Project 1 (EAS502 : Intro Prob Theory)

Rishi Shanthan Bhagavatham

2024-03-19

Submitted to : Dr. Vochita Mihai

I. Tossing a coin T

The probability of getting a Heads or a Tails on a coin toss is both 0.5. We can use R to simulate an experiment of flipping a coin several times and compare our results with the theoretical probability. First let fix the conventon:

- 0 = Tails and 1 = Heads
- We can use the following command to tell R to flip a coin 15 times:

```
e0 = sample(0:1,15,rep=T)
e0
```

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

This gives 7 Tails and 8 heads.

In fact, we can write a function to flip a coin n times:

```
FlipCoin = function(n) {
   sample(0:1,n,rep=T)
}
e1=FlipCoin(30)
e1
```

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

Now we can use the sum command to compare the results from this experiment to the theoretical probabilities. For example, in the above experiment of flipping a coin 30 times, we can count the heads and tails as:

```
sum(e1==0)
```

```
## [1] 17
```

```r
sum(e1==0)/30
```

```
## [1] 0.5666667
```

```r
sum(e1==1)
```

```
## [1] 13
```

```r
sum(e1==1)/30
```

```
## [1] 0.4333333
```

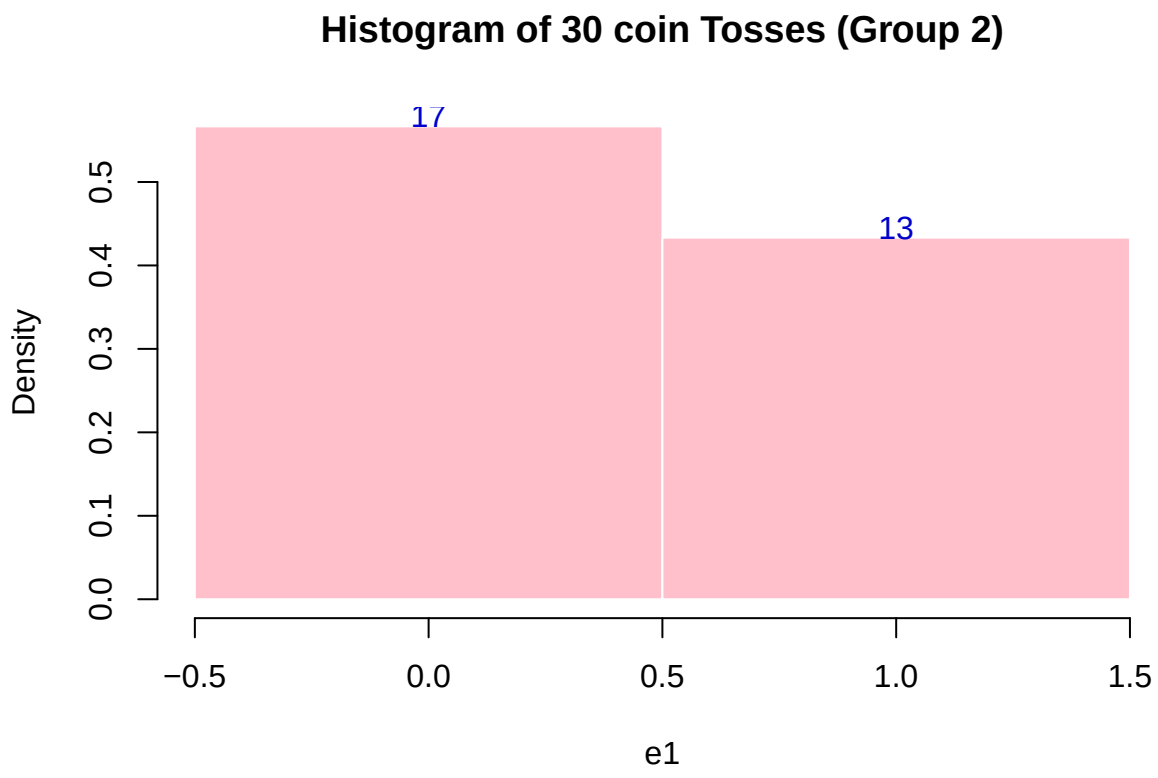
This gives us 17 Tails and 13 Heads. The "probability" or relative frequency of a Tail in this experiment is 0.5666667 and a Head is 0.4333333.

