1. Dice Experiment

   (a) Rolling a die 1500 Times.

   ```
   > x = c(1,2,3,4,5,6)
   > probx= c(1/8,1/8,1/8,1/8,3/8,1/8)
   > F1500=sample(x, size=1500,  replace=T, prob=probx)
   ```

   i. Describe what each `R` command is performing in the above.

   ii. Using the `mean` and `var` command find the mean and variance of `F1500`. From this information alone what would you conclude is the range of the random variable `F1500`.

   iii. Does the mean and variance from the sample generated compare closely with the true mean and variance of `F1500`.

   (b) (Sums of Rolls) Suppose we wish to simulate in R the experiment of Rolling a die 5 times and noting down its sum. We can use the `sample`, `matrix` and `apply`.

   ```
   > x = c(1,2,3,4,5,6)
   > probx= c(1/6,1/6,1/6,1/6,1/6,1/6)
   > Rolls=sample(x, size=1500,  replace=T, prob=probx)
   > Rollm=matrix(Rolls, nrow = 5)
   > Rollsums = apply(Rollm, 2, sum)
   ```

   i. Describe the commands `matrix` and `apply`

   ```
   > library(ggplot2)
   > density = function(x,a,s){ (1/((2*pi)^(0.5)*s ))* exp(-(x-a)^2/(2*s^2))}
   > dfrolls = data.frame(Rollsums)
   > mu = mean(dfrolls$Rollsums)
   > sigma= sd(dfrolls$Rollsums)
   > ggplot(data=dfrolls) +
   +   geom_histogram(mapping=aes(x=Rollsums,y=..density..),
   +                  color="#00846b", fill=NA, binwidth=1) +
   +   xlim(5,30)+geom_function(fun=density, args=list(a=mu, s= sigma), color="black")
   ```

   i. From the picture what does $\int_{12}^{21} \texttt{density}(x, \mu, \sigma)dx$ approximate ?

   ii. If

   $$\text{Area under the histogram between 12 and 21} \approx \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx,$$

   then what would be your guess for a and b

2. Coin Toss Experiment

   (a) Tossing a coin 10 times.

   ```
   > b1 = rbinom(1000,10,0.5)
   > b2 = rbinom(1000,10,0.25)
   > b3 = rbinom(1000,10,0.75)
   ```

   i. Using the `?rbinom` explain what each of the above commands is performing in `R`

   ii. Using the `mean` and `var` command find the mean and variance of `b1,b2,b3`. Compare them with the true mean and variance of the respective Binomial distribution.

   (b) `geom_hist` command.

```
> library(ggplot2)
> df1=data.frame(b1)
> p11= ggplot(df1) + geom_histogram(mapping=aes(x=b1), color="#00846b", fill="NA", binwidth=1)
> p21= ggplot(df1) +
+   geom_histogram(mapping=aes(x=b1, y=..density..),color="#00846b", fill="NA", binwidth=1)
```

i. Explain what are the plots p11,p21 providing.

ii. Rewrite the code to provide the plots for `b2` and `b3`.

iii. What can you say about the three plots ?

(c) (Density Approximation.) The below code plots the function `density` function in the interval $(0, 10)$ with $a = 5, s = \sqrt{2.5}$ along with the plot `p21`.

```
> library(ggplot2)
> density = function(x,a,s){ (1/((2*pi)^(0.5)*s ))* exp(-(x-a)^2/(2*s^2))}
> df1=data.frame(b1)
> ggplot(df1) +
+          geom_histogram(mapping=aes(x=b1, y=..density..),
+                         color="#00846b", fill="NA", binwidth=1) +
+          xlim(0,10) +
+          geom_function(fun=density, args=list(a=5,s=(2.5)^(0.5)))
```

i. From the picture what does $\int_3^6 \texttt{density}(x, 5, \sqrt{2.5})dx$ approximate ?

ii. If

$$\text{Area under the histogram between 3 and 7} \approx \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx,$$

then what would be your guess for a and b

iii. How would you try the same idea for `b2` and `b3` ? Would you get the same result ?