Concentration Inequalities

Lecturer: Dr. Hong-Gwa Yeh
Department of Mathematics
National Central University
hgyeh@math.ncu.edu.tw

A Simple Start

Thm (Markov's Inequality)

For $t > 0$, we have

$$P(|X| \geq t) \leq \frac{E|X|}{t}$$

pf:

$$\mathcal{E}|X| = \int_{\Omega} |X|\, dP \geq \int_{|X| \geq t} |X|\, dP$$

$$\geq t\, P(|X| \geq t)$$

QED

Chernoff's Inequality

Thm: $X_1, X_2, \cdots, X_n \overset{\text{i.i.d.}}{\sim} B(1,p)$.

Let $\lambda = np$ and $S_n = \sum_{i=1}^{n} X_i$.

Then, for $0 \le t \le n-\lambda$,

$$P\left\{ S_n \ge ES_n + t \right\} \le \left(\frac{\lambda}{\lambda+t}\right)^{\lambda+t} \left(\frac{n-\lambda}{n-\lambda-t}\right)^{n-\lambda-t}.$$

pf: For $u \geq 0$

$$\Pr\{S_n \geq ES_n + t\} \leq e^{-u(\lambda+t)} \prod_{i=1}^{n} Ee^{uX_i}$$

$$= e^{-u(\lambda+t)} (1-p+pe^{u})^{n} \quad \bigstar$$

let $\quad f(x) = x^{-(\lambda+t)} (1-p+px)^{n}$

$$f'(x) = -(\lambda+t)\, x^{-(\lambda+t+1)} (1-p+px)^{n} + np\, x^{-(\lambda+t)} (1-p+px)^{n-1}$$

and $\quad f'(x) = 0 \Rightarrow \hat{x} = \dfrac{(1-p)(\lambda+t)}{(n-\lambda-t)p}$

$f(x)$ attains its minimum at $\hat{x}$. assume $n-\lambda-t > 0$

i.e. $\bigstar$ attains its minimum at $e^{u} = \hat{x}$ "

This yields

$$P\{S_n \geq ES_n + t\} \leq \left(\frac{(\lambda+t)(1-p)}{(n-\lambda-t)\,p}\right)^{-(\lambda+t)} \left(1-p+\frac{(1-p)(\lambda+t)}{(n-\lambda-t)}\right)^{n}$$

$$= \left(\frac{\lambda+t}{n-\lambda-t}\right)^{-(\lambda+t)} \left(\frac{1-p}{p}\right)^{-(\lambda+t)} (1-p)^{n} \left(\frac{n}{n-\lambda-t}\right)^{n}$$

Next we use the fact i.e. $p=\frac{\lambda}{n}$ to get

$$\Pr\{S_n \geq ES_n + t\} \leq \left(\frac{\lambda}{\lambda+t}\right)^{\lambda+t} \left(\frac{n-\lambda}{n-\lambda-t}\right)^{n-\lambda-t}$$

as $0 \leq t \leq n-\lambda$.

QED

Chernoff's bounding technique

If s is an arbitrary positive number then for any r.v. X and any $t>0$,

$$P(X \geq t) = P(e^{sX} \geq e^{st}) \leq \frac{E e^{sX}}{e^{st}}.$$

In chernoff's method, we find $s>0$ that makes the upper bound small.

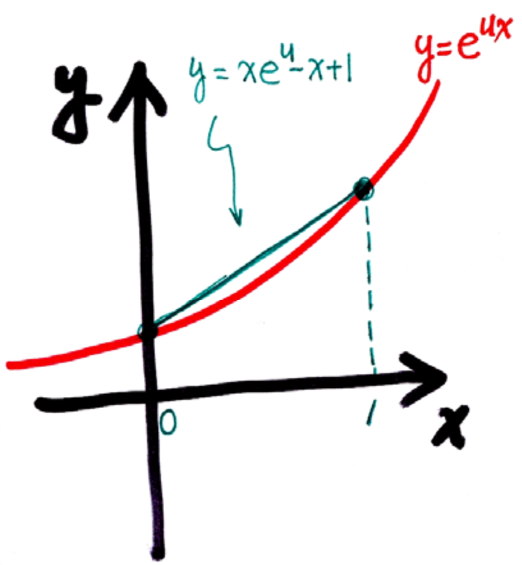
Extensions of Chernoff's Ineq. can be derived from....

