Probabilistic Methods in Combinatorics
Lecture 8

Linyuan Lu

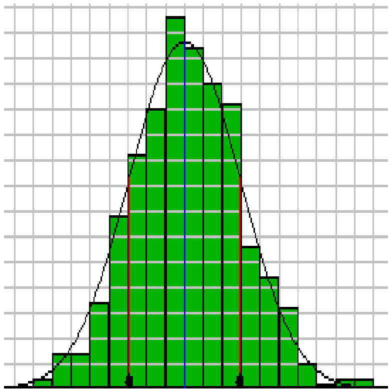
University of South Carolina

Mathematical Sciences Center at Tsinghua University

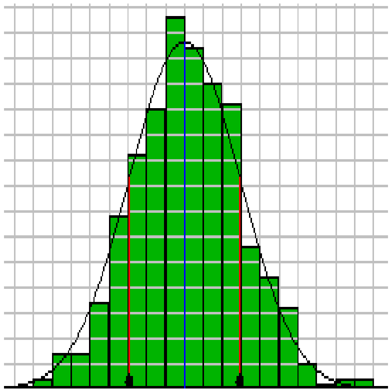
November 16, 2011 – December 30, 2011

Binomial distribution $B(n, p)$

- n is the number of variables.
- $0 < p < 1$ is the probability.



Binomial distribution $B(n,p)$

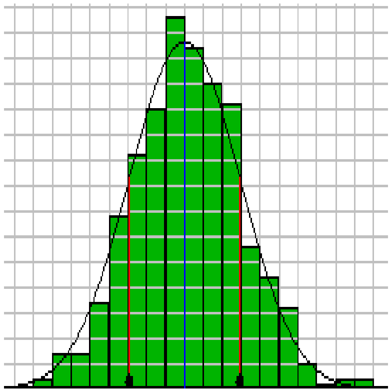


- n is the number of variables.
- $0 < p < 1$ is the probability.

$X = \sum_{i=1}^{n} X_i$. X_i, independent 0-1 random variables.

$$Pr(X_i = 1) = p, \quad Pr(X_i = 0) = 1 - p.$$

Binomial distribution $B(n, p)$



- n is the number of variables.
- $0 < p < 1$ is the probability.

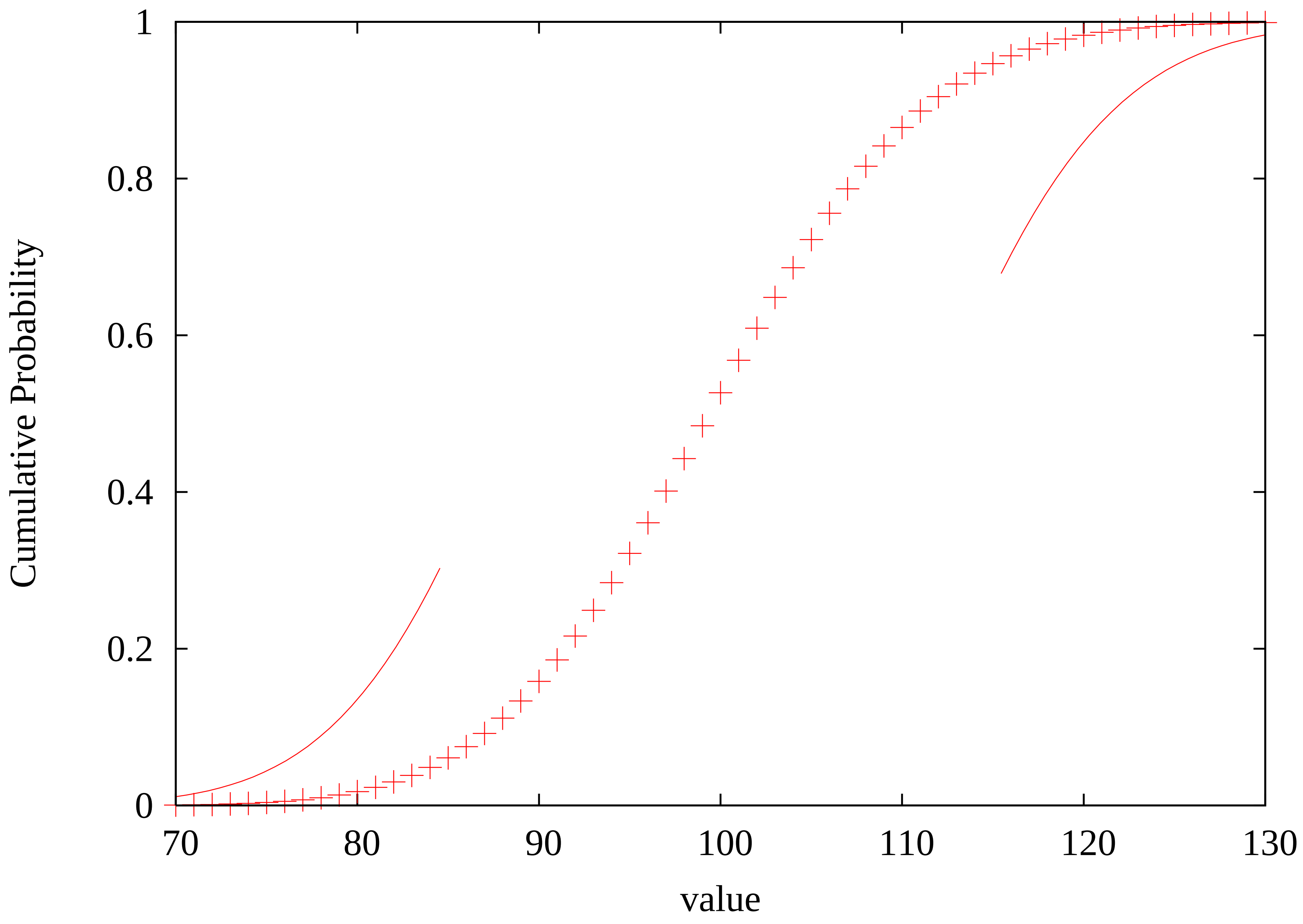
$X = \sum_{i=1}^{n} X_i$. X_i, independent 0-1 random variables.

