Randomized Algorithms
Week 2: Tail Inequalities

Rao Kosaraju

In this section, we study three ways to estimate the tail probabilities of random variables. Please note that the more information we know about the random variable, the better the estimate we can derive about a given tail probability.

2.1 Markov Inequality

Theorem 1. Markov Inequality

If $X$ is a non-negative valued random variable with an expectation of $\mu$, then for any $c > 0$, $P[X \geq c\mu] \leq \frac{1}{c}$.

Proof. By definition,

$\mu = \sum_a aP[X = a] = \sum_{a < c\mu} aP[X = a] + \sum_{a \geq c\mu} aP[X = a] \geq 0 + \sum_{a \geq c\mu} c\mu P[X = a]$ as $X$ is non-negative valued

$= c\mu \sum_{a \geq c\mu} P[X = a] = c\mu P[X \geq c\mu]$

Hence, $P[X \geq c\mu] \leq \frac{\mu}{c\mu} = \frac{1}{c}$.

An equivalent form for the theorem is $P(X \geq c) \leq \frac{\mu}{c}$.

The knowledge of the standard deviation of the random variable $X$ would most often give a better bound.

2.2 Chebychev Inequality

Theorem 2. Chebychev Inequality

Let $X$ be a random variable with expectation $\mu_X$ and standard deviation $\sigma_X$. Then, $P[|X - \mu_X| \geq c\sigma_X] \leq \frac{1}{c^2}$. 
Proof. Let random variable \( Y = (X - \mu_X)^2 \). Then, \( E[Y] = E[(X - \mu_X)^2] = \sigma_X^2 \) by definition of \( \sigma_X \). Note that \( Y \) is a non-negative valued random variable.

Now, \( P[|X - \mu_X| \geq c\sigma_X] = P[(X - \mu_X)^2 \geq c^2\sigma_X^2] = P[Y \geq c^2\sigma_X^2] \).

Applying Markov Inequality to the random variable \( Y \), \( \Pr[Y \geq c^2\sigma_X^2] \leq \frac{1}{c^2} \).

Note that the random variable \( X \) need not be non-negative valued for the Chebychev inequality to hold.

### 2.3 Chernoff Bounds

The tail estimates given by Theorem 1 and Theorem 2 work for random variables in general. However, if the random variable \( X \) can be expressed as a sum of \( n \) independent random variables each of which is \( 0,1 \)-valued, then we can obtain very tight bounds on the tail estimates. This is expressed in the following theorem and the bounds are commonly called Chernoff Bounds.

**Theorem 3. Chernoff Bound for upper tail**

Let \( X \) be a random variable defined as \( X = X_1 + X_2 + \cdots + X_n \) where each \( X_i \), \( 1 \leq i \leq n \), is a \( 0,1 \)-valued random variable and all \( X_i \)'s are independent. Also, let \( E[X] = \mu \) and \( P[X_i = 1] = p_i, 1 \leq i \leq n \). Then for any \( \delta > 0 \), \( P[X \geq \mu(1 + \delta)] \leq \left( \frac{e^{\delta}}{(1+\delta)^{1+\delta}} \right)^\mu \).

**Proof.** While proving Chebychev inequality (Theorem 2), we made use of a second-order moment. Application of higher order moments would generally improve the bound on the tail inequality. We establish the tail estimates of the sums of independent random variables by utilizing the exponential function, which essentially captures a weighted sum of all the moments.

Let \( Y = e^{tX} \), for an appropriate positive value of \( t \) to be chosen later on. If we let \( Y_i = e^{tX_i} \), \( Y_i \)'s are also independent as \( X_i \)'s are. Note that \( Y = Y_1Y_2\cdots Y_n \). Note also that

\[
E[Y_i] = E[e^{tX_i}] = p_ie^t + (1-p_i)e^0 = 1 - p_i + p_ie^t \tag{1}
\]

\[
E[Y] = E[Y_1Y_2\cdots Y_n] = \prod_{i=1}^n E[Y_i] = \prod_{i=1}^n (1 - p_i + p_ie^t) \tag{2}
\]

in which the second equality follows from the independence of \( Y_i \)'s.

Observe that,

\[
\mu = \sum_{i=1}^n p_i \tag{3}
\]
by definition of $X$ and using the linearity of expectation. Note also that $Y$ is non-negatively
valued.

