Chernoff Bounds: what are them and why we are using them

Chernoff Bounds are a set of powerful techniques used to provide tight bounds on the tail probabilities of sums of independent random variables.

* They are particularly useful for assessing the likelihood that the sum deviates significantly from its expected value
* Unlike simpler bounds like Markov's, Chernoff bounds take advantage of the distribution's specific characteristics to offer sharper estimates, especially useful for understanding the decay of tail probabilities exponentially fast

Immagine che contiene testo, Carattere, schermata, linea

Descrizione generata automaticamenteConsider the following footprint exercise:

What is important in this set of exercises is the set of following steps (always like this):

* Characterize the event (dependent on the type of problem you are dealing with)
  + And find the probability of success
* Characterize the expected value
* Use it to find
* Apply the bound given by the exercise

Consider we are usually bounding a precise value: apart from some *strange* cases, normally you have to get exactly the number given by the exercise.

So, here, we have to first consider . We know they are independent. Now, we simply need to find the expected value. We already have here the probabilty of success given by . This applies for all events since they are independent so:

Now, we find and this is done according to value we have to bound, given by the exercise or explicitly told here like :

Now that we found , let’s plug it in back in the original bound:

To infinity, it dominates the second factor, so we’d have

Immagine che contiene testo, Carattere, bianco, ricevuta

Descrizione generata automaticamente. In this case, consider if coin is tail, otherwise. These are all independent.

We have . Now find with

So, we do the following:

Now, we apply the bound as follows:

as the exercise wanted.

Immagine che contiene testo, Carattere, schermata, algebra

Descrizione generata automaticamenteConsider since we can assign the ball to any possible bin. So, this is uniform at random. The load of the specific bin is given by

First, we have to find :

Since each ball is assigned to a bin chosen uniformly at random, we have

To apply the Chernoff bound, we set equal to so:

Now, we apply the bound:

Recall the property of exponentials and logarithms there, so:

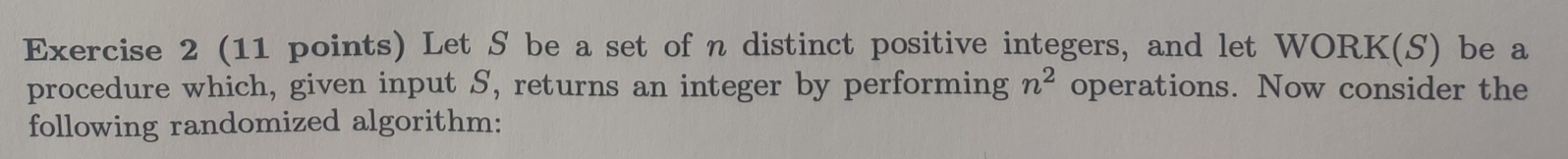
Recall from the exercise hint that

So, we have:

as the exercise wanted. We showed with high probability the bin with maximum load containing at most balls. We applied this for *one* bin, so we have to use now the union bound; simply use the previous result multiplying by all bins, so :

To characterize the *no bin will exceed*, use the complement event 🡪

Immagine che contiene testo, menu, documento

Descrizione generata automaticamente

Recalling the analysis of Randomized QuickSort: the event can be characterized as “in the first nodes of there have been lucky choices”. We are studying this specific event:

* if at the vertex of there is a lucky choice of the pivot
* are independent

We want the probability of to bound . Given , its expected value is as follows:

Now, let’s apply the following Chernoff bound (the first):

We then apply the Chernoff lemma as follows:

Immagine che contiene testo, Carattere, bianco, algebra

Descrizione generata automaticamenteHere, we use Karger since the hint uses that inequality and the only point in the program we saw that is exactly that analysis – so, that’s why we use that in place of a “normal” Chernoff Bound.

Characterize the event of getting a probability *at least* using Karger’s analysis; run different times the analysis and fix a constant , :

Since the probability is at least , so we characterize using Karger. Here, we will characterize the probability of failure (the complement with respect to the previous, so , running times to reduce the error probability.

We want to find a value for such that . In this case, it’s standard the use of this inequality:

This inequality is derived from the exponential function and the binomial expansion. It represents an upper bound on the expression , showing that it grows slower than .

Now, we use Karger’s analysis, in place of using we use and everything comes naturally.

By choosing it follows that (it holds coming from Karger):

is not in the form

Recall the following from the Karger analysis (choice of and rest of reasoning is that):

Continuing using Karger and the inequality:

Let’s wrap up:

For reference from theory, Karger’s analysis is this.

for some constant

The previous one is the probability of an unsuccessful event, so we want it very low, something like .

We want to find a value for such that . In this case, it’s standard the use of this inequality:

This inequality is derived from the exponential function and the binomial expansion. It represents an upper bound on the expression , showing that it grows slower than . The probability of not contracting the minimum cut in each iteration needs to be bounded and manipulated to ensure the overall algorithm's success probability is high.

By choosing it follows that:

Given I am curious, I asked myself: why exactly that value for ?

Consider the probability of success if while the failure is, by complement, which, amplified by runs, becomes . The constant is the desired level of confidence to keep the wanted threshold (in this case ) as low as possible. Then, using some good old GPT-4:

Immagine che contiene testo, schermata, Carattere, algebra

Descrizione generata automaticamente

Moving on:

is not in the form

Recall the following:

Let’s apply that:

Let’s wrap up (here, in the prof. notes, magically disappears, but I assume it to be so this works):

Then, by choosing that value for the Karger’s algorithm succeeds with high probability:

So, in the end, the Karger algorithm accumulates the size of the min-cut with probability at least .

Immagine che contiene testo, Carattere, schermata, algebra

Descrizione generata automaticamente

Here we have already and we need to find Now, we find :

Now, we use the bound:

Which is then verified for the constant as showed.

There are jobs to be assigned randomly to processors (Note: remember . Consider a processor and show that, with high probability on , processor does not receive more than jobs (Hint: define an appropriate indicator variable and apply the following Chernoff Bound).

*Theorem 1*: Let be independent indicator random variables with . Let and . Then, for all ,

*Solution*: Let (with jobs and when the job gets assigned to processor . Here, with independent between each other. The number of jobs received by the processor is then given it holds for each processor.

Now, find

Then, we find :

To apply the Chernoff bound, we set equal to so:

Now, we apply the bound:

Recall the property of exponentials and logarithms there, so:

Recall from the exercise hint that

So, we have:

as the exercise wanted. We showed with high probability the bin with maximum load containing at most jobs. We applied this for *one* bin, so we have to use now the union bound; simply use the previous result multiplying by all jobs, so :

To characterize the *no job will exceed*, use the complement event 🡪