Lecture Notes CS:5360 Randomized Algorithms
Lecture 7: Sept 11, 2018
Scribe: Tanmay Inamdar

1 Markov's Inequality

Recall that our general theme is to upper bound tail probabilities, i.e., probabilities of the form $\Pr(X \geq c \cdot E[X])$ or $\Pr(X \leq c \cdot E[X])$. The first tool towards that end is Markov's Inequality.
Note. This is a simple tool, but it is usually quite weak. It is mainly used to derive stronger tail bounds, such as Chebyshev's Inequality.

Theorem 1 (Markov's Inequality) *Let X be a non-negative random variable. Then,*

$$\Pr(X \geq a) \leq \frac{E[X]}{a}, \qquad \text{for any } a > 0.$$

Before we discuss the proof of Markov's Inequality, first let's look at a picture that illustrates the event that we are looking at.

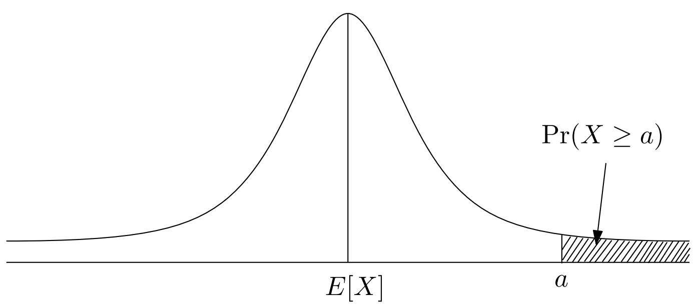


Figure 1: Markov's Inequality bounds the probability of the shaded region.

Proof:[1] Suppose X is a discrete random variable, for simplicity.

$$\begin{aligned} E[X] &= \sum_x x \cdot \Pr(X = x) \\ &\geq \sum_{x \geq a} x \cdot \Pr(X = x) \\ &\geq a \cdot \sum_{x \geq a} \cdot \Pr(X = x) \\ &= a \cdot \Pr(X \geq a) \end{aligned}$$

Rearranging, we get

$$\Pr(X \geq a) \leq \frac{E[X]}{a}.$$

□

Proof:[2] Define a random variable Y as follows. $Y = \begin{cases} 1 & \text{if } X \geq a \\ 0 & \text{otherwise.} \end{cases}$

Now, if $X < a$, $Y = 0$. Otherwise, $X \geq a$, in which case $Y = 1$. In both cases, we have that $Y \leq \frac{X}{a}$. Note that we use the fact that X is a non-negative random variable in the first case.

Therefore, $E[Y] \leq \frac{E[X]}{a}$. However, since Y is an indicator random variable, $E[Y] = \Pr(Y = 1) = \Pr(X \geq a)$. This implies that $\Pr(X \geq a) \leq \frac{E[X]}{a}$. □

Example. Let X be a random variable that denotes the number of heads, when n *fair* coins are tossed independently. Using Linearity of Expectation, we get that $E[X] = \frac{n}{2}$.

Plugging in $a = \frac{3n}{4}$ in Markov's Inequality, we get that $\Pr(X \geq \frac{3n}{4}) \leq \frac{n/2}{3n/4} = \frac{2}{3}$. This is a quite weak bound on the tail probability using Markov's Inequality, since we intuitively know that X should be concentrated very tightly around its mean. (If we toss 10,000 fair coins, we have a sense that the probability of getting 7,500 or more heads is going to be very small.)

To illustrate this point further, consider $\Pr(X \geq n)$. Plugging in $a = n$, we get $\Pr(X \geq n) \leq \frac{n/2}{n} = \frac{1}{2}$. However, we know that $\Pr(X \geq n) = \Pr(X = n) = \frac{1}{2^n}$, since outcomes of *all* n coin tosses must be heads, when $X = n$.

□

The example above illustrates that often, the bounds given by Markov's Inequality are quite weak. This should not be surprising, however, since this bound only makes use of the expected value of a random variable.

2 Chebyshev's Inequality

In order to get more information about a random variable, we can use *moments* of a random variable.

Definition 2 (Moment) *The* k^{th} moment *of a random variable* X *is* $E[X^k]$.

Higher moments often reveal more information about a random variable, which, in turn helps us derive better bounds. However, there is a trade-off. It is often difficult to compute higher moments in practical cases, e.g., while analyzing randomized algorithms. Now, let us look at the variance of a random variable.