Hence,

$$P[X \geq \mu(1 + \delta)] = P[Y \geq e^{\mu(1+\delta)}] \leq \frac{E[Y]}{e^{\mu(1+\delta)}} = \frac{\prod_{i=1}^{n}(1 - p_i + p_i e^t)}{e^{\mu(1+\delta)}}$$  \hspace{1cm} (4)

by using Markov inequality from Theorem 1 and equation (2). The first equality follows from the fact that $t$ is positive.

We now make use of the inequality $1 + x \leq e^x$ in equation (4) resulting in

$$P[X \geq \mu(1 + \delta)] \leq \prod_{i=1}^{n} \frac{e^{-p_i(1-e^t)}}{e^{\mu(1+\delta)}} = \frac{e^{-\mu(1-e^t)}}{e^{\mu(1+\delta)}} = e^{\mu(1-e^t-\mu(1+\delta))}$$  \hspace{1cm} (5)

where the first equality follows from equation (3). In (5), we can choose a value of $t$ that minimizes the probability estimate. To find the minimum let $f(t) = \ln e^{\mu(1-e^t-\mu(1+\delta))} = -\mu(1-e^t) - t\mu(1+\delta)$. Differentiating $f(t)$ with respect to $t$ and equating it to zero gives us $\mu e^t - \mu(1+\delta) = 0$ or $t = \ln(1+\delta)$. (Note that we don’t need to establish that $f(t)$ gets minimized for this value of $t$ since any positive value can be chosen for $t$.) Substituting this value of $t$ in (5),

$$P[X \geq \mu(1 + \delta)] \leq \frac{e^{-\mu(1-(1+\delta))}}{(1+\delta)^{\mu(1+\delta)}} = \frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \mu = e^{\mu \delta/2} \left(1 + \delta\right)^{\mu}. \hspace{1cm} (6)$$

But the form of the inequality in equation (6) is not very convenient to handle. In addition, this form is hard to invert, i.e. given the probability bound, choose an appropriate $\delta$. Instead, we often use the following more convenient form:

**Theorem 4.** Let $X$ be defined as in Theorem 3. Then,

$$P[X \geq \mu(1 + \delta)] \leq \begin{cases} e^{-\mu \delta^2/3} & \text{if } 0 < \delta \leq 1 \\ e^{-\mu \delta \ln \delta} & \text{if } \delta > 1 \end{cases}$$

**Proof.** For the interval $0 < \delta \leq 1$, we prove the weaker inequality $\left(\frac{e^\delta}{(1+\delta)^{\mu}}\right)^{\mu} \leq e^{-\mu \delta^2/4}$, but the claimed bound can be established by refining the proof. (Problem zzz )

Taking logs and dividing by $\mu$, it suffices to prove

$$\delta - (1+\delta) \ln(1+\delta) \leq \frac{\delta^2}{4}.$$
That is, it suffices if we prove
\[ f(\delta) = \delta - (1 + \delta) \ln(1 + \delta) + \frac{\delta^2}{2} \leq 0. \]

Differentiating \( f \) twice, we have
\[
\begin{align*}
    f'(\delta) &= -\ln(1 + \delta) + \frac{\delta}{2}, \\
    f''(\delta) &= \frac{\delta - 1}{2(1+\delta)}.
\end{align*}
\]

Note that \( f''(\delta) \leq 0 \) for \( 0 < \delta \leq 1 \). Hence \( f'(\delta) \) is monotonically non-increasing as \( \delta \) varies from 0 to 1.

Since \( f'(1) < 0 \), \( f'(\delta) < 0 \) for any \( 0 < \delta \leq 1 \). Hence \( f(\delta) \) is monotonically decreasing as \( \delta \) varies from 0 to 1. Since \( f(0) = 0 \), \( f(\delta) \leq 0 \) for any \( 0 < \delta \leq 1 \).

