Mini Project #1
Nikhil Pareek, Tanushri Singh
Contribution of each group member:
Both worked together to finish the two questions. Collaborated to learn R and then write the scripts. Nikhil worked to check the accuracy of the script and Tanu worked to Document and report all the findings. Both partners worked efficiently to complete the project requirements!

Question 1:

a. Analytical Problem

From the given information the cumulative distribution of function can be inferred as ($f_T(t)$).

Now P (T > 15) = 1 – P (T <= 15)

= 1 – F (T <= 15)

= 1 – integration (f_T (t)) over the values 0 to 15

= 1 – integration $_0^{15}$ ($0.2e^{-0.1t} - 0.2e^{-0.2t}$)

= 1 – $_0^{15}$ [0.2 ((($e^{-0.1t}$) / (-0.1)) – (($e^{-0.2t}$) / (-0.2)))

= 1 – $_0^{15}$ [$-2e^{-0.1t} + e^{-0.2t}$]

= 1 – [($-2e^{-0.1x15} + e^{-0.2x15}$) – ($-2e^{-0.1x0} + e^{-0.2x0}$)]

= 1 – [$-2e^{-1.5} + e^{-3} + 2e^{0} - e^{0}$]

= 1 – [$e^{-3} - 2e^{-1.5} + 1$]

= 1- [0.049787 – 0.446260 + 1]

= 1 – 0.603527

= 0.396473

b.

i.

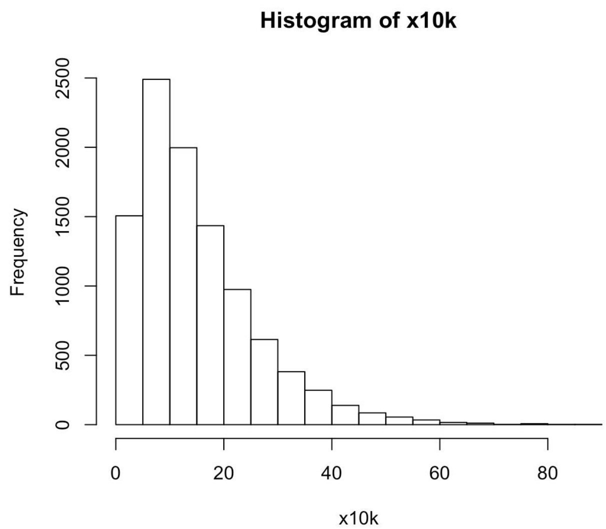
```
> #Question 1
> #Test 1, Sample 10000
> pdf.T <- function(x){ return (0.2*exp(-0.1*x)-0.2*exp(-0.2*x)) }
```

ii.

```
> x10k = replicate(10000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
```

```
x10k        num [1:10000] 22.16 9.32 13.14 8.56 32.87 ...
```

iii.



iv. Analytically derived value for the probability density function was 15 and the value gained from the Monte Carlo simulation is 14.8463 which is comparable!

```
> mean(x10k)
[1] 14.8463
```

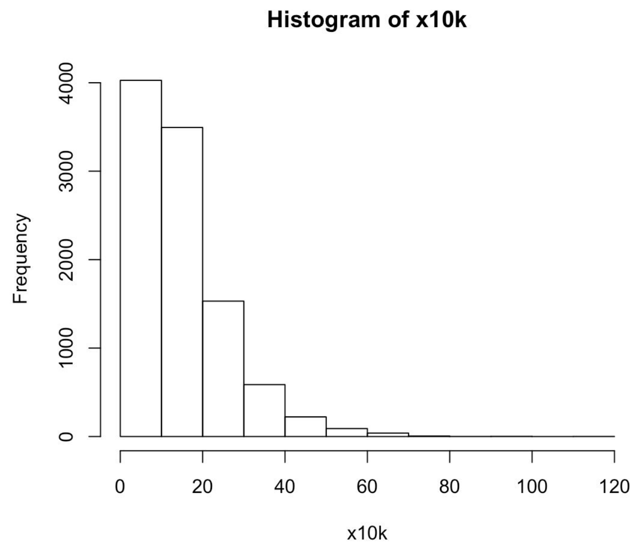
v. The mean is different and the current sample size is 10,000 random variable hence there is a slight difference in probabilities.

```
> 1-pexp(15, rate = 1 / mean(x10k))
[1] 0.3640905
>
```

vi.

