1 Chernoff bound

The Chernoff bound gives exponentially decreasing bounds on tail distributions of sums of independent random variables. This means the bound is asymptotically stronger than Markov’s or Chebyshev’s inequality. However, independence of the random variables is required.

**Theorem 1 (Chernoff bound)** Let \( X = x_1 + x_2 + \cdots + x_n \) where \( X_1, X_2, \ldots, X_n \) are \( n \) independent 0/1 variables, each having probability \( p_i \). Then

\[
Pr [X \geq (1 + \epsilon)\mu] \leq \left( \frac{e^{\epsilon}}{(1 + \epsilon)^{1+\epsilon}} \right)^\mu
\]

where \( \mu = E[X] \).

**Proof:** Since \( e^x \) is a convex function, for all \( t > 0, X \geq (1 + \epsilon)\mu \) is equivalent with \( e^{tX} \geq e^{t(1+\epsilon)\mu} \). Here, \( t \) is a variable that we will optimize later. We can then apply Markov’s inequality which gives

\[
Pr[X \geq (1 + \epsilon)\mu] = Pr[e^{tX} \geq e^{t(1+\epsilon)\mu}] \leq \frac{E[e^{tX}]}{e^{t(1+\epsilon)\mu}}.
\]

All \( X_i \) are independent so we can split the expectation into \( n \) parts.

\[
E[e^{tX}] = E\left[e^{t\sum_{i=1}^n X_i}\right] = E\left[e^{t\sum_{i=1}^n tX_i}\right] = E\left[\prod_{i=1}^n e^{tX_i}\right] = \prod_{i=1}^n E\left[e^{tX_i}\right] = \prod_{i=1}^n (1 - p_i) + p_i e^t
\]

Picking \( t = \ln(1 + \epsilon) \) gives

\[
E[e^{tX}] = \prod_{i=1}^n (1 - p_i) + p_i e^{\ln(1 + \epsilon)} = \prod_{i=1}^n (1 - p_i) + p_i (1 + \epsilon) = \prod_{i=1}^n 1 + \epsilon p_i.
\]

Now we use the bound \( 1 + x \leq e^x \) to get

\[
\prod_{i=1}^n 1 + \epsilon p_i \leq \prod_{i=1}^n e^{\epsilon p_i} = e^{\epsilon \sum_{i=1}^n p_i} = e^{\mu}.
\]

Combining this with equation 2 gives

\[
Pr[X \geq (1 + \epsilon)\mu] \leq \frac{e^{\mu}}{e^{\ln(1+\epsilon)(1+\epsilon)\mu}} = \left( \frac{e^{\epsilon}}{(1 + \epsilon)^{1+\epsilon}} \right)^\mu.
\]

\( \square \)
Corollary 2  We can simplify the Chernoff bound by distinguishing two cases.

If $\epsilon \leq 2e - 1$

$$
Pr[X \geq (1 + \epsilon)\mu] \leq \left( \frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}} \right)^\mu \leq e^{-\epsilon\mu/4}
$$

(7)

and if $\epsilon > 2e - 1$

$$
Pr[X \geq (1 + \epsilon)\mu] \leq \left( \frac{e^\epsilon}{(1 + \epsilon)^{1+\epsilon}} \right)^\mu \leq \left( \frac{e}{\epsilon} \right)^{\epsilon\mu}.
$$

(8)

2 Some remarks

- In our proof, the $X_i$ were 0/1 random variables but this is not strictly necessary. The bound also holds if $X_i \in [0, 1]$.