$$Pr(X_i = 1) = p, \quad Pr(X_i = 0) = 1 - p.$$

$$B(n, p) \sim N(\mu, \sigma)$$

$N(\mu, \sigma)$: normal distribution with $\mu = np$ and $\sigma = \sqrt{np(1-p)}$.

Large deviations



Large deviation inequality I

Chernoff inequalities: Suppose $X = \sum_{i=1}^{n} X_i$, where X_i are independent 0-1 random variables with

$$\Pr(X_i = 1) = p_i, \quad \Pr(X_i = 0) = 1 - p_i.$$

Then we have

$$\begin{aligned} Pr(X < E(X) - \lambda) &\leq e^{-\frac{\lambda^2}{2E(X)}} \\ Pr(X > E(X) + \lambda) &\leq e^{-\frac{\lambda^2}{2(E(X)+\lambda/3)}} \end{aligned}$$

Large deviation inequality II

A weighted version of Chernoff's inequality:

- $X = \sum_{i=1}^{n} a_i X_i$

Large deviation inequality II

A weighted version of Chernoff's inequality:

- $X = \sum_{i=1}^{n} a_i X_i$
- $0 \leq a_1, \ldots, a_n \leq M$

Large deviation inequality II

A weighted version of Chernoff's inequality:

- $X = \sum_{i=1}^{n} a_i X_i$
- $0 \leq a_1, \ldots, a_n \leq M$
- $X_1, \ldots, X_n$: independent, 0-1, with $Pr(X_i = 1) = p_i$

Large deviation inequality II

A weighted version of Chernoff's inequality:

- $X = \sum_{i=1}^{n} a_i X_i$
- $0 \leq a_1, \ldots, a_n \leq M$
- $X_1, \ldots, X_n$: independent, 0-1, with $Pr(X_i = 1) = p_i$
- $E(X) = \sum_{i=1}^{n} a_i p_i$

Large deviation inequality II

A weighted version of Chernoff's inequality:

- $X = \sum_{i=1}^{n} a_i X_i$
- $0 \leq a_1, \ldots, a_n \leq M$
- $X_1, \ldots, X_n$: independent, 0-1, with $Pr(X_i = 1) = p_i$
- $E(X) = \sum_{i=1}^{n} a_i p_i$
- $\nu = \sum_{i=1}^{n} a_i^2 p_i$

Large deviation inequality II

A weighted version of Chernoff's inequality:

- $X = \sum_{i=1}^{n} a_i X_i$
- $0 \leq a_1, \ldots, a_n \leq M$
- $X_1, \ldots, X_n$: independent, 0-1, with $Pr(X_i = 1) = p_i$
- $E(X) = \sum_{i=1}^{n} a_i p_i$
- $\nu = \sum_{i=1}^{n} a_i^2 p_i$

Theorem [Chung,Lu] We have

$$Pr(X < E(X) - \lambda) \leq e^{-\lambda^2/2\nu} \quad (1)$$

$$Pr(X > E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\nu + M\lambda/3)}}. \quad (2)$$