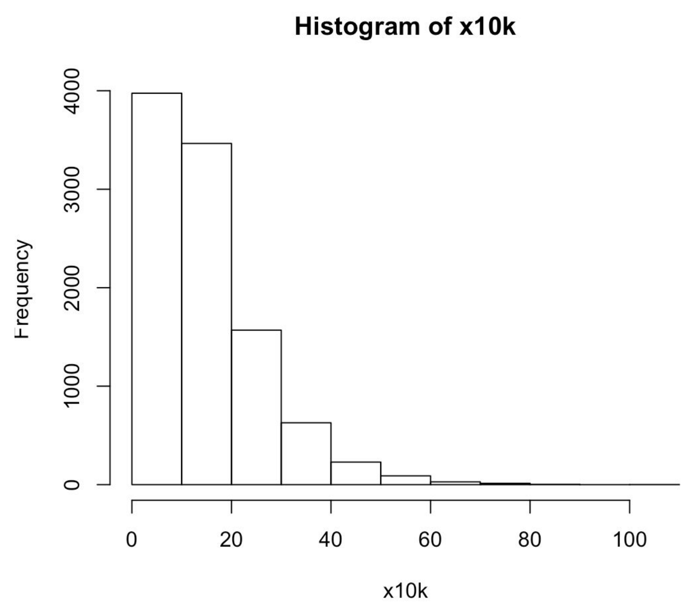
Test 2:

```
> #Test 2
> x10k = replicate(10000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x10k)
> mean(x10k)
[1] 14.88382
> 1-pexp(15, rate = 1 / mean(x10k))
[1] 0.3650191
>
```



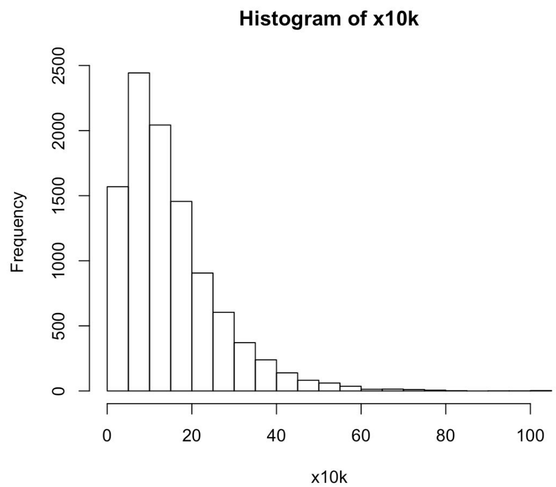
Test 3:

```
> #Test 3
> x10k = replicate(10000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x10k)
> mean(x10k)
[1] 15.01919
> 1-pexp(15, rate = 1 / mean(x10k))
[1] 0.3683497
>
```



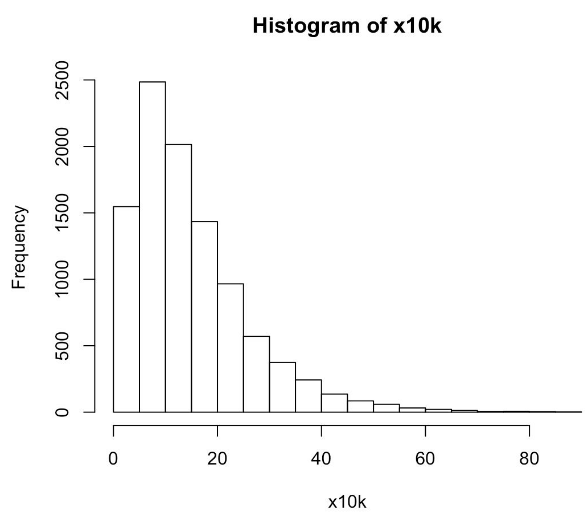
Test 4:

```
> #Test 4
> x10k = replicate(10000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x10k)
> mean(x10k)
[1] 15.02079
> 1-pexp(15, rate = 1 / mean(x10k))
[1] 0.3683891
>
```