Lemma*: Let $X_1, \cdots, X_n$ be independent with $0 \le X_i \le 1$ for each i. Let $p = \frac{\mathcal{E}S_n}{n}$, where $S_n = \sum_{i=1}^n X_i$.

Then for any $0 \le t < 1-p$,

$$P(S_n - \mathcal{E}S_n \ge nt) \le \left(\left(\frac{p}{p+t}\right)^{p+t} \left(\frac{1-p}{1-p-t}\right)^{1-p-t} \right)^n$$

pf: $P(S_n \ge np + nt) \le e^{-u(np+nt)} \prod_{i=1}^{n} E e^{uX_i}$, where $u > 0$



$$\le e^{-u(np+nt)} \prod_{i=1}^{n} E\,(1 - X_i + X_i e^u)$$

$$\le e^{-u(np+nt)} \left(\frac{n - ES_n + (ES_n)e^u}{n} \right)^n$$

geometric mean $\le$ arithmetic mean

$$= \left[e^{-u(p+t)} (1 - p + pe^u) \right]^n$$

$=$ **RHS** by letting

$$e^u = \frac{(p+t)(1-p)}{p(1-p-t)}$$

QED

Weaker but more useful bounds

Thm: Let $X_1, X_2, \cdots, X_n$ be independent rvs s.t. $0 \le X_i \le 1$ for each i. Let $p = \frac{ES_n}{n}$, where $S_n = \sum_{i=1}^{n} X_i$. Then

(a) For any $\epsilon > 0$,

$$P(S_n \ge (1+\epsilon)np) \le \exp\left(-np\left((1+\epsilon)\ln(1+\epsilon) - \epsilon\right)\right)$$
$$\le \exp\left(-\frac{\epsilon^2 np}{2(1+\epsilon/3)}\right)$$

(b) For any $\epsilon > 0$,

$$P(S_n \le (1-\epsilon)np) \le \exp\left(-\frac{1}{2}\epsilon^2 np\right)$$

pf of (a) In the proof of Lemma *,

let $t = \epsilon p$ and $e^u = (1+\epsilon)$, then we have

$$P(S_n \geq (1+\epsilon)np)$$

$$= P(S_n \geq np + nt)$$

$$\leq \left[(1+\epsilon)^{-(1+\epsilon)} (1+\epsilon p) \right]^n \quad \text{(by the proof of Lemma *)}$$

$$\leq \left[(1+\epsilon)^{-(1+\epsilon)} (1+\epsilon p)^{\frac{1}{p}} \right]^{np}$$

$$\leq \left[(1+\epsilon)^{-(1+\epsilon)} e^{\epsilon} \right]^{np} \quad (\because 1+\epsilon p \leq e^{\epsilon p})$$

this proves the first inequality in (a).

Pf of (a): Claim For all $x \geq 0$, $(1+x)\ln(1+x) - x \geq \dfrac{3x^2}{6+2x}$.

Note that $* = e^{-np[(1+\epsilon)\ln(1+\epsilon)-\epsilon]}$

$\leq e^{-np\frac{3\epsilon^2}{6+2\epsilon}}$, done.

Pf of (b): Let $f(t) = \ln\left(\left(\frac{p}{p+t}\right)^{p+t}\left(\frac{1-p}{1-p-t}\right)^{1-p-t}\right)$

let $h(x) = f(-xp)$ for $0 \leq x < 1$.