For the interval \( \delta > 1 \), we now establish \( \left( \frac{e^{\delta}}{(1+\delta)^{(1+\delta)}} \right)^\mu \leq e^{-\frac{\mu \ln(1+\delta)}{2}} \). Simplifying as above, it suffices to prove \( g(\delta) = \delta - (1+\delta) \ln(1+\delta) + \frac{\delta \ln \delta}{2} \leq 0 \) when \( \delta > 1 \). Once again differentiating twice,
\[
\begin{align*}
    g'(\delta) &= \frac{1}{2} - \ln(1 + \delta) + \frac{\ln \delta}{2}, \\
    g''(\delta) &= -\frac{1}{1+\delta} + \frac{1}{1+\delta} + \frac{1}{2\delta} = \frac{1-\delta}{2\delta(1+\delta)}. \quad \text{Note that } g''(\delta) < 0 \text{ when } \delta > 1. 
\end{align*}
\]

Hence, \( g'(\delta) \) is monotonically decreasing, \( g'(1) \) is negative. Hence, \( g'(\delta) < 0 \) when \( \delta \geq 1 \). Consequently, \( g(\delta) \) is monotonically decreasing as \( \delta \) increases. Since \( g(1) < 0 \), \( g(\delta) < 0 \) for any \( \delta > 1 \).

\[ \square \]

### 2.4 Application of Tail Inequalities

We now apply tail inequalities for two problems.
2.4.1 $n$ Balls and $n$ Bins

Consider throwing $n$ balls, independently and uniformly at random, into $n$ bins. We are interested in the probability that bin 1 contains more than 7 balls. Define $n$-valued random variables $X_i$, $1 \leq i \leq n$, defined as $X_i = 1$ if ball $i$ falls into bin 1 and 0 otherwise. By uniformity, $P[X_i = 1] = \frac{1}{n}$. Define the random variable $X = X_1 + X_2 + \cdots + X_n$. Thus $X$ denotes the number of balls that fall in bin 1. By the linearity of expectation, $E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = n \frac{1}{n} = 1$.

Using the Markov inequality from Theorem 1, we get

$$P[X \geq 7] \leq \frac{1}{7}$$

(7)

For Chebychev inequality (Theorem 2), we first compute the standard deviation of $X$

$$\text{Var}(X_i) = E[X_i^2] - E[X_i]^2 = \frac{1}{n} - \frac{1}{n^2}, \text{ and}$$

(8)

$$\text{Var}(X) = \sum_{i=1}^{n} \text{Var}(X_i) = n \left( \frac{1}{n} - \frac{1}{n^2} \right) = 1 - \frac{1}{n},$$

(9)

where the first equality in (9) follows from the independence of $X_i$’s. Hence,

$$\sigma_X = \sqrt{1 - \frac{1}{n}}.$$  

(10)

Applying Chebychev inequality to the random variable $X$,

$$P[X \geq 7] = P[X - 1 \geq 6] \leq P[|X - 1| \geq 6] \leq \frac{1}{\left( \frac{6}{\sqrt{1 - \frac{1}{n}}} \right)^2} = \frac{1 - \frac{1}{n}}{36} \leq \frac{1}{36}$$

(11)

Using Chernoff bound from Theorem 4,

$$P[X \geq 7] = P[X \geq (1 + 6)1] \leq e^{-\frac{6 \ln n}{2}} = \frac{1}{216},$$

(12)

Now comparing equations (7), (11) and (12) we can see that using the Chebychev inequality gives a better bound than that of the Markov inequality, and a much better bound is obtained using Chernoff bounds. As another example, let us consider the probability that bin 1 has more than $1 + 10 \ln n$ balls. Using the Markov inequality, we get $P[X \geq 1 + 10 \ln n] \leq \frac{1}{1 + 10 \ln n}$. Using the Chebychev inequality, we get that $P[|X - 1| \geq 10 \ln n] \leq \frac{1}{100 \ln^2 n} \leq \frac{1}{100 \ln^2 n}$ whereas, using Chernoff bounds, we
get $P[X \geq 1 + 10 \ln n] \leq e^{\frac{10 \ln n}{2} \ln(10 \ln n)} = n^{-5 \ln n \ln(10 \ln n)} \leq \frac{1}{n^{10}}$, for any $n \geq 2$.

Out of the 3 bounds, only the Chernoff bound permits us to prove that with high probability $X < 1 + 10 \ln n$. In fact, we can show that even if we replace $10 \ln n$ by the smaller $\frac{c \ln n}{\ln \ln n}$ and for an appropriate positive constant, $P(X \geq 1 + \frac{c \ln n}{\ln \ln n}) \leq \frac{1}{n^{10}}$.