Test 5:

```
> #Test 5
> x10k = replicate(10000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x10k)
> mean(x10k)
[1] 14.97844
> 1-pexp(15, rate = 1 / mean(x10k))
[1] 0.3673504
>
```



Comparison Table:

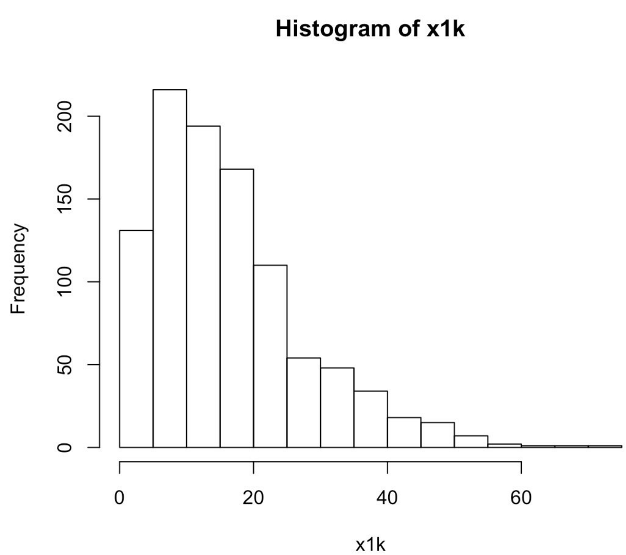
Test for Sample size 10,000	E(T)	P(T>15)
Test 1	14.8463	0.3640905
Test 2	14.88382	0.3650191
Test 3	15.01919	0.3683497
Test 4	15.02079	0.3683891
Test 5	14.97844	0.3673504

It is clear that with a sample size of 10,000 the E(T) value is close to the numerical value 15 with mild variation and the P(T>15) has similar variation. These values coincide with the values that were analytically calculated in 1a. Hence the definition of the Central Limit Theorem continues to be proven correct thus far.

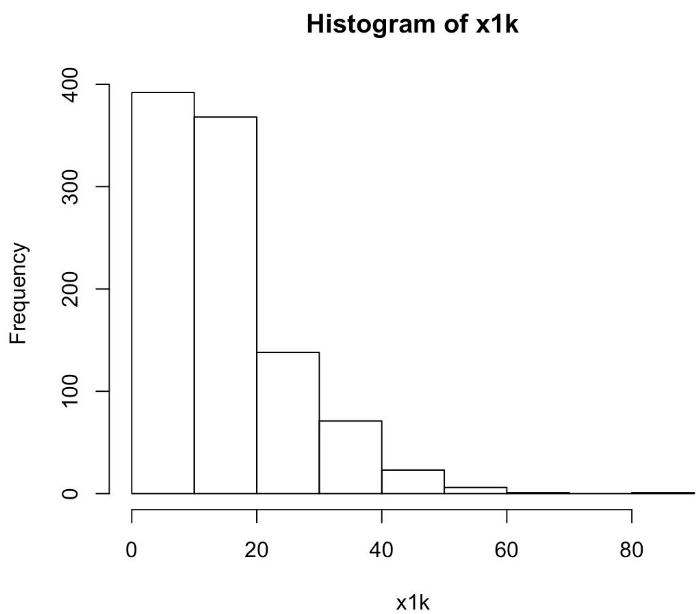
c.

Sample Size : 1,000

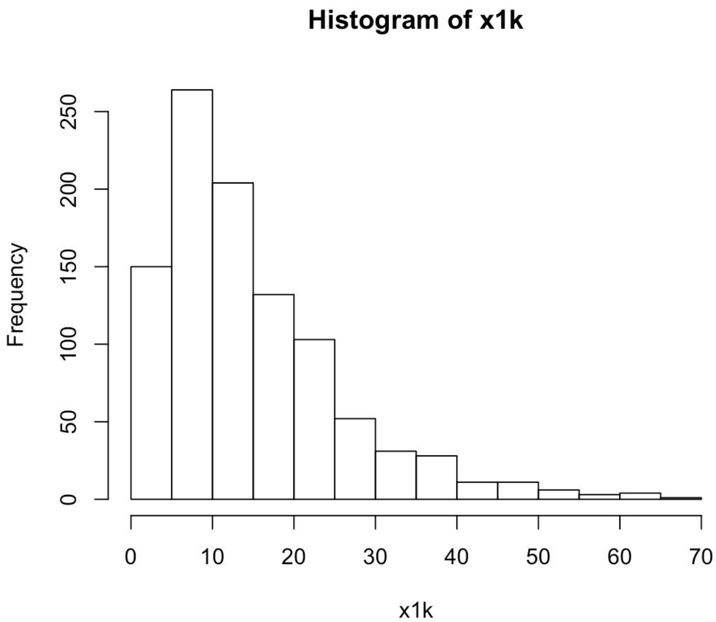
```
> #Test 1, Sample size 1,000
> x1k = replicate(1000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x1k)
> mean(x1k)
[1] 16.18445
> 1-pexp(15, rate = 1 / mean(x1k))
[1] 0.3958121
>
```



