Goel, Ajay: Nu Id (001897443)

INFO 6205
Program Structures & Algorithms
Summer Full 2018
Assignment 4

In this Assignment, I have to verify the Birthday Problem and the Coupon Collector's Problem. Run experiments to validate the two expressions:

$$C_1(m) \sim \sqrt{\pi\, m / 2}$$

$$B_0(m) \sim m \ln m$$

where

- m is the number of bins/slots; and
- $C_1(m)$ is the (average) number of hashes/throws before the first collision in encountered; and
- $B_0(m)$ is the (average) number of hashes/throws before all bins/slots are filled (i.e. there are zero empty bins).

Hash Function:

It is done by calculation of modulus with number of bins.

➔ Value of index where value to be stored = hashCode() of the random number generated & 0x7fffffff % size

Here, I have used **0x7FFFFFFF is a number in hexadecimal (2,147,483,647 in decimal) that represents the maximum positive value for a 32-bit signed binary integer, '&' with this mask effectively sets the signed bit to 0, which means the number will always be positive.**

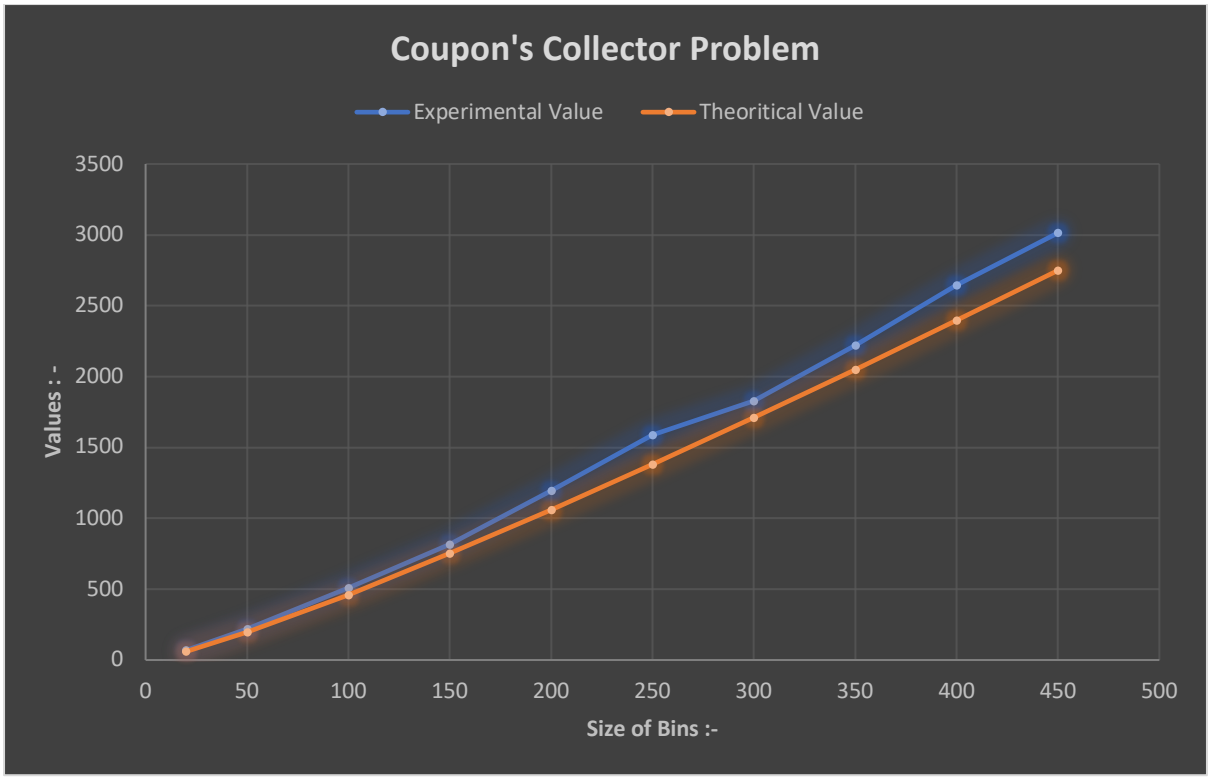
Analysis:

Coupon Collector:

In Coupon Collector problem, we are supposed to calculate the hashes/throws before the first collision in encountered:

I am generating random numbers and counting the hashes till full array is filled.