Large deviation inequality III

Theorem [McDiarmid]: Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with $X_i - E(X_i) \leq M$ for a positive constant M. Let $X = \sum_{i=1}^n X_i$. Then

$$Pr(X - E(X) > \lambda) \leq e^{-\frac{\lambda^2}{2(Var(X) + M\lambda/3)}}.$$

Large deviation inequality III

Theorem [McDiarmid]: Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with $X_i - E(X_i) \leq M$ for a positive constant M. Let $X = \sum_{i=1}^{n} X_i$. Then

$$Pr(X - E(X) > \lambda) \leq e^{-\frac{\lambda^2}{2(Var(X)+M\lambda/3)}}.$$

Note: If $Pr(X_i = a_i) = p_i$ and $Pr(X_i = 0) = 1 - p_i$, then $\mathrm{Var}(X) = a_i^2 p_i (1 - p_i) \leq \nu$. Thus

$$Pr(X - E(X) > \lambda) \leq e^{-\frac{\lambda^2}{2(\nu+M\lambda/3)}}.$$

Large deviation inequality III

Theorem [McDiarmid]: Suppose $X_1, X_2, \ldots, X_n$ are independent random variables with $X_i - E(X_i) \leq M$ for a positive constant M. Let $X = \sum_{i=1}^{n} X_i$. Then

$$Pr(X - E(X) > \lambda) \leq e^{-\frac{\lambda^2}{2(Var(X)+M\lambda/3)}}.$$

Note: If $Pr(X_i = a_i) = p_i$ and $Pr(X_i = 0) = 1 - p_i$, then $\mathrm{Var}(X) = a_i^2 p_i(1 - p_i) \leq \nu$. Thus

$$Pr(X - E(X) > \lambda) \leq e^{-\frac{\lambda^2}{2(\nu+M\lambda/3)}}.$$

This theorem implies inequality of upper tail in previous Theorem.

Large deviation inequality IV

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \leq M$, for $1 \leq i \leq n$. Let $X = \sum_{i=1}^n X_i$ and $\|X\| = \sqrt{\sum_{i=1}^n \mathrm{E}(X_i^2)}$. Then we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\|X\|^2 + M\lambda/3)}}.$$

Large deviation inequality IV

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \leq M$, for $1 \leq i \leq n$. Let $X = \sum_{i=1}^{n} X_i$ and $\|X\| = \sqrt{\sum_{i=1}^{n} \mathrm{E}(X_i^2)}$. Then we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\|X\|^2 + M\lambda/3)}}.$$

This theorem implies McDiarmid's Theorem.

Large deviation inequality IV

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \leq M$, for $1 \leq i \leq n$. Let $X = \sum_{i=1}^{n} X_i$ and $\|X\| = \sqrt{\sum_{i=1}^{n} \mathrm{E}(X_i^2)}$. Then we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\|X\|^2 + M\lambda/3)}}.$$

This theorem implies McDiarmid's Theorem.

$$\text{Let } X_i' = X_i - \mathrm{E}(X_i), \text{ and } X' = X - \mathrm{E}(X).$$

$$X - \mathrm{E}(X) = X' - \mathrm{E}(X')$$

$$\|X'\|^2 = \sum_{i=1}^{n} \mathrm{E}(X_i'^2) = \mathrm{Var}(X).$$

Lower tail

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \geq 0$, for $1 \leq i \leq n$. Let $X = \sum_{i=1}^{n} X_i$ and $\|X\| = \sqrt{\sum_{i=1}^{n} \mathrm{E}(X_i^2)}$. Then we have

$$\Pr(X \leq E(X) - \lambda) \leq e^{-\frac{\lambda^2}{2\|X\|^2}}.$$

Lower tail

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \geq 0$, for $1 \leq i \leq n$. Let $X = \sum_{i=1}^n X_i$ and $\|X\| = \sqrt{\sum_{i=1}^n \mathrm{E}(X_i^2)}$. Then we have

$$\Pr(X \leq E(X) - \lambda) \leq e^{-\frac{\lambda^2}{2\|X\|^2}}.$$

Proof: Let $X_i' = -X_i$ and $X' = -X$. Applying the upper tail to X' with $M = 0$, we get

$$\begin{aligned}\Pr(X \leq E(X) - \lambda) &= \Pr(X' \geq E(X') + \lambda) \\ &\leq e^{-\frac{\lambda^2}{2\|X'\|^2}} = e^{-\frac{\lambda^2}{2\|X\|^2}}.\end{aligned}$$

A special function

$$g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} = \frac{2(e^y - 1 - y)}{y^2}.$$

A special function

$$g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} = \frac{2(e^y - 1 - y)}{y^2}.$$

Facts:

- $g(0) = 1.$

A special function

$$g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} = \frac{2(e^y - 1 - y)}{y^2}.$$

Facts:

- $g(0) = 1$.
- $g(y) \le 1$, for $y < 0$.

