Random Dropping

Khudayar Farmanli
Politecnico Di Torino
Data Science and Engineering
Computer Aided Simulations and Performance Evaluation
Torino, Italy
s276310@studenti.polito.it

***Abstract*—In this homework we apply different random dropping policies with given equal number of bins and balls and based on different conditions. At the end we measure average maximum bin occupancy and compare it with theoritical ones.**

I. Assumptions

- The number of Bins and Balls are equal
- There are 2 different dropping policies:
  1) Random Dropping policy - where a ball will be delivered to randomly picked bin.
  2) Random Load Balancing policy - where a ball will be delivered to one of randomly picked d number of bins, exactly which has least occupation.

II. Input Parameters

- N - number of Bins and Balls
- D - Random Dropping condition where d = [1,2,4]
- Number of experiments
- Confidence Level
- Seed

III. Output parameters

- Average Maximum Bin Occupancy
- Upper and Lower bound of Confidence Interval
- Theoritical maximum occupancies:
  1) For d = 1 condition:

$$\frac{logn}{loglogn}(1 + O(1))$$

  2) For d = [2, 4] condition:

$$\frac{loglogn}{logd} + O(1)$$

  where n is the number of bins
- Relative Error

$$\varepsilon = \frac{Up - mean}{mean}$$

  where "Up" is upper bound of particular confidence interval and "mean" is average maximum bin occupancy of particular bunch of bins

IV. Development of Simulation

I enter the input parameters and based on them fill the bins according to the required condition. For instance, in the case where d = 1, a ball should be taken and a bin should be picked uniformly in a random way with replacement then throw that ball to that bin. On the other hand, if the case is one of d = [2, 4], again a ball should be taken and d number of bins picked uniformly in a random way with replacement from a bunch of bins and then the ball should be thrown to the one with least occupancy. This experiment is done repeatedly with pre-defined number of experiments and at the end maximum bin occupancy calculated with respect to the number of bins, in my particular case [10, 20, 30, 40, 50]. We get one maximum for each experiment run and based on that we calculate Average maximum bin occupancy accordingly. Then Confidence intervals and Relative errors are calculated based on these results and we compare our results with theoritical ones.

V. Results

As I mentioned already, I used different number of bins and balls such as [10, 20, 30, 40, 50] with conditions d = [1, 2, 4]. I defined number of experiments as 15 and confidence interval as 0.95. After running my code with these parameters I got results in particular which are shown below figures visually.

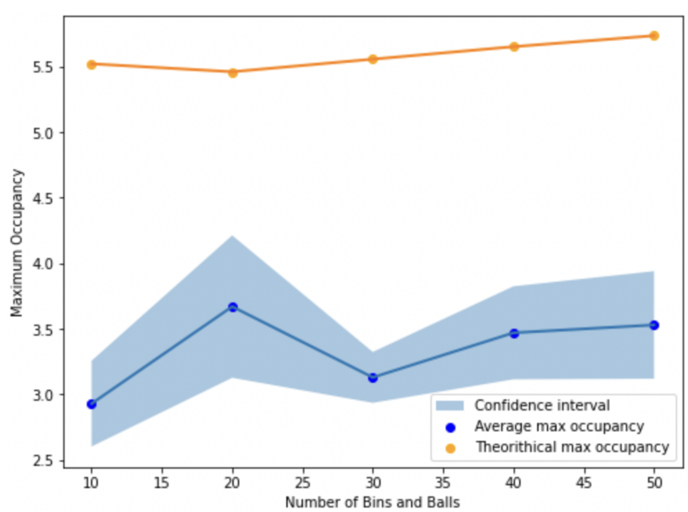

Fig. 1: Random Dropping where d = 1.