Plotting Histogram

```r
r<-hist(e1,breaks=c(-0.5,0.5,1.5),
        prob=T,main="Histogram of 30 coin Tosses (Group 2)",
        col="pink",
        border="white")

label_pos <- r$density - 0.015

text(r$mids,label_pos, r$counts, adj = c(.5, -.5), col = "blue3")
```



Questions

1. Use R to simulate an experiment of tossing a coin 100 times. Print the relative histogram as above with your name on it.

```
e2 = FlipCoin(100)
e2
```

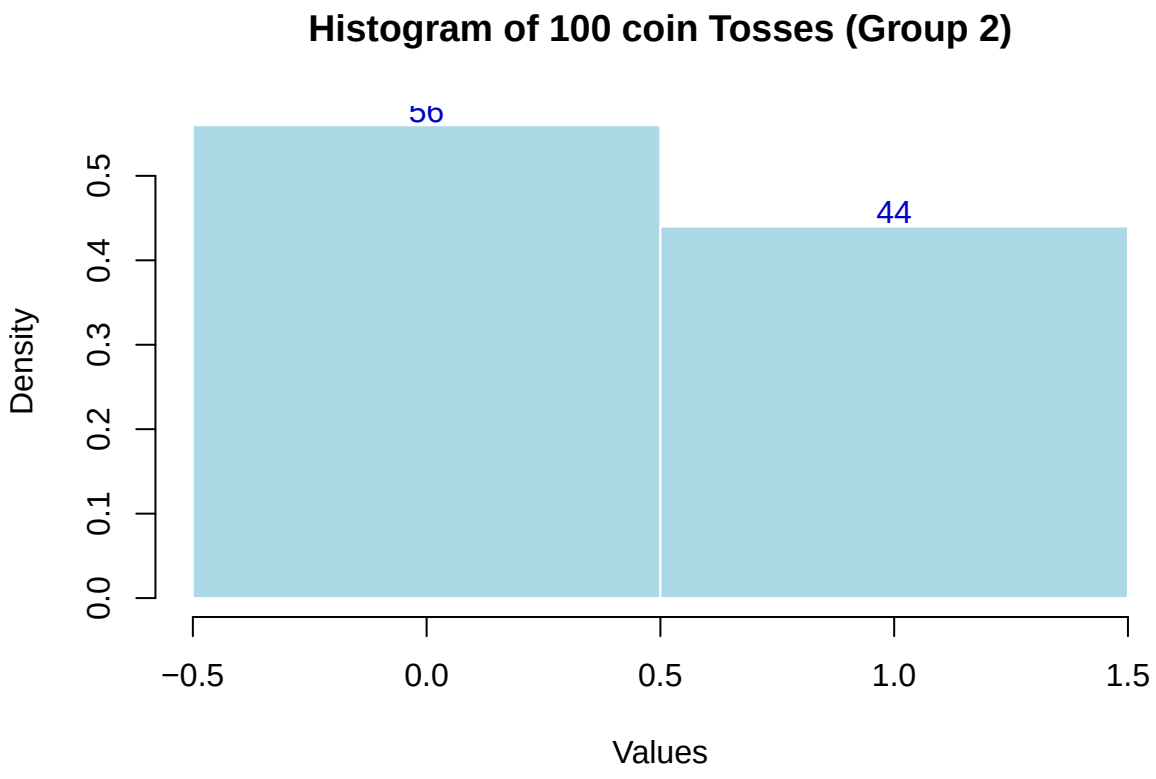
```
##   [1] 1 0 0 1 0 0 0 0 1 1 0 1 1 0 1 1 1 1 1 1 0 0 1 0 0 0 0 0 0 1 0 1 0 0 0 0 0
##  [38] 0 1 1 1 0 1 1 1 1 0 1 1 0 1 1 0 1 1 0 0 0 1 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0
##  [75] 0 0 1 0 1 1 0 0 0 1 1 0 1 0 0 1 1 0 1 1 1 0 0 0 0 0
```