Definition 3 (Variance) *The* variance *of a random variable* X*, denoted as* $\mathrm{Var}[X]$*, is* $E[(X - E[X])^2]$.

The variance of a random variable can be seen as the expected square of the distance of X, from its expected value $E[X]$. Another way to look at $\mathrm{Var}[X]$ is as follows.

$$\begin{aligned} \mathrm{Var}[X] &= E\left[(X - E[X])^2\right] \\ &= E\left[X^2 - 2 \cdot X \cdot E[X] + E[X]^2\right] \\ &= E[X^2] - E[2E[X] \cdot X] + E\left[E[X]^2\right] && \text{(Linearity of Expectation)} \\ &= E[X^2] - 2E[X] \cdot E[X] + E[X]^2 && \text{(2 and } E[X] \text{ are constants)} \\ &= E[X^2] - E[X]^2. \end{aligned}$$

That is, the variance of X equals the difference between the second moment of X, and the square of the expected value of X (i.e., the square of the first moment of X).

Now, we can derive Chebyshev's Inequality, which often gives much stronger bounds than the Markov's Inequality.

Theorem 4 (Chebyshev's Inequality) *For any $a > 0$,*

$$\Pr(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}.$$

Again, let us look at a picture that illustrates Chebyshev's Inequality.

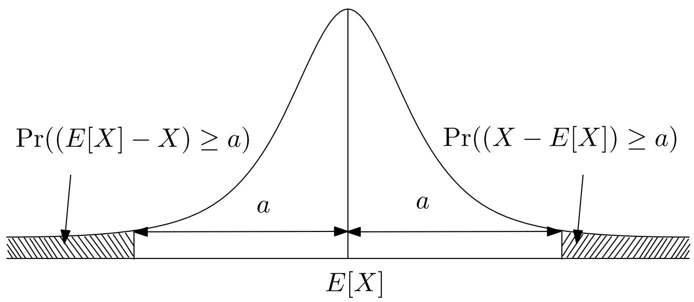


Figure 2: Chebyshev's Inequality bounds the probability of the shaded regions.

Proof:

$$\Pr\left(|X - E[X]| \geq a\right) = \Pr\left((X - E[X])^2 \geq a^2\right) = \Pr(Y \geq a^2)$$

Where, $Y = (X - E[X])^2$. Note that Y is a non-negative random variable. Therefore, using Markov's Inequality,

$$\Pr(Y \geq a^2) \leq \frac{E[Y]}{a^2} = \frac{E\left((X - E[X])^2\right)}{a^2} = \frac{\text{Var}[X]}{a^2}.$$

□

Example. Again consider the fair coin example. Recall that X denotes the number of heads, when n *fair* coins are tossed independently. We saw that $\Pr(X \geq \frac{3n}{4}) \leq \frac{2}{3}$, using Markov's Inequality. Let us see how Chebyshev's Inequality can be used to give a much stronger bound on this probability. First, notice that:

$$\Pr\left(X \geq \frac{3n}{4}\right) = \Pr\left(X - \frac{n}{2} \geq \frac{n}{4}\right) \leq \Pr\left(\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right) = \Pr\left(|X - E[X]| \geq \frac{n}{4}\right).$$

That is, we are interested in bounding the upper tail probability. However, as seen before, Chebyshev's Inequality upper bounds probabilities of both tails. In order to use Chebyshev's Inequality, we must first calculate $\text{Var}[X]$. First, we must characterize what kind of random variable X is.

Definition 5 (Binomial Random Variable) *A random variable X is* Binomial *with parameters n and p (denoted as $X \sim \text{Bin}(n,p)$) if X takes on values $0, 1, \ldots, n-1, n$, with the following distribution.*

$$\Pr(X = j) = \binom{n}{j} p^j (1-p)^{n-j}.$$

A binomial random variable $X \sim \text{Bin}(n,p)$ denotes the number of successes (heads), when n independent coins are tossed, with each coin having success (heads) probability of p. In our example, X is a binomial random variable with parameters n and $\frac{1}{2}$. However, we consider the more general case.

To compute $\text{Var}[X]$, we need $E[X]$ and $E[X^2]$. For the case of Binomial Random variable, $E[X] = np$ can be computed easily, as seen before. However, computing $E[X^2]$ directly is quite tedious. Therefore, we decompose X is the following manner.

