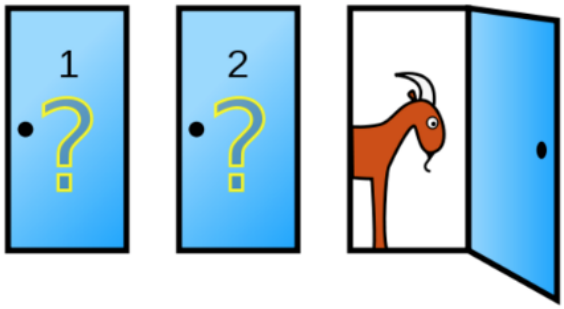
Monty Hall Simulation

데이터테크전공 20173204 곽명빈

2020-11-23

4. 세 개의 문 중에 하나를 선택하여 문 뒤에 있는 선물을 가질 수 있는 게임쇼에 참가했다. 한 문 뒤에는 자동차가 있고, 나머지 두 문 뒤에는 염소가 있다. 이때 어떤 참가자가 1번 문을 선택했을 때, 게임쇼 진행자는 3번 문을 열어 문 뒤에 염소가 있음을 보여주면서 1번 대신 2번을 선택하겠냐고 물었다. 이 때 어떤 선택을 해야 하는가?



가. 선택을 바꾼다.
나. 선택을 고수한다.

Single Trial

```
set.seed(1)
monty_hall <- function() {
  key <- 1:3 %>%
    sample(size = 1)
  goat <- 1:3 %>%
    setdiff(key)
  contestant <- 1:3 %>%
    sample(size = 1)
  monty <- contestant %>%
    `==` (key) %>%
    ifelse(goat %>% sample(size = 1),
           goat %>% setdiff(contestant))
  switch <- 1:3 %>%
    setdiff(c(contestant, monty))
  result <- switch %>%
    `==`(key) %>%
    ifelse("Switching wins", "Staying wins")
  c("Key" = key,
    "Contestant" = contestant,
    "Monty" = monty,
    "Switch" = switch,
    "Result" = result)
}
monty_hall()
```

```
##             Key       Contestant            Monty           Switch           Result
##             "1"              "3"              "2"              "1" "Switching wins"
```

N trials

```
N <- 30
monty_result <-
  replicate(N, monty_hall()) %>%
  t
monty_result
```

```
##       Key Contestant Monty Switch Result
##  [1,] "1" "2"        "3"   "1"    "Switching wins"
##  [2,] "1" "3"        "2"   "1"    "Switching wins"
##  [3,] "3" "2"        "1"   "3"    "Switching wins"
##  [4,] "2" "3"        "1"   "2"    "Switching wins"
##  [5,] "3" "1"        "2"   "3"    "Switching wins"
##  [6,] "1" "1"        "3"   "2"    "Staying wins"
##  [7,] "2" "2"        "3"   "1"    "Staying wins"
##  [8,] "3" "1"        "2"   "3"    "Switching wins"
##  [9,] "3" "1"        "2"   "3"    "Switching wins"
## [10,] "1" "1"        "2"   "3"    "Staying wins"
## [11,] "2" "1"        "3"   "2"    "Switching wins"
## [12,] "1" "2"        "3"   "1"    "Switching wins"
## [13,] "2" "2"        "1"   "3"    "Staying wins"
## [14,] "3" "1"        "2"   "3"    "Switching wins"
## [15,] "3" "2"        "1"   "3"    "Switching wins"
## [16,] "2" "2"        "3"   "1"    "Staying wins"
## [17,] "3" "2"        "1"   "3"    "Switching wins"
## [18,] "1" "3"        "2"   "1"    "Switching wins"
## [19,] "2" "1"        "3"   "2"    "Switching wins"
## [20,] "1" "3"        "2"   "1"    "Switching wins"
## [21,] "2" "2"        "1"   "3"    "Staying wins"
## [22,] "3" "2"        "1"   "3"    "Switching wins"
## [23,] "2" "2"        "3"   "1"    "Staying wins"
## [24,] "1" "2"        "3"   "1"    "Switching wins"
## [25,] "2" "2"        "3"   "1"    "Staying wins"
## [26,] "1" "3"        "2"   "1"    "Switching wins"
## [27,] "3" "2"        "1"   "3"    "Switching wins"
## [28,] "3" "3"        "2"   "1"    "Staying wins"
## [29,] "2" "3"        "1"   "2"    "Switching wins"
## [30,] "3" "1"        "2"   "3"    "Switching wins"
```

```
table(monty_result[, 5])
```

```
##
##   Staying wins Switching wins
##              9             21
```

```
sum(monty_result[, 5] == "Switching wins") / N
```

```
## [1] 0.7
```

```
cumsum(monty_result[, 5] == "Switching wins")
```

```
##  [1]  1  2  3  4  5  5  5  6  7  7  8  9  9 10 11 11 12 13 14 15 15 16 16 17 17 18 19 19 20
21
```

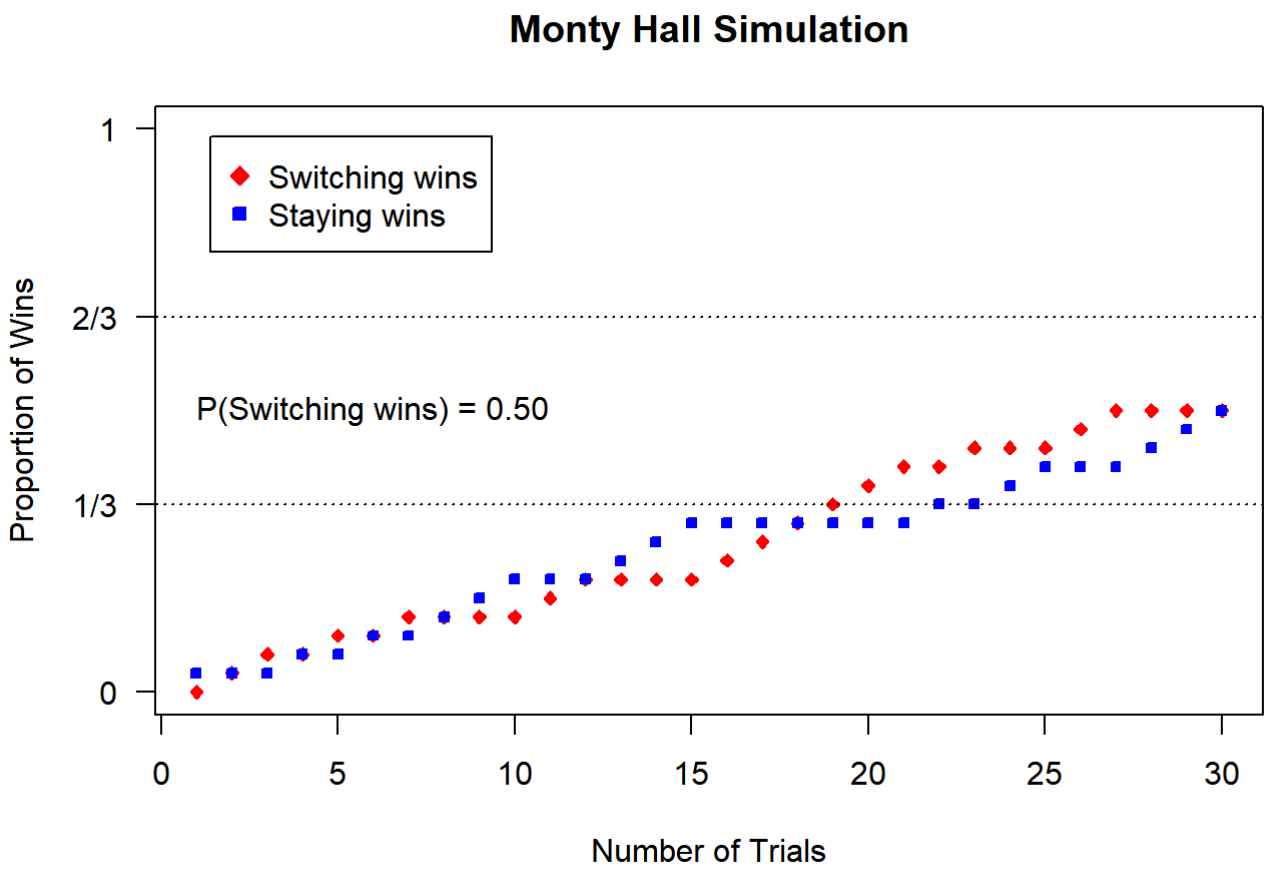
```
cumsum(monty_result[, 5] == "Staying wins")
```

```
##  [1] 0 0 0 0 0 1 2 2 2 3 3 3 4 4 4 5 5 5 5 5 6 6 7 7 8 8 8 9 9 9
```

```
y_switch <- cumsum(monty_result[, 5] == "Switching wins")
# y_stay <- cumsum(monty_result[, 5] == "Staying wins")
y_stay <- 1:N - y_switch
```

plot

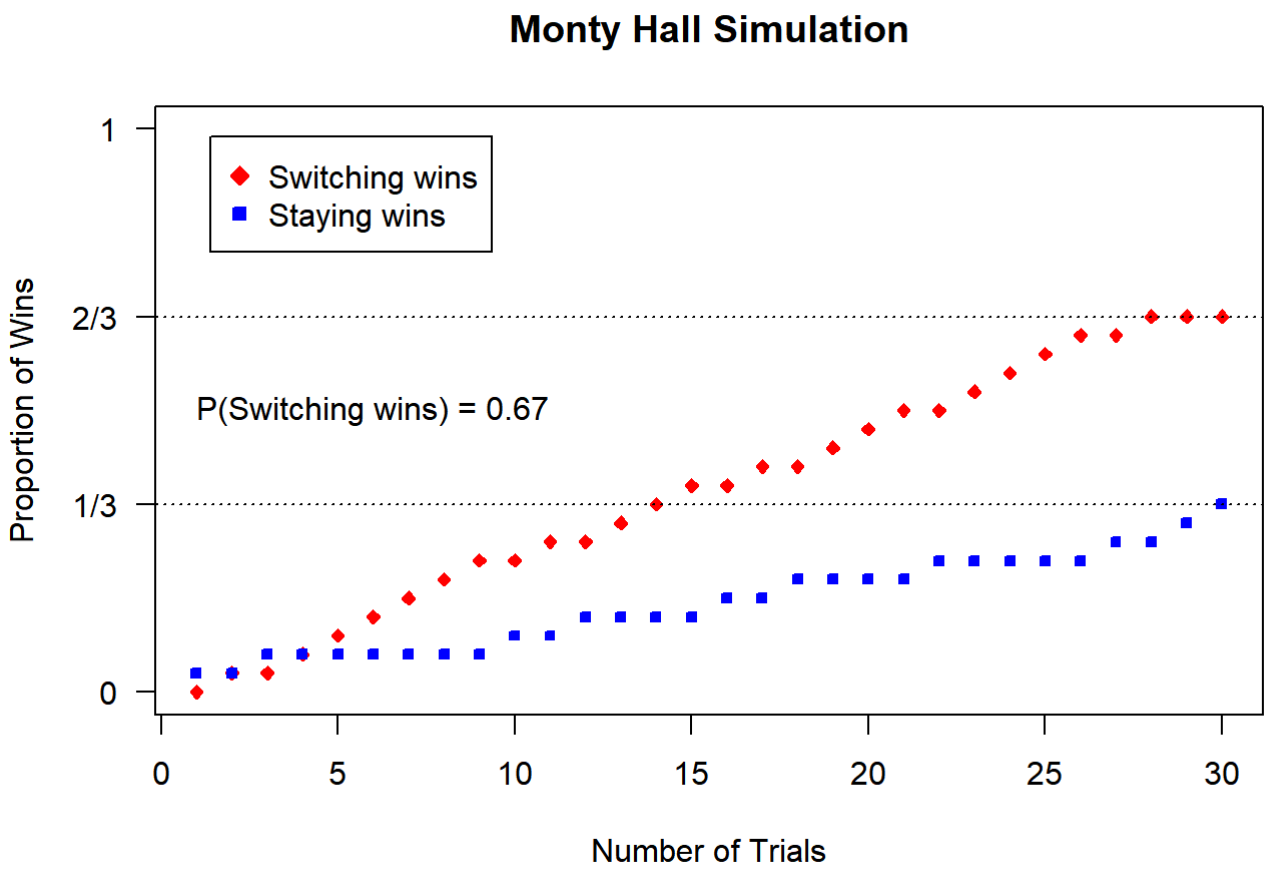
```
monty_plot <- function(N) {
  monty_result <-
    replicate(N, monty_hall()) %>%
    t
  y_switch <- cumsum(monty_result[, 5] == "Switching wins")
  y_stay <- 1:N - y_switch
#  y_stay <- cumsum(monty_result[, 5] == "Staying wins")
  p_wins <- sum(monty_result[, 5] == "Switching wins") / N
plot(x = 1:N,
     y = y_switch / N,
     pch = 23,
     col = "red",
     bg = "red",
     ylim = c(0, 1),
     xlab = "Number of Trials",
     ylab = "Proportion of Wins",
     yaxt = "n",
     cex = 0.7)
axis(side = 2,
     at = c(0, 1/3, 2/3, 1),
     labels = c("0", "1/3", "2/3", "1"), las = 2)
points(x = 1:N,
       y = y_stay / N,
       pch = 22,
       col = "blue",
       bg = "blue",
       cex = 0.7)
abline(h = c(1/3, 2/3), lty = 3)
title(main = "Monty Hall Simulation")
legend("topleft",
       inset = 0.05,
       legend = c("Switching wins", "Staying wins"),
       pch = c(23, 22),
       col = c("red", "blue"),
       pt.bg = c("red", "blue"))
text(x = N / 5, y = 1 / 2,
     labels = paste0("P(Switching wins) = ", format(p_wins, digits = 2, nsmall = 2)))
p_wins
}
monty_plot(30)
```