Relative Histogram

```
r <- hist(e2, breaks = c(-0.5, 0.5, 1.5),
          prob = TRUE,
          main = "Histogram of 100 coin Tosses (Group 2)",
          col = "lightblue",
          xlab = "Values",
          border = "white",)

label_pos <- r$density - 0.01

text(r$mids,label_pos, r$counts, adj = c(.5, -.5), col = "blue3")
```



QUESTION 2 IS HANDLED AFTER QUESTION 3

3. Repeat 2 for tossing a coin 500 times (do not print histogram).

```
e3 = FlipCoin(500) #Flipping Coin 500 times
e3
```

```
##   [1] 1 0 0 0 0 0 0 0 0 1 1 1 1 0 1 1 1 1 1 0 1 0 0 1 1 0 0 1 0 0 1 0 0 0 1 0 0
##  [38] 0 1 0 1 1 1 1 1 1 1 1 0 1 1 1 0 0 0 0 0 1 1 1 0 0 1 1 0 1 0 0 1 0 1 1 1 0
##  [75] 1 1 0 0 0 0 1 0 1 1 0 0 0 1 0 0 0 0 1 0 0 0 0 1 0 0 1 1 0 1 1 1 0 1 1 0 1
## [112] 0 0 1 1 1 1 1 1 0 1 1 1 0 1 0 0 1 0 1 0 1 1 0 1 1 0 1 1 0 1 0 1 1 1 1 0 1
## [149] 1 1 1 1 0 1 0 1 0 0 1 0 0 1 1 0 1 1 1 0 1 0 0 1 0 0 1 0 0 1 1 0 1 1 0 0 0
## [186] 1 1 0 1 0 0 0 1 1 1 0 0 0 1 0 0 0 1 0 1 0 0 0 0 0 1 1 1 1 0 1 0 0 1 0 0 0
## [223] 0 0 1 1 0 1 1 1 1 1 0 0 0 1 1 0 1 1 1 0 0 1 0 0 0 1 0 0 1 1 0 1 1 0 0 0 1
## [260] 1 0 1 1 0 1 1 0 1 1 0 1 1 1 0 0 0 0 0 0 1 1 0 1 0 1 0 0 1 0 1 1 0 0 1 0 1
## [297] 1 0 0 0 1 0 0 1 0 1 0 1 1 1 0 0 1 0 0 0 1 1 0 0 1 0 1 1 0 0 1 1 0 0 0 0 0
## [334] 1 0 1 0 1 1 1 0 1 0 1 0 1 1 1 0 1 1 1 0 0 0 1 0 1 1 1 1 0 0 0 0 0 1 0 1 1
## [371] 0 1 1 0 0 0 0 0 0 1 1 1 0 1 0 0 0 0 1 1 1 0 1 1 0 0 0 1 1 1 1 1 0 1 1 1 1
## [408] 1 1 1 0 1 1 1 0 1 0 0 0 0 1 1 0 0 1 1 1 0 0 1 1 1 1 1 1 1 1 0 0 1 0 1 0 0
## [445] 0 0 0 0 1 0 1 0 0 0 1 1 0 1 1 1 1 0 0 1 1 0 1 0 1 0 0 1 0 0 0 1 0 0 0 1 0
## [482] 1 1 1 0 1 0 1 1 0 0 1 1 1 0 0 0 0 0 1
```