For each coin toss $i = 1, \ldots, n$, define an indicator r.v. $X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p. \end{cases}$

That is, X_i is 1 if the i^{th} coin toss is heads, and 0 otherwise. It is easy to see that $X = \sum_{i=1}^{n} X_i$. Before we show how the variance of X can be decomposed, we need the following definition.

Definition 6 (Covariance) *The Covariance of random variables X_i and X_j, denoted as $\text{Cov}(X_i, X_j)$, is $E\left[(X_i - E[X_i]) \cdot (X_j - E[X_j])\right]$.*

$\text{Cov}(X_i, X_j)$ is a measure of correlation between X_i and X_j. It immediately follows from the definition that $\text{Cov}(X_i, X_j) = \text{Cov}(X_j, X_i)$. Another way to look at $\text{Cov}(X_i, X_j)$ is as follows.

$$\begin{aligned}
\text{Cov}(X_i, X_j) &= E\left[(X_i - E[X_i]) \cdot (X_j - E[X_j])\right] \\
&= E\left[X_i X_j - X_i E[X_j] - X_j E[X_i] + E[X_i]E[X_j]\right] \\
&= E[X_i X_j] - E[X_j] \cdot E[X_i] - E[X_i] \cdot E[X_i] + E[E[X_i] \cdot E[X_j]] \\
&\qquad \text{(Using Linearity of Expectation)} \\
&= E[X_i X_j] - E[X_i] \cdot E[X_j]
\end{aligned}$$

Now, we state the following theorem without proof.

Theorem 7

$$\text{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \text{Var}[X_i] + \sum_{\substack{i,j \\ i \neq j}} \text{Cov}(X_i, X_j)$$

Consider the case where all X_i's are mutually independent. Then, for any X_i, X_j, $E[X_i X_j] = E[X_i] \cdot E[X_j]$, which implies that $\text{Cov}(X_i, X_j) = 0$. That is,

Theorem 8 (Linearity of Variance) *If $X_1, X_2, \ldots, X_n$ are all mutually independent, then*

$$\text{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \text{Var}[X_i].$$

Notes.

1. Linearity of Variance requires the independence of the random variables, whereas Linearity of Expectation does not.

2. We do not need mutual independence between the random variables for Linearity of Variance. A weaker notion called *pairwise independence* suffices. That is, for any distinct X_i, X_j, it is sufficient to require X_i, X_j be independent.

Example: (Continued) In our coin toss example, all X_i's are in fact mutually independent. Therefore, $\text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i]$.

For any X_i, $\text{Var}[X_i] = E[X_i^2] - E[X_i]^2$. $E[X_i] = \Pr(X_i = 1) = p$. Note that X_i^2 also has the same distribution as X_i, and therefore, $E[X_i^2] = p$. So, $\text{Var}[X_i] = p - p^2 = p(1-p)$.

And therefore, $\text{Var}[X] = np(1-p)$.

Theorem 9 (Variance of a Binomial Random Variable) *If* $X \sim \text{Bin}(n,p)$, *then* $\text{Var}[X] = np(1-p)$.

Therefore, in the case of *fair* coin tosses, $\text{Var}[X] = \frac{n}{4}$. By Chebyshev's Inequality,

$$\Pr\left(\left|X - \frac{n}{2}\right| \geq \frac{n}{4}\right) \leq \frac{(n/4)}{(n/4)^2} = \frac{4}{n}.$$

Recall that Markov's Inequality gave us a much weaker bound of $\frac{2}{3}$ on the same tail probability. Later on, we will discover that using Chernoff Bounds, we can get an even stronger bound of $O\left(\frac{1}{\exp(n)}\right)$ on the same probability. However, Chernoff Bounds require mutual independence, whereas even the weaker notion of pairwise independence suffices for an application of Chebyshev's Inequality.

□

Lecture Notes CS:5360 Randomized Algorithms
Lectures 8 and 9: Sept 13 & Sept 18, 2018
Scribe: Tanmay Inamdar

Example 1: (Coupon Collector's Problem)
Recall that in the Coupon Collector's Problem, we defined X_i as the number of cereal boxes bought while having $i-1$ distinct coupons, and the random varaible $X = \sum_{i=1}^{n} X_i$ denoted the total number of boxes required to obtain all n distinct coupons. We had shown the following:

1. Each X_i is a geometric random variable with parameter $p_i = \frac{n-i+1}{n}$. So, $E[X_i] = \frac{n}{n-i+1}$.
2. $E[X] = n \ln n + \Theta(n)$, using Linearity of Expectation.