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

Repetitions

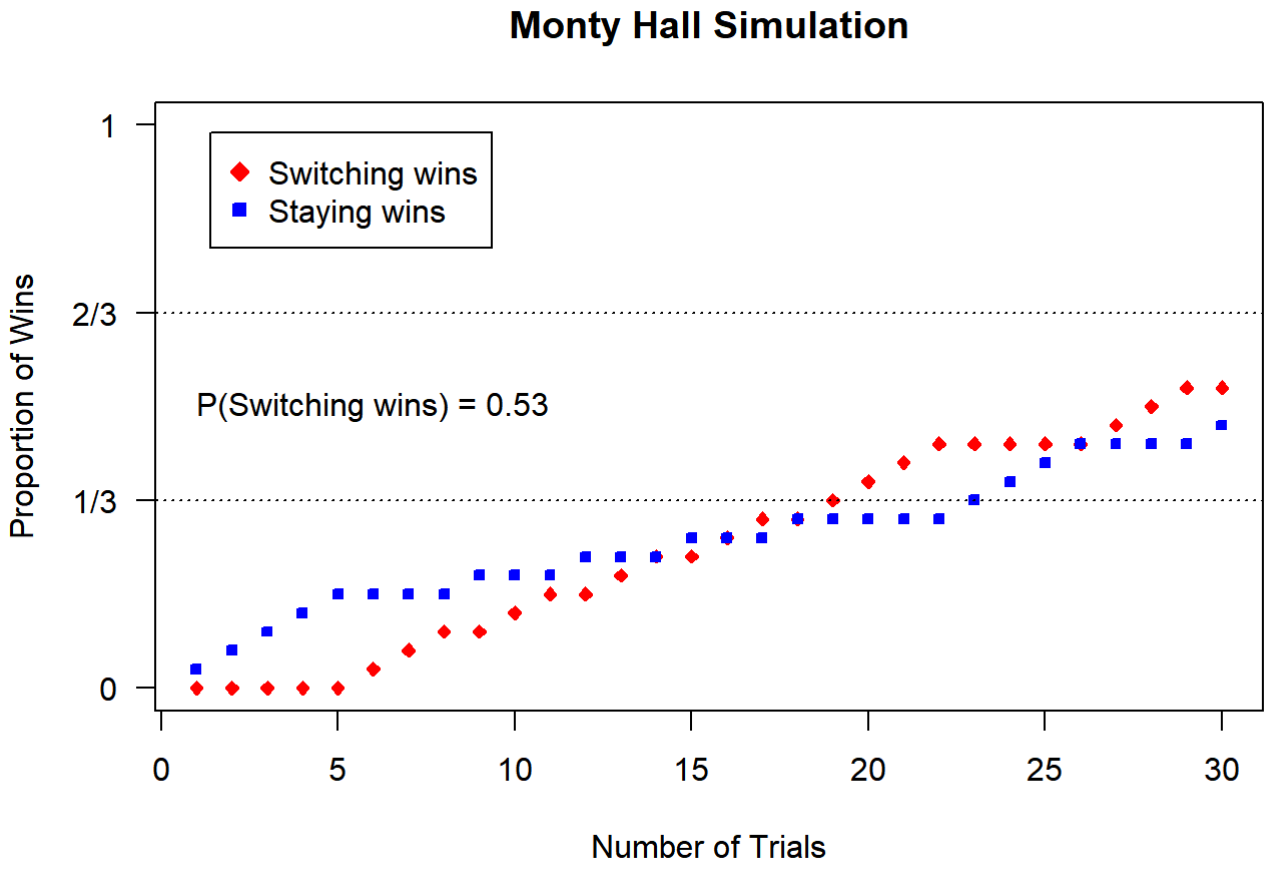
```
Prop_Switching_wins <- monty_plot(30)
```

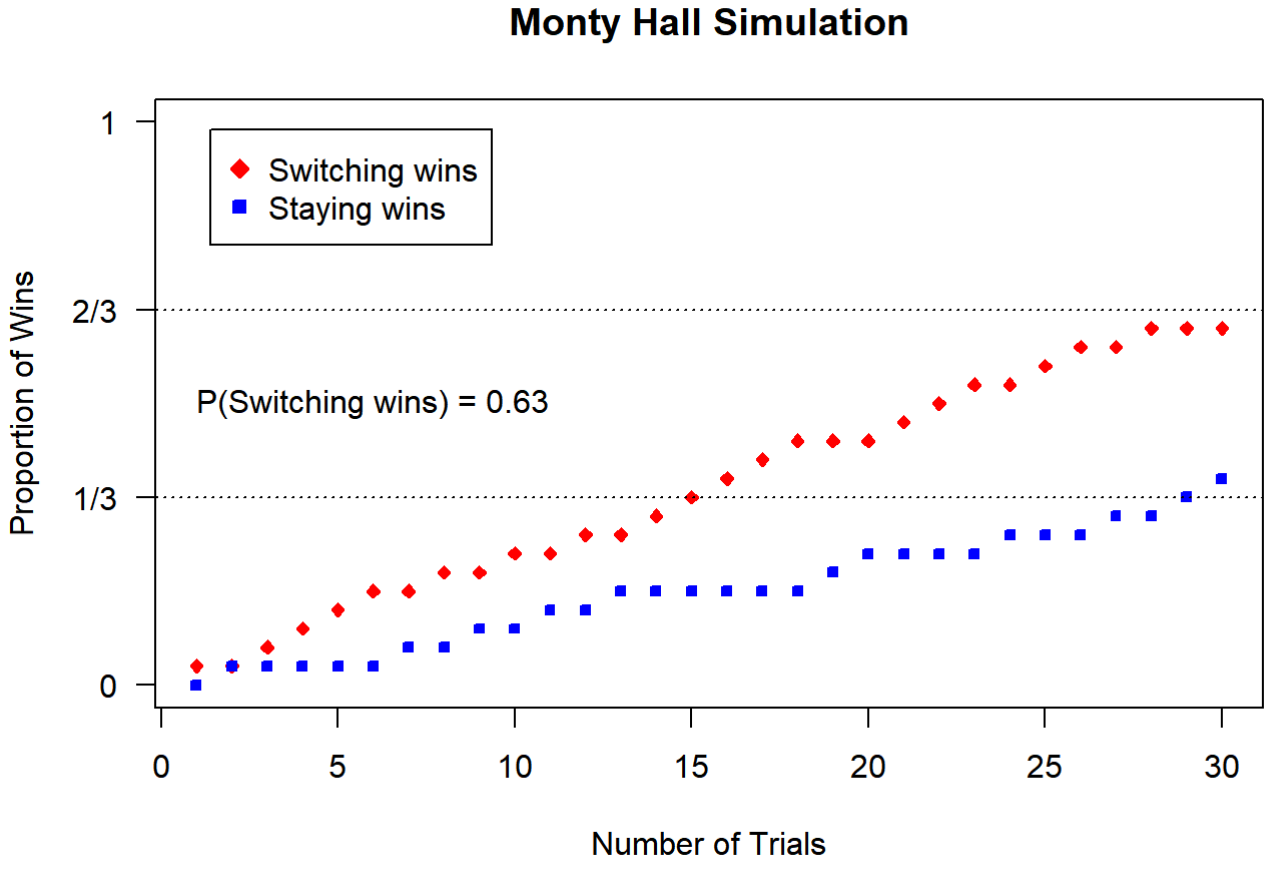



```
Prop_Switching_wins
```

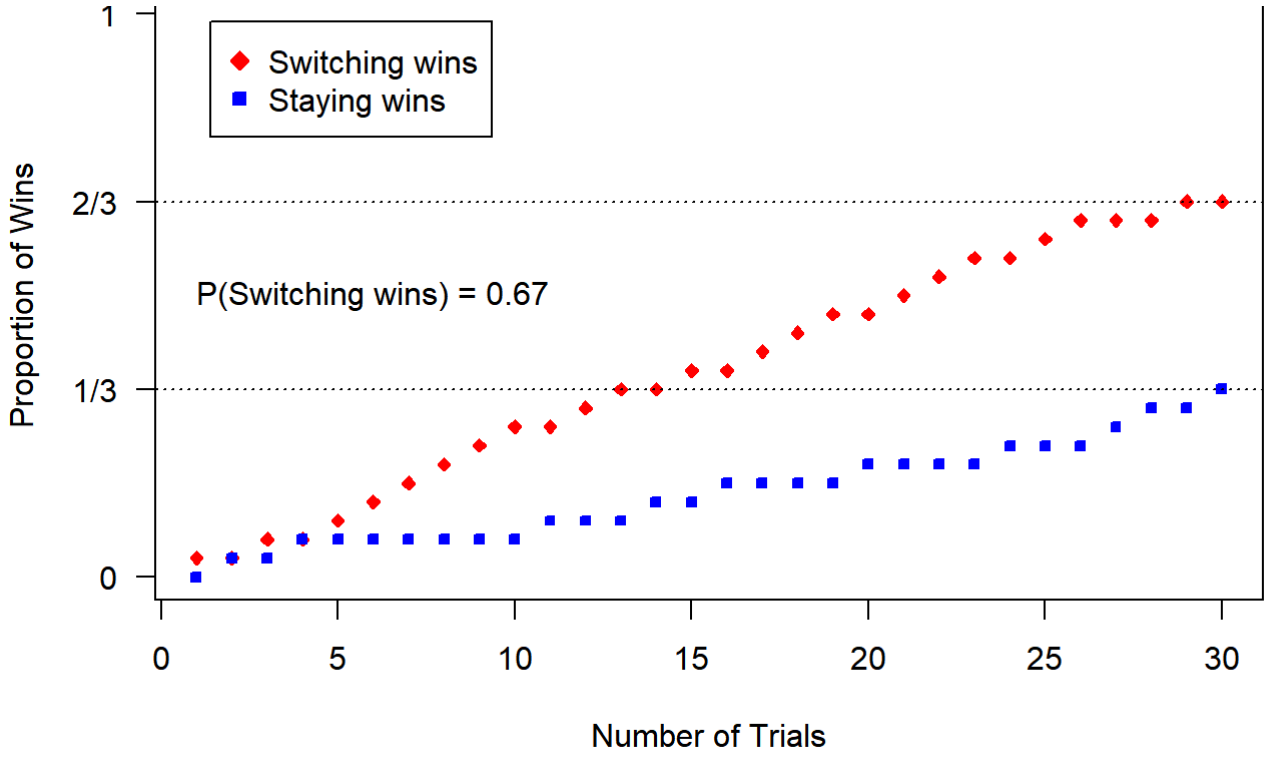
```
## [1] 0.6666667
```

```
Prop_Switching_wins_10 <- replicate(10, monty_plot(30))
```