2. Find the relative frequency of a Tail and Head in your experiment and fill in the table on the next page.

```
# Create table of frequencies
freq_table <- matrix(NA, nrow = 2, ncol = 2,
                     dimnames = list(c("Relative frequency of Tails",
                                       "Relative frequency of Heads"),
                                     c("100 tosses", "500 tosses")))

# Frequencies
freq_table["Relative frequency of Tails", "100 tosses"] <- sum(e2 == 0)/100
freq_table["Relative frequency of Heads", "100 tosses"] <- sum(e2 == 1)/100
freq_table["Relative frequency of Tails", "500 tosses"] <- sum(e3 == 0)/500
freq_table["Relative frequency of Heads", "500 tosses"] <- sum(e3 == 1)/500

print(freq_table, box = TRUE,
      quote = FALSE, digits = 2,
      col.just = "center",
      row.just = "center",
      hline.after = TRUE,
      vline.after = TRUE,
      print.gap = 2,
      box = TRUE)
```

```
##                               100 tosses  500 tosses
## Relative frequency of Tails       0.56        0.49
## Relative frequency of Heads       0.44        0.51
```

Probability

100 tosses

500 tosses

Relative frequency of Tails

0.56

0.494

Relative frequency of Heads

0.44

0.506

II. Rolling dice

The probability of getting a number between 1 to 6 on a roll of a die is $1/6 = 0.1666667$. We can use R to simulate an experiment of rolling a die several times and compare our results with the theoretical probability. We can use the following command to tell R to roll a die 20 times:

```
sample(1:6,20,rep=T)
```

```
##  [1] 1 1 1 2 2 3 6 1 1 5 3 5 4 1 5 2 4 4 1 3
```

We can write a function to roll a die n times

```
RollDie = function(n) {
  sample(1:6,n,rep=T)
}
d1=RollDie(50)
d1
```

```
##  [1] 2 4 2 2 4 2 4 4 1 3 1 4 3 6 1 2 3 4 6 2 1 6 4 5 4 3 2 3 3 1 3 3 5 1 6 4 2 1
## [39] 1 2 6 6 5 4 5 2 6 6 6 1
```

Now we can use the sum command to compare the results from this experiment to the theoretical probabilities. For example, in the above experiment the number of 3's and its relative frequency is:

```
sum(d1==3)
```

```
## [1] 8
```

```
sum(d1==3)/50
```

```
## [1] 0.16
```

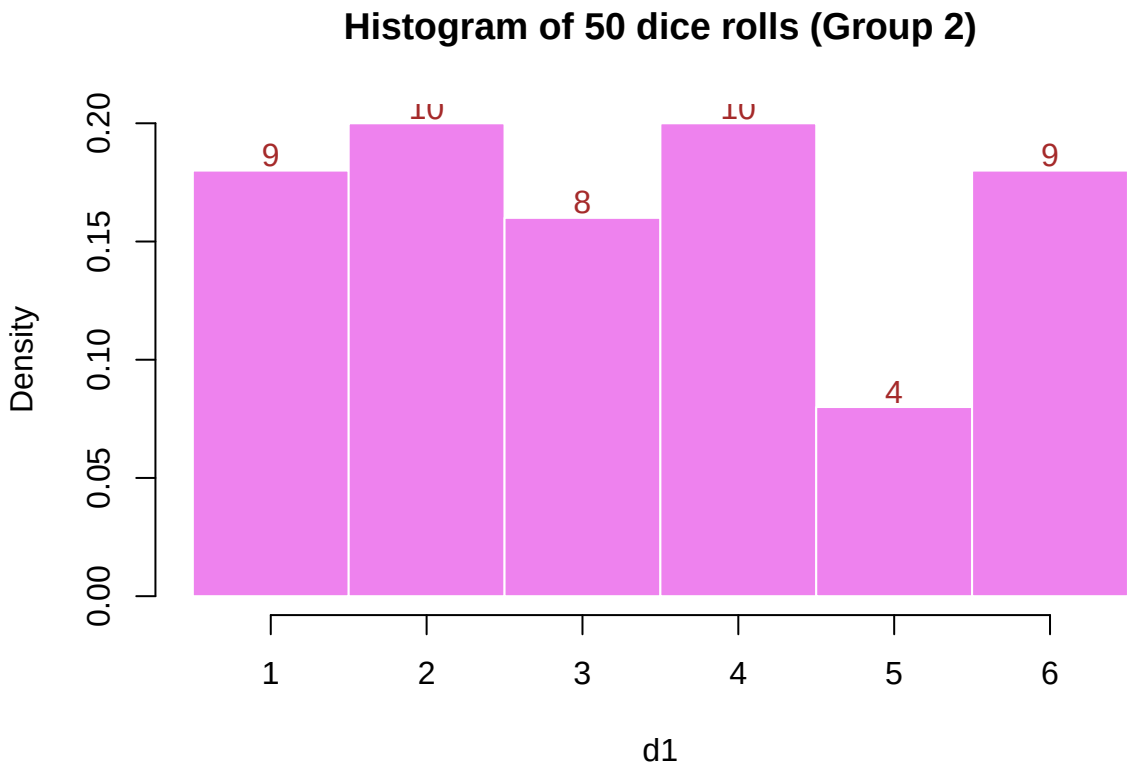
This number 3 occurs 8 times and it's relative frequency is 0.16 which is quite close to 1/6 .

