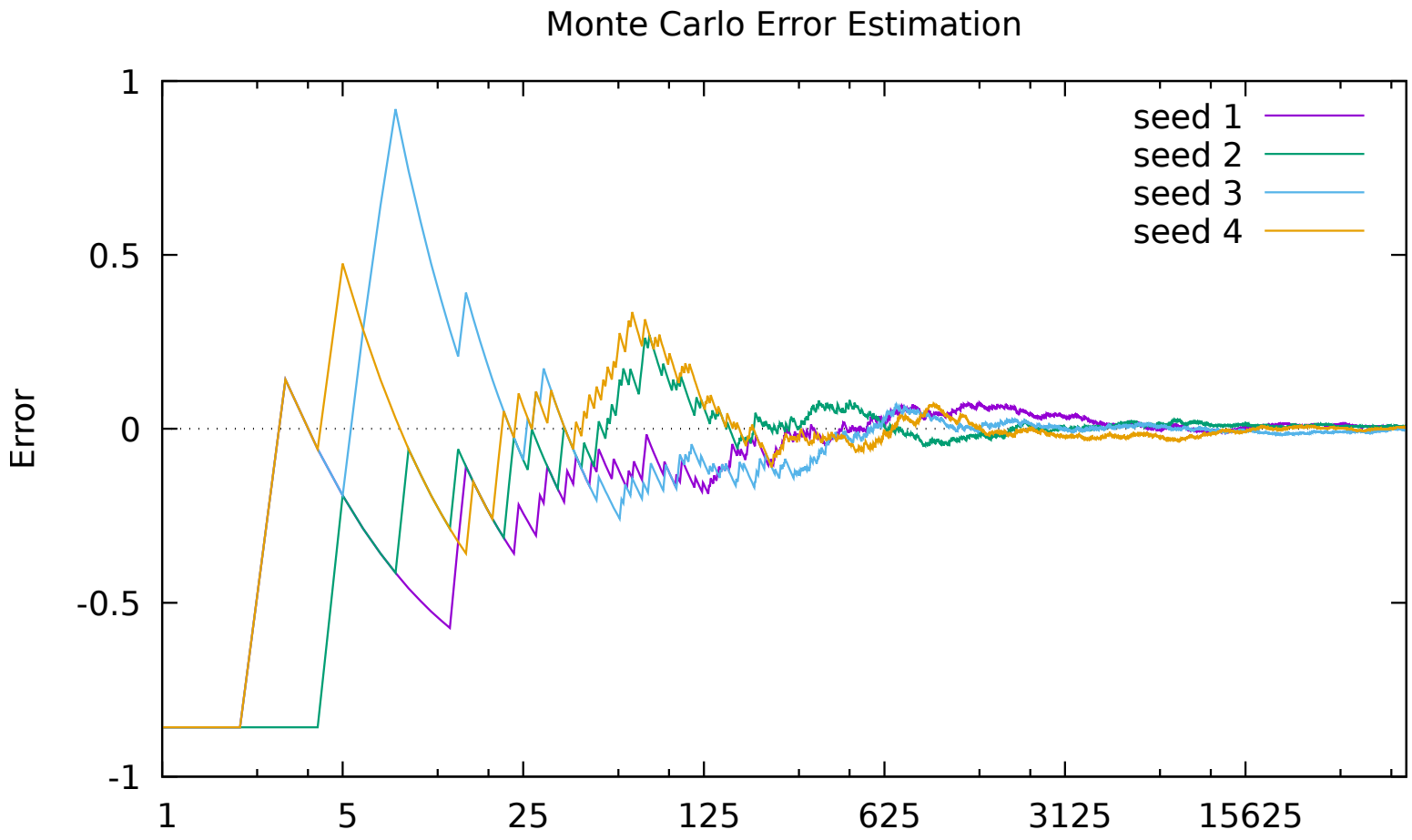
ASSIGNMENT 1 WRITEUP

Chris Moon

January 2023

1 GRAPHS: see below


Monte Carlo Error Estimation
Error
1
0.5
0
-0.5
-1
1
5
25
125
625
3125
15625
seed 1
seed 2
seed 3
seed 4

GRAPH ESTIMATION ERROR

The above graph shows how accurate the Monte Carlo simulation program is at estimating the value of PI, as more points are used.

Each line represents a random number generator seed used in the Monte Carlo simulation.

The graphed pattern clearly shows, that regardless of the seed, the estimated value of PI gets more accurate as more points are used.

I initially created four temporary data files to gather the data for this graph.

I ran the Monte Carlo program with 65536 points for each file, changing the seed between each.

Then, I used the pipeline command to take the Monte Carlo output and run it into the awk command.

I used the awk command in order to get the first and second columns of the Monte Carlo output, which were the number of points at the iteration and the estimate for PI at the iteration.