2.4.2 n ln n balls and n bins

In this case, there are $n \ln n$ balls and $n$ bins where balls are thrown into the bins independently and u.a.r. Define the random variables $X_i, 1 \leq i \leq n \ln n$, and $X$ as above. We have $P[X_i = 1] = \frac{1}{n}$ and $E[X] = \ln n$. Let us estimate the probability that bin 1 has more than $10 \ln n$ balls. Using the Markov inequality, we get that $P[X \geq 10 \ln n] \leq 1/10$. Instead, using the Chernoff bound, we get $P[X \geq 10 \ln n] = P[X \geq (1 + 9) \ln n] \leq e^{-9 \ln n \ln 9} \leq 1/\ln n^{10}$. Hence with high probability $X < 10 \ln n$. In this case, $10 \ln n$ is only a constant multiple of the $E(X)$. In the $n$ balls $n$ bins case, the corresponding multiple is $\frac{c \ln n}{\ln \ln n}$.

In general, if the expectation of the random variable is small, then we pay a higher penalty in the multiplicative factor to derive a w.h.p. bound.

2.5 Chernoff Bound for lower tail

In the above Chernoff bound, we have calculated the upper tail: $P[X \geq \mu(1 + \delta)]$. Here we outline the derivation of the lower tail: $P[X \leq \mu(1 - \delta)]$.

**Theorem 5.** Chernoff Bounds for lower tail

Let $X$ be a random variable defined as in Theorem 3. Then $P[X \leq \mu(1 - \delta)] \leq \left(\frac{e^{\delta}}{1 + \delta}ight)^{\mu}$.

**Proof.** Note that $P[X \leq \mu(1 - \delta)] = P[-X \geq -\mu(1 - \delta)]$. Now define $Y = e^{-tX}$, and proceed as in the proof of Theorem 3. \qed

We use the following convenient form of Theorem 5.

**Theorem 6.** Let $X$ be defined as in Theorem 3. Then, for any $0 < \delta \leq 1$ $P[X \leq \mu(1 - \delta)] \leq e^{-\mu\delta^2/2}$. Note that $\delta > 1$ is not of interest since $X$ cannot take negative values.
2.6 Special Case of Tail Inequalities

2.6.1 Independent and identical \{-1,+1\} valued random variables

**Theorem 7.** Let \(X_i, 1 \leq i \leq n,\) be \(n\) independent and identically distributed \{-1,+1\} valued random variables such that \(P[X_i = +1] = P[X_i = -1] = 1/2.\) Let the random variable \(X\) be defined by \(X = \sum_{i=1}^{n} X_i.\) Then, \(P[X \geq \delta] = P[X \leq -\delta] \leq e^{-\delta^2/2n}\) for any \(\delta > 0.\)

**Proof.** By Symmetry we have \(P[X \geq \delta] = P[X \leq -\delta].\) Here we prove a weaker form of the theorem: \(P[X \geq \delta] \leq e^{-\delta^2/6n},\) making use of Theorem 4. By making use of the exponential function \(e^{tX},\) a direct derivation results in the claimed bound. Observe that \(E[X_i] = 0\) and \(E[Y_i] = 1/2.\) Hence \(E[Y_i] = n/2.\) Thus, \(P[X \geq \delta] = P[Y \geq n + \delta] = P[Y \geq \frac{n}{2} + \frac{\delta}{2}] = e^{-\frac{\delta^2}{2n} + \frac{\delta}{n}} = e^{-\frac{\delta^2}{n}}.\)

When \(\delta > n,\) there is no need to apply this result since we know that \(X\) can never take a value greater than \(n.\) An alternative form of the above Theorem is \(P[X \geq \delta n] \leq e^{-\delta^2 n/2}.\)

3 Set Balancing Problem

In this section, we apply Chernoff bounds to another problem known as the Set Balancing Problem, which is defined as follows. Given an \(n \times n\) \(\{0,1\}\) matrix \(A,\) find a \{-1,+1\} valued column vector \(X\) such that the product \(AX\) has the smallest maximum absolute entry, i.e. minimize \(\|AX\|_{\infty}.

**Example 1.** Let the matrix \(A\) be as given below.

\[
A = \begin{bmatrix}
0 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0
\end{bmatrix}.
\]

For \(X = \begin{bmatrix} 1 & -1 & -1 & -1 \end{bmatrix}^T, (AX)^T = \begin{bmatrix} -2 & -1 & 0 & 1 \end{bmatrix}.\)

Thus the maximum absolute entry is 2.

