CS5314
Randomized Algorithms

Lecture 11: Chernoff Bounds
(Application)
Objectives

• Apply Chernoff bounds to bound the tail and the head distributions of sum of Poisson trials
  - General case of Binomial random variable

• Revisit the Coin Flip example
Sum of Bernoulli Trials

Recall that: $X = \text{Bin}(n,p)$ is a random variable for the sum of $n$ independent indicators, each with success probability $p$

Each indicator is called a Bernoulli trial

$n = \# \text{ trials} \quad X = \text{total} \# \text{ success}$
Suppose that we allow each of these indicators to choose its own success probability (instead of a fixed $p$).

Each indicator is called a Poisson trial.

$n = \# \text{trials} \quad X = \text{total } \# \text{success}$
Bounding Sum of Poisson Trials

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent Poisson trials, such that

$$\Pr( X_i = 1 ) = p_i$$

Let $X = X_1 + X_2 + \ldots + X_n$

Then,

$$\mu = E[X] = E[X_1 + X_2 + \ldots + X_n]$$

$$= p_1 + p_2 + \ldots + p_n$$
Bounding Sum of Poisson Trials

We shall use Chernoff bound to bound the tail distribution \( \Pr(X \geq (1+\delta)\mu) \):

Let \( M_X \) be the MGF of \( X \), and
\[ M_{X_i} \] be the MGF of \( X_i \).

Since \( X_i \)'s are independent, we have
\[ M_X(t) = \prod_i M_{X_i}(t) \quad \text{... (why?)} \]
Bounding Sum of Poisson Trials

Question: What is $M_{X_i}(t)$?

By definition,

$$M_{X_i}(t) = E[e^{tX_i}]$$

$$= p_i e^{t(1)} + (1-p_i) e^{t(0)}$$

$$= p_i e^t + (1-p_i)$$

$$= 1 + p_i (e^t - 1)$$

$$\leq e^{pi} (e^t - 1) \quad \cdots \text{(why?)}$$
Bounding Sum of Poisson Trials

So,

\[ M_X(t) = \prod_i M_{X_i}(t) \]

\[ \leq \prod_i e^{p_i(e^t - 1)} \quad \text{... (why?)} \]

\[ = \exp \{ \sum_i p_i(e^t - 1) \} \]

\[ = e^{\mu(e^t - 1)} \]
Bounding Sum of Poisson Trials

Theorem: Let $X_1, X_2, \ldots, X_n$ be independent Poisson trials such that $\Pr(X_i = 1) = p_i$.

Let $X = X_1 + X_2 + \ldots + X_n$ and $\mu = \mathbb{E}[X]$.

Then, for all $\delta > 0$,

$$
\Pr(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}
$$

How to prove?
Proof: For any \( t > 0 \), we have

\[
\Pr(X \geq (1+\delta)\mu) = \Pr(e^{tX} \geq e^{t(1+\delta)\mu}) \\
\leq \frac{\mathbb{E}[e^{tX}]}{e^{t(1+\delta)\mu}} \\
= \mathcal{M}_X(t) / e^{t(1+\delta)\mu} \\
\leq e^{\mu(e^t-1)} / e^{t(1+\delta)\mu}
\]

... (by which inequality?)

To get the best bound for \( \Pr(X \geq (1+\delta)\mu) \), we now choose \( t \) so as to minimize the term

\[
e^{\mu(e^t-1)} / e^{t(1+\delta)\mu}
\]
Proof (cont)

Question: Which $t$ should we choose?

Observation:

To minimize $e^{\mu(e^t-1)} / e^{t(1+\delta)\mu}$

$\iff$ to minimize $\log_e (e^{\mu(e^t-1)} / e^{t(1+\delta)\mu})$

So, we want to choose $t$ so as to minimize the term $\log_e (e^{\mu(e^t-1)} / e^{t(1+\delta)\mu})$, which is

$\mu(e^t-1) - t(1+\delta)\mu$

By calculus, the best $t$ will be $\log_e (1+\delta)$
Proof (cont)

So, by substituting $t = \log_e (1+\delta)$ in the previous inequality:

$$\Pr(X \geq (1+\delta)\mu) \leq e^{\mu(e^t-1)/e^t}$$

we get:

$$\Pr(X \geq (1+\delta)\mu) \leq e^{\mu(\delta)/(1+\delta)} / (1+\delta)^{(1+\delta)\mu}$$

$$= \left(e^\delta/(1+\delta)^{(1+\delta)}\right)^\mu$$
Two Weaker but Easier Bounds

Theorem: Let $X_1, X_2, \ldots, X_n$ be independent Poisson trials such that $\Pr(X_i = 1) = p_i$.

