

Birthdays and picture cards

Vikram Bharadwaj Ramesh

Problem Statement:

Verify the **Birthday Problem** and the **Coupon Collector's Problem**. So, you will need to run experiments to validate the following two expressions:

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

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

Hash Function:

A Hash function is implemented by calculating a modulo of random value with the number of bins or slots i.e. Index number where the value needs to be stored = random Integer value % bin/slot value.

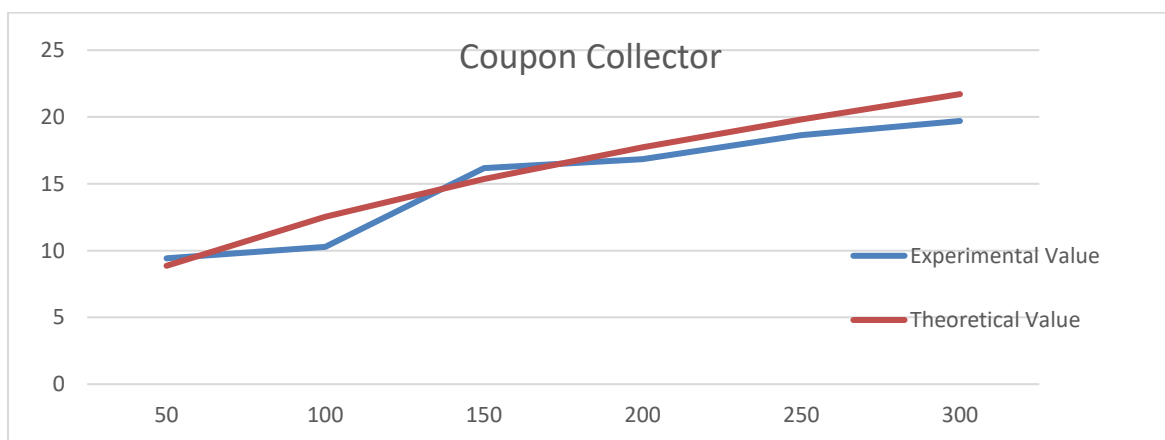
Analysis:

Coupon Collector:

According to the theorem, the expected number of coupons needed to complete a collection of size M is $\sim M \ln M$. So here we are supposed to calculate the hashes/throws before the first collision is encountered. To calculate this, I am using the "initializeArray1" method. In this method, count is incremented until the first collision is occurred. When the first collision occurs, the loop is stopped, and the value of count is stored.

Observation:

| Number of Runs | Bins/Slots | Experimental Value | Theoretical Value |
|----------------|------------|--------------------|-------------------|
| 50 | 50 | 9.42 | 8.86226925 |
| | 100 | 10.28 | 12.5331414 |
| | 150 | 16.18 | 15.3499006 |
| | 200 | 16.84 | 17.7245385 |
| | 250 | 18.64 | 19.8166365 |
| | 300 | 19.7 | 21.7080376 |



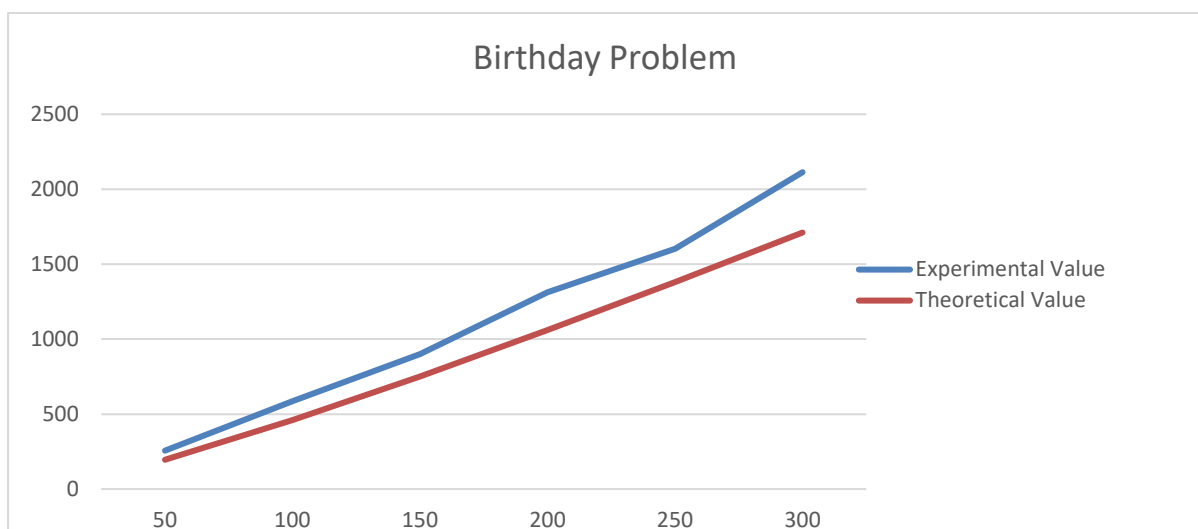
From the above analysis, we can see that the experimental value is almost equal to $\sqrt{(\pi * M/2)}$

Birthday Problem:

In the Birthday Problem, we need to calculate the number of hashes prior to the values being filled in the array. To calculate this, I am using the “**InitializeArray2**” method. This method will be called until all the values are filled in the array. For each iteration, all the elements of the array are checked.

Observation:

| Number of Runs | Bins/Slots | Experimental Value | Theoretical Value |
|----------------|------------|--------------------|-------------------|
| 50 | 50 | 256.42 | 195.6011503 |
| | 100 | 586.28 | 460.5170186 |
| | 150 | 901.18 | 751.5952941 |
| | 200 | 1312.84 | 1059.663473 |
| | 250 | 1603.64 | 1380.365229 |
| | 300 | 2113.7 | 1711.134742 |



From the above analysis, we can verify that that the Experimental value is almost equal to: no. of bins*ln (no of bins)
i.e $M * \ln M$

Conclusion:

From the above experiment, we can conclude that Hashing function is used to solve and calculate the Birthday and Coupon Collector problems. The experimental value and theoretical value is nearly same and therefore the two expressions which are given can be verified.