Then $h(0) = f(0) = 0$, $h'(0) = f'(0)(-p) = 0$ $\left(\because f'(t) = \ln\left(\frac{p(1-p-t)}{(p+t)(1-p)}\right)\right)$

$h''(x) = f''(-xp)p^2 = -\frac{p}{(1-x)(1-p+xp)} \leq -p$ $(\because 0 \leq 1-x \leq 1,\ 0 \leq 1-p(1-x) \leq 1)$

$\left(\text{Note } f''(t) = -\frac{1}{(p+t)(1-p-t)}\right)$

Taylor's Thm $\Rightarrow$ $h(x) = h(0) + h'(0)x + \frac{h''(\theta)x^2}{2}$, $0 \leq \theta \leq x$

$\leq 0 + 0 - \frac{px^2}{2}$

Pf of (b) Lemma* implies that

$$P\big((n-S_n)-(n-ES_n)\geq nt\big)\leq e^{f(t)n}$$

let $t=\varepsilon p$. we have

$$P(S_n\leq(1-\varepsilon)ES_n)=P(S_n\leq ES_n-nt)\leq e^{f(t)n}$$

$$=e^{h(-\frac{t}{p})n}\leq e^{-\frac{p}{2}(\frac{t}{p})^2 n}=e^{-\frac{1}{2}\epsilon^2 np}.$$

$(\because h(x)\leq-\frac{px^2}{2},\ \text{for } 0\leq x<1)$

QED

Remark: The first inequality in (a) implies

$$P(S_n \ge 2ES_n) \le e^{-(0.386)ES_n}$$ and

$$P(S_n \ge \delta ES_n) \le e^{-\delta \ln(\frac{\delta}{e}) ES_n}$$

The second inequality in (a) implies

$$P(S_n \ge (1+\epsilon)ES_n) \le e^{-\frac{1}{3}\epsilon^2 ES_n}$$

Hoeffding's Lemma

Lemma: $\mathcal{E}X=0$, $a \le X \le b$.

Then for any $s>0$,

$$\mathcal{E}(e^{sX}) \le e^{\frac{s^2(b-a)^2}{8}}$$

pf: Note that $e^{sx} \le \left(\frac{x-a}{b-a}\right)(e^{sb}-e^{sa})+e^{sa}$ for $a \le x \le b$. It follows that

$$\mathcal{E} e^{sX} \le \frac{-a}{b-a}(e^{sb}-e^{sa})+e^{sa}$$

$= e^{\phi(u)}$, where $u \stackrel{def}{=} s(b-a)$, $\phi(u) \stackrel{def}{=} -pu+\ln(1-p+pe^{u})$

and $p \stackrel{def}{=} -\frac{a}{b-a}$

Note that $\phi'(u) = -p+\frac{p}{p+(1-p)e^{-u}}$, $\phi''(u) = \frac{p(1-p)e^{-u}}{(p+(1-p)e^{-u})^2} \le \frac{p(1-p)\frac{p}{1-p}}{\left(p+(1-p)\frac{p}{1-p}\right)^2} = \frac{1}{4}$

Taylor's Thm $\Rightarrow$ for some $\theta \in [0,u]$,

$$\phi(u) = \underbrace{\phi(0)}_{=0} + \underbrace{\phi'(0)}_{=0}u + \frac{\phi''(\theta)u^2}{2} \le \frac{u^2}{8} = \frac{s^2(b-a)^2}{8}$$

QED

Hoeffding's Inequality

Thm Let $X_1, X_2, \cdots, X_n$ be independent rvs s.t.

$a_i \le X_i \le b_i$ with probability one, $1 \le i \le n$.