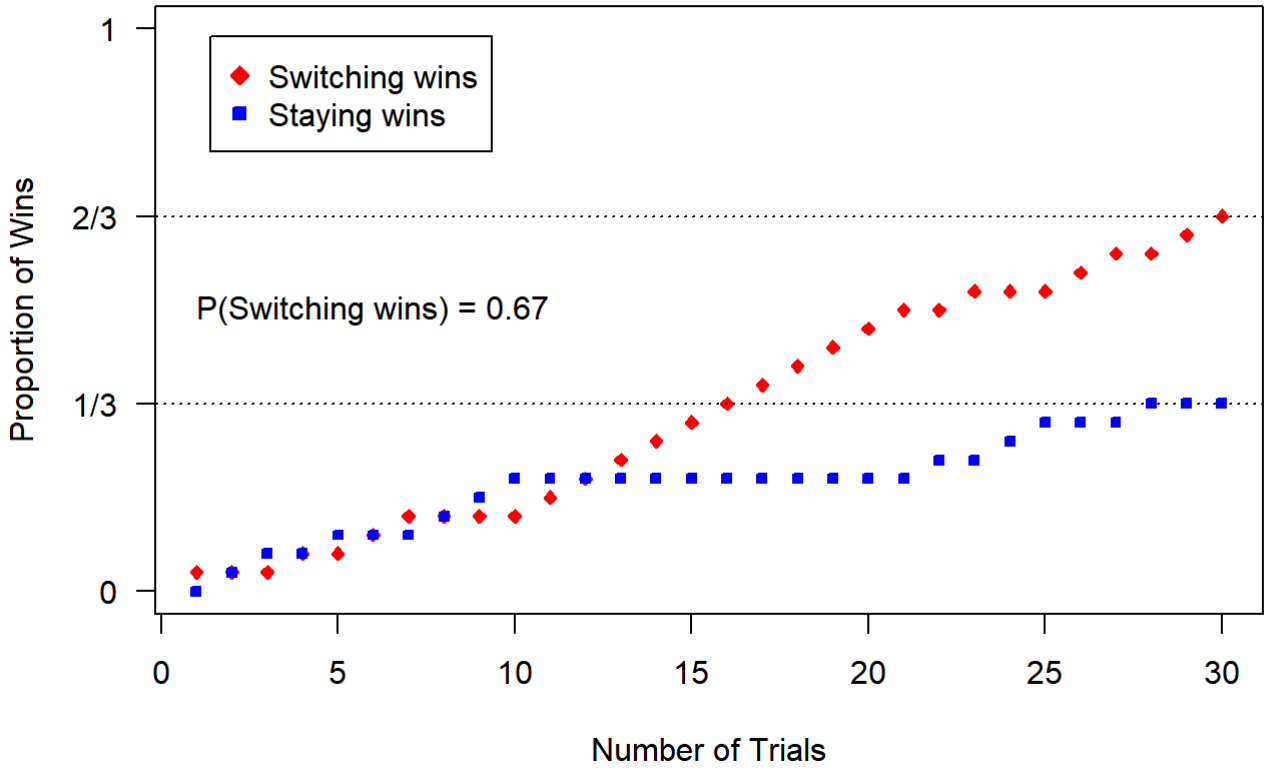

Monty Hall Simulation
Switching wins
Staying wins
P(Switching wins) = 0.53
Proportion of Wins
1
2/3
1/3
0
0
5
10
15
20
25
30
Number of Trials

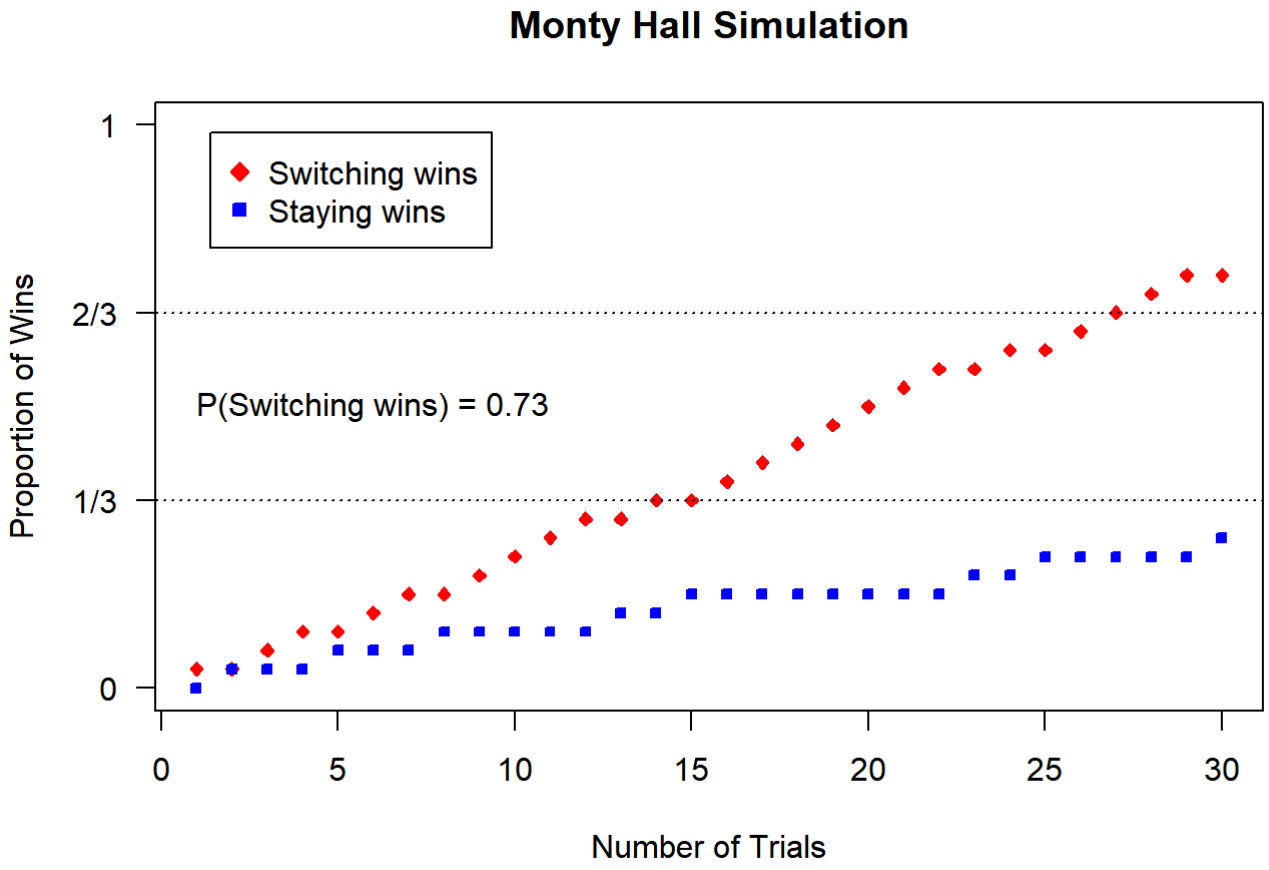

Monty Hall Simulation
Switching wins
Staying wins
P(Switching wins) = 0.63
Proportion of Wins
1
2/3
1/3
0
0
5
10
15
20
25
30
Number of Trials

