This lecture note closely follows the presentation of Chapter 1 and 2 of “Concentration,” by Colin McDiarmid (link)
Theorem (Chernoff Bound)

Let \( X_1, \ldots, X_n \) be independent binary random variables with \( \Pr[X_k = 1] = p \), for every \( 1 \leq k \leq n \). Let \( S_n = \sum X_k \). Then for any \( t \geq 0 \),

\[
\Pr [S_n - np \geq nt] \leq \exp \left( -2nt^2 \right)
\]
Proof of Chernoff Bound

\[ \Pr[S_n \geq np + nt] = \Pr[\exp(hS_n) \geq \exp(hn(p + t))], \text{ for any } h \geq 0 \]
\[ \leq \frac{\mathbb{E}[\exp(hS_n)]}{\exp(hn(p + t))} \]
\[ \leq \frac{u(h)}{\exp(hn(p + t))} \]

- Let \( u(h) \) be an upper bound on \( \mathbb{E}[\exp(hS_n)] \)
- Let \( h^* \) be the value of \( h > 0 \) that minimizes \( \frac{u(h)}{\exp(hn(p + t))} \)

\[ \Pr[S_n - np \geq nt] \leq \frac{u(h^*)}{\exp(h^*n(p + t))} \]
Bound on Expectation

$$E[\exp(hS_n)] = \prod \exp(hS_k)$$

$$= \prod (1 - p + p \exp(h))$$

$$= (1 - p + p \exp(h))^n =: u(h)$$

$$h^* = \arg\min_h \left( \frac{1 - p + p \exp(h)}{\exp(h(p + t))} \right)^n$$

Set $$\exp(h^*) = \frac{(p+t)(1-p)}{p(1-p-t)}$$ to get the bound
A Useful Lemma

**Lemma**

Let $X_1, \ldots, X_n$ be independent such that $0 \leq X_k \leq 1$ for each $k$. Let $\mu = \mathbb{E}[S_n]$, $p = \mu/n$ and $\bar{p} = 1 - p$. Then for any $0 < t < \bar{p}$

$$\Pr(S_n - np \geq nt) \leq \left( \left( \frac{p}{p + t} \right)^{p + t} \left( \frac{\bar{p}}{\bar{p} - t} \right)^{\bar{p} - t} \right)^n$$

- Let $p_k = \mathbb{E}[X_k]$
- $\mathbb{E}[\exp(hX_k)] \leq 1 - p_k + p_k \exp(h)$, using Jensen’s Inequality
- $\mathbb{E}[\exp(hS_n)] = \prod \mathbb{E}[\exp(hX_k)] \leq \prod (1 - p_k + p_k \exp(h)) \leq_{AM-GM} (1 - p + p \exp(h))^n =: u(h)$
Theorem (Hoeffding’s Bound)

Let $X_1, \ldots, X_n$ be independent such that $a_k \leq X_k \leq b_k$ for each $k$. Let $S_n = \sum X_k$ and $np = \mathbb{E}[S_n]$. Then, for any $t \geq 0$,

$$\Pr(S_n - np \geq nt) \leq \exp \left( -\frac{2n^2 t^2}{\sum (b_k - a_k)^2} \right)$$
Proof

Left as an exercise. Use the following lemma on the random variable \((X_k - \mathbb{E}[X_k])\) and apply AM-GM inequality:

**Lemma**

Let \(X\) be a random variable such that \(\mathbb{E}[X] = 0\) and \(a \leq X \leq b\). Then for any \(h > 0\),

\[
\mathbb{E}(\exp(hX)) \leq \exp(h^2(b - a)^2/8)
\]
Additional materials on the course website provide references to the following intuitions:

- Identical bounds also hold for $\Pr[\max S_k - kp \geq nt]$ (this uses Doob’s maximal inequality for submartingales)
- Identical bounds also hold for random variables with “slightly lesser” independence
- Bounds for $k$-wise independent also exit
- Concentration bound for hypergeometric distribution (sampling with replacement) is tighter
**Theorem**

Let \( \mathbf{X} = (X_1, \ldots, X_m) \) be a family of independent random variables with \( X_k \) taking values in \( \Omega_k \), for each \( k \). For a real valued function \( f \) defined on \( \prod \Omega_k \), the following holds:

\[
|f(x) - f(x')| \leq c_k,
\]

whenever \( x \) and \( x' \) differ only in the \( k \)-th coordinate. Let \( \mu = \mathbb{E}[f(\mathbf{X})] \). Then for any \( t \geq 0 \), we have:

\[
\Pr[f(\mathbf{X}) - \mu \geq nt] \leq \exp(-2n^2t^2/\sum c_k^2)
\]

Think: Concentration of longest common subsequence
Hamming Distance

- $d_H(x, y)$ is the number of coordinates where $x$ and $y$ differ
- $d_H(x, A)$ is the minimum distance $d_H(x, y)$, where $y \in A$

**Theorem**

Let $X = (X_1, \ldots, X_n)$ be a family of independent random variables with $X_k$ taking values in $\Omega_k$, for each $k$. Let $A$ be a subset of the product space $\prod \Omega_k$. Then for any $t \geq 0$,

$$\Pr[X \in A] \cdot \Pr[d_H(X, A) \geq nt] \leq \exp(-nt^2/2)$$
For $\alpha = (\alpha_1, \ldots, \alpha_n) \succeq \mathbf{0}$ and $\|\alpha\|_2 = 1$, define
\[
d_\alpha(x, y) = \sum_{k : x_k \neq y_k} \alpha_k
\]
d_\alpha(x, A) is the minimum distance $d_\alpha(x, y)$, where $y \in A$

**Theorem**

Let $X = (X_1, \ldots, X_n)$ be a family of independent random variables with $X_k$ taking values in $\Omega_k$, for each $k$. Let $A$ be a subset of the product space $\prod \Omega_k$. Then for any $t \geq 0$ and $\alpha \succeq \mathbf{0}$ and $\|\alpha\|_2 = 1$,
\[
\Pr[X \in A] \cdot \Pr[d_\alpha(X, A) \geq nt] \leq \exp\left(-\frac{n^2 t^2}{2}\right)
\]

- Idea is to use a “dense set $A$”
- Talagrand inequality will generalize this further