Relative Histogram

```r
r<-hist(d1,breaks=c(0.5,1.5,2.5,3.5,4.5,5.5,6.5),
        prob=T,col="violet",
        main="Histogram of 50 dice rolls (Group 2)",
        border="white")

label_pos <- r$density - 0.003

text(r$mids,label_pos, r$counts, adj = c(.5, -.5), col = "brown")
```



Questions

1. Use R to simulate an experiment of rolling a die 200 times. Print the relative histogram and write your name on it.

```r
d2 = RollDie(200)
d2
```

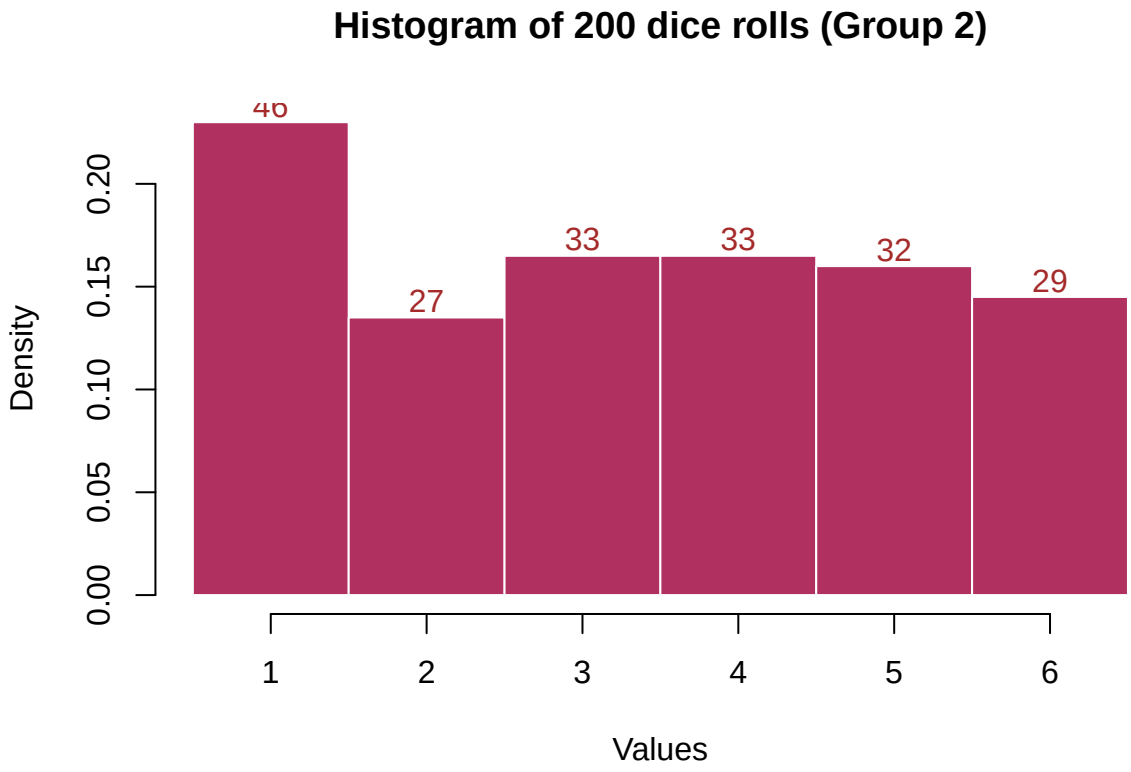
```
##   [1] 3 5 4 6 1 1 6 4 5 6 4 1 4 6 1 1 1 2 3 1 6 1 2 2 3 3 3 5 4 4 3 5 6 1 4 6 3
##  [38] 5 5 5 6 5 5 4 5 5 5 6 5 1 2 5 5 1 4 6 3 6 4 1 2 4 1 2 3 1 4 4 5 2 2 5 1 4
##  [75] 1 1 2 5 3 2 2 4 6 2 5 6 6 5 1 5 6 4 1 2 3 5 6 2 1 2 1 6 3 3 4 4 2 5 1 3 4
## [112] 1 1 1 4 1 3 6 6 5 1 3 4 3 1 4 1 5 5 1 4 5 5 4 2 6 4 2 4 3 3 2 4 6 1 1 2 6
## [149] 3 5 6 3 1 1 3 4 1 1 3 2 3 2 3 3 3 6 3 3 2 1 4 4 1 3 2 6 1 3 1 2 6 1 6 3 5
## [186] 2 1 4 3 4 5 1 1 1 1 4 2 6 6 5
```