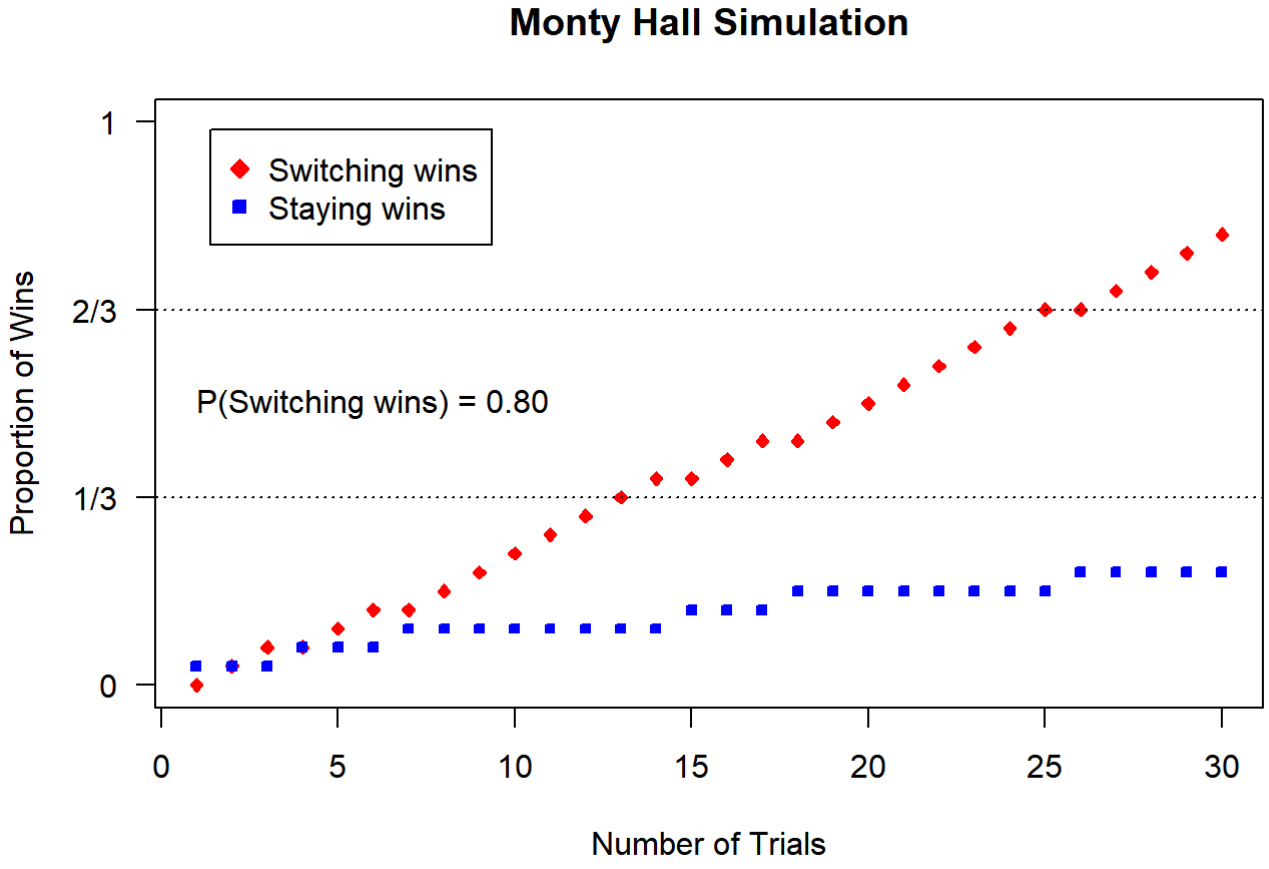

Monty Hall Simulation

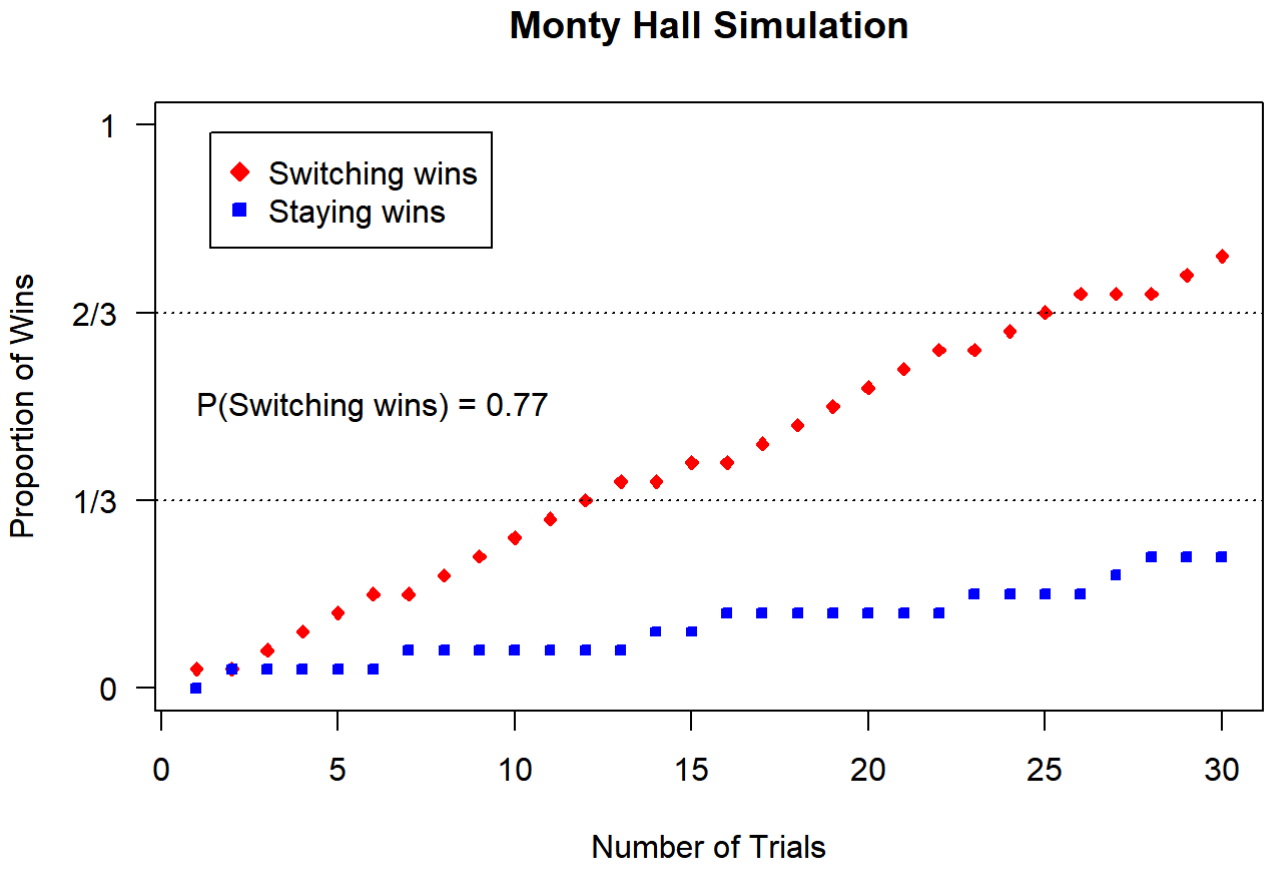


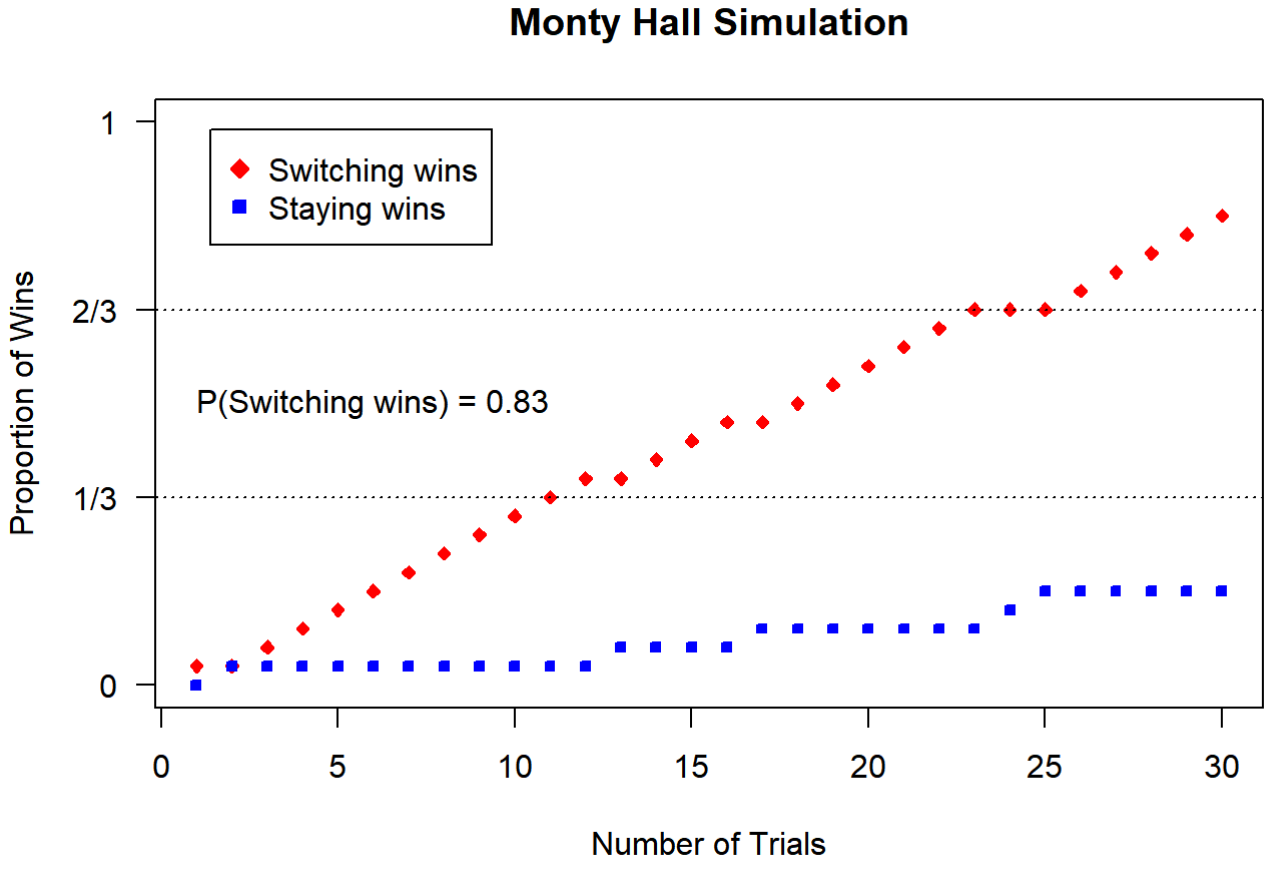
Monty Hall Simulation

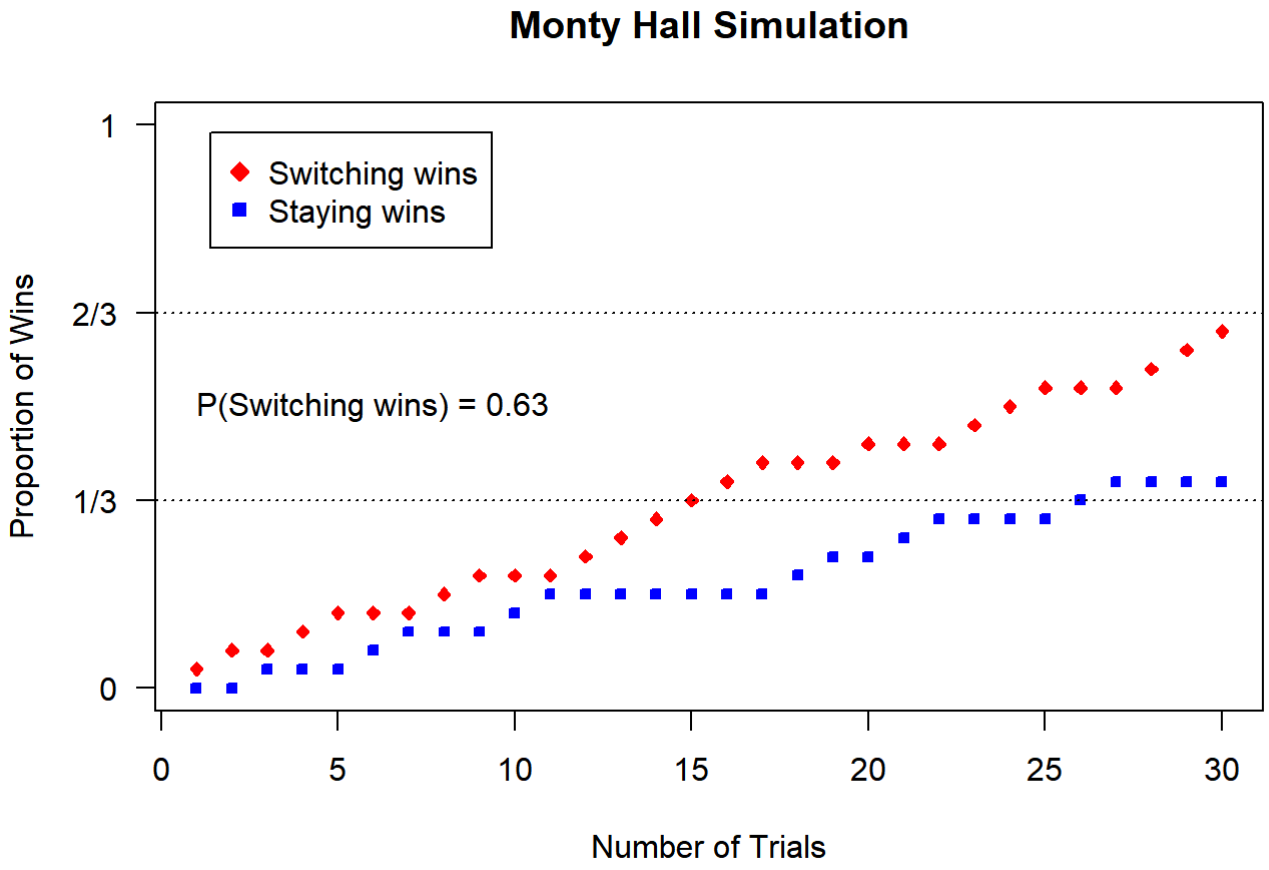


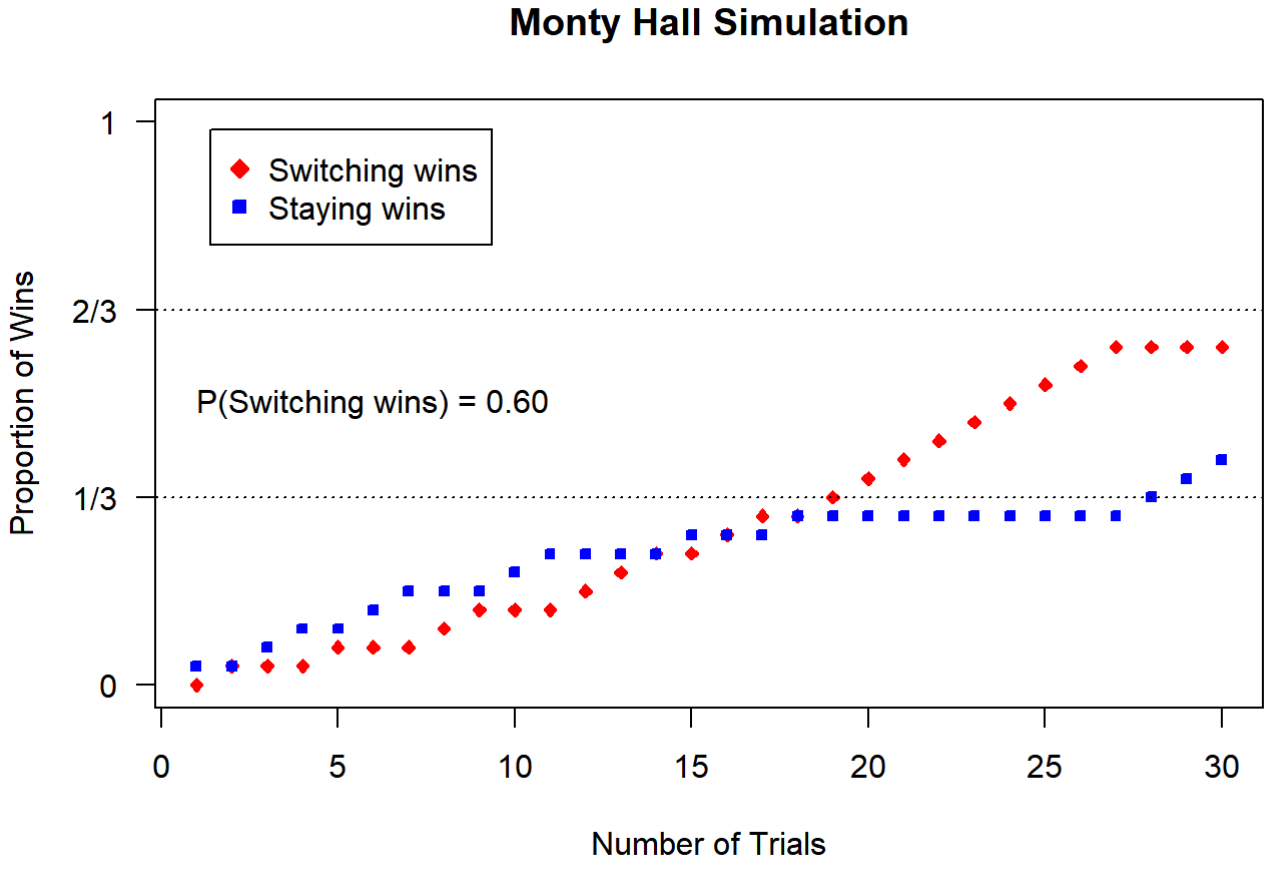

Monty Hall Simulation
Switching wins
Staying wins
P(Switching wins) = 0.73
Proportion of Wins
1
2/3
1/3
0
0
5
10
15
20
25
30
Number of Trials