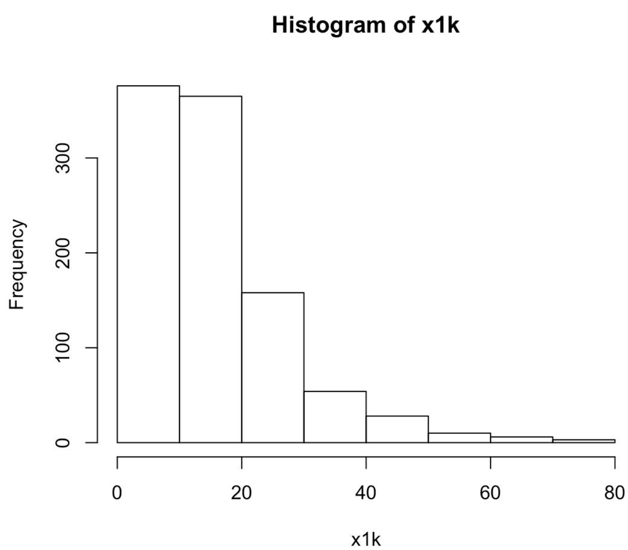
```
> #Test 2, Sample size 1,000
> x1k = replicate(1000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x1k)
> mean(x1k)
[1] 14.81066
> 1-pexp(15, rate = 1 / mean(x1k))
[1] 0.3632063
>
```



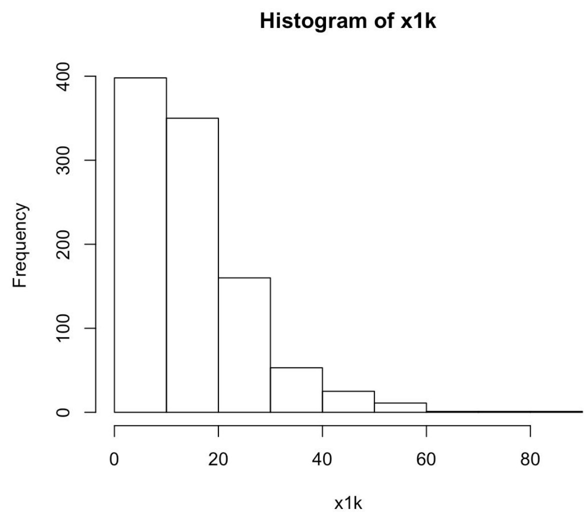

```
> #Test 3, Sample size 1,000
> x1k = replicate(1000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x1k)
> mean(x1k)
[1] 14.87652
> 1-pexp(15, rate = 1 / mean(x1k))
[1] 0.3648385
>
```




```
> #Test 4, Sample size 1,000
> x1k = replicate(1000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x1k)
> mean(x1k)
[1] 15.43804
> 1-pexp(15, rate = 1 / mean(x1k))
[1] 0.3784673
>
```