Relative Histogram

```
r <- hist(d2, breaks = c(0.5,1.5,2.5,3.5,4.5,5.5,6.5),
          prob = TRUE,
          main = "Histogram of 200 dice rolls (Group 2)",
          col = "maroon",
          xlab = "Values",
          border = "white")

label_pos <- r$density - 0.003

text(r$mids,label_pos, r$counts, adj = c(.5, -.5), col = "brown")
```



QUESTION 2 IS HANDLED AFTER QUESTION 3

3. Repeat 2 for rolling a die 1000 @mes (do not print histogram).

```
d3 = RollDie(1000)
d3
```

```
##    [1] 6 2 5 6 4 6 3 6 5 4 1 5 6 3 1 4 1 1 6 3 2 3 4 1 4 6 3 1 5 6 2 4 5 4 1 2 6
##   [38] 2 4 3 6 5 2 4 5 4 3 4 6 4 1 3 3 5 4 2 2 4 3 4 6 2 2 6 2 6 1 1 5 1 3 1 3 1
##   [75] 5 5 4 5 5 6 3 4 4 5 1 3 5 6 1 5 5 4 6 6 4 5 3 4 1 2 3 3 6 5 1 2 6 1 3 3 5
##  [112] 5 2 6 3 1 2 3 2 3 2 4 5 6 4 1 1 1 5 6 6 6 6 2 5 4 4 1 4 5 6 1 4 6 6 2 6 1
##  [149] 2 1 3 6 4 2 6 1 5 3 1 1 5 1 3 4 5 2 1 2 6 2 6 6 2 3 3 2 1 5 6 4 3 5 3 5 3
##  [186] 2 6 1 3 2 6 2 1 6 3 6 1 5 2 2 6 5 1 6 1 1 3 5 2 3 1 3 2 3 2 3 5 4 2 4 4 4
```

```
##  [223] 1 4 3 5 5 2 1 6 2 3 6 3 6 4 2 5 2 6 4 2 1 1 3 4 6 1 2 1 6 2 1 3 2 1 6 1 1
##  [260] 1 3 1 5 4 2 1 4 6 1 6 3 2 4 5 4 4 6 4 6 2 4 4 6 3 4 6 4 6 3 5 2 2 5 5 6 4
##  [297] 1 5 4 1 4 2 6 5 5 4 2 6 5 3 6 1 1 3 6 1 2 5 3 1 2 6 5 2 1 2 4 5 2 6 2 6 1
##  [334] 5 5 4 5 4 6 4 2 3 5 2 3 6 4 6 1 4 5 1 1 3 5 5 1 1 4 3 4 1 4 3 4 2 6 4 2 2
##  [371] 6 1 6 2 6 3 3 6 5 1 2 2 2 5 3 4 4 5 2 5 5 5 