Randomness and Computation
or, “Randomized Algorithms”

Mary Cryan

School of Informatics
University of Edinburgh
Chernoff Bounds from the book

*Poisson trials* - sequence of Bernoulli variables $X_i$ with varying $p_i$s.

**Theorem (4.4)**

Let $X_1, \ldots, X_n$ be independent 0/1 Poisson trials such that

$\Pr[X_i = 1] = p_i$ for all $i \in [n]$. Let $X = \sum_{i=1}^n X_i$, and $\mu = \mathbb{E}[X]$. We have the following Chernoff bounds:

1. For any $\delta > 0$,

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

2. For any $0 < \delta \leq 1$,

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

3. For $R \geq 6\mu$,

   \[
   \Pr[X \geq R] \leq 2^{-R} .
   \]
Chernoff Bounds from the book (other direction)

Theorem (4.5)

Let $X_1, \ldots, X_n$ be independent 0/1 Poisson trials such that $\Pr[X_i = 1] = p_i$ for all $i \in [n]$. Let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}[X]$. For any $0 < \delta < 1$, we have the following Chernoff bounds:

1. $$\Pr[X \leq (1 - \delta)\mu] \leq \left( \frac{e^{-\delta}}{(1 - \delta)^{1-\delta}} \right)^{\mu};$$

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

- Proof is similar to Thm 4.4.
- Bound of 2. is slightly better than for the $\geq (1 + \delta)\mu$ bound.
- No 3. Why?
Concentration

Corollary (4.6)

Let $X_1, \ldots, X_n$ be independent 0/1 Poisson trials such that $\Pr[X_i = 1] = p_i$ for all $i \in [n]$. Let $X = \sum_{i=1}^n X_i$, and $\mu = E[X]$. Then for any $\delta, 0 < \delta < 1$,

$$\Pr[|X - \mu| \geq \delta \mu] \leq 2e^{-\mu \delta^2/3}.$$ 

- For almost all applications, we will want to work with a symmetric version like the Corollary.
- We “threw away” a bit in moving from the $(e^{\pm \delta} (1 \pm \delta)^{1 \pm \delta})^\mu$ versions, but they are tricky to work with.
Unbiased $+1/-1$ variables

In fact, for the case of unbiased variables, we can do even better than $2e^{-\mu \delta^2/3}$. We first (like the book) switch to $+1/-1$ variables.

**Theorem (4.7)**

Let $X_1, \ldots, X_n$ be independent random variables with
\[ \Pr[X_i = 1] = 1/2 = \Pr[X_i = -1] \] for all $i \in [n]$. Let $X = \sum_{k=1}^n X_k$. Note $\mu = E[X] = 0$. Then for any $a > 0$,
\[ \Pr[X \geq a] \leq e^{-a^2/2n}. \]

Proof is in the book.
(uses Taylor series expansions for $e^t, e^{-t}$).

Constant is just a bit better than with Theorem 4.6. I will do the details of this on the visualiser.
Unbiased 0/1 variables

Now consider $Y_1, \ldots, Y_n$ such that $\Pr[Y_i = 1] = 1/2$ for every $i \in [n]$. Define $X_i = 2Y_i - 1$ for every $i \in [n]$. Then

$$X_i = \begin{cases} 1 & |Y_i = 1 \\ -1 & |Y_i = 0 \end{cases}$$

Note also that for any $t \in \mathbb{Z}$, that

$$\sum_{i=1}^{n} Y_i = t \iff \sum_{i=1}^{n} X_i = 2t + n$$

Corollary (4.9, 4.10)

For $Y = \sum_{i=1}^{n} Y_i$, $X = \sum_{i=1}^{n} X_i$, we have

$$\Pr[Y \geq \frac{n}{2} + a] = \Pr[X \geq 2a] \leq e^{-2a^2/n}$$

$$\Pr[Y \leq \frac{n}{2} - a] = \Pr[X \leq -2a] \leq e^{-2a^2/n}$$

where the $\leq$ bounds come from applying Thm 4.7 to $X$ and to $-X$. 

RC (2017/18) – Lecture 8 – slide 6
Set Balancing for statistical experiments

We have an $n \times m$ binary matrix $A$ (entries from $\{0, 1\}$). We consider the value of

$$A \cdot \bar{b} = \bar{c},$$

when $\bar{b} \in \{-1, +1\}^m$ (note $\bar{c}$ will then be $n$-dimensional).