Can we bound the tail probability $\Pr(X \geq 2 \ln n)$?
Using Markov's Inequality, $\Pr(X \geq 2\ln n) \leq \frac{n \ln n + \Theta(n)}{2 \ln n} = \frac{1}{2} + \Theta\left(\frac{1}{\ln n}\right) = \frac{1}{2} + o(1)$. For sufficiently large n, this bound is arbitrarily close to $\frac{1}{2}$.
What do we require for using Chebyshev's Inequality?

1. We need to set $a = n \ln n$, ignoring the lower order $\Theta(n)$ term in $E[X]$.
2. We need to compute $\text{Var}(X)$. The variables X_i's are mutually independent. For example, the number of cereal boxes bought while having 3 distinct coupons does not affect the number of cereal boxes bought while having 4 distinct coupons. Therefore, we can use Linearity of Variance, provided that we can compute $\text{Var}[X_i]$ for each X_i.

We state the following result without proof.

Theorem 10 *If* $Y \sim \text{Geom}(p)$, *then* $\text{Var}[Y] = \frac{1-p}{p^2}$.

Using this, for any X_i, we have that $\text{Var}[X_i] = \frac{1-p_i}{p_i^2} \leq \frac{1}{p_i^2} = \frac{n^2}{(n-i+1)^2}$. Therefore,

$$\begin{aligned}
\text{Var}[X] &= \sum_{i=1}^{n} \text{Var}[X_i] \\
&\leq \sum_{i=1}^{n} \frac{n^2}{(n-i+1)^2} \\
&= n^2 \sum_{j=1}^{n} \frac{1}{j^2} \\
&\leq n^2 \sum_{j=1}^{\infty} \frac{1}{j^2} \\
&= n^2 \cdot \frac{\pi^2}{6} && \left(\because \sum_{j=1}^{\infty} \frac{1}{j^2} = \frac{\pi^2}{6}.\right)
\end{aligned}$$

Therefore, plugging $a = n \ln n$, $E[X] = n \ln n$ and $\text{Var}[X] = \frac{n^2\pi^2}{6}$ in Chebyshev's Inequality,

$$\Pr(X \geq 2 \ln n) \leq \Pr(|X - E[X]| \geq n \ln n) \leq \frac{n^2\pi^2}{6n^2(\ln n)^2} = \Theta\left(\frac{1}{(\ln n)^2}\right).$$

Note that this bound is much better than the bound obtained via Markov's Inequality. □

A note on error probabilities.

Error probability (upper bound)	Comment
$\frac{1}{c}$	Weakest useful probability bound. Can use probability amplification.
$\frac{1}{\text{poly}(\log n)}$	Approaches 0 as n increases, but rate of approach may not be satisfactory.
$\frac{1}{n^c}, 0 < c < 1.$	Gray area. Sometimes referred to as "high probability".
$\frac{1}{n}$	Standard meaning of "high probability".

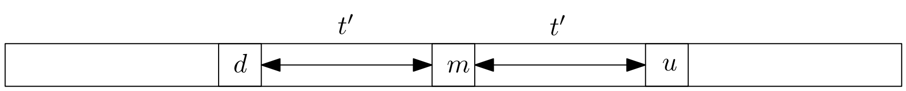
3 Median Finding by Sampling

We are going to discuss a Monte Carlo algorithm with deterministic $O(n)$ running time, with failure probability at most $\frac{1}{n^{1/4}}$. This algorithm is notable for its use of sampling, a probabilistic technique that has many applications. The analysis of the algorithm uses Chebyshev's inequality. A different randomized median find algorithm appears in the homework – it is Las Vegas with expected $O(n)$ running time.

The input is a list $L[1..n]$ of n distinct numbers. We are required to find the median element, i.e., an element of rank $\frac{n}{2}$ in the sorted version of L, which we denote by $m = \text{median}(L)$. First, we describe the idea of the algorithm informally.

Informal Description.