Then for any $t \ge 0$ we have

$$P(S_n - ES_n \ge t) \le \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

$$P(S_n - ES_n \le -t) \le \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

where $S_n = \sum_{i=1}^n X_i$.

pf: $P(S_n - ES_n \geq t) = P\left(e^{s(S_n - ES_n)} \geq e^{st}\right)$

$$\leq e^{-st} E e^{s(S_n - ES_n)}$$

$$= e^{-st} \prod_{i=1}^{n} E e^{s(X_i - EX_i)}$$

$$\leq e^{-st} \prod_{i=1}^{n} e^{\frac{s^2 (b_i - a_i)^2}{8}}$$ by Hoeffding's Lemma

$(\because (b_i - EX_i) - (a_i - EX_i) = b_i - a_i)$

$$= e^{-st} e^{\frac{s^2}{8} \sum_{i=1}^{n} (b_i - a_i)^2}$$

$$= e^{-\frac{2t^2}{\sum_{i=1}^{n} (b_i - a_i)^2}}$$ (by choosing $s = \frac{4t}{\sum_{i=1}^{n} (b_i - a_i)^2} > 0$)

Chebyshev's Ineq. is Inadequate

Let $X_1, X_2, \cdots \overset{iid}{\sim} B(1,p)$

Chebyshev's Inequality says that $P(|S_n - ES_n| \geq n\varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}$

However, as $n \to \infty$, Central Limit Thm suggests that

$$P(S_n - ES_n \geq n\varepsilon) = P\left(\frac{S_n - ES_n}{\sqrt{Var S_n}} \geq \frac{n\varepsilon}{\sqrt{np(1-p)}}\right) \sim 1 - \Phi\left(\frac{n\varepsilon}{\sqrt{np(1-p)}}\right)$$

$$\leq \exp\left(-\frac{n\varepsilon^2}{2p(1-p)}\right) \quad \left(\because \sqrt{2\pi}\{1 - \Phi(x)\} < \frac{1}{x}\exp\left(-\frac{x^2}{2}\right)\right)$$

It seems Hoeffding Ineq. is adequate.....

However....

Hoeffdingó bound is independent of p, so it may loose if p is small (or large).

That is this inequality ignores information about the variance of the X_i's.

The **Bernstein's Inequality** will give an answer.....

A Lemma first

Lemma: If $|X| \le c$, $\mathcal{E}X = 0$ and $Var(X) = \sigma^2$,

then $\mathcal{E}(e^{sX}) \le \exp\left(\frac{\sigma^2}{c^2}\left(e^{sc} - 1 - sc\right)\right)$

pf:

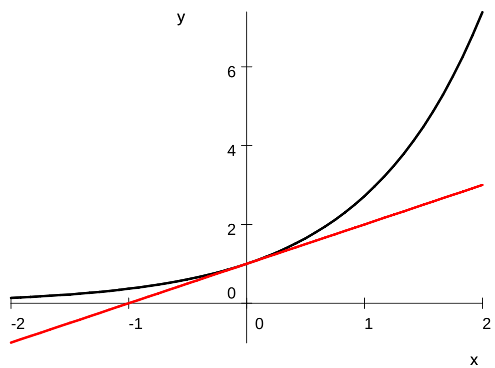
$$e^{sx} = 1 + sx + \sum_{r=2}^{\infty} \frac{(sx)^r}{r!}$$

$$\Rightarrow \quad E\, e^{sX} = 1 + \sum_{r=2}^{\infty} \frac{s^r E X^r}{r!} \qquad (\because EX = 0)$$

$$\le 1 + \sum_{r=2}^{\infty} \frac{s^r c^{r-2} \sigma^2}{r!} \qquad \left(\because EX^r \le E|X|^{r-2}X^2 \le c^{r-2}\sigma^2 \text{ for } r \ge 2\right)$$

$$= 1 + \frac{\sigma^2}{c^2} \sum_{r=2}^{\infty} \frac{(sc)^r}{r!}$$

$$= 1 + \frac{\sigma^2}{c^2}\left(e^{sc} - 1 - sc\right)$$



$$\le e^{\frac{\sigma^2}{c^2}\left(e^{sc} - 1 - sc\right)} \qquad (\because e^x \ge 1 + x \quad \forall x)$$