6 2 3 3 5 2 3 2 1 1 1 4 5 2
##  [408] 1 5 1 1 2 1 1 4 3 6 6 2 6 6 6 4 4 6 1 4 3 4 6 5 1 3 1 3 2 6 5 5 5 5 3 5 2
##  [445] 6 6 3 2 2 3 4 3 4 4 2 1 4 4 3 6 4 6 4 1 6 1 3 1 5 5 1 5 2 1 1 4 2 3 2 6 2
##  [482] 5 4 4 3 6 3 4 5 6 1 4 3 3 3 3 2 5 6 3 5 2 4 2 5 3 1 5 2 3 6 4 5 6 6 4 1 5
##  [519] 5 6 2 6 5 2 3 4 3 4 4 6 4 5 3 3 2 2 5 1 1 6 6 6 6 5 6 6 5 3 1 6 1 5 3 5 4
##  [556] 2 4 6 3 3 6 3 1 1 1 4 2 4 5 6 3 5 1 4 2 6 6 3 4 1 4 2 6 3 2 2 3 2 5 5 2 3
##  [593] 6 5 5 3 4 1 4 6 5 3 4 1 2 3 2 4 4 4 4 4 4 2 1 1 5 5 1 4 6 2 5 4 6 6 1 1 1 5
##  [630] 4 6 2 3 2 2 4 6 2 2 5 6 1 2 3 4 4 6 5 1 1 1 5 5 5 6 1 1 4 4 1 6 4 1 4 5 3
##  [667] 1 2 4 1 3 3 3 6 5 6 5 3 1 3 3 5 6 6 1 4 6 4 4 4 1 2 4 2 5 4 3 2 3 5 4 4 3
##  [704] 5 5 1 5 4 3 6 6 3 6 2 6 4 2 5 4 3 6 1 3 4 2 3 2 5 5 4 6 3 5 5 2 3 4 1 3 6
##  [741] 6 1 6 4 4 1 4 2 5 3 5 5 2 4 5 4 3 6 6 6 2 1 3 2 1 6 1 5 2 1 3 2 1 6 6 4 2
##  [778] 4 5 1 6 1 2 6 2 3 1 5 3 4 3 2 4 6 3 2 1 6 5 1 3 3 4 4 5 3 3 5 1 2 4 2 3 3
##  [815] 3 6 3 1 1 6 3 6 5 6 5 2 6 2 2 1 3 6 6 5 3 6 1 6 3 5 6 5 3 4 6 5 5 4 2 3 4
##  [852] 2 6 1 6 3 5 2 3 4 1 5 2 3 3 4 1 6 6 2 2 2 1 4 6 5 3 3 5 4 6 3 1 3 1 2 4 3
##  [889] 3 3 6 2 4 6 3 6 2 5 4 6 4 2 5 3 5 5 4 1 6 1 5 1 6 1 1 4 1 5 2 6 6 5 1 6 2
##  [926] 6 3 5 1 5 2 5 2 6 3 6 1 2 3 3 1 6 6 2 5 3 5 3 5 3 6 4 2 3 5 3 5 3 5 3 1 3
##  [963] 5 1 4 5 4 1 1 2 3 3 2 2 3 6 5 1 5 6 3 6 2 4 1 3 2 6 1 3 5 2 2 5 2 6 6 4 6
## [1000] 3
```