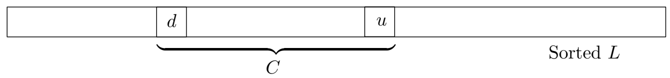
- We sample L to get a sublist S of t elements, where $t \ll n$ is a parameter to be fixed later. Sampling a smaller subset of the input is often referred to *down sampling*.
- The intuition is that median(S) should be "close to" median(L) in the sorted version of L. We can find median(S) by any standard sorting algorithm in $O(t \log t)$ time, which would be $O(n)$, if we choose t small enough.
- Let d be the element in Sorted S of rank $\frac{t}{2} - t'$, and u be the element in Sorted S of rank $\frac{t}{2} + t'$. Again, t' is a parameter to be fixed later. See the following figure for an illustration.



Sorted S

The intuition is that median(L) should lie between d and u in Sorted L.

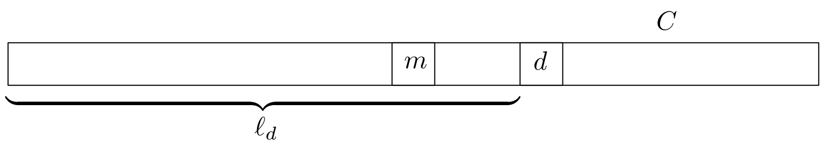
- Then, we find the set $C = \{x \in L \mid d \le x \le u\}$ – this takes $O(n)$ time.



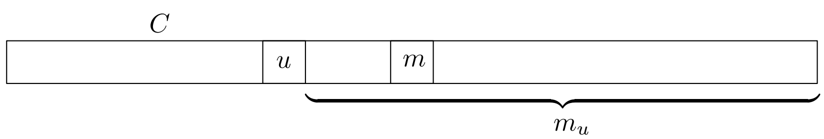
- We sort C and find the median of L.

- In what ways can the things go wrong?
  Define $\ell_d = |\{x \in L \mid x < d\}|$, and $m_u = |\{x \in L \mid x > u\}|$. ($\ell_d$ stands for *less than* d and m_u stands for *more than* u.) The "bad events" are as follows.

  1. $B_1 \equiv \ell_d > \frac{n}{2}$. In this case, m is smaller than d, and so $m \notin C$. See the figure for illustration.



  2. $B_2 \equiv m_u > \frac{n}{2}$. In this case, m is greater than u, and so $m \notin C$. See the figure for illustration.



  3. $B_3 \equiv C$ is "too large" to be sorted!

Now, we describe the algorithm formally.

Algorithm 1: MonteCarloMedianFind($L[1..n]$)

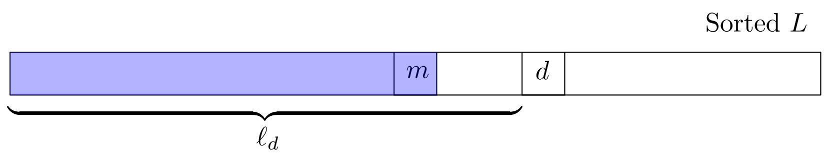
1 Pick $t = \lceil n^{3/4} \rceil$ elements from L by sampling independently, *with replacement*, uniformly at random. The resulting S is a *multiset*.
2 Sort S and let
   $d =$ element of rank $\lfloor \frac{n^{3/4}}{2} - \sqrt{n} \rfloor$ in Sorted S,
   $u =$ element of rank $\lceil \frac{n^{3/4}}{2} + \sqrt{n} \rceil$ in Sorted S.
3 By scanning L and comparing each element in L with d and u, compute $C = \{x \in L \mid d \le x \le u\}$, $\ell_d = |\{x \in L \mid x < d\}|$, and $m_u = |\{x \in L \mid x > u\}|$.
4 **if** $\ell_d > \frac{n}{2}$ **then** FAIL
5 **if** $m_u > \frac{n}{2}$ **then** FAIL
6 **if** $|C| > 4 \cdot n^{3/4}$ **then** FAIL
7 Sort C and return the element with rank $\frac{n}{2} - \ell_d$ in Sorted C.

Lines 4 and 5 correspond to the bad events B_1 and B_2 as defined earlier. Line 6 corresponds to the bad event B_3, which we now define formally as $B_3 \equiv |C| > 4 \cdot n^{3/4}$. Note that if none of the bad events B_1, B_2 and B_3 happens, then the algorithm is guaranteed to return the correct median of the list L, in $O(n)$ time (because if B_3 does not happen, then S and C are small enough to be sorted by a deterministic algorithm, say merge sort, in $O(n)$ time). We will omit ceilings and floors in the analysis for simplicity.