Monty Hall Simulation
Switching wins
Staying wins
P(Switching wins) = 0.80
Proportion of Wins
1
2/3
1/3
0
0
5
10
15
20
25
30
Number of Trials


Monty Hall Simulation
Switching wins
Staying wins
P(Switching wins) = 0.77
Proportion of Wins
1
2/3
1/3
0
0
5
10
15
20
25
30
Number of Trials


Monty Hall Simulation
Switching wins
Staying wins
P(Switching wins) = 0.83
Proportion of Wins
1
2/3
1/3
0
0
5
10
15
20
25
30
Number of Trials





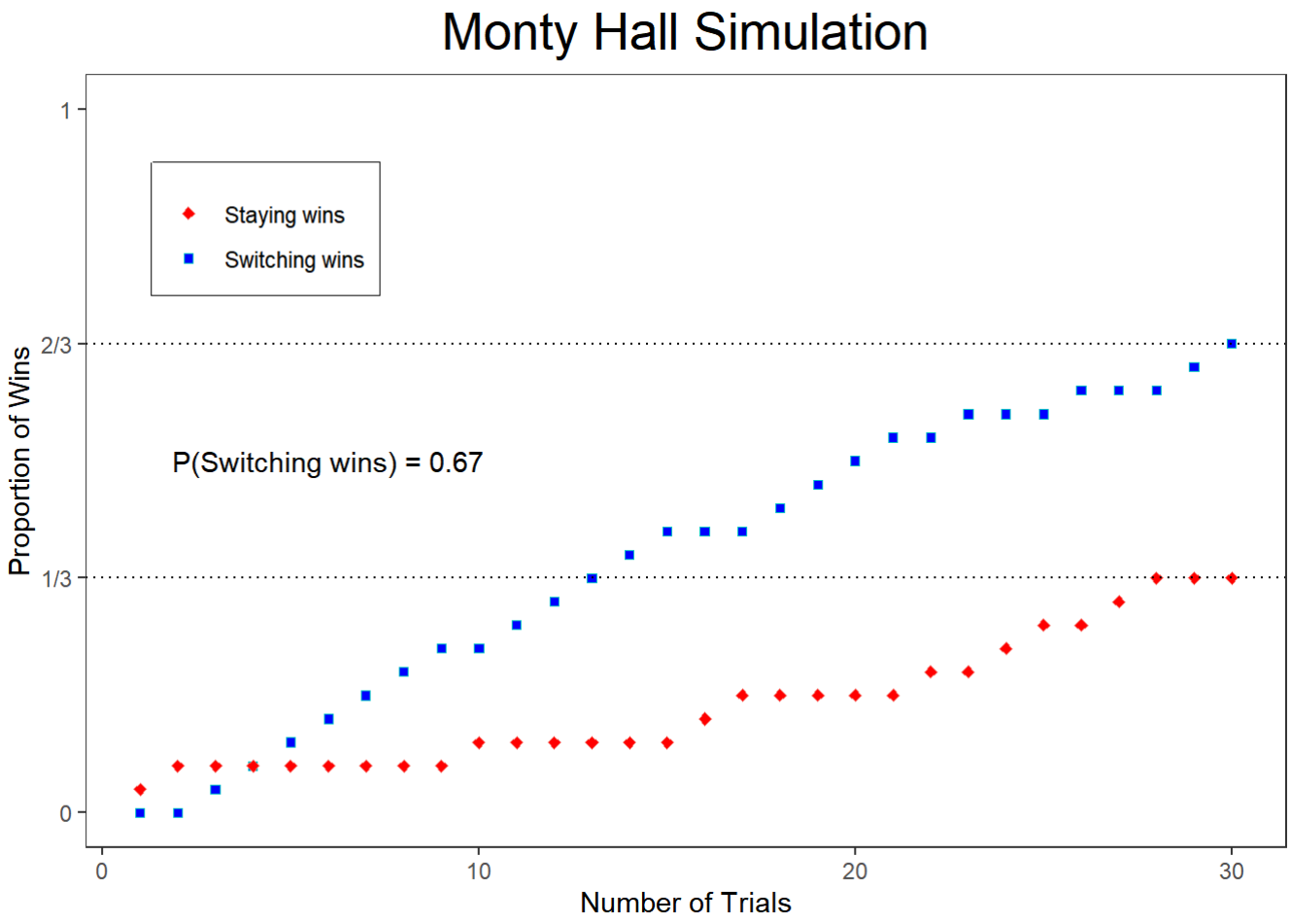
Stem and Leaf

```
stem(Prop_Switching_wins_10)
```

```
##
##   The decimal point is 1 digit(s) to the left of the |
##
##   5 | 3
##   6 | 03377
##   7 | 37
##   8 | 03
```

ggplot

```
library(ggplot2)
monty_ggplot <- function(N) {
  monty_result <-
    replicate(N, monty_hall()) %>%
    t %>%
    data.frame
  y_switch <- cumsum(monty_result[, 5] == "Switching wins")
#  y_stay <- cumsum(monty_result[, 5] == "Staying wins")
  y_stay <- 1:N - y_switch
  y_df <- data.frame(x = rep(1:N, times = 2),
                     Result = c(y_switch, y_stay),
                     Decision = rep(c("Switching wins","Staying wins"), each = N))
  p_wins <- sum(monty_result[, 5] == "Switching wins") / N
monty <-
  ggplot(data = y_df,
         mapping = aes(x = x,
                       y = Result / N,
                       colour = Decision,
                       shape = Decision,
                       fill = Decision)) +
    geom_point() +
    scale_shape_manual(values = c(23, 22)) +
    scale_fill_manual(values = c("red", "blue")) +
    scale_y_continuous(name = "Proportion of Wins",
                       limits = c(0, 1),
                       breaks = c(0, 1/3, 2/3, 1),
                       labels = c("0", "1/3", "2/3", "1")) +
    geom_hline(yintercept = c(1/3, 2/3),
               linetype = "dotted") +
    theme_bw() +
    labs(title = "Monty Hall Simulation",
         x = "Number of Trials") +
    annotate("text",
             x = N / 5,
             y = 1 / 2,
             label = paste0("P(Switching wins) = ", format(p_wins, digits = 2, nsmall = 2))) +
    theme(legend.position = c(0.15, 0.8),
          legend.title = element_blank(),
          legend.box.background = element_rect(fill = "transparent"),
          panel.grid = element_blank(),
          plot.title = element_text(hjust = 0.5, size = 20))
list(monty = monty, p_wins = p_wins)
}
monty_ggplot(30)$monty
```



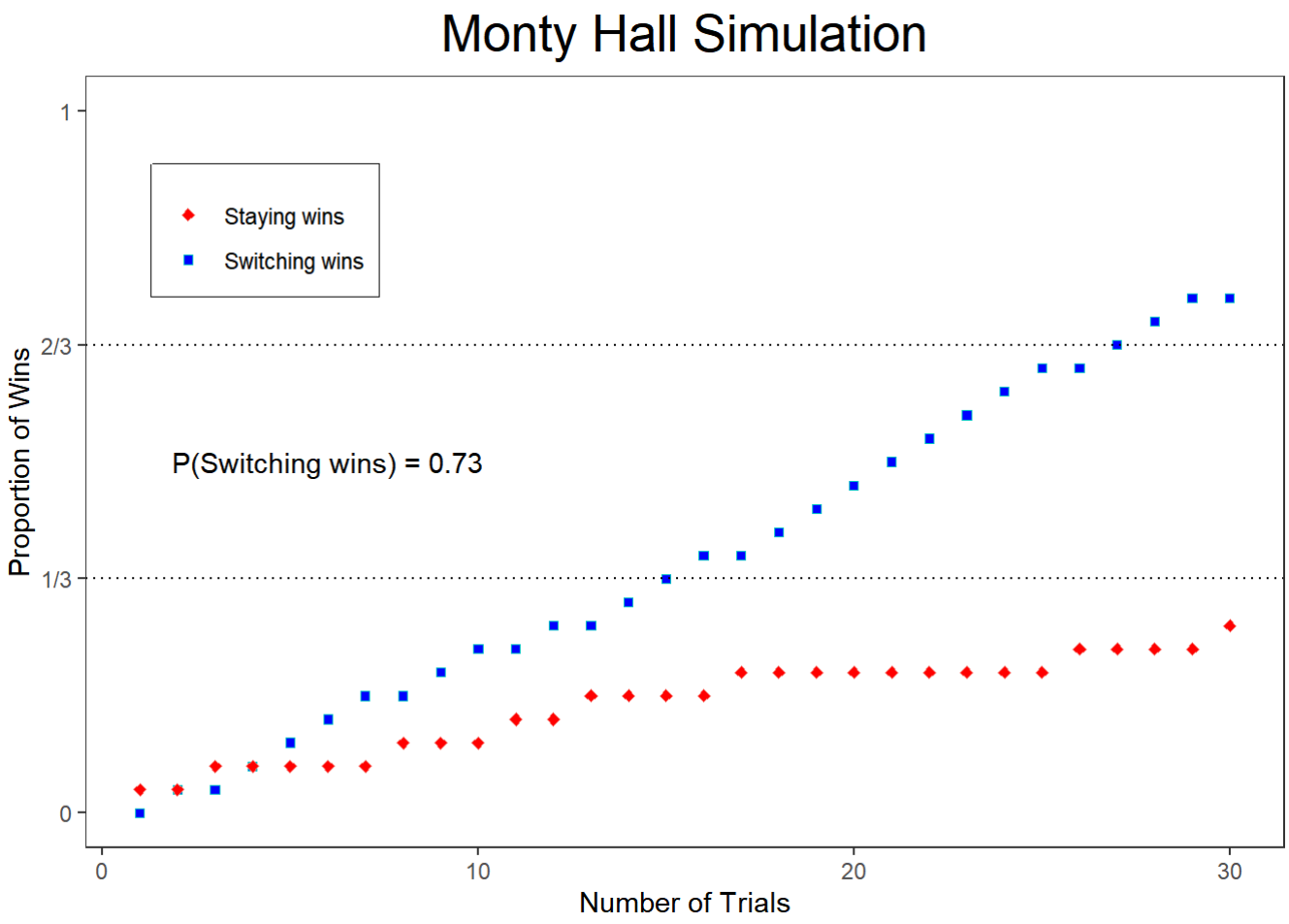

```
monty_ggplot(30)$p_wins
```

```
## [1] 0.6
```