A special function

$$g(y) = 2\sum_{k=2}^{\infty}\frac{y^{k-2}}{k!} = \frac{2(e^y - 1 - y)}{y^2}.$$

Facts:

- $g(0) = 1.$
- $g(y) \leq 1$, for $y < 0.$
- $g(y)$ is monotone increasing, for $y \geq 0.$

A special function

$$g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} = \frac{2(e^y - 1 - y)}{y^2}.$$

Facts:

- $g(0) = 1.$
- $g(y) \leq 1$, for $y < 0.$
- $g(y)$ is monotone increasing, for $y \geq 0.$
- For $y < 3$, we have

$$g(y) = 2\sum_{k=2}^{\infty} \frac{y^{k-2}}{k!} \leq \sum_{k=2}^{\infty} \frac{y^{k-2}}{3^{k-2}} = \frac{1}{1 - y/3}.$$

Proof of upper tail

$$\mathrm{E}(e^{tX}) = \prod_{i=1}^{n} \mathrm{E}(e^{tX_i})$$

Proof of upper tail

$$\begin{aligned}
\mathrm{E}(e^{tX}) &= \prod_{i=1}^{n} \mathrm{E}(e^{tX_i}) \\
&= \prod_{i=1}^{n} \mathrm{E}\Big(\sum_{k=0}^{\infty} \frac{t^k X_i^k}{k!}\Big)
\end{aligned}$$

Proof of upper tail

$$\begin{aligned} E(e^{tX}) &= \prod_{i=1}^{n} E(e^{tX_i}) \\ &= \prod_{i=1}^{n} E(\sum_{k=0}^{\infty} \frac{t^k X_i^k}{k!}) \\ &= \prod_{i=1}^{n} E(1 + tE(X_i) + \frac{1}{2} t^2 X_i^2 g(tX_i)) \end{aligned}$$

Proof of upper tail

$$
\begin{aligned}
\mathrm{E}(e^{tX}) &= \prod_{i=1}^{n} \mathrm{E}(e^{tX_i}) \\
&= \prod_{i=1}^{n} \mathrm{E}\Big(\sum_{k=0}^{\infty} \frac{t^k X_i^k}{k!}\Big) \\
&= \prod_{i=1}^{n} \mathrm{E}\Big(1 + t\mathrm{E}(X_i) + \frac{1}{2} t^2 X_i^2 g(tX_i)\Big) \\
&\leq \prod_{i=1}^{n} \Big(1 + t\mathrm{E}(X_i) + \frac{1}{2} t^2 \mathrm{E}(X_i^2) g(tM)\Big)
\end{aligned}
$$

Proof of upper tail

$$
\begin{aligned}
\mathrm{E}(e^{tX}) &= \prod_{i=1}^{n} \mathrm{E}(e^{tX_i}) \\
&= \prod_{i=1}^{n} \mathrm{E}(\sum_{k=0}^{\infty} \frac{t^k X_i^k}{k!}) \\
&= \prod_{i=1}^{n} \mathrm{E}(1 + t\mathrm{E}(X_i) + \frac{1}{2} t^2 X_i^2 g(tX_i)) \\
&\leq \prod_{i=1}^{n} (1 + t\mathrm{E}(X_i) + \frac{1}{2} t^2 \mathrm{E}(X_i^2) g(tM)) \\
&\leq \prod_{i=1}^{n} e^{t\mathrm{E}(X_i) + \frac{1}{2} t^2 \mathrm{E}(X_i^2) g(tM)}
\end{aligned}
$$

Proof of upper tail

$$
\begin{aligned}
\mathrm{E}(e^{tX}) &= \prod_{i=1}^{n} \mathrm{E}(e^{tX_i}) \\
&= \prod_{i=1}^{n} \mathrm{E}(\sum_{k=0}^{\infty} \frac{t^k X_i^k}{k!}) \\
&= \prod_{i=1}^{n} \mathrm{E}(1 + t\mathrm{E}(X_i) + \frac{1}{2} t^2 X_i^2 g(tX_i)) \\
&\leq \prod_{i=1}^{n} (1 + t\mathrm{E}(X_i) + \frac{1}{2} t^2 \mathrm{E}(X_i^2) g(tM)) \\
&\leq \prod_{i=1}^{n} e^{t\mathrm{E}(X_i) + \frac{1}{2} t^2 \mathrm{E}(X_i^2) g(tM)} \\
&= e^{t\mathrm{E}(X) + \frac{1}{2} t^2 g(tM) \|X\|^2}.
\end{aligned}
$$