Lemma 11

$$\Pr(B_1) \le \frac{1}{4n^{1/4}}.$$

Proof: Recall that $B_1 \equiv \ell_d > \frac{n}{2}$. Let X denote the number of elements in S that are $\le m$. Note that d has a rank $\frac{n^{3/4}}{2} - \sqrt{n}$ in S, i.e., the number of elements in S that are less than d, is $\le \frac{n^{3/4}}{2} - \sqrt{n}$. Now, since $m \le d$, the number of elements sampled from the shaded blue region in the following picture, is at most $\frac{n^{3/4}}{2} - \sqrt{n}$.



Therefore, $\Pr(B_1) \le \Pr(X < \frac{n^{3/4}}{2} - \sqrt{n})$. Now, we upper bound the latter probability.

For analyzing X, we can look at the sampling process as a sequence of $t = n^{3/4}$ trials, with "success" corresponding to an element being chosen from the left of m. The probability of "success" is $\frac{\text{number of elements} \le m}{n} = \frac{\frac{n-1}{2}+1}{n} = \frac{1}{2} + \frac{1}{2n}$, assuming n is odd.

Therefore, $E[X] = n^{3/4} \cdot \left(\frac{1}{2} + \frac{1}{2n}\right) \ge \frac{n^{3/4}}{2}$, and $\text{Var}[X] = n^{3/4} \cdot \left(\frac{1}{2} + \frac{1}{2n}\right) \cdot \left(\frac{1}{2} - \frac{1}{2n}\right) \le \frac{n^{3/4}}{4}$, using the formulas for the expected value and the variance of a binomial random variable.

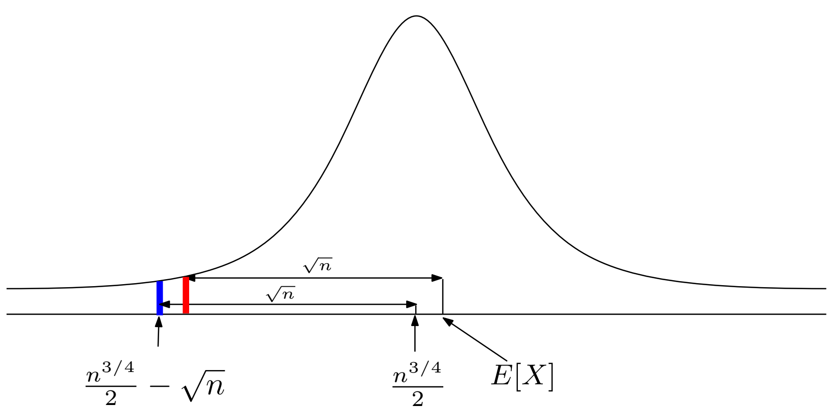


Figure 3: Area to the left of blue line: $X < \frac{n^{3/4}}{2} - \sqrt{n}$, Area to the left of red line: $X < E[X] - \sqrt{n}$

Now, we want to analyze $\Pr\left(X < \frac{n^{3/4}}{2} - \sqrt{n}\right)$, which is the area left of the blue line in the picture above. We upper bound it by $\Pr(X < E[X] - \sqrt{n})$, which is the area left of the red line. Therefore,

$$\Pr(B_1) \le \Pr\left(X < \frac{n^{3/4}}{2} - \sqrt{n}\right) \quad \text{(As argued earlier)}$$
$$\le \Pr(X < E[X] - \sqrt{n}) \quad \text{(From Figure 3.)}$$
$$\le \Pr\left(|X - E[X]| > \sqrt{n}\right) \quad \text{(Bounding lower tail probability by both tails)}$$
$$\le \frac{n^{3/4}}{4 \cdot (\sqrt{n})^2} \quad \text{(Plugging in } \mathrm{Var}[X] \le \tfrac{n^{3/4}}{4} \text{ and } a = \sqrt{n} \text{ in Chebyshev's inequality)}$$
$$= \frac{1}{4n^{1/4}}.$$

□

We have a similar lemma about the bad event $B_2 \equiv m_u > n/2$. The proof of this lemma is entirely symmetric, so we state it here without proof.