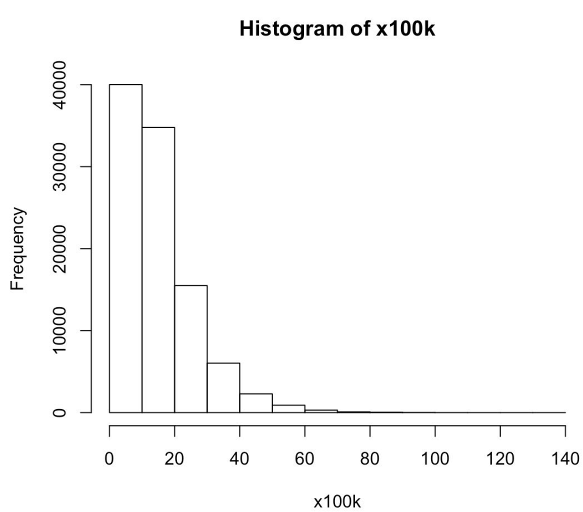


```
> #Test 5, Sample size 1,000
> x1k = replicate(1000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x1k)
> mean(x1k)
[1] 15.07782
> 1-pexp(15, rate = 1 / mean(x1k))
[1] 0.369783
>
```

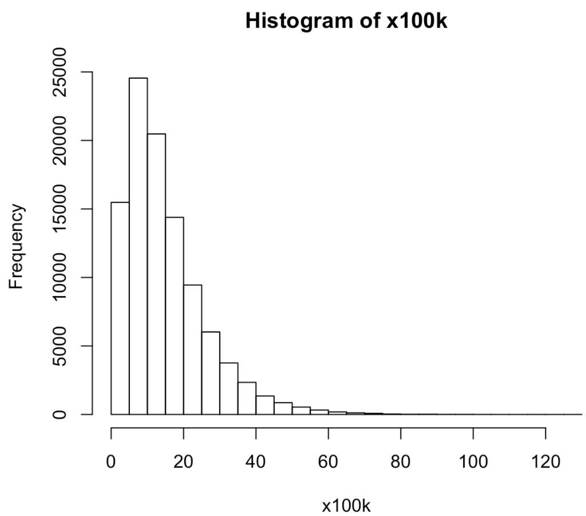


Sample Size : 100,000

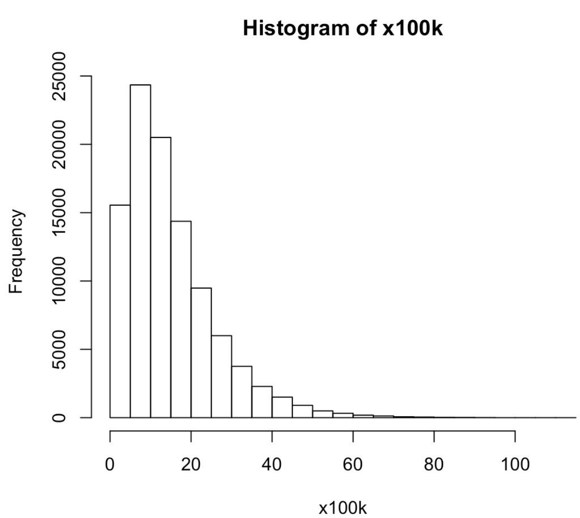
```
> #Test 1, Sample size 100,000
> x100k = replicate(100000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x100k)
> mean(x100k)
[1] 14.99187
> 1-pexp(15, rate = 1 / mean(x100k))
[1] 0.36768
>
```



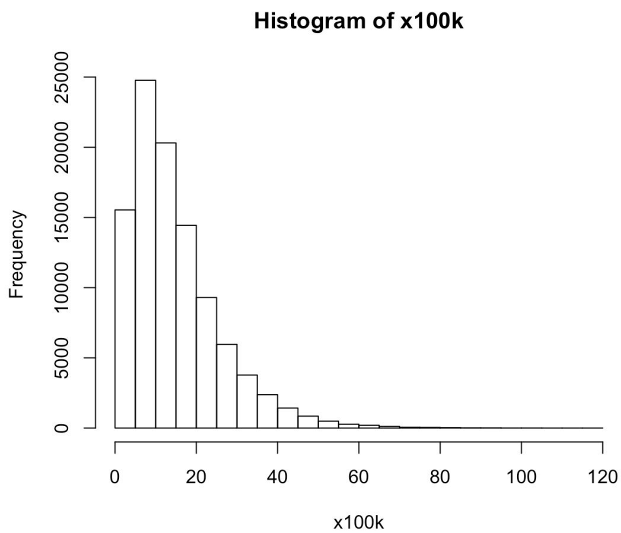
```
> #Test 2, Sample size 100,000
> x100k = replicate(100000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x100k)
> mean(x100k)
[1] 14.96088
> 1-pexp(15, rate = 1 / mean(x100k))
[1] 0.3669189
>
```