Let $X = X_1 + X_2 + \ldots + X_n$ and $\mu = \mathbb{E}[X]$. Then,

(1) for all $0 < \delta \leq 1$,

$$\Pr(X \geq (1+\delta)\mu) \leq e^{-\mu\delta^2/3}$$

(2) For $R \geq 6\mu$,

$$\Pr(X \geq R) \leq 2^{-R}$$

How to prove?
Proof of (1)

To prove (1), it is sufficient to show whenever $0 < \delta \leq 1$,

$$e^\delta/(1+\delta)^{(1+\delta)} \leq e^{-\delta^2/3}$$

Equivalently, we want to show

$$\delta - (1+\delta) \log_e (1+\delta) \leq -\delta^2/3$$

(this is obtained by taking log on both sides)

Let $f(\delta) = \delta - (1+\delta) \log_e (1+\delta) + \delta^2/3$

**Target:** to show $f(\delta) \leq 0$
Proof of (1)

Then, \( f'(\delta) = 1 - (1+\delta)/(1+\delta) - \log_e (1+\delta) + 2\delta/3 \)
\[ = - \log_e (1+\delta) + 2\delta/3 \]

and

\( f''(\delta) = -1/(1+\delta) + 2/3 \)

So, we see that

\( f''(\delta) < 0 \quad \text{for } 0 \leq \delta < 1/2 \)

\( f''(\delta) \geq 0 \quad \text{for } 1/2 \leq \delta \leq 1 \)
Proof of (1)

This implies that

\[ f'(\delta) \text{ is decreasing for } 0 \leq \delta < 1/2 \]
\[ f'(\delta) \text{ is increasing for } 1/2 \leq \delta \leq 1 \]

Next, from \( f'(\delta) = -\log_e (1+\delta) + 2\delta/3 \),
we see that \( f'(0) = 0 \) and \( f'(1) < 0 \)
Together with the above, we conclude that
\[ f'(\delta) \leq 0 \quad \text{in the interval } [0,1] \]
Proof of (1)

Since $f'(\delta) \leq 0$ in the interval $[0,1]$, $f$ is decreasing in the interval $[0,1]$. 

$\Rightarrow f(0)$ is the maximum point of $f$ in the interval $[0,1]$.

Thus, for $0 < \delta \leq 1$,

$$f(\delta) = \delta - (1+\delta) \log_e (1+\delta) + \frac{\delta^2}{3} \leq f(0) = 0$$

This completes the proof of (1).
Proof of (2)

For (2), we want to show for $R \geq 6\mu$, 
\[
\Pr(X \geq R) \leq 2^{-R}
\]

Let $R = (1+\delta)\mu$, so that $\delta \geq 5 > 0$

Then, we have
\[
\Pr(X \geq (1+\delta)\mu) \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}\mu} \right)
\]
\[
\leq \left(\frac{e^{(1+\delta)}}{(1+\delta)^{(1+\delta)}}\right)^\mu
\]
\[
= \left(\frac{e}{(1+\delta)}\right)^\mu(1+\delta)
\]
\[
\leq \left(\frac{e}{6}\right)^R \leq 2^{-R}
\]
Bounding the Head Distribution

Theorem: Let $X_1, X_2, ..., X_n$ be independent Poisson trials such that $\Pr(X_i = 1) = p_i$.

Let $X = X_1 + X_2 + ... + X_n$ and $\mu = E[X]$.

Then, for $0 < \delta < 1$,

1. $\Pr(X \leq (1-\delta)\mu) \leq (e^{-\delta}/(1-\delta)^{(1-\delta)})^\mu$
2. $\Pr(X \leq (1-\delta)\mu) \leq e^{-\mu\delta^2/2}$

How to prove?
Proof of (1): For any \( t < 0 \), we have

\[
\begin{align*}
\Pr(X \leq (1-\delta)\mu) &= \Pr(e^{+X} \geq e^{+(1-\delta)\mu}) \\
&\leq E[e^{+X}] / e^{+(1-\delta)\mu} \\
&= \mathcal{M}_X(t) / e^{+(1-\delta)\mu} \\
&\leq e^{\mu(e^{+t-1})} / e^{+(1-\delta)\mu}
\end{align*}
\]

... (by which inequality?)