Then, I subtracted the estimate for PI from the actual value of PI to get the error value.

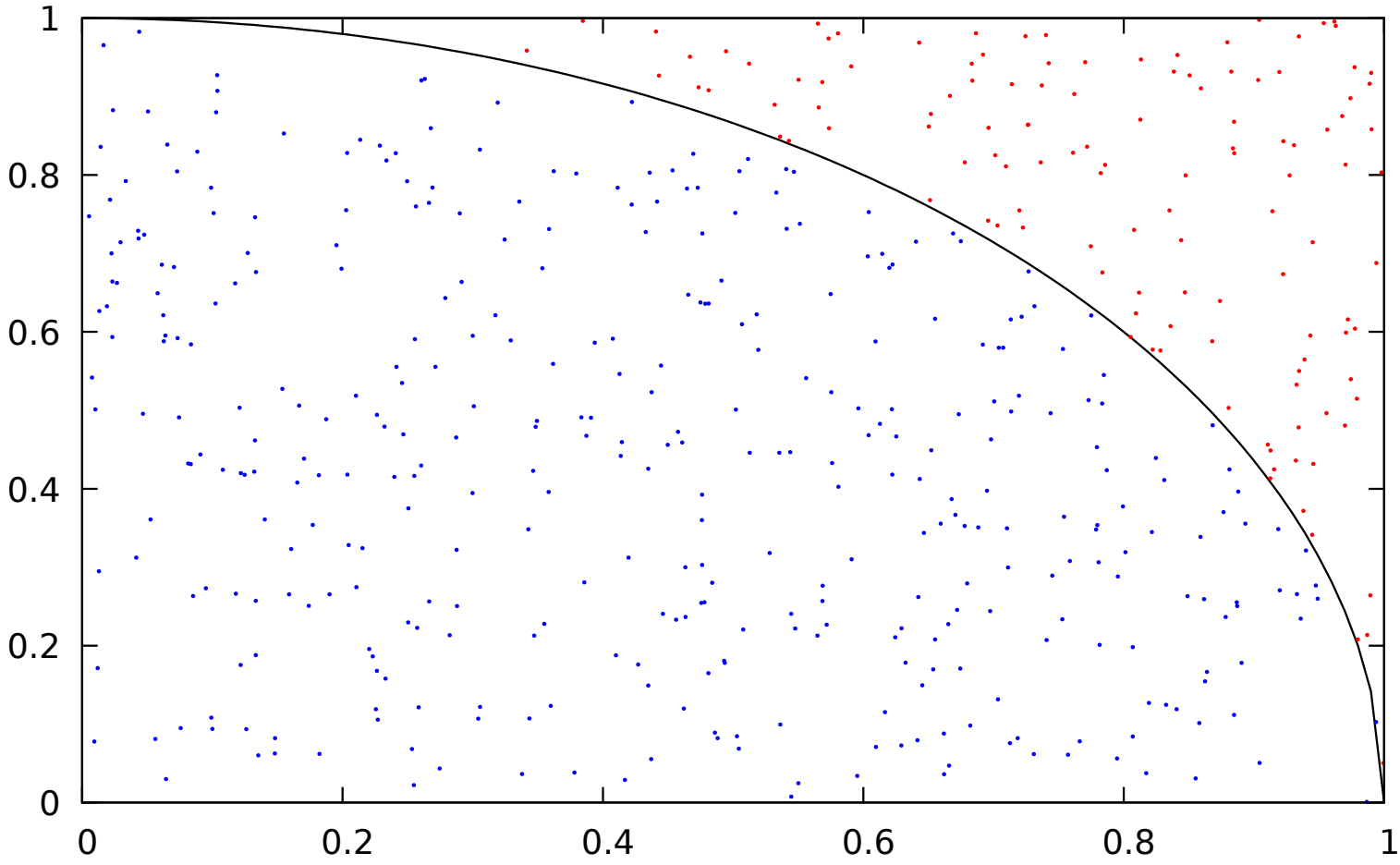
Then, I used the pipeline command to take the output of the awk command and run it into the tail command. This was so I could remove the header labels in the Monte Carlo output, since, being letters, they couldn't be graphed. I used tail -n 65536-1 in order to get every line in the output except for the first (the headers).

Finally, I pipelined the result into the corresponding temporary data file using the echo command.

The process simplified was as following:

Monte Carlo data into awk to get only the number of points and error into tail to remove the column headers into the file using echo.

After I had filled up all 4 temporary files with data, I graphed them using Gnuplot.


1
0.8
0.6
0.4
0.2
0
0
0.2
0.4
0.6
0.8
1

GRAPH COORDINATES

The above graph shows the coordinates of each point in a Monte Carlo simulation run, and whether they fall within a certain area marked by a circle area.

Similar to the previous graph, I created temporary data files. This time, one would hold the points falling outside the circle area and one would hold the points falling inside the circle area.

