DATA 252 / DATA 551: Homework 9

- This homework is due by April 13, 2020 at the beginning of class. **You need to submit your answers on Moodle in a pdf document**. In addition, there will be a short quiz at the beginning of class, which might contain contents from this homework (including the assigned video), in addition to contents from the most recent lecture.

1. Watch this video on using darts to estimate pi: https://www.youtube.com/watch?v=M34TO71SKGk&t=261s. Answer the following questions.

   (1) Is this a Monte Carlo method? Briefly explain.
   Yes, because of random throwing of darts.

   (2) A few different techniques of throwing darts were attempted in the video (e.g.: throwing with eyes covered or throwing a few darts at once). Discuss one technique that does not work very well and one technique that works better.
   A board with only one circle covered in whole area did not worked well.
   A board with multiple figures (circles with squares). + changing the orientation of board while throwing darts.

2. Use Monte Carlo integration to calculate $\int_0^3 x^2 dx$. Provide your code.

```
m = 10000
x = runif(m, min=0, max=3)
mean(x^2) * 3
 [1] 9.040013
```

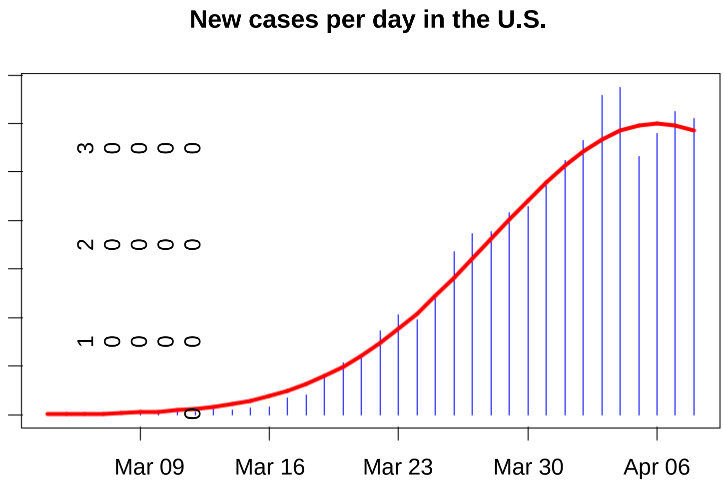
3. If you run the algorithm in 2 again, you would get a different estimate. Use a for loop to run the algorithm 10000 times and store the resulting 10000 estimates. Plot a histogram of these 10000 estimates and construct a 95% confidence interval of the true value (hint: use the quantile function). Provide your code.

4.

```
nsim=10000
result=rep(NA,nsim)
for(i in 1:nsim){
  x = runif(m, min=0, max=3)
  result[i]=mean(x^2) * 3
}
quantile(result, probs=c(0.025,0.975))

    2.5%    97.5%
8.843486 9.158240
```

5. Suppose the number of new COVID-19 cases per day in the U.S. can be modeled by the exponential function $30000 \cdot e^{-(x-34)^2/160}$. Below is a plot of the actual number of new cases in the U.S., from March 4 to April 8 (36 days), with the given function overlaid (data source: https://covidtracking.com/data/us-daily).



(1) In this setup, $\int_0^{36} 30000 \cdot e^{-(x-34)^2/160}dx$ represents the *total* number of cases in the U.S. by April 8. Use Monte Carlo integration to estimate this number (the actual number is 423164, for comparison). Provide your code.

```
m = 1000
x = runif(m, min=0, max=36)
mean(30000 * exp(1)^-(((x-34)^2)/160)) * 36
```

```
[1] 371037.2
```

```
m = 1000
nsim=10000
result=rep(NA,nsim)
for(i in 1:nsim){
  x = runif(m, min=0, max=36)
  result[i]=mean(30000 * exp(1)^-(((x-34)^2)/160)) * 36
}
quantile(result, probs=c(0.025,0.975))
```

```
    2.5%    97.5%
370871.5 420527.5
```

(2) Suppose we can use the given exponential function to model the number of new cases in the future. Then, for instance, $\int_0^{40} 30000 \cdot e^{-(x-34)^2/160}dx$ would represent the total number of cases by April 12. Use Monte Carlo integration to estimate this number. Provide your code.

```
m = 1000
x = runif(m, min=0, max=40)
mean(30000 * exp(1)^-(((x-34)^2)/160)) * 40
  [1] 536307.8
```

```
m = 1000
nsim=10000
result=rep(NA,nsim)
for(i in 1:nsim){
  x = runif(m, min=0, max=40)
  result[i]=mean(30000 * exp(1)^-(((x-34)^2)/160)) * 40
}
quantile(result,probs=c(0.025,0.975))
```

```
    2.5%    97.5%
474468.5 532875.3
```