Lemma 12

$$\Pr(B_2) \le \frac{1}{4n^{1/4}}.$$

Now, we have a bad event $B_3 \equiv |C| > 4 \cdot n^{3/4}$. We have the following lemma.

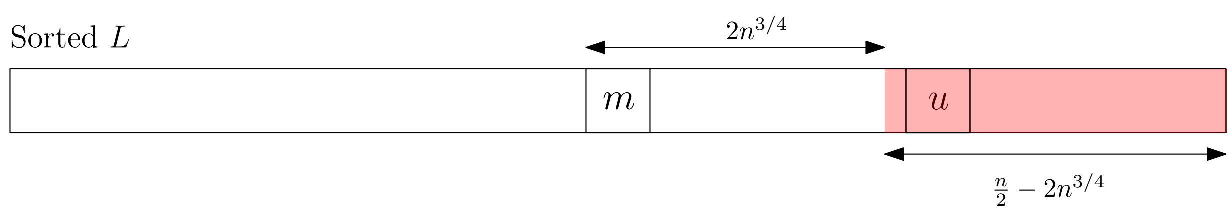
Lemma 13

$$\Pr(B_3) \le \frac{1}{2n^{1/4}}.$$

Proof: We decompose B_3 into two events B_{31} and B_{32}, where

$$B_{31} \equiv \text{number of elements in } C \text{ that are} \ge m \text{ is} > 2n^{3/4}$$
$$B_{32} \equiv \text{number of elements in } C \text{ that are} \le m \text{ is} > 2n^{3/4}$$

We have that $B_3 \equiv B_{31} \cup B_{32}$. We will show that $\Pr(B_{31}) \le \frac{1}{4n^{1/4}}$, and the proof for $\Pr(B_{32}) \le \frac{1}{4n^{1/4}}$ is symmetric. Note that by union bound, these two bounds imply the lemma.



If the event B_{31} occurs, then there are $> 2n^{3/4}$ elements between m and u (including both). However, the number of elements in S that are $> u$ is $\frac{n^{3/4}}{2} - \sqrt{n}$, all of which must be sampled from the red region (in fact, the subset of the red region that lies to the right of u). But we know that $< \frac{n^{3/4}}{2} - 2\sqrt{n}$ elements are sampled from the red region.

Let X denote the number of elements sampled in S from the red region, i.e., the number of sampled elements with rank $\ge \frac{n}{2} + 2n^{3/4}$ in L. Similar to earlier, X is a binomial random variable, with $n^{3/4}$ trials, and success probability $\frac{\frac{n}{2}-2n^{3/4}}{n}$. Therefore, $E[X] = \frac{n^{3/4}}{2} - 2\sqrt{n}$, and $\text{Var}[X] = n^{3/4}\left(\frac{1}{2} - \frac{2}{n^{1/4}}\right)\left(\frac{1}{2} + \frac{2}{n^{1/4}}\right) \leq \frac{n^{3/4}}{4}$.

$$
\begin{aligned}
\Pr(B_{31}) &= \Pr\left(X \geq \frac{n^{3/4}}{2} - \sqrt{n}\right) \\
&\leq \Pr\left(X - \frac{n^{3/4}}{2} - \sqrt{n} \geq \sqrt{n}\right) \\
&\leq \Pr\left(|X - E[X]| \geq \sqrt{n}\right) \\
&\leq \frac{n^{3/4}}{4(\sqrt{n})^2} \qquad \text{(Plugging in Var}[X] \leq \tfrac{n^{3/4}}{4} \text{ and } a = \sqrt{n} \text{ in Chebyshev's inequality)} \\
&= \frac{1}{4n^{1/4}}.
\end{aligned}
$$

□

Now, we conclude with the following theorem about Algorithm 1.

Theorem 14 *Algorithm 1 runs in $O(n)$ deterministic time and returns a median with probability at least $1 - \frac{1}{n^{1/4}}$.*

Proof: We skip the arguments about the running time and the correctness of the algorithm. We have already argued that the algorithm can make an error in four ways, corresponding to the bad events B_1, B_2 and B_{31}, B_{32}. We have argued that the probability of each of these bad events can be upper bounded by $\frac{1}{4n^{1/4}}$. By union bound over these four events, the probability that the algorithm makes an error is at most $\frac{1}{n^{1/4}}$. □