Our goal is to make every entry of \(AX\) as close to 0 as possible. In a way, we are measuring the discrepancy using the maximum absolute value of the entries of \(AX.\) In general, it is not possible to make every entry of \(AX\) to be 0, for example when \(A\) has a row with an odd number of 1’s. A brute force solution for choosing \(X\) involves trying all possible column vectors of size \(n,\) which would take \(\Omega(2^n)\) time. Instead, we develop a very simple randomized algorithm that guarantees an expected discrepancy of \(O(\sqrt{n\ln n}).\) In a subsequent chapter, we derandomize the algorithm and obtain a deterministic polynomial time
algorithm with the same guarantee on the discrepancy. It is interesting to note that for this problem, Spencer \( \| \) proved that for any matrix \( A \) there is a column vector \( X \) such that the discrepancy is at most \( 6\sqrt{n} \). It is not known whether there is a polynomial time randomized algorithm that guarantees a discrepancy of \( O(\sqrt{n}) \).

The randomized algorithm works as follows. Let \( X = [X_1 X_2 \cdots X_n]^T \). Choose each \( X_i \) independently and u.a.r. with \( P[X_i = +1] = P[X_i = -1] = 1/2, 1 \leq i \leq n \). Before analyzing the performance guarantee of the randomized algorithm we state the classic Boole’s inequality.

**Fact 1. Boole’s Inequality:**

Let \( E_1, E_2, \cdots, E_n \) be \( n \) events. Then, \( P[E_1 \cup E_2 \cup \cdots \cup E_n] \leq P[E_1] + P[E_2] + \cdots + P[E_n] \).

Boole’s inequality has the following application. Let the events \( E_i \) be bad events. Then the union of all these bad events defines the event where at least one bad event occurs. To be able to prove that no bad event occurs with high probability, we can bound the probability that some bad event occurs by estimating the probability of each bad event (individually even though we are not given that all bad events are independent) and summing the probabilities.

Let the product \( AX \) be \( Y = [Y_1 Y_2 \cdots Y_n]^T \). Consider any \( Y_i \) wlog let it be, say \( Y_1 \). By the definition of matrix multiplication, \( Y_1 = A_{11} X_1 + A_{12} X_2 + \cdots + A_{1n} X_n \) where the \( A_{ij} \) denotes the element of \( A \) at \( i^{th} \) row and \( j^{th} \) column. Note that \( E[X_i] = 0 \) and by linearity of expectation, \( E[Y_1] = 0 \). For \( \delta = 2\sqrt{n \ln n} \), using Theorem 7 we get \( P[Y_1 \geq \delta] = P[Y_1 \leq -\delta] \leq e^{(-\frac{4n \ln n}{\delta^2})} \leq e^{-2 \ln n} = \frac{1}{n^2} \).

Hence, \( P[|Y_1| \geq \delta] = P[Y_1 \geq \delta] + P[Y_1 \leq -\delta] \leq \frac{2}{n^2} \).

Let us interpret each event \( |Y_i| \geq 2\sqrt{n \ln n} \), designated \( E_i \), as a “bad event”. Thus using Boole’s inequality, \( P[\text{for some } i, |Y_i| \geq 2\sqrt{n \ln n}] \leq \sum_{i=1}^{n} P[|Y_i| \geq 2\sqrt{n \ln n}] \leq n \frac{2}{n^2} = \frac{2}{n} \).

Thus with probability greater than \( 1 - \frac{2}{n} \), every entry in \( Y \) has absolute value at most \( 2\sqrt{n \ln n} \).

Hence with high probability, \( ||AX||_{\infty} < 2\sqrt{n \ln n} \). We can even upperbound \( E(||AX||_{\infty}) \) by observing that when \( ||AX||_{\infty} \) is not less than \( 2\sqrt{n \ln n} \), it can have a value at most \( n \). Hence, the expected value of the maximum absolute value is at most \((1 - \frac{2}{n})2\sqrt{n \ln n} + \frac{2}{n^2}n \).

### 3.1 Analysis of Randomized Quicksort

In this section, we revisit the RandQuickSort algorithm. We will use the Chernoff bound formula to bound a tail probability of the execution time.