2. Find the relative frequency of the numbers 1 to 6 in your experiment and fill in the table.

```r
# Create table of frequencies
dice_table <- matrix(NA, nrow = 6, ncol = 2,
                     dimnames = list(c("Relative frequency of 1",
                                       "Relative frequency of 2",
                                       "Relative frequency of 3",
                                       "Relative frequency of 4",
                                       "Relative frequency of 5",
                                       "Relative frequency of 6"),
                                     c("200 rolls", "1000 rolls")))

# Frequencies
dice_table["Relative frequency of 1", "200 rolls"] <- sum(d2 == 1)/200
dice_table["Relative frequency of 2", "200 rolls"] <- sum(d2 == 2)/200
dice_table["Relative frequency of 3", "200 rolls"] <- sum(d2 == 3)/200
dice_table["Relative frequency of 4", "200 rolls"] <- sum(d2 == 4)/200
dice_table["Relative frequency of 5", "200 rolls"] <- sum(d2 == 5)/200
dice_table["Relative frequency of 6", "200 rolls"] <- sum(d2 == 6)/200

dice_table["Relative frequency of 1", "1000 rolls"] <- sum(d3 == 1)/1000
dice_table["Relative frequency of 2", "1000 rolls"] <- sum(d3 == 2)/1000
dice_table["Relative frequency of 3", "1000 rolls"] <- sum(d3 == 3)/1000
dice_table["Relative frequency of 4", "1000 rolls"] <- sum(d3 == 4)/1000
dice_table["Relative frequency of 5", "1000 rolls"] <- sum(d3 == 5)/1000
dice_table["Relative frequency of 6", "1000 rolls"] <- sum(d3 == 6)/1000

print(dice_table, box = TRUE,
      quote = FALSE,
      digits = 2,
      col.just = "center",
```

```
      row.just = "center",
      hline.after = TRUE,
      vline.after = TRUE,
      print.gap = 2)
```

```
##                          200 rolls  1000 rolls
## Relative frequency of 1       0.23        0.16
## Relative frequency of 2       0.14        0.16
## Relative frequency of 3       0.17        0.17
## Relative frequency of 4       0.17        0.16
## Relative frequency of 5       0.16        0.16
## Relative frequency of 6       0.14        0.18
```

Probability

200 rolls

1000 rolls

Relative frequency of 1

0.23

0.164

Relative frequency of 2

0.27

0.316

Relative frequency of 3

0.165

0.171

Relative frequency of 4

0.165

0.16

Relative frequency of 5

0.16

0.164

Relative frequency of 6

0.145

0.183

III. Conditional Probabilities

Questions

1. Suppose that you are dealt 2 cards sequentially (face down) from a standard deck (52 cards). What is the probability that the first and second cards are Kings. Write a code in R to compute this probability

```r
deck_cards <- 52 # Total Cards
kings <- 4 # Total Kings
probability_of_first_king <- kings/deck_cards
probability_of_second_king <- (kings-1)/(deck_cards - 1)

# Final Probability
final_probability <- probability_of_first_king * probability_of_second_king
final_probability
```

```
## [1] 0.004524887
```

Final Probability that the first and second cards are kings -> `0.0045249`

Project 1

Group 2

Team Members

- Bandari Shivani
- Bansal Vaibhav
- Bhagavatham Rishi Shanthan
- Bhosale Samidha

Coin Toss

Probability

100 tosses

500 tosses

Relative frequency of Tails

`0.56`

`0.494`

Relative frequency of Heads

`0.44`

`0.506`

Rolling Dice

Probability

200 rolls

1000 rolls

Relative frequency of 1

`0.23`

`0.164`

Relative frequency of 2

0.27

0.316

Relative frequency of 3

0.165

0.171

Relative frequency of 4

0.165

0.16

Relative frequency of 5

0.16

0.164

Relative frequency of 6

0.145

0.183