Continuous

Hence, for t satisfying $tM < 3$, we have

Continuous

Hence, for t satisfying $tM < 3$, we have

$$\Pr(X \geq \mathrm{E}(X) + \lambda) \;=\; \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda})$$

Continuous

Hence, for t satisfying $tM < 3$, we have

$$\begin{aligned} \Pr(X \geq \mathrm{E}(X) + \lambda) &= \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda}) \\ &\leq e^{-t\mathrm{E}(X)-t\lambda}\mathrm{E}(e^{tX}) \end{aligned}$$

Continuous

Hence, for t satisfying $tM < 3$, we have

$$\begin{aligned}\Pr(X \geq \mathrm{E}(X) + \lambda) &= \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda}) \\ &\leq e^{-t\mathrm{E}(X)-t\lambda}\mathrm{E}(e^{tX}) \\ &\leq e^{-t\lambda+\frac{1}{2}t^2 g(tM)\|X\|^2}\end{aligned}$$

Continuous

Hence, for t satisfying $tM < 3$, we have

$$\begin{aligned}
\Pr(X \geq \mathrm{E}(X) + \lambda) &= \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda}) \\
&\leq e^{-t\mathrm{E}(X)-t\lambda}\mathrm{E}(e^{tX}) \\
&\leq e^{-t\lambda+\frac{1}{2}t^2 g(tM)\|X\|^2} \\
&\leq e^{-t\lambda+\frac{1}{2}t^2\|X\|^2\frac{1}{1-tM/3}}.
\end{aligned}$$

Continuous

Hence, for t satisfying $tM < 3$, we have

$$\begin{aligned}
\Pr(X \geq \mathrm{E}(X) + \lambda) &= \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda}) \\
&\leq e^{-t\mathrm{E}(X)-t\lambda}\mathrm{E}(e^{tX}) \\
&\leq e^{-t\lambda+\frac{1}{2}t^2 g(tM)\|X\|^2} \\
&\leq e^{-t\lambda+\frac{1}{2}t^2\|X\|^2\frac{1}{1-tM/3}}.
\end{aligned}$$

Choose $t = \frac{\lambda}{\|X\|^2+M\lambda/3}$. We have $1 - \frac{Mt}{3} = \frac{\|X\|^2}{\|X\|^2+M\lambda/3}$.

Continuous

Hence, for t satisfying $tM < 3$, we have

$$
\begin{aligned}
\Pr(X \geq \mathrm{E}(X) + \lambda) &= \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda}) \\
&\leq e^{-t\mathrm{E}(X)-t\lambda}\mathrm{E}(e^{tX}) \\
&\leq e^{-t\lambda+\frac{1}{2}t^2 g(tM)\|X\|^2} \\
&\leq e^{-t\lambda+\frac{1}{2}t^2\|X\|^2\frac{1}{1-tM/3}}.
\end{aligned}
$$

Choose $t = \frac{\lambda}{\|X\|^2+M\lambda/3}$. We have $1 - \frac{Mt}{3} = \frac{\|X\|^2}{\|X\|^2+M\lambda/3}$.

$$
Pr(X > E(X) + \lambda) \leq e^{-t\lambda+t^2\|X\|^2\frac{1}{2(1-Mt/3)}}
$$

Continuous

Hence, for t satisfying $tM < 3$, we have

$$\begin{aligned}
\Pr(X \geq \mathrm{E}(X) + \lambda) &= \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda}) \\
&\leq e^{-t\mathrm{E}(X)-t\lambda}\mathrm{E}(e^{tX}) \\
&\leq e^{-t\lambda+\frac{1}{2}t^2 g(tM)\|X\|^2} \\
&\leq e^{-t\lambda+\frac{1}{2}t^2\|X\|^2\frac{1}{1-tM/3}}.
\end{aligned}$$

Choose $t = \frac{\lambda}{\|X\|^2+M\lambda/3}$. We have $1 - \frac{Mt}{3} = \frac{\|X\|^2}{\|X\|^2+M\lambda/3}$.

$$\begin{aligned}
Pr(X > E(X) + \lambda) &\leq e^{-t\lambda+t^2\|X\|^2\frac{1}{2(1-Mt/3)}} \\
&= e^{-\frac{\lambda^2}{\|X\|^2+M\lambda/3}+\frac{\lambda^2}{(\|X\|^2+M\lambda/3)^2}\|X\|^2\frac{\|X\|^2+M\lambda/3}{2\|X\|^2}}
\end{aligned}$$