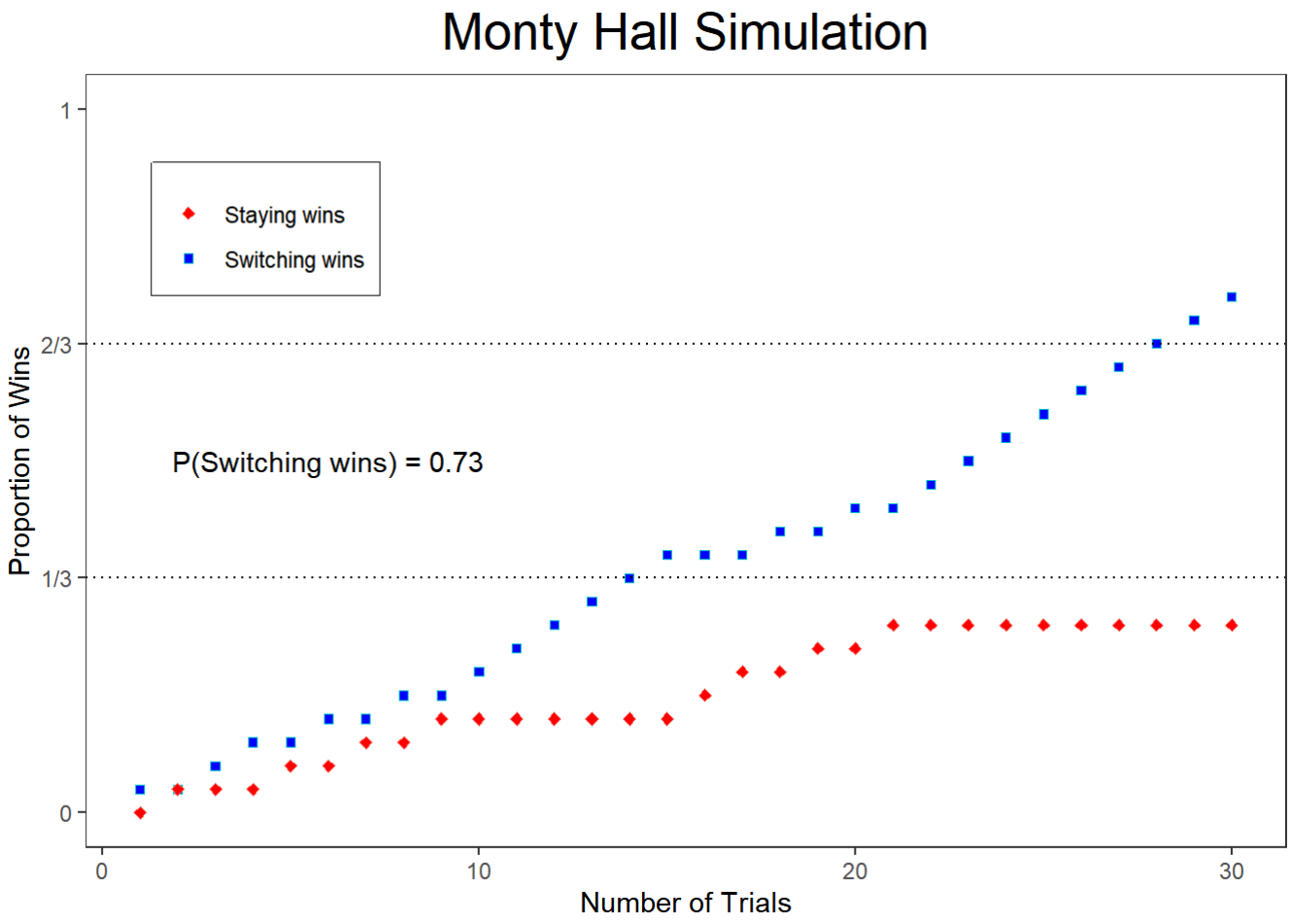
Repetitions

```
x <- 10
m <- list()
while (x > 0) {
m <- list(m, monty_ggplot(30)$monty)
  x <- x - 1
}
m
```

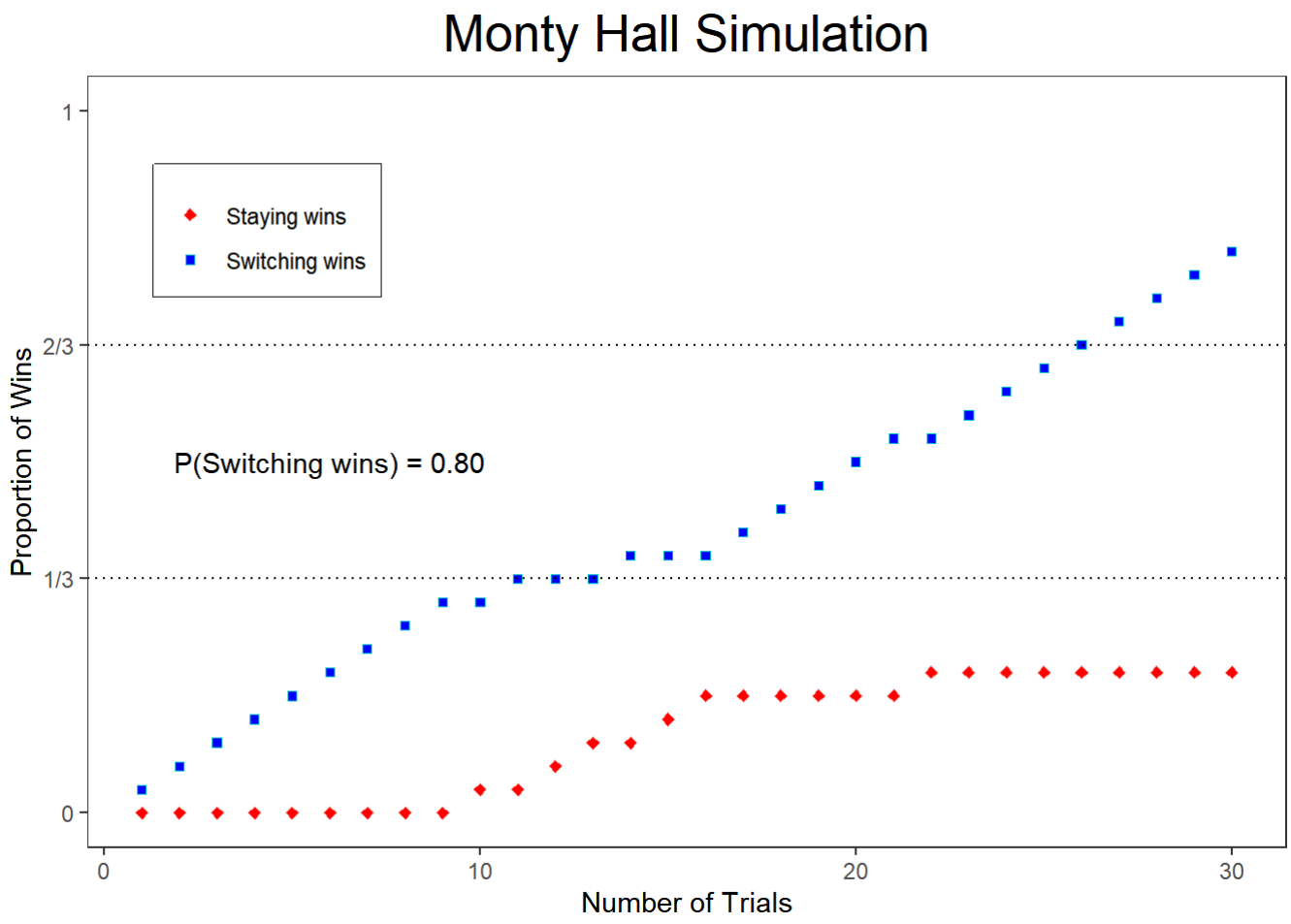
```
## [[1]]
## [[1]][[1]]
## [[1]][[1]][[1]]
## [[1]][[1]][[1]][[1]]
## [[1]][[1]][[1]][[1]][[1]]
## [[1]][[1]][[1]][[1]][[1]][[1]]
## [[1]][[1]][[1]][[1]][[1]][[1]][[1]]
## [[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]]
## [[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]]
## [[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]]
## list()
##
## [[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[2]]
```



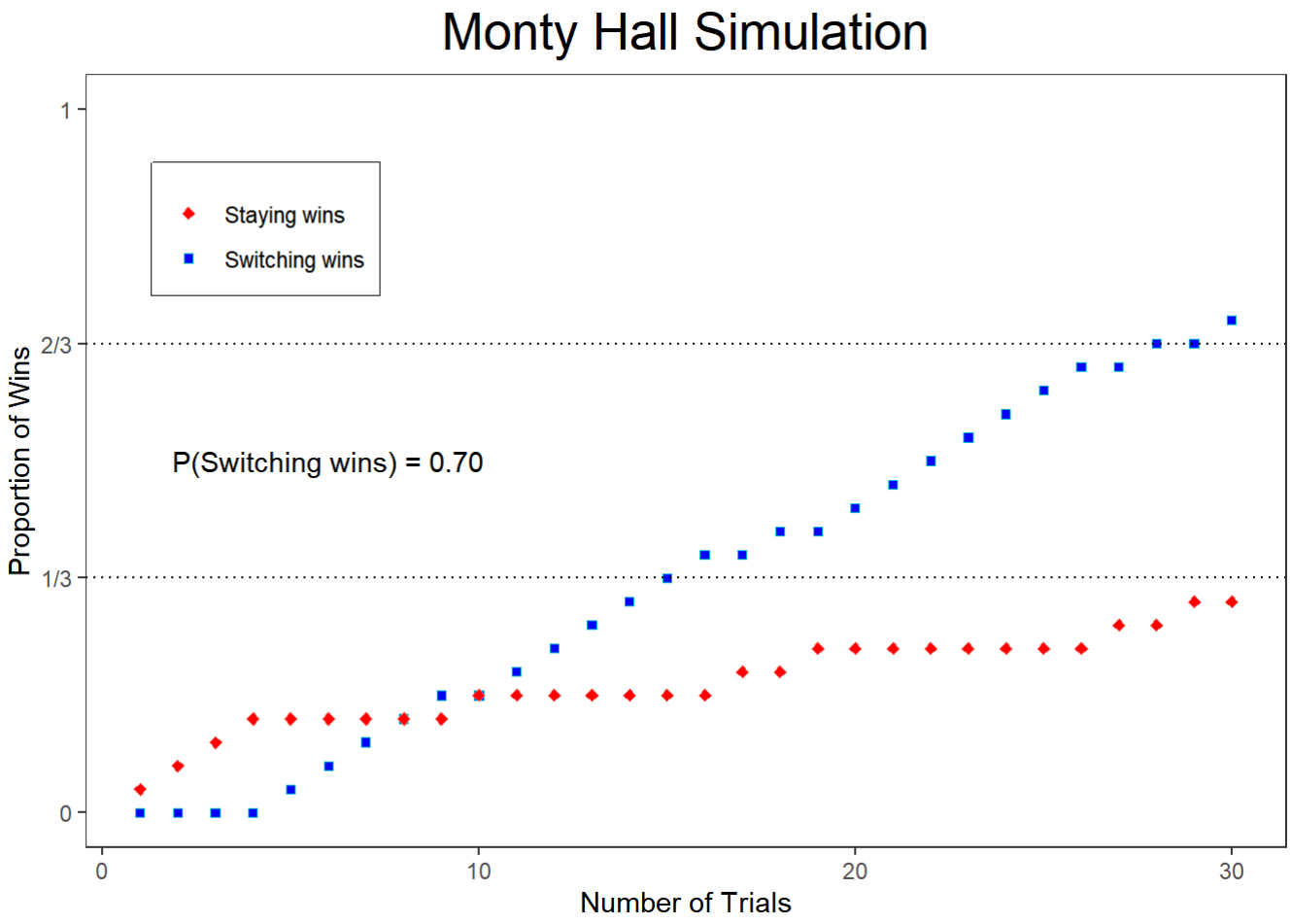

```
##
##
## [[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.73
Proportion of Wins
Number of Trials
1
2/3
1/3
0
0
10
20
30

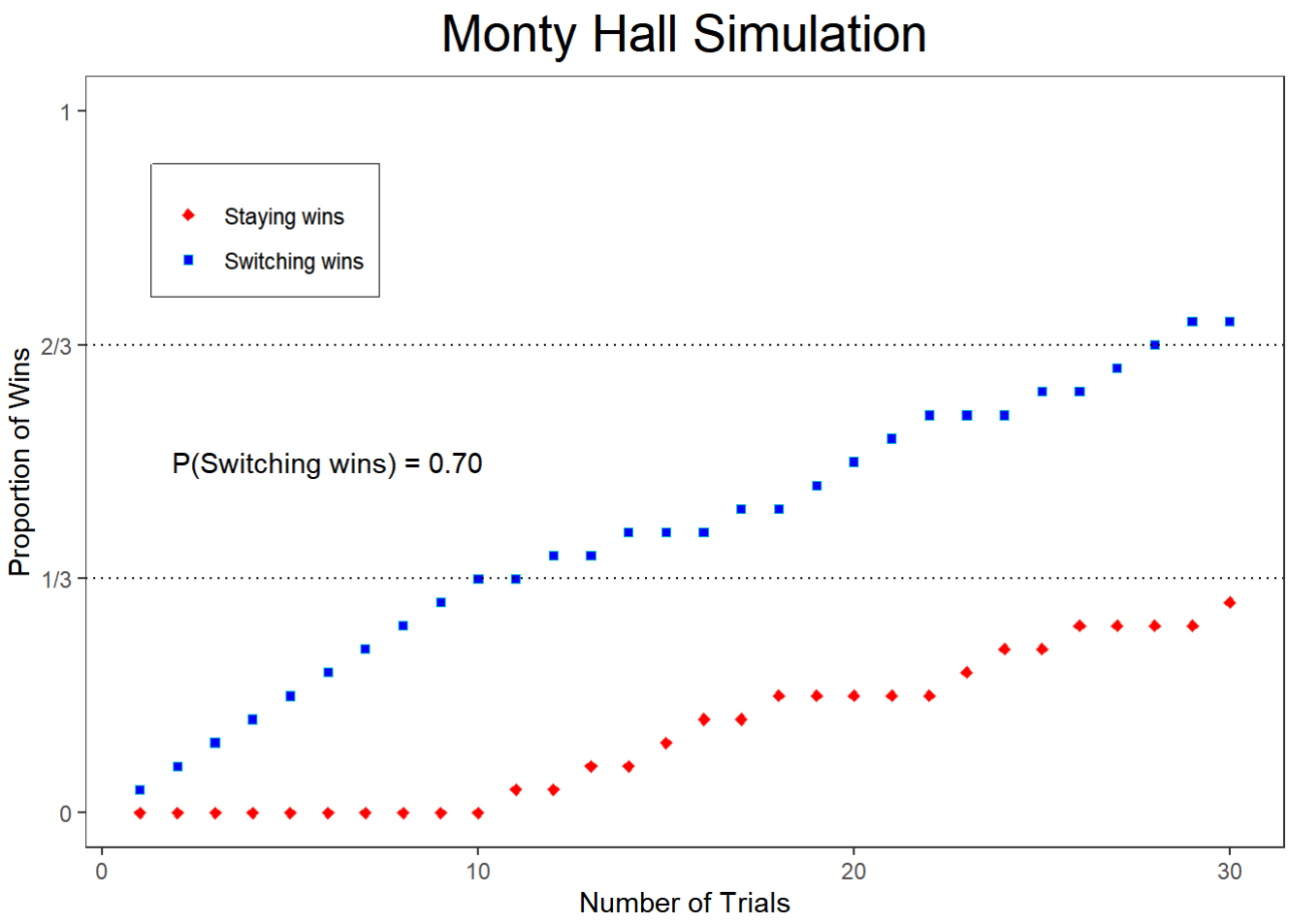
```
##
##
## [[1]][[1]][[1]][[1]][[1]][[1]][[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.80
Proportion of Wins
Number of Trials
1
2/3
1/3
0
0
10
20
30