To get the best bound for \( \Pr(X \leq (1-\delta)\mu) \), we now choose \( t \) so as to minimize the term

\[ e^{\mu(e^{+t-1})} / e^{+(1-\delta)\mu} \]
Proof of (1)

By calculus, the best $t$ is equal to $\log_e (1-\delta)$

So, by substituting $t = \log_e (1-\delta)$ in the previous inequality, we get

$$\Pr(X \leq (1-\delta)\mu) \leq e^{\mu(-\delta)/(1-\delta)(1-\delta)\mu}$$

$$= \left(e^{-\delta}/(1-\delta)^{(1-\delta)}\right)^{\mu}$$
Proof of (2)

To prove (2), it is sufficient to show whenever $0 < \delta < 1$,

$$e^{-\delta} / (1-\delta)^{(1-\delta)} \leq e^{-\delta^2/2}$$

Equivalently, we want to show

$$- \delta - (1-\delta) \log_e (1-\delta) \leq -\delta^2/2$$

(this is obtained by taking log on both sides)

Let $g(\delta) = - \delta - (1-\delta) \log_e (1-\delta) + \delta^2/2$

Target: to show $g(\delta) \leq 0$
Proof of (2)

Then,

\[ g'(\delta) = -1 + \frac{1-\delta}{1-\delta} + \log_e (1-\delta) + \frac{2\delta}{2} \]

\[ = \log_e (1-\delta) + \delta \]

and

\[ g''(\delta) = -1/(1-\delta) + 1 \]

So, we see that

\[ g''(\delta) < 0 \quad \text{for} \ 0 \leq \delta < 1 \]
This implies that
\[ g'(\delta) \text{ is decreasing for } 0 \leq \delta < 1 \]

Next, from \[ g'(\delta) = - \log_e (1+\delta) + \frac{2\delta}{3}, \]
we see that \[ g'(0) = 0 \]
Together with the above, we conclude that
\[ g'(\delta) \leq 0 \text{ in the interval } [0,1) \]
Proof of (2)

Since \( g'(\delta) \leq 0 \) in the interval \([0,1)\),
\( g \) is decreasing in the interval \([0,1)\)

\( \Rightarrow g(0) \) is the maximum point of \( g \) in the interval \([0,1)\)

Thus, for \( 0 < \delta < 1 \),
\[
g(\delta) = -\delta - (1-\delta) \log_e (1-\delta) + \delta^2/2
\]
\[
\leq g(0) = 0
\]
This completes the proof of (2)
Useful Corollary

Corollary: Let $X_1, X_2, ..., X_n$ be independent Poisson trials such that $\Pr(X_i = 1) = p_i$.

Let $X = X_1 + X_2 + ... + X_n$ and $\mu = \mathbb{E}[X]$. Then, for all $0 < \delta < 1$,

$$ \Pr(|X - \mu| \geq \delta \mu) \leq 2e^{-\mu \delta^2/3} $$

How to prove?
Example: Coin Flip

Let \( X = \# \) heads in \( n \) fair coin flips

By Markov: \( \Pr( X \geq 3n/4 ) \leq 2/3 \)

By Chebyshev: \( \Pr( X \geq 3n/4 ) \leq 4/n \)

By Chernoff:

\[
\Pr(X \geq 3n/4) = \Pr(X \geq (1.5)\mu) \\
\leq e^{-\mu(0.5)^2/3} = e^{-n/24}
\]
In fact, w.h.p., #heads is around the mean:

\[
\Pr(|X - n/2| \geq (6n \log_e n)^{0.5}/2 )
\]

\[
= \Pr(|X - n/2| \geq ((6n \log_e n)^{0.5}/n)(n/2) )
\]

\[
\leq 2 \exp \left\{ -(1/3)(n/2)((6n \log_e n)/n^2) \right\}
\]

\[
= 2 \exp \{-\log_e n\} = 2/n
\]