Thm: (Chernoff's bound) Let X_i be independent random variables such that $\mathbb{E}[X_i]=0$ and $|X_i| \leq 1$ almost surely. Define $\sigma_i^2 = \mathbb{E}[X_i^2]$ and $\sigma^2 = \sum_{i=1}^{n} \sigma_i^2$. Then

$$\mathbb{P}\left(\sum_i X_i \geq t \right) \leq \max\left(e^{-t^2/4\sigma^2}, e^{-t/2} \right)$$

Note: Let $X = \sum_i X_i$. If applying Chebyshev's inequality:

$$\mathbb{P}(X > t) \leq \frac{\sigma^2}{t^2} \qquad \text{"weak"}$$

Proof of Chernoff's bound:

$$\mathbb{P}\left(\sum_i X_i \geq t\right) = \mathbb{P}\left(\lambda \sum_i X_i \geq \lambda t\right) \quad \text{for } \lambda \geq 0.$$

$$= \mathbb{P}\left(e^{\lambda \sum_i X_i} \geq e^{\lambda t}\right)$$

because e^x is monotone

$$\leq \frac{\mathbb{E}[e^{\lambda \sum_i X_i}]}{e^{\lambda t}} \quad \text{by Markov's}$$

$$= \frac{\prod_i \mathbb{E}[e^{\lambda X_i}]}{e^{\lambda t}}.$$

Now, for $x \in [0,1]$, we have

$$e^x \leq 1 + x + x^2 \qquad \text{(why?)}$$

$$\Rightarrow \quad \mathbb{E}[e^{\lambda X_i}] \leq 1 + \mathbb{E}[\lambda X_i] + \mathbb{E}[\lambda^2 X_i^2]$$

$$= 1 + \lambda^2 \mathbb{E}[X_i^2]$$

$$= 1 + \lambda^2 \sigma_i^2$$

Since $1 + x \leq e^x$,

$$1 + \lambda^2 \sigma_i^2 \leq e^{\lambda^2 \sigma_i^2}.$$

$$\therefore \quad \frac{\prod_i \mathbb{E}[e^{\lambda X_i}]}{e^{\lambda t}} \leq \frac{\prod_i \mathbb{E}[e^{\lambda^2 \sigma_i^2}]}{e^{\lambda t}}$$

$$= e^{\lambda^2 \sigma^2 - \lambda t}.$$

$\Rightarrow$ optimizing over $\lambda \in [0, 1]$, we get that

$$\lambda = \min\left(1, \frac{t}{2\sigma^2}\right).$$

which completes the proof.

* Other useful forms:

Chernoff's inequalities:

Let $X_1, X_2, \ldots, X_n$ be independent $\{0,1\}$-valued r.v's.

Each $X_i = \begin{cases} 1 & \text{with prob. } p_i \\ 0 & 1 - p_i \end{cases}$

Let $X = \sum_{i=1}^{n} X_i$ and $\mu = \mathbb{E}[X] = \sum_i p_i$. Then

$$\mathbb{P}(X \geq (1+\varepsilon)\mu) \leq e^{-\mu\varepsilon^2/4}$$

$$\mathbb{P}(X \leq (1-\varepsilon)\mu) \leq e^{-\mu\varepsilon^2/2}$$

or $\mathbb{P}(|X - \mu| \geq \varepsilon\mu) \leq 2e^{-\mu\varepsilon^2/4}.$

- Example application: 1) coin tossing.

Suppose we have a fair coin.

Let S_n = the numbers of heads from the first n tosse

$\Rightarrow$ $\mathbb{E}[S_n] = \frac{n}{2}$ and $Var[S_n] = \frac{n}{4}$.

Applying Chebyshov's inequality:

$$\mathbb{P}\left(\left| S_n - \overset{\mathbb{E}[S_n]}{\frac{n}{2}} \right| \geq \varepsilon \right) \leq \frac{n}{4\varepsilon^2}$$

$$\Rightarrow \quad \mathbb{P}\left(\left| \frac{S_n}{n} - \frac{1}{2} \right| \geq \frac{\varepsilon}{n} \right) \leq \frac{n}{4\varepsilon^2}.$$

For example, take $\varepsilon = \frac{n}{4}$.

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \frac{1}{2} \right| \geq \frac{1}{4} \right) \leq \frac{4}{n}.$$

But how about Chernoff's bound?

$$\mathbb{P}\left(\left| S_n - \frac{n}{2} \right| \geq \delta \frac{n}{2} \right) \leq 2e^{-\frac{n}{2}\delta^2/4}.$$

$$\Rightarrow \quad \mathbb{P}\left(\left| \frac{S_n}{n} - \frac{1}{2} \right| \geq \frac{\delta}{2} \right) \leq 2e^{-\frac{n\delta^2}{8}}.$$

Take $\delta = \frac{1}{2}$,

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \frac{1}{2} \right| \geq \frac{1}{4} \right) \leq 2e^{-\frac{n}{32}}.$$

Take $\delta = 2\sqrt{\ln n / n}$

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \frac{1}{2} \right| \geq \sqrt{\frac{\ln n}{n}} \right) \leq 2e^{-\frac{\ln n}{4}} = \frac{2}{n^{1/4}}.$$

2) Set balancing.

Given an $n \times m$ matrix A with entries in $\{0, 1\}$, look for a vector $\vec{b}$ with entries in $\{-1, 1\}$ that minimizes

$$\|A\vec{b}\|_\infty .$$

In designing statistical experiments, each column of the matrix A represents a subject and each row represents a feature. Goal: the vector $\vec{b}$ partitions the subject into two disjoint groups, so that each feature is roughly as balanced as possible between two groups.

randomized algo: choose entries of $\vec{b}$ with

$$\mathbb{P}(b_i = 1) = \mathbb{P}(b_i = -1) = 1/2$$

and b_i are independent.

Thm: For a random vector $\vec{b}$ with entries chosen independently and equal prob. from the set $\{-1, 1\}$,

$$\mathbb{P}\left(\|A\vec{b}\|_\infty \geq \sqrt{4m \ln n}\right) \leq \frac{2}{n}.$$

Pf: Exercise.