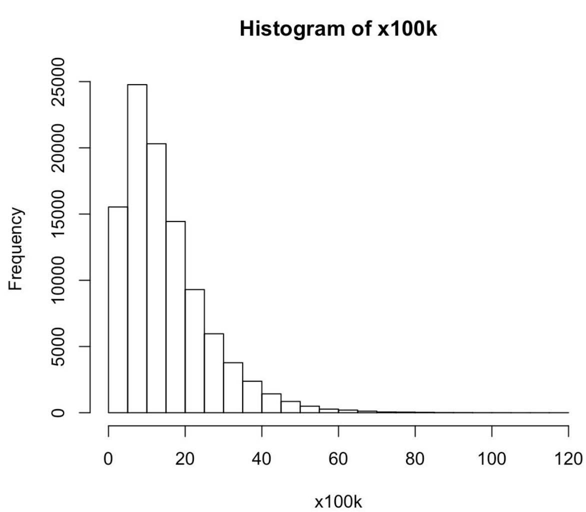

```
> #Test 3, Sample size 100,000
> x100k = replicate(100000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x100k)
> mean(x100k)
[1] 15.00287
> 1-pexp(15, rate = 1 / mean(x100k))
[1] 0.3679498
>
```



```
> #Test 4, Sample size 100,000
> x100k = replicate(100000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x100k)
> mean(x100k)
[1] 14.93113
> 1-pexp(15, rate = 1 / mean(x100k))
[1] 0.3661865
>
```




```
> #Test 5, Sample size 100,000
> x1000k = replicate(100000, max(rexp(n=1, rate=1/10), rexp(n=1,rate=1/10)))
> hist(x=x100k)
> mean(x100k)
[1] 14.93113
> 1-pexp(15, rate = 1 / mean(x100k))
[1] 0.3661865
>
```



Comparison Table:

Test for Sample size 1,000	E(T)	P(T>15)
Test 1	16.18445	0.3958121
Test 2	14.81066	0.3632063
Test 3	14.87652	0.3648385
Test 4	15.43804	0.3784673
Test 5	15.07782	0.369783

Test for Sample size 10,000	E(T)	P(T>15)
Test 1	14.8463	0.3640905
Test 2	14.88382	0.3650191
Test 3	15.01919	0.3683497
Test 4	15.02079	0.3683891
Test 5	14.97844	0.3673504

Test for Sample size 100,000	E(T)	P(T>15)
Test 1	14.99187	0.36768
Test 2	14.96088	0.3669189
Test 3	15.00287	0.3679498
Test 4	14.93113	0.3661865
Test 5	14.93113	0.3661865

It is evident through the results that as the sample size becomes larger the variation begins to reduce and this correlates directly with what the Central Limit Theorem says! If you compare the sample size between the five tests conducted on the 1,000 versus the 100,000 and 10,000 it is evident that higher sample sizes results in less variation for the E(T) as well as P(T>15).

Question 2:

To find the value of pi it is important to understand that the probability of a point falling in the circle under the space of a square = Area of circle / Area of square = pi/ 4. Now a number needs to be generated that is between 0 and 1 for both x and y, this should be iterated 10,000 times. Lastly we need to check if a number falls within the range of the circle or not.

Sample code and its output values:

```
> #Question 2
> iterations <- 10000
> x <- runif(iterations, min=0, max=1)
> y <- runif(iterations, min=0, max=1)
> in.circle <- (x-0.5)^2 + (y-0.5)^2 <= 0.5^2
> mc.pi <- (sum(in.circle)/iterations)*4
> mc.pi
[1] 3.1456
>
```

Values	
in.circle	logi [1:10000] TRUE TRUE FALSE TR…
iterations	10000
mc.pi	3.1456
num.cols	Named logi [1:33] FALSE FALSE TRU…
x	num [1:10000] 0.5267 0.8236 0.049…
y	num [1:10000] 0.44 0.797 0.176 0.…

The simulated value of pi is 3.1456 which is close to the 3.14159 that is the true value of pi that is used in various mathematical calculations.