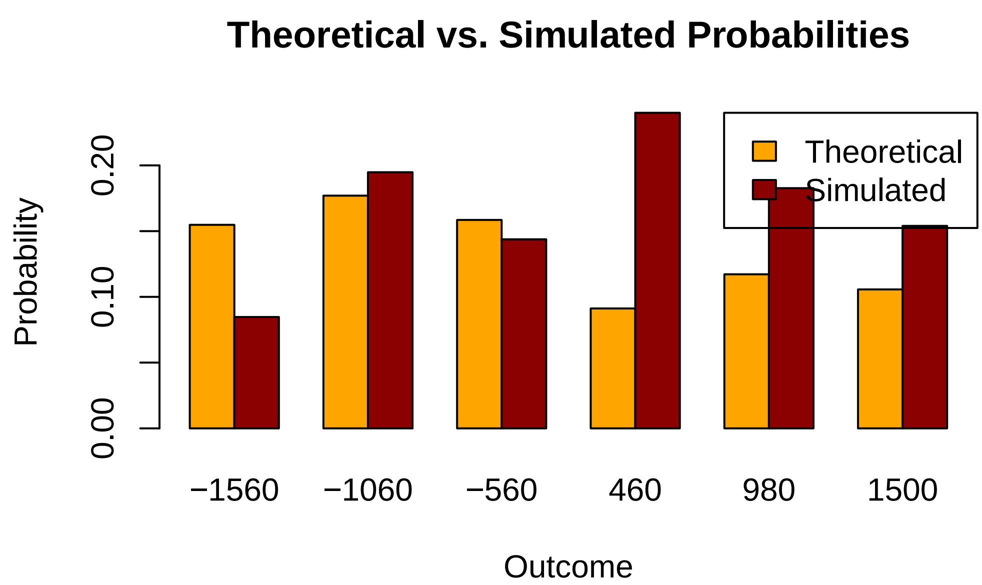
As shown in the visualization below, there are more scenarios that result in a net positive the negative from this betting strategy, but it is far from a certainty,

```
hist(simulated_winnings, breaks=10,
     main="Frequency Distribution of Simulated Net Winnings",
     xlab="Net Winnings", ylab="Frequency", col="orange")
```



Also, in comparing the simulated model to the theoretical demonstrates that there are differences. Some of which can be explained by a small sample size of only up to 5 games but is also indicative of close probabilities that lead to this being a risky strategy.

```r
# Visualization 3: Theoretical vs. Simulated Probabilities
theoretical_probs <- c(p_win_3, p_win_4, p_win_5, p_lose_3, p_lose_4, p_lose_5)
simulated_probs <- table(simulated_winnings) / n_simulations

barplot(rbind(theoretical_probs, simulated_probs), beside=TRUE,
        col=c("orange", "darkred"),
        main="Theoretical vs. Simulated Probabilities",
        xlab="Outcome", ylab="Probability",
        legend.text=c("Theoretical", "Simulated"), args.legend=list(x="topright"))
```



Conclusion

This study evaluated the betting strategy of always betting on the Boston Red Sox to win the game in three different playoff series formats. This is a strategy often employed by fans of a particular team, they are putting money on their preferred outcome rather than trying to maximize their potential winnings.

In this case, based on the winnings potential ($500 for a win and loss of $520 for a loss), it would appear that the teams are fairly well matched in the opinion of the betting market. Also given then the winning probabilities given being close, but flipping in the favor of each road team - it is clear that this risky strategy will perform best in scenarios where Boston has home field advantage and plays more home than away games.

References

1. Albright, S. (2016) Business Analytics. Sixth Edition. Cengage Learning. Boston, MA.
2. Evans, J. R. (2013). Statistics, data analysis, and decision modeling: International Edition. Pearson Higher Ed.