We can view the execution of the algorithm on any input of \( n \) numbers as a binary tree of pivots. Every node has an associated sets of elements and a pivot element. For
the root, the associated set is the set of given elements and the pivot is the chosen element. For any node in the tree, the associated set is the set of elements passed down to it from its parents and the pivot is the chosen element out of that set. Let the input \( X = 1, 5, 9, 12, 27, 4, 34, 22, 20 \). A possible tree for the set of elements is given below. For the right child of the root the associated set is \( \{27,34,22,20\} \) and the pivot is 22.

![Tree Diagram]

Claim: The run-time of an execution of RandQuickSort is \( \leq \) sum of the depths of the nodes in the tree.

Proof: This follows from a simple amortization argument. Fix any element \( \alpha \). It is involved in one comparison at each of the nodes of the tree, starting at the root, along a path until \( \alpha \) is chosen as a pivot element or \( \alpha \) becomes a leaf, thus the charge to that node is the depth of the node. Charge all these comparisons to that node at which \( \alpha \) becomes a pivot element or a leaf. Thus the charge to the node is the depth of the node. For example, in the tree shown above item 34 is involved in 2 comparisons (12 : 34, 22 : 34) and the depth of the node with 34 as the pivot is 2. The total number of charges to all the nodes is equal to the sum of depths of the nodes in the tree. Hence the claim holds.

Claim: With high probability, there is no node with depth \( \geq 24 \ln n \).

We concentrate on the \( i \)th smallest element, \( \alpha_i \). (In above set \( X \), the 4th smallest element is 9.) For each \( 1 \leq i \leq n \), we calculate the probability that the \( i \)th smallest element, when it becomes a pivot has a depth > \( 24 \ln n \). We consider the first \( 24 \ln n \) nodes on the path from the root to the \( i \)th smallest element.

At a typical node in the tree, let its set of element \( S \) be of size \( m \). If the pivot is in the middle \( m/2 \) elements (i.e. one of \( \frac{m}{4} \)th to \( \frac{3m}{4} \)th smallest elements of \( S \)), then the maximum size of a child’s set is \( \leq 3m/4 \).
In the path from the root to the $i$th smallest element, we define a 0,1 valued r.v. $X_i$ for the $i$th as follows: $X_i$ takes the value 1 if its pivot is in the middle half of the elements.

Let $\alpha$ be the number of $X_i$'s that equal 1. If the array at the $24 \ln n$th level has more than one element, then

$$(\frac{3}{4})^{\alpha n} > 1, \text{ Then } \alpha < 3 \ln n \text{ (the exact value is not important) } \Rightarrow \alpha \ln(3/4) + \ln n > 0$$

$$\Rightarrow \alpha < \frac{\ln n}{\ln(4/3)} \Rightarrow \alpha < 3 \ln n$$

Thus, if more than $3 \ln n$ of the $X_i$'s are 1, then path of pivots from the root to the $i$th element is shorter than $24 \ln n$.

Let $X = X_1 + X_2 + \cdots + X_{24\ln n}$ (this is the number of times the $X_i$'s are 1). Note that $P(X_i = 1) = 1/2$, hence $E(X) = 12 \ln n$.

$$P(\text{the path of pivots from the root to the $i$th element is longer than } 24 \ln n) \leq P(X \leq 3 \ln n).$$

(13)

$$P[X \leq 3\ln n] = P[X \leq 12\ln n (1 - \frac{3}{4})] \leq e^{-\frac{(12 \ln n)(3/4)^2}{2}} \leq \frac{1}{n^2}.$$  

The first inequality follows from the chernoff bound for the lower tail. Using the following version of Chernoff Bound: $P(X \leq (1-\delta)\mu) \leq e^{-\mu \delta^2/2}, \text{ (??) } \leq e^{-\frac{(12 \ln n)(3/4)^2}{2}} \leq \frac{1}{n^{2(\ln n)}} \leq \frac{1}{n^n}.$

Now, using Boole’s inequality, $P(\text{The path of pivots from the root to any one of the } n \text{ elements is longer than } 24 \ln n) \leq \frac{1}{n^2}.$

### 3.2 Conclusion

Previously, we have determined that the expected run-time of RandQuickSort is $\leq 2n \ln n$. Now, we have established a high probability bound by sacrificing a constant multiplier in the runtime.