- Independence of the $X_i$ is crucial. Consider $X = \sum_{i=1}^n X + i$ where

$$
X_1 = \begin{cases} 
1 & \text{with probability } \frac{1}{2} \\
0 & \text{with probability } \frac{1}{2}
\end{cases}
$$

(9)

and $X_1 = X_2 = \cdots = X_n$. Then

$$
X = \begin{cases} 
n & \text{with probability } \frac{1}{2} \\
0 & \text{with probability } \frac{1}{2}
\end{cases}
$$

(10)

However, the Chernoff bound (equation 7) would imply

$$
\frac{1}{2} = Pr[X = n] = Pr\left[ X \geq \left( 1 + \frac{1}{2} \right) \frac{n}{2} \right] \leq e^{-\frac{n}{16}}
$$

(11)

- Boundedness is also crucial. Consider the case where $X_1, X_2, \ldots, X_n$ are independent random variables with $E[X_i] = \frac{1}{n}$ and a random variable $X_0$ with

$$
X_0 = \begin{cases} 
n & \text{with probability } \frac{1}{n} \\
0 & \text{with probability } 1 - \frac{1}{n}
\end{cases}
$$

(12)

Then

$$
Pr\left[ X \geq \frac{n}{2} \right] \geq Pr[X_0 = n] = \frac{1}{n}
$$

(13)

but the Chernoff bound (equation 8) would imply

$$
Pr\left[ X \geq \frac{n}{2} \right] \leq \left( \frac{e}{n/4} \right)^{n/4} \leq \left( \frac{1}{2} \right)^{n/4}
$$

(14)
3 Edge-disjoint paths

We can find an application of Chernoff bounds in the problem of finding edge-disjoint paths in a graph.

**Problem 1 (Edge-disjoint paths)** Given a directed graph \( G = (V, E) \), and \( k \) pairs \((s_1, t_1), (s_2, t_2), \ldots, (s_k, t_k)\), find edge disjoint paths between \((s_i, t_i)\) for all \( i \in \{1, 2, \ldots, k\}\).

This problem is NP-hard. One might think this problem could be modeled into a max flow problem as follows. Create two auxiliary vertices \( s \) and \( t \) and connect all \( s_i \) to \( s \) with capacity 1, and connect all \( t_i \) to \( t \) with capacity 1. Then find a flow of \( k \) units from \( s \) to \( t \). However, this model can return a solution with a path from \( s_i \) to \( t_j \), \( i \neq j \). For an example, we refer to figure 1.

```
Figure 1: In this graph, a flow will be found from \( s_1 \) to \( t_2 \) and from \( s_2 \) to \( t_1 \). Edge-disjoint flows from \( s_1 \) to \( t_1 \) and from \( s_2 \) to \( t_2 \) are not possible here.
```

Even though this problem is NP-hard, we can still say something useful about it. For this reason we introduce the term *Congestion*.

**Definition 3 (Congestion)** The congestion \( C_e \) of an edge \( e \) in a set of paths \( P \) is the amount of paths in \( P \) that use edge \( e \).

**Theorem 4 (Congestion bound)** There exists a polynomial-time algorithm that finds paths from \( s_i \) to \( t_i \) for \( i \in \{1, 2, \ldots, k\} \) such that for all \( e \)

\[
C_e = O \left( \frac{\log n}{\log \log n} \right)
\]

provided edge-disjoint paths exist.

We will prove this theorem, but first we need to do some work. We model this problem into a flow LP defined by equations 16 through 20.
\[
\sum_{i=1}^{k} f^{(i)}_e \leq 1 \quad \forall e \in E \quad \text{(congestion constraint)} \quad (16)
\]
\[
f^{(i)}_e \geq 0 \quad \forall e \in E, i = 1, 2, \ldots, k \quad (17)
\]
\[
\sum_{e \in N^+(v)} f^{(i)}_e = \sum_{e \in N^-(v)} f^{(i)}_e \quad \forall v \in V, i = 1, 2, \ldots, k \quad (18)
\]
\[
\sum_{e \in N^+(s_i)} f^{(i)}_e = 1 \quad (19)
\]
\[
\sum_{e \in N^-(t_i)} f^{(i)}_e = 1 \quad (20)
\]

In this LP, \(i = 1, 2, \ldots, k\) are commodities, \(f^{(i)}_e\) represents how much commodity \(i\) is routed on edge \(i\) and \(N^-(v)\) and \(N^+(v)\) are the amount of incoming and outgoing edges of vertex \(v\).