Goal is to find $\bar{b} \in \{-1, +1\}^m$ such that the value of

$$\|A \cdot \bar{b}\|_\infty = \max_{j=1}^n |c_j|$$

is minimized.

Solution: choose $\bar{b} \in \{-1, +1\}^m$ by generating $b_i$ independently and uniformly from $\{-1, +1\}$. We can show

**Theorem (4.11)**

*For $\bar{b}$ chosen uar from $\{-1, +1\}^m$,*

$$\Pr[\|A\bar{b}\|_\infty \geq \sqrt{4m \ln(n)}] \leq \frac{2}{n}.$$
Set Balancing for statistical experiments

We have an $n \times m$ binary matrix $A$ (entries from $\{0, 1\}$). We consider the value of

$$A \cdot \bar{b} = \bar{c},$$

when $\bar{b} \in \{-1, +1\}^m$ (note $\bar{c}$ will then be $n$-dimensional).

Goal is to find $\bar{b} \in \{-1, +1\}^m$ such that the value of

$$\|A \cdot \bar{b}\|_{\infty} = \max_{j=1}^n |c_j|$$

is minimized.

Solution: choose $\bar{b} \in \{-1, +1\}^m$ by generating $b_i$ independently and uniformly from $\{-1, +1\}$. We can show

Theorem (4.11)

For $\bar{b}$ chosen uar from $\{-1, +1\}^m$,

$$\Pr[\|A\bar{b}\|_{\infty} \geq \sqrt{4m \ln(n)}] \leq \frac{2}{n}.$$
Set Balancing for statistical experiments

We have an $n \times m$ binary matrix $A$ (entries from $\{0, 1\}$). We consider the value of

$$A \cdot \bar{b} = \bar{c},$$

when $\bar{b} \in \{-1, +1\}^m$ (note $\bar{c}$ will then be $n$-dimensional).

Goal is to find $\bar{b} \in \{-1, +1\}^m$ such that the value of

$$\|A \cdot \bar{b}\|_\infty = \max_{j=1}^n |c_j|$$

is minimized.

Solution: choose $\bar{b} \in \{-1, +1\}^m$ by generating $b_i$ independently and uniformly from $\{-1, +1\}$. We can show

**Theorem (4.11)**

For $\bar{b}$ chosen uar from $\{-1, +1\}^m$,

$$\Pr[\|A\bar{b}\|_\infty \geq \sqrt{4m \ln(n)}] \leq \frac{2}{n}.$$
Set Balancing for statistical experiments

We have an \( n \times m \) binary matrix \( A \) (entries from \{0, 1\}). We consider the value of
\[
A \cdot \bar{b} = \bar{c},
\]
when \( \bar{b} \in \{-1, +1\}^m \) (note \( \bar{c} \) will then be \( n \)-dimensional).

Goal is to find \( \bar{b} \in \{-1, +1\}^m \) such that the value of
\[
\|A \cdot \bar{b}\|_\infty = \max_{j=1}^n |c_j|
\]
is minimized.

Solution: choose \( \bar{b} \in \{-1, +1\}^m \) by generating \( b_i \) independently and uniformly from \{−1, +1\}. We can show

**Theorem (4.11)**

*For \( \bar{b} \) chosen uar from \( \{-1, +1\}^m \),*

\[
\Pr[\|A\bar{b}\|_\infty \geq \sqrt{4m \ln(n)}] \leq \frac{2}{n}.
\]
Set Balancing for statistical experiments (added post-lecture)

- \| \cdot \|_\infty is the absolute value of the largest entry of the tuple. We want to show that with high probability, every entry of \( A \cdot \bar{b} \) has absolute value \( \leq \sqrt{4m \ln(n)} \).

- There are \( n \) different entries of \( \bar{c} = A \cdot \bar{b} \); we will show that for each entry, we are “small enough” with probability \( \geq 1 - \frac{2}{n^2} \). Then Union Bound shows that \( \| A \cdot \bar{b} \|_\infty \) is bounded with prob \( \geq 1 - \frac{2}{n} \).

