A short probability puzzle:

Suppose you have two bags, A and B. There is a probability distribution over A and B, on where one of n toys should go. Let the toy go in bag A with probability a, and bag B with probability b.

The objective is: if only t toys are placed in all bags, then bucket B should concentrate around ε of what you'd expect. That is, for some error threshold ε, bag B should contain $b(1 \pm \varepsilon)$ percentage of the toys.

What should the number of toys t be, in terms of a, b, ε? If you fix t, what should the probabilities a, b be to ensure ε concentration in bag B?

0.1 Notes

Here, if $a = b = \frac{1}{2}$, then the number of toys t should be approximately ε^{-2} (unproven).

Suppose $t = \varepsilon^{-1}$. Then what should b be?

1 Relation to Sampling, and obtaining Effective Resistance in $O(\varepsilon^{-1})$ time.

Suppose you have a given vertex v, and a cut with two parts: A and B. Let v be in part A, and let it have a' neighbors in A and b' neighbors in B.

Here, we model degree-preservation on v as placing down $a' + b'$ edges, either in A or in B.

Now, if we do leverage score sampling and sample E edges, the probability any edge appears is equal to the (leverage-score)$*|E|/n$. If we set $E = n/\varepsilon$, then it becomes $(leverage-score) * \varepsilon^{-1}$.

For any vertex of a unit graph, any edge leading into it has leverage score at least $\frac{1}{d}$, and thus for each vertex the sum of leverage scores is at least 1. Therefore, the expected number of edges is at least $\frac{1}{\varepsilon}$.

Assuming the worst case, the number of edges is equal to $\frac{1}{\varepsilon}$.