Any flow \(f^{(i)}\) from \(s_i\) to \(t_i\) can be decomposed into a convex combination of paths. We formulate this more formally in the following lemma.

**Lemma 5 (Path decomposition)** Let \(P^{(i)}\) be the set of paths from \(s_i\) to \(t_i\). For any flow \(f^{(i)}\) from \(s_i\) to \(t_i\) there exist \(\alpha^{(i)}_j, i = 1, 2, \ldots, |P^{(i)}|\) such that

\[
f^{(i)} = \sum_j \alpha^{(i)}_j P^{(i)}_j, \quad \sum_j \alpha^{(i)}_j = 1.
\]

**Proof:** Let \(e\) be the edge with minimal flow in \(f^{(i)}\) and let \(f^{(i)}_e = \delta\). We can then extend this edge to a path from \(s_i\) to \(t_i\) through \(e\) with flow \(\delta\) and remove this from the flow. We can keep repeating this process until no flow remains which gives our convex combination of paths.

We can now prove the theorem.

**Proof: (Congestion bound)** Let \(P_i\) be the collection of \(s_i\)-\(t_i\) paths. For all commodities \(i = 1, 2, \ldots, k\) we pick a path \(p \in P_i\) with probability \(\alpha_p\). Now the expected congestion of an edge \(e\) is given by

\[
\mathbb{E}[C_e] = \sum_i \sum_{p \in P_i, e \in p} \alpha^{(i)}_p = \sum_i f^{(i)}_e \leq 1. \quad (22)
\]

Now the congestion of an edge can be written as

\[
C_e = \sum_i X^{(i)}_e \quad (23)
\]

where

\[
X^{(i)}_e = \begin{cases} 
1 & \text{if path for } i \text{ uses edge } e \\
0 & \text{else}.
\end{cases} \quad (24)
\]
Now \( Pr[X_e^{(i)} = 1] = f_e^{(i)} \). The Chernoff bound gives

\[
Pr \left[ C_e \geq \frac{c \log n}{\log \log n} \right] \leq \frac{1}{n^c}
\]

(25)

for all edges \( e \) and \( c > 0 \). As there are always less than \( n^2 \) edges, the maximal congestion is bounded by

\[
Pr \left[ \max_{e \in E} C_e \geq \frac{c \log n}{\log \log n} \right] \leq n^2 \frac{1}{n^c} = \frac{1}{n^{c-2}}
\]

(26)

which proves the theorem.

Now let \( G = (V, E) \) be a graph where the optimal maximum congestion is \( C \) where \( C \) is a large constant. An example of this problem could be a network where all paths have to cross a certain service point.

**Claim 6** The same algorithm as before now returns a congestion of \( C + O(\sqrt{C \log n}) \).

**Proof:** We again define random variables \( X_e^{(i)} \) such that

\[
X_e^{(i)} = \begin{cases} 
1 & \text{if path for } i \text{ uses edge } e \\
0 & \text{else.}
\end{cases}
\]

(27)

We find \( E[C_e] = E \left[ \sum_i X_e^{(i)} \right] \leq C \).

\[
Pr \left[ \sum_i X_e^{(i)} \geq C + O \left( \sqrt{C \log n} \right) \right] = Pr \left[ \sum_i X_e^{(i)} \geq \left( 1 + O \left( \frac{\sqrt{C \log n}}{C} \right) \right) C \right]
\]

(28)

Since \( C \) is large, we can assume \( C > n \log n \) and thus \( C \gg \sqrt{C \log n} \). Therefore we can apply equation [7] of the Chernoff bound.

\[
Pr \left[ \sum_i X_e^{(i)} \geq \left( 1 + O \left( \frac{\sqrt{C \log n}}{C} \right) \right) C \right] \leq \exp \left[ -C \left( \frac{C \log n}{C^2} \right) \frac{C}{4} \right] = e^{-O(\log n)} = \frac{1}{\text{poly}(n)}
\]

(29)

\qed