Continuous

Hence, for t satisfying $tM < 3$, we have

$$\begin{aligned}
\Pr(X \geq \mathrm{E}(X) + \lambda) &= \Pr(e^{tX} \leq e^{t\mathrm{E}(X)+t\lambda}) \\
&\leq e^{-t\mathrm{E}(X)-t\lambda}\mathrm{E}(e^{tX}) \\
&\leq e^{-t\lambda+\frac{1}{2}t^2 g(tM)\|X\|^2} \\
&\leq e^{-t\lambda+\frac{1}{2}t^2\|X\|^2 \frac{1}{1-tM/3}}.
\end{aligned}$$

Choose $t = \frac{\lambda}{\|X\|^2+M\lambda/3}$. We have $1 - \frac{Mt}{3} = \frac{\|X\|^2}{\|X\|^2+M\lambda/3}$.

$$\begin{aligned}
Pr(X > E(X) + \lambda) &\leq e^{-t\lambda+t^2\|X\|^2 \frac{1}{2(1-Mt/3)}} \\
&= e^{-\frac{\lambda^2}{\|X\|^2+M\lambda/3} + \frac{\lambda^2}{(\|X\|^2+M\lambda/3)^2}\|X\|^2 \frac{\|X\|^2+M\lambda/3}{2\|X\|^2}} \\
&= e^{-\frac{\lambda^2}{2(\|X\|^2+M\lambda/3)}}.
\end{aligned}$$

□

Large deviation inequality V

Theorem [Chung, Lu] Let X_i denote independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + a_i + M$, for $1 \leq i \leq n$. For, $X = \sum_{i=1}^{n} X_i$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=1}^{n} a_i^2 + M\lambda/3)}}.$$

Proof: Let $X_i' = X_i - \mathrm{E}(X_i) - a_i$ and $X' = \sum_{i=1}^{n} X_i'$. We claim

Large deviation inequality V

Theorem [Chung, Lu] Let X_i denote independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + a_i + M$, for $1 \leq i \leq n$. For, $X = \sum_{i=1}^{n} X_i$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=1}^{n} a_i^2 + M\lambda/3)}}.$$

Proof: Let $X_i' = X_i - \mathrm{E}(X_i) - a_i$ and $X' = \sum_{i=1}^{n} X_i'$. We claim

- $X_i' \leq M$ for $1 \leq i \leq n$.

Large deviation inequality V

Theorem [Chung, Lu] Let X_i denote independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + a_i + M$, for $1 \leq i \leq n$. For, $X = \sum_{i=1}^{n} X_i$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=1}^{n} a_i^2 + M\lambda/3)}}.$$

Proof: Let $X_i' = X_i - \mathrm{E}(X_i) - a_i$ and $X' = \sum_{i=1}^{n} X_i'$. We claim

- $X_i' \leq M$ for $1 \leq i \leq n$.
- $X' - \mathrm{E}(X') = X - \mathrm{E}(X)$.

Large deviation inequality V

Theorem [Chung, Lu] Let X_i denote independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + a_i + M$, for $1 \leq i \leq n$. For, $X = \sum_{i=1}^n X_i$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=1}^n a_i^2 + M\lambda/3)}}.$$

Proof: Let $X_i' = X_i - \mathrm{E}(X_i) - a_i$ and $X' = \sum_{i=1}^n X_i'$. We claim

- $X_i' \leq M$ for $1 \leq i \leq n$.
- $X' - \mathrm{E}(X') = X - \mathrm{E}(X)$.
- $\|X'\|^2 = \mathrm{Var}(X) + \sum_{i=1}^n a_i^2$.

Large deviation inequality V

Theorem [Chung, Lu] Let X_i denote independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + a_i + M$, for $1 \leq i \leq n$. For, $X = \sum_{i=1}^{n} X_i$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=1}^{n} a_i^2 + M\lambda/3)}}.$$

Proof: Let $X_i' = X_i - \mathrm{E}(X_i) - a_i$ and $X' = \sum_{i=1}^{n} X_i'$. We claim

- $X_i' \leq M$ for $1 \leq i \leq n$.
- $X' - \mathrm{E}(X') = X - \mathrm{E}(X)$.
- $\|X'\|^2 = \mathrm{Var}(X) + \sum_{i=1}^{n} a_i^2$.

continue

$$\Pr(X \geq E(X) + \lambda) \;=\; \Pr(X' \geq E(X') + \lambda)$$

continue

$$\begin{aligned}\Pr(X \geq E(X) + \lambda) &= \Pr(X' \geq E(X') + \lambda) \\ &\leq e^{-\frac{\lambda^2}{2(\|X'\|^2 + M\lambda/3)}}\end{aligned}$$

continue

$$\begin{aligned}\Pr(X \geq E(X) + \lambda) &= \Pr(X' \geq E(X') + \lambda) \\ &\leq e^{-\frac{\lambda^2}{2(\|X'\|^2 + M\lambda/3)}} \\ &= e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=1}^n a_i^2 + M\lambda/3)}}.\end{aligned}$$