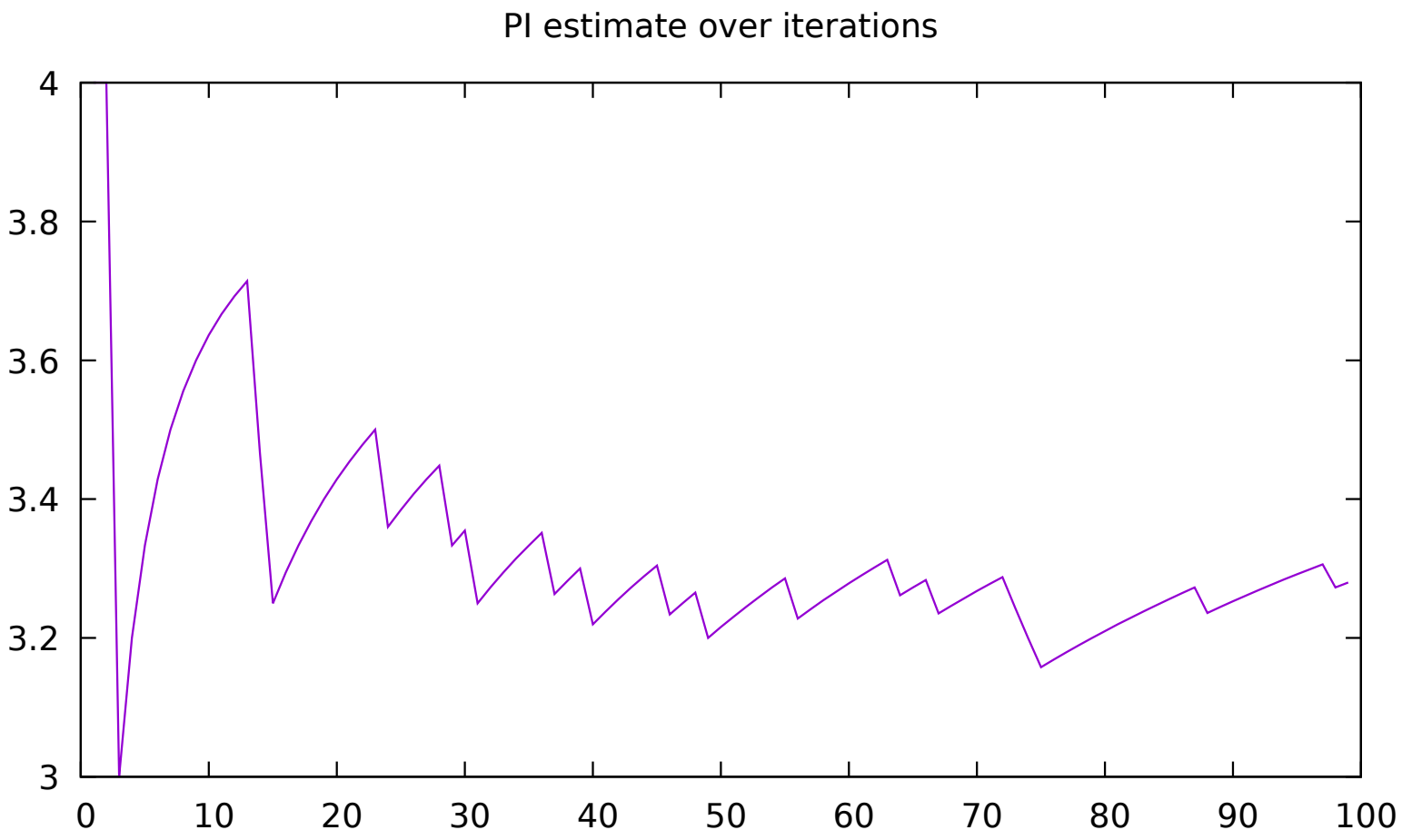
I ran the Monte Carlo program with 500 points and the same seed for each file.

As before I pipelined the resulting data output into the tail command (using -n 500-1) to remove the headers.

Then, I pipelined the result into the awk command and used an if statement inside to check if the value of the 5th column of each line was one or zero. A one represented a point inside the circle; a zero, outside. Then, I pipelined the x and y coordinates into the corresponding files.

This was so each file contained only points falling outside the circle and vice versa.

Finally, I graphed both files using Gnuplot.


PI estimate over iterations
4
3.8
3.6
3.4
3.2
3
0
10
20
30
40
50
60
70
80
90
100

GRAPH ESTIMATE VS INDEX

The above graph shows the Monte Carlo estimation for PI, as more points are added.

The graph clearly shows the estimations get closer to 3.1415... as the number of points increases.

The graph was created using an almost identical method to the error graph from before. However, only one file was graphed (since only one seed was tested), and the estimation values were left untouched instead of being subtracted from PI to find the error.