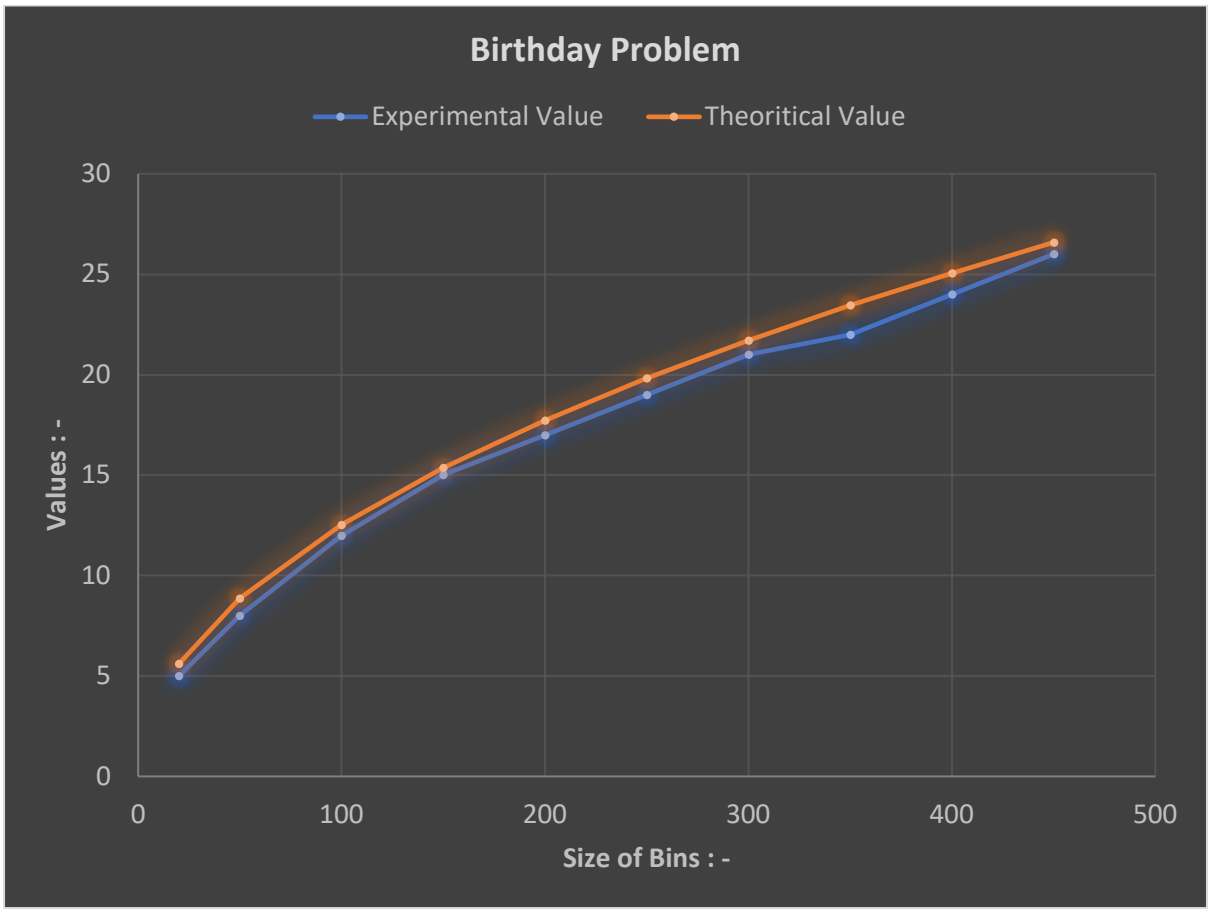
Coupon's collector Problem - Number of Runs - 100		
Bins - Size	**Experimental Value**	**Theoritical Value**
20	**69**	**59.9146**
50	222	195.6011
100	510	460.5170186
150	818	751.5952941
200	1196	1059.663473
250	1587	1380.365229
300	1830	1711.134742
350	2220	2050.2766
400	2647	2396.5858
450	3014	2749.1614



From the graph, we can see that value is almost equal to **Sqrt(pi * m /2).**
Hence Proved

Birthday Problem:
In Birthday Problem, we have to calculate the number of hashes before all the values are filled in the array.

Birthday's Problem - Number of Runs - 100		
Bins - Size	**Experimental Value**	**Theoritical Value**
20	**5**	**5.6049**
50	8	8.8622
100	12	12.53314137
150	14	15.34990062
200	16	17.72453851
250	20	19.81663649
300	24	21.70803764
350	22	23.4473
400	21	25.06628275
450	25	26.5868



From the above graph, we can see that value is almost equal to: **number of bins * ln (number of bins).**

SCREENSHOTS: TEST CASES

1. Hash function
2. Birthday formula
3. Coupon's collector formula
4. logic of counting balls in a bin

Please find screenshot below:

Goel, Ajay: Nu Id (001897443)