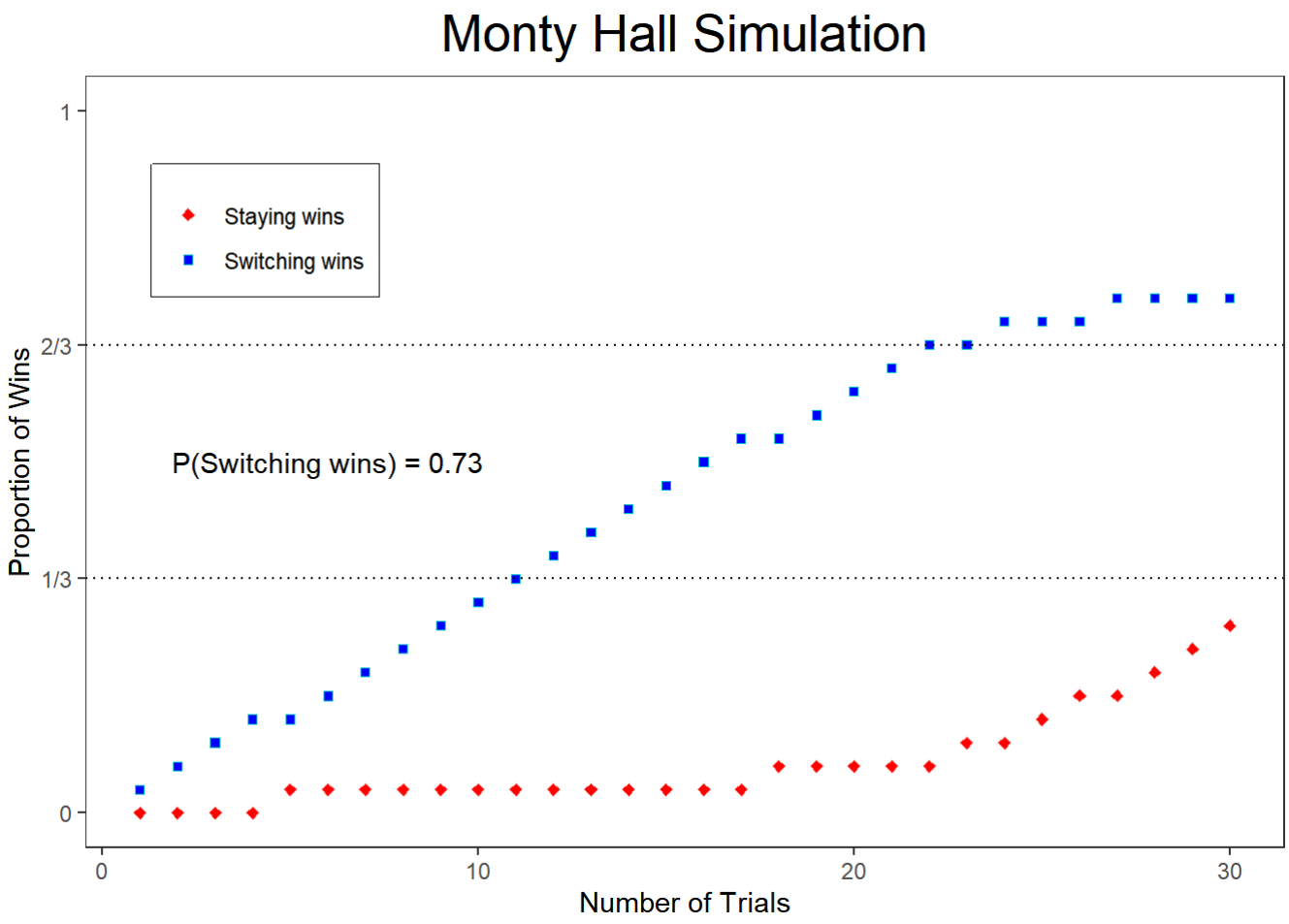
```
##
##
## [[1]][[1]][[1]][[1]][[1]][[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.70
Proportion of Wins
Number of Trials
1
2/3
1/3
0
0
10
20
30

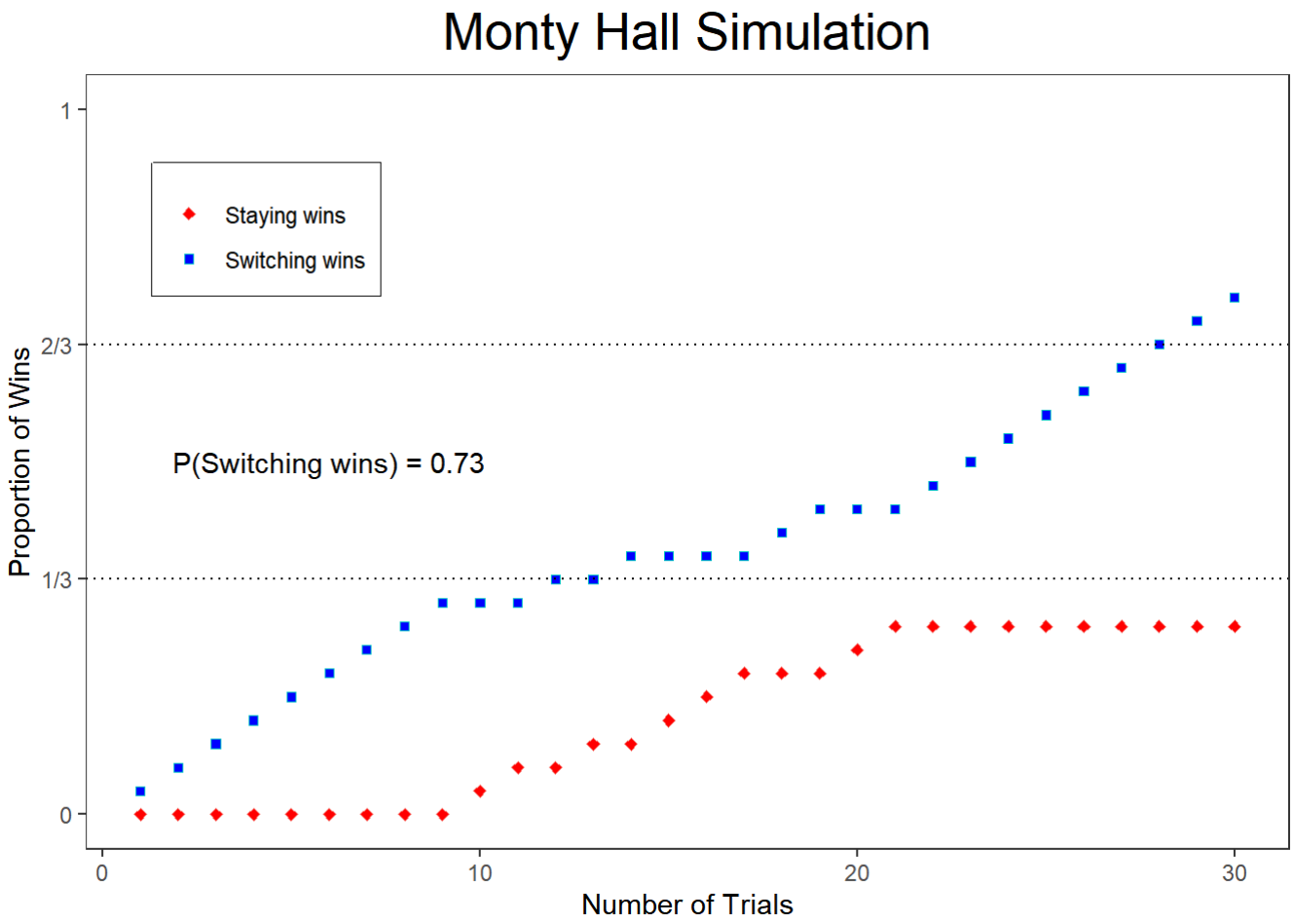
```
##
##
## [[1]][[1]][[1]][[1]][[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.70
Proportion of Wins
Number of Trials
1
2/3
1/3
0
0
10
20
30