It remains to verify

$$X' - \mathrm{E}(X') = X - \mathrm{E}(X).$$

$$\|X'\|^2 = \mathrm{Var}(X) + \sum_{i=1}^n a_i^2.$$

□

Large deviation inequality VI

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + M_i$, for $0 \leq i \leq n$. We order X_i's so that M_i are in an increasing order. Let $X = \sum_{i=1}^n X_i$. Then for any $1 \leq k \leq n$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X)+\sum_{i=k}^n (M_i - M_k)^2 + M_k\lambda/3)}}.$$

Large deviation inequality VI

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \le \mathrm{E}(X_i) + M_i$, for $0 \le i \le n$. We order X_i's so that M_i are in an increasing order. Let $X = \sum_{i=1}^n X_i$. Then for any $1 \le k \le n$, we have

$$\Pr(X \ge E(X) + \lambda) \le e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=k}^n (M_i - M_k)^2 + M_k \lambda/3)}}.$$

Compared with McDiarmid's inequality

Large deviation inequality VI

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + M_i$, for $0 \leq i \leq n$. We order X_i's so that M_i are in an increasing order. Let $X = \sum_{i=1}^{n} X_i$. Then for any $1 \leq k \leq n$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=k}^{n}(M_i - M_k)^2 + M_k\lambda/3)}}.$$

Compared with McDiarmid's inequality

- M is replaced by M_k.

Large deviation inequality VI

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + M_i$, for $0 \leq i \leq n$. We order X_i's so that M_i are in an increasing order. Let $X = \sum_{i=1}^{n} X_i$. Then for any $1 \leq k \leq n$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=k}^{n}(M_i - M_k)^2 + M_k\lambda/3)}}.$$

Compared with McDiarmid's inequality

- M is replaced by M_k.
- Additional cost $\sum_{i=k}^{n}(M_i - M_k)^2$.

Large deviation inequality VI

Theorem [Chung, Lu] Suppose X_i are independent random variables satisfying $X_i \leq \mathrm{E}(X_i) + M_i$, for $0 \leq i \leq n$. We order X_i's so that M_i are in an increasing order. Let $X = \sum_{i=1}^n X_i$. Then for any $1 \leq k \leq n$, we have

$$\Pr(X \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2(\mathrm{Var}(X) + \sum_{i=k}^n (M_i - M_k)^2 + M_k \lambda/3)}}.$$

Compared with McDiarmid's inequality

- M is replaced by M_k.
- Additional cost $\sum_{i=k}^n (M_i - M_k)^2$.
- McDiarmid's inequality is a special case with $k = n$.

Proof

For fixed k, we choose $M = M_k$ and

$$a_i = \begin{cases} 0 & \text{if } 1 \leq i \leq k \\ M_i - M_k & \text{if } k \leq i \leq n \end{cases}$$

Proof

For fixed k, we choose $M = M_k$ and

$$a_i = \begin{cases} 0 & \text{if } 1 \le i \le k \\ M_i - M_k & \text{if } k \le i \le n \end{cases}$$

$$X_i - \mathrm{E}(X_i) \le M_i \le a_i + M_k. \quad \text{for } 1 \le k \le n.$$

Proof

For fixed k, we choose $M = M_k$ and

$$a_i = \begin{cases} 0 & \text{if } 1 \le i \le k \\ M_i - M_k & \text{if } k \le i \le n \end{cases}$$

$$X_i - \mathrm{E}(X_i) \le M_i \le a_i + M_k. \quad \text{for } 1 \le k \le n.$$

$$\sum_{i=1}^{n} a_i^2 = \sum_{i=k}^{n} (M_i - M_k)^2.$$

Apply previous theorem with these a_i's. □

An application

Example: Consider the sum $X = \sum_{i=1}^{n} X_i$.

An application

Example: Consider the sum $X = \sum_{i=1}^{n} X_i$.

- $X_1, X_2, \ldots, X_n$: independent random variables.

An application

Example: Consider the sum $X = \sum_{i=1}^{n} X_i$.

- $X_1, X_2, \ldots, X_n$: independent random variables.
- For $1 \leq i \leq n-1$, we have

$$\Pr(X_i = 0) = 1 - p \quad \text{and} \quad \Pr(X_i = 1) = p.$$

An application

Example: Consider the sum $X = \sum_{i=1}^{n} X_i$.