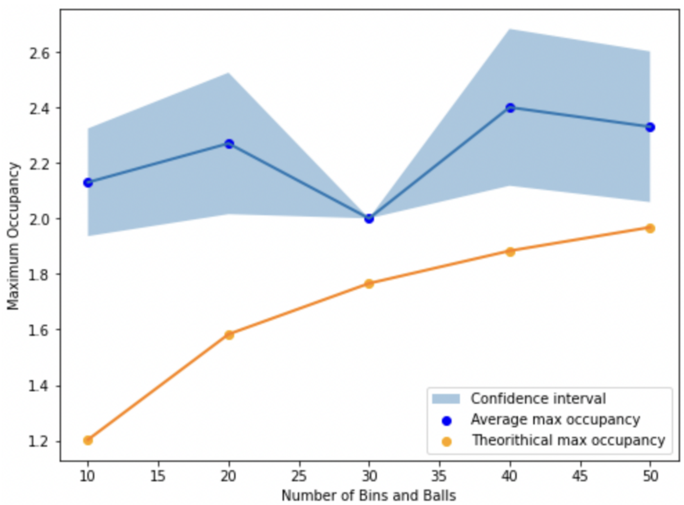

Fig. 2: Random Load Balancing where d = 2.

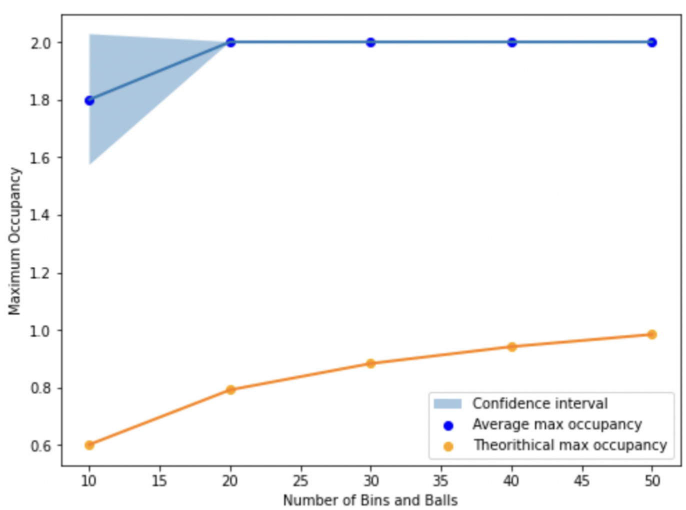

Fig. 3: Random Load Balancing where d = 4.

It can be easily seen from the above figures that with the increasing number of d value we distribute balls more equally between bins and create more balance. Also by increasing d value the maximum number of distributed balls into a single bin decreased so, greater d value smaller load per bin.

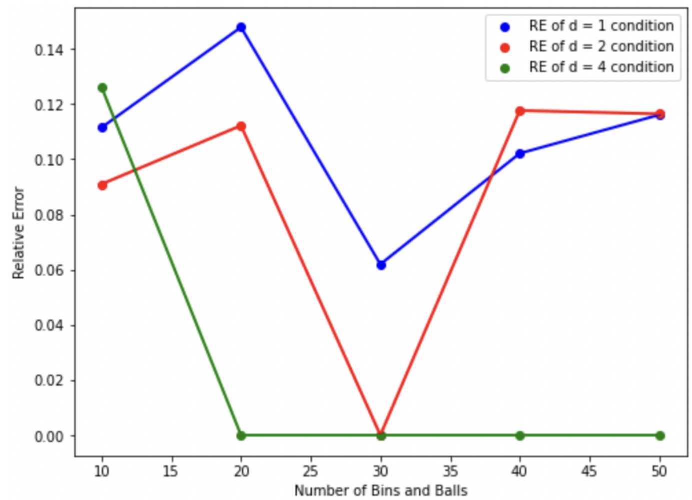

Fig. 4: Comparison of Relative Errors according to d value.

Also from Figure 4 we see how with the increasing number of d value the relative error decreases. It means the average maximum load occupancy value we have got from different experiments becomes maximum load for almost all experiments.

VI. Conclusion

In conclusion, based on the experiments done we can easily say that Random Load Balancing condition provide better results, so the number of d, even with smaller increase creates big impact on distribution of balls.