Lab #7: Introduction to Simulations

Name

Date of lab session

Lab report

Exercises:

Exercise 1: 180/1000 or approximately 9 in every 50 times, we would get exactly 4 in 10 would be heads. (18%)

```
coin <- c("heads","tails")
sample(coin, size = 1, replace = TRUE)
```

```
## [1] "tails"
```

```
sample(coin, size = 10, replace = TRUE)
```

```
##  [1] "heads" "tails" "tails" "tails" "tails" "tails" "tails" "tails" "tails"
## [10] "heads"
```

```
rbinom(n=1, size=1, prob=0.5)
```

```
## [1] 0
```

```
rbinom(n=1, size=10, prob=0.5)
```

```
## [1] 7
```

```
rbinom(n=10, size=1, prob=0.5)
```

```
##  [1] 0 1 1 0 1 0 0 0 1 0
```

```
nheads <- vector() # empty vector to hold number of heads
for (i in 1 : 1000){
  num <- sum(rbinom(10, size=1, prob = 0.5)) # prob of getting a head
  nheads <- c(nheads, num) # add new observation to the vector
}
sum(nheads == 4)
```

```
## [1] 191
```

Exercise 2: According to the formula it is roughly 20.5%. This is 2.5% greater than the result of the simulation, but it is roughly similar.

```
nx <- factorial(10)/(factorial(6)*factorial(4))
bi <- nx*((1/2)^4)*((1/2)^6)
bi
```

```
## [1] 0.2050781
```

```
dbinom(4,size = 10,prob=0.5)
```

```
## [1] 0.2050781
```

Exercise 3: Since the p-value is 0.8779, which is very large, we cannot say that there is a significant difference in the treatment and control groups.

```
control <- rnorm(n = 150, mean = 0, sd = 5)
treatment <- rnorm(n = 15, mean = 2, sd = 5)
t.test(treatment, control)
```

```
##
##  Welch Two Sample t-test
##
## data:  treatment and control
## t = 3.4596, df = 16.966, p-value = 0.003003
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  1.605691 6.627477
## sample estimates:
##  mean of x  mean of y
##  3.8302747 -0.2863093
```

Exercise 4: Power is 29%

```
alpha = 0.05
numrejects <- 0

for(i in 1:1000){
  # generate observations for control and treatment group
  # with mean difference of 2
  control <- rnorm(n = 150, mean = 0, sd = 5)
  treatment <- rnorm(n = 15, mean = 2, sd = 5)

  # use t test to compare two sample means
  ttest <- t.test(treatment, control)

  # reject the equal sample mean hypothesis if p value is small
  reject <- ttest$p.value < alpha

  # count for rejections
  numrejects <- numrejects + reject
}
numrejects/1000
```

```
## [1] 0.304
```

Exercise 5: On average you need to buy around 250

```r
source('https://lgpcappiello.github.io/teaching/stat100b/harrypotter.R')
cards <- c("Dumbledore","McGonagall","Grindewald","Lestrange","Snape","Scamander",
         "Moody","Flitwick","Sprout","Flamel","Others")

length(cards)
```

```
## [1] 11
```

```r
p1 <- .005 # probability of Dumbledore card
p2 <- .01  # probability of McGonagall card
p3 <- .04  # etc...
p4 <- .05
p5 <- .025
p6 <- .02
p7 <- .03
p8 <- .025
p9 <- .02
p10<- .02
#Create a vector with all probabilities and probability of others
probs.pre<-c(p1,p2,p3,p4,p5,p6,p7,p8,p9,p10)
probs<-c(probs.pre,1-sum(probs.pre))
MC_geom(10000, cards, probs)
```

```
## $mean
## [1] 251.2399
##
## $sd
## [1] 183.291
```

On Your Own:

OYO 1: According to the simulation, about 3 in 1000 times, you get 1 3 times in a row.(0.3%) The binomial distribution formula says 0.46% so the simulation was slightly less likely then the the formula.

```r
dice <- c("one","two","three","four","five","six")
sample(dice, size = 1, replace = TRUE)
```

```
## [1] "six"
```

```r
sample(dice, size = 10, replace = TRUE)
```

```
##  [1] "five"  "five"  "one"   "two"   "five"  "four"  "one"   "three" "one"
## [10] "five"
```

```r
rbinom(n=1, size=1, prob=0.5)
```

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

```r
rbinom(n=1, size=10, prob=0.5)
```

```
## [1] 4
```

```r
rbinom(n=10, size=1, prob=0.5)
```

```
##  [1] 1 1 1 0 0 1 0 1 1 0
```

```r
n1 <- vector() # empty vector to hold number of heads
for (i in 1 : 1000){
  num <- sum(rbinom(3, size=1, prob = 0.1667)) # prob of getting a head
  n1 <- c(n1, num) # add new observation to the vector
}
sum(n1 == 3)
```

```
## [1] 3
```

```r
dbinom(3 ,size = 3,prob = 0.1667)
```

```
## [1] 0.004632408
```

OYO 2: Power is 74.8%

```r
alpha = 0.05
numrejects <- 0

for(i in 1:1000){
  # generate observations for control and treatment group
  # with mean difference of 2
  control <- rnorm(n = 100, mean = 0, sd = 2.5)
  treatment <- rnorm(n = 75, mean = 1, sd = 2.5)

  # use t test to compare two sample means
  ttest <- t.test(treatment, control)

  # reject the equal sample mean hypothesis if p value is small
  reject <- ttest$p.value < alpha

  # count for rejections
  numrejects <- numrejects + reject
}
numrejects/1000
```

```
## [1] 0.74
```

OYO 3: Would need to buy roughly need to buy 123 chocolate frogs.

```
cards <- c("Luna","McGonagall","Neville","Harry","others")

length(cards)
```

```
## [1] 5
```

```
p1 <- .024 # probability of Luna card
p2 <- .01  # probability of McGonagall card
p3 <- .026  # etc...
p4 <- .02

#Create a vector with all probabilities and probability of others
probs.pre<-c(p1,p2,p3,p4)
probs<-c(probs.pre,1-sum(probs.pre))
MC_geom(100, cards, probs)
```

```
## $mean
## [1] 127.31
##
## $sd
## [1] 111.5043
```

```


sample_mean <- 122.36
se <- 86.95947 / sqrt(100)
lower <- sample_mean - 1.96 * se
upper <- sample_mean + 1.96 * se
c(lower, upper)
```

OYO 4:

```
## [1] 105.3159 139.4041
```