- For any row \( i \) of \( A \), there are some entries \( S_i, |S_i| \leq m \) which are non-0 (ie, 1). The absolute value of \( A_i \cdot \bar{b} \) is the (absolute) weighted sum of these 1s, randomly weighted by +1 or -1 ... so we have \( S_i \) random trials of unbiased +1/-1. Setting \( a = \sqrt{4m \ln(n)} \), Thm 4.7 says the probability we exceed this is at most

\[
2e^{-4m\ln(n)/2|S_i|} = 2n^{-2m/|S_i|} \leq \frac{2}{n^2},
\]

as required.

RC (2017/18) – Lecture 8 – slide 8
Set Balancing for statistical experiments (added post-lecture)

- $\|\cdot\|_{\infty}$ is the absolute value of the largest entry of the tuple. We want to show that with high probability, every entry of $A \cdot \bar{b}$ has absolute value $\leq \sqrt{4m \ln(n)}$.

- There are $n$ different entries of $\bar{c} = A \cdot \bar{b}$; we will show that for each entry, we are “small enough” with probability $\geq 1 - \frac{2}{n^2}$. Then Union Bound shows that $\|A \cdot \bar{b}\|_{\infty}$ is bounded with prob $\geq 1 - \frac{2}{n}$.

- For any row $i$ of $A$, there are some entries $S_i, |S_i| \leq m$ which are non-0 (ie, 1). The absolute value of $A_i \cdot \bar{b}$ is the (absolute) weighted sum of these 1s, randomly weighted by $+1$ or $-1$ … so we have $S_i$ random trials of unbiased $+1/-1$. Setting $a = \sqrt{4m \ln(n)}$, Thm 4.7 says the probability we exceed this is at most

$$2e^{-4m \ln(n)/2|S_i|} = 2n^{-2m/|S_i|} \leq \frac{2}{n^2},$$

as required.
Set Balancing for statistical experiments (added post-lecture)

▸ $\|\cdot\|_\infty$ is the absolute value of the largest entry of the tuple. We want to show that with high probability, *every entry* of $A \cdot \bar{b}$ has absolute value $\leq \sqrt{4m \ln(n)}$.

▸ There are $n$ different entries of $\bar{c} = A \cdot \bar{b}$; we will show that for each entry, we are “small enough” with probability $\geq 1 - \frac{2}{n^2}$. Then Union Bound shows that $\|A \cdot \bar{b}\|_\infty$ is bounded with prob $\geq 1 - \frac{2}{n}$.

▸ For any row $i$ of $A$, there are some entries $S_i, |S_i| \leq m$ which are non-0 (ie, 1). The absolute value of $A_i \cdot \bar{b}$ is the (absolute) weighted sum of these 1s, randomly weighted by $+1$ or $-1$ ... so we have $S_i$ random trials of unbiased $+1/-1$. Setting $a = \sqrt{4m \ln(n)}$, Thm 4.7 says the probability we exceed this is at most

$$2e^{−4m \ln(n)/2|S_i|} = 2n^{−2m/|S_i|} \leq \frac{2}{n^2},$$

as required.
More General Chernoff Bounds

Many many variations. My favourite general statement is McDiarmid’s presentation:

Theorem
Let $X_1, \ldots, X_n$ be independent random variables, $X_k$ taking values in a set $A_k$, for every $k \in [n]$. Suppose that the (measurable) function $f : \prod_{k=1}^{n} A_k \rightarrow \mathbb{R}$ satisfies

$$|f(\bar{x}) - f(\bar{x}')| \leq c_k$$

whenever $\bar{x}, \bar{x}'$ only differ in the $k$-th coordinate.

Let $Y$ be the random variable $f[X_1, \ldots, X_n]$. Then for any $t > 0$,

$$\Pr[|Y - E[Y]| \geq t] \leq 2 \exp \left[ -\frac{2t^2}{\sum_{k \in [n]} c_k^2} \right].$$
The rest of the course

Lects 9-10  The “birthday paradox” and (more generally) “balls in bins”.

Lects 11-12  The Probabilistic method, derandomization via Conditional expectation (bit more than half Chapter 6)

Lect 13  The Lovasz Local Lemma and its application to proving existence (6.7, 6.8)

Lects 14-15  Markov chain basics and application to 2SAT (7.1, 7.2)

Lects 16-17  The Monte Carlo method, DNF counting (some of Chapter 11)

Lects 18-19  Mixing time bounds for Markov chains (Chapter 11)
References

▶ Chapter 4 of “Probability and Computing"
▶ We don’t have time to cover the packet routing analysis of 4.5. It’s worth reading (but not examinable in the exam).