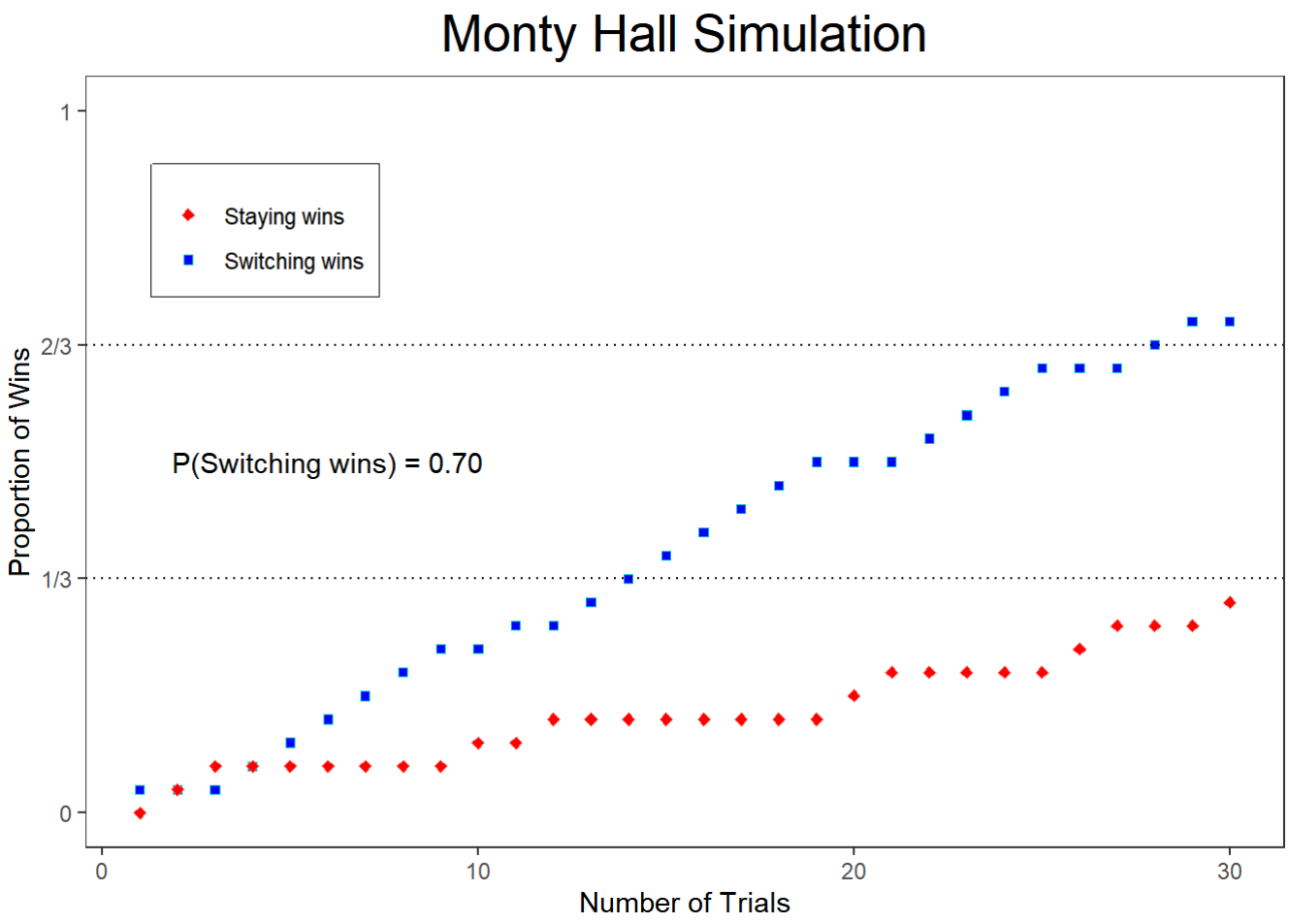
```
##
##
## [[1]][[1]][[1]][[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.73
Proportion of Wins
Number of Trials
1
2/3
1/3
0
0
10
20
30

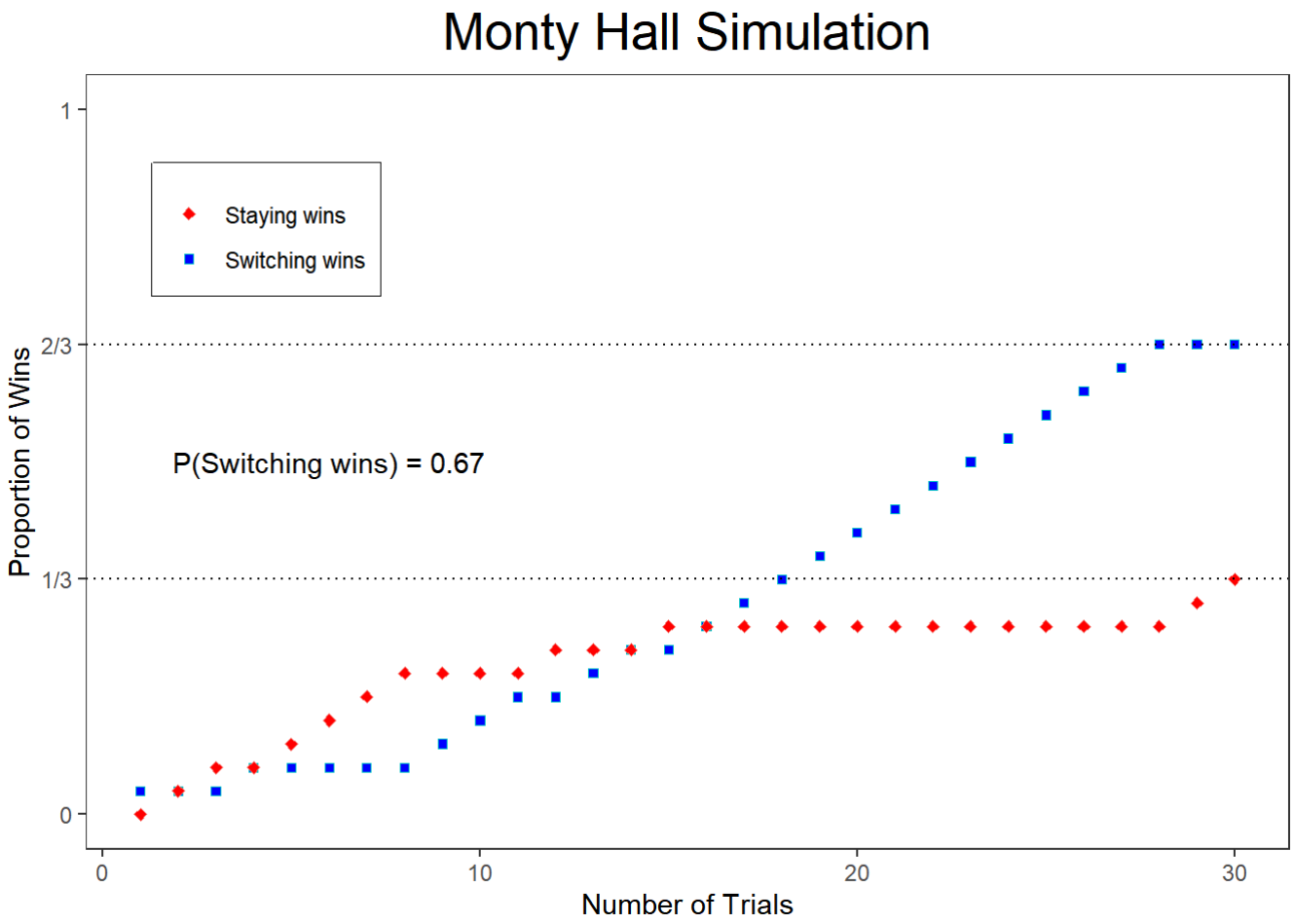
```
##
##
## [[1]][[1]][[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.73
Proportion of Wins
1
2/3
1/3
0
0
10
20
30
Number of Trials

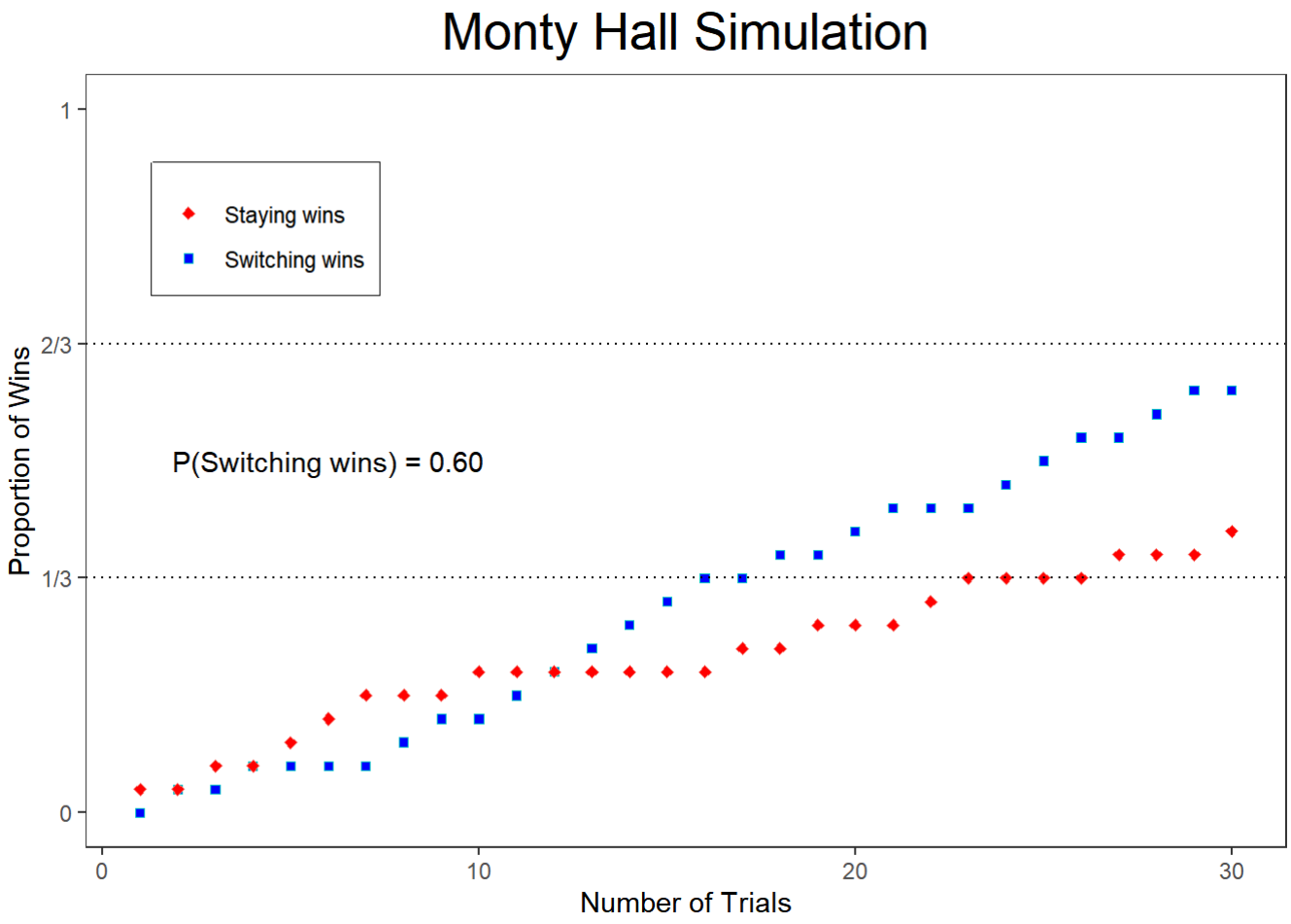
```
##
##
## [[1]][[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.70
Proportion of Wins
Number of Trials
1
2/3
1/3
0
0
10
20
30

```
##
##
## [[1]][[2]]
```


Monty Hall Simulation
Staying wins
Switching wins
P(Switching wins) = 0.67
Proportion of Wins
Number of Trials
1
2/3
1/3
0
0
10
20
30

```
##
##
## [[2]]
```




```
Prop_Switching_wins_100 <- replicate(100, monty_ggplot(30)$p_wins)
```

Stem and Leaf

```
stem(Prop_Switching_wins_100)
```

```
##
##   The decimal point is 1 digit(s) to the left of the |
##
##   4 | 77
##   5 | 033333
##   5 | 77777777
##   6 | 00000000000333333333333333333333
##   6 | 7777777777
##   7 | 00000000000000033333333333333
##   7 | 7777777
##   8 | 00033
```

Comments

이번시간에는 Monty_Hall쇼를 통해서 시뮬레이션을 해보는 작업을 할수있었습니다. 기존의 생각에서는 바꿨을때 확률이 1/2인줄 알았지만 그렇지 않다는 것을 알게되었습니다. 이해가 잘 안되는 부분이 있었지만 모의 실험을 한 결과 를 보았을때 바꾸었을때 2/3에 가깝다는 것을 데이터로 확인할수있었습니다. ggplot과 base plot 을 이용하여 2가지로 만드는 방법을 알 수 있었습니다. 개요부분을 통해 Monty_Hall을 만드는 식을 알수있었습니다. Monty_Hall replicate를 이용하여 반복하여 30번의 실험 결과를 확인할수 있었습니다. pch를 통해 모양을 바

꿔 더 쉽게 알아볼수있었습니다. r로 실험을 한 결과 대부분의 바꾼다를 선택했을때 2/3 에 수렴한다는 것을 확인하였습니다. 이해가 안되는 것들을 데이터시각화를 통해 알아내 쉽게 이해할수 있다는 것을 깨닫는 수업이였습니다.