- $X_1, X_2, \ldots, X_n$: independent random variables.
- For $1 \leq i \leq n-1$, we have

$$\Pr(X_i = 0) = 1 - p \quad \text{and} \quad \Pr(X_i = 1) = p.$$

- X_n is special.

$$\Pr(X_n = 0) = 1 - p \quad \text{and} \quad \Pr(X_n = \sqrt{n}) = p.$$

An application

Example: Consider the sum $X = \sum_{i=1}^{n} X_i$.

- $X_1, X_2, \ldots, X_n$: independent random variables.
- For $1 \leq i \leq n-1$, we have

$$\Pr(X_i = 0) = 1 - p \quad \text{and} \quad \Pr(X_i = 1) = p.$$

- X_n is special.

$$\Pr(X_n = 0) = 1 - p \quad \text{and} \quad \Pr(X_n = \sqrt{n}) = p.$$

Expectation and Variance

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_{i=1}^{n} \mathrm{E}(X_i) \\
&= (n-1)p + \sqrt{n}p.
\end{aligned}
$$

Expectation and Variance

$$
\begin{aligned}
\mathrm{E}(X) &= \sum_{i=1}^{n} \mathrm{E}(X_i) \\
&= (n-1)p + \sqrt{n}p.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}(X) &= \sum_{i=1}^{n} \mathrm{Var}(X_i) \\
&= (n-1)p(1-p) + np(1-p) \\
&= (2n-1)p(1-p).
\end{aligned}
$$

Comparison

Applying McDiarmid's Theorem

Comparison

Applying McDiarmid's Theorem

- $M = (1 - p)\sqrt{n}$

Comparison

Applying McDiarmid's Theorem

- $M = (1-p)\sqrt{n}$

We have

$$\Pr(X \geq \mathrm{E}(X) + \lambda) \leq e^{-\frac{\lambda^2}{2((2n-1)p(1-p)+(1-p)\sqrt{n}\lambda/3)}}.$$

Comparison

Applying McDiarmid's Theorem

- $M = (1-p)\sqrt{n}$

We have

$$\Pr(X \geq \mathrm{E}(X) + \lambda) \leq e^{-\frac{\lambda^2}{2((2n-1)p(1-p)+(1-p)\sqrt{n}\lambda/3)}}.$$

In particular, for constant $p \in (0,1)$ and $\lambda = \Theta(n^{\frac{1}{2}+\epsilon})$, we have

$$\Pr(X \geq \mathrm{E}(X) + \lambda) \leq e^{-\Theta(n^{\epsilon})}.$$

Comparison

Applying last Theorem

Comparison

Applying last Theorem

- $M_1 = \ldots = M_{n-1} = (1-p)$
- $M_n = \sqrt{n}(1-p)$

Comparison

Applying last Theorem

- $M_1 = \ldots = M_{n-1} = (1-p)$
- $M_n = \sqrt{n}(1-p)$

We choose $k = n-1$,

$$\mathrm{Var}(X) + (M_n - M_{n-1})^2 \leq (1-p^2)n.$$

Comparison

Applying last Theorem

- $M_1 = \ldots = M_{n-1} = (1-p)$
- $M_n = \sqrt{n}(1-p)$

We choose $k = n-1$,

$$\mathrm{Var}(X) + (M_n - M_{n-1})^2 \leq (1-p^2)n.$$

$$\Pr(X_i \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2((1-p^2)n + (1-p)^2\lambda/3)}}.$$

Comparison

Applying last Theorem

- $M_1 = \ldots = M_{n-1} = (1-p)$
- $M_n = \sqrt{n}(1-p)$

We choose $k = n-1$,

$$\mathrm{Var}(X) + (M_n - M_{n-1})^2 \leq (1-p^2)n.$$

$$\Pr(X_i \geq E(X) + \lambda) \leq e^{-\frac{\lambda^2}{2((1-p^2)n + (1-p)^2\lambda/3)}}.$$

For constant $p \in (0,1)$ and $\lambda = \Theta(n^{\frac{1}{2}+\epsilon})$, we have

$$\Pr(X \geq \mathrm{E}(X) + \lambda) \leq e^{-\Theta(n^{2\epsilon})}.$$

Reference

- C. McDiarmid. Concentration. In Probabilistic Methods for Algorithmic Discrete Mathematics, edited by M. Habib, C. McDiarmid, J. Ramirez- Alfonsin, and B. Reed, pp. 195248, Algorithms and Combinatorics 16. Berlin: Springer, 1998.
- Chung and Lu, Concentration inequalities and martingale inequalities — a survey, Internet Mathematics, 3 (2006), No. 1, 79-127.
- Chung and Lu, Complex Graphs and Networks, (2006) published by AMS, ISBN-10: 0-8218-3657-9, ISBN-13: 978-0-8218-3657-6.