The Bennett's Inequality

Thm Let $X_1, X_2, \cdots, X_n$ be independent r.v.s with $EX_i = 0$ and $|X_i| \le c$, $1 \le i \le n$. Let $S_n = \sum_{i=1}^n X_i$ and $\sigma^2 = \frac{1}{n}\sum_{i=1}^n Var(X_i)$. Then for any $t > 0$,

$$P(S_n > t) \le \exp\left(-\frac{t}{c}\left(\left(1+\frac{n\sigma^2}{ct}\right)\ln\left(1+\frac{ct}{n\sigma^2}\right)-1\right)\right)$$

The Proof

Proof:

$$P(S_n > t) \le e^{-st} E e^{s S_n} \qquad \text{(Markov's ineq)}$$

$$= e^{-st} \prod_{i=1}^{n} E e^{s X_i}$$

$$\le e^{-st} \prod_{i=1}^{n} e^{\frac{\sigma_i^2}{c^2}(e^{sc} - 1 - sc)} \qquad \text{by above lemma}$$

$$= e^{-st} e^{\frac{n\sigma^2}{c^2}(e^{sc} - 1 - sc)}$$

The last term is minimized for

$$s = \frac{1}{c} \ln\left(1 + \frac{tc}{n\sigma^2}\right).$$

Resubstituting this value, we obtain Bennett's ineq.

Bernstein's Inequality

Thm Under the conditions of Bennett's Ineq, for $\varepsilon > 0$, we have

$$P(S_n > n\varepsilon) \leq \exp\left(-\frac{n\varepsilon^2}{2\sigma^2 + c\varepsilon}\right)$$

pf: Applying the following elementary ineq. to Bennett's bound: $\ln(1+x) \geq \frac{2x}{2+x}, \; x \geq 0$

ln(1+x)= x-(1/2)x²+(1/3)x³+O(x⁴)

(2x)/(2+x)= x-(1/2)x²+(1/4)x³+O(x⁴)

Bennett's bound

$$\leq \exp\left(-\frac{n\varepsilon}{c}\left(\left(1+\frac{n\sigma^2}{c\varepsilon n}\right)\left(\frac{2\,\frac{cn\varepsilon}{n\sigma^2}}{2+\frac{cn\varepsilon}{n\sigma^2}}\right)-1\right)\right)$$

$$=\exp\left(-\frac{n\varepsilon}{c}\left(\frac{c\varepsilon+\sigma^2}{c\varepsilon}\,\frac{2c\varepsilon}{c\varepsilon+2\sigma^2}-1\right)\right)$$

$$=\exp\left(-\frac{n\varepsilon}{c}\,\frac{c\varepsilon}{2\sigma^2+c\varepsilon}\right)=\exp\left(-\frac{n\varepsilon^2}{2\sigma^2+c\varepsilon}\right)$$

QED

What's after Bernstein's ineq.....

Poisson-type Inequality

Thm Let $X_1, X_2, \dots, X_n$ be independent r.v.s with $0 \le X_i \le 1 \ \forall i$ and $m = \mathcal{E} S_n$.

Then for any $t \ge m$,

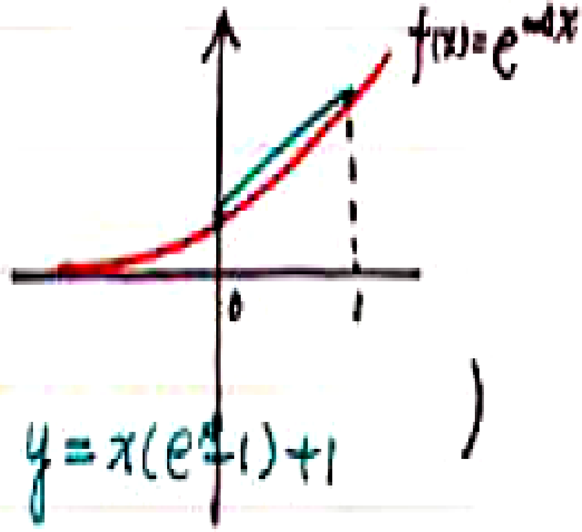
$$\mathcal{P}(S_n \ge t) \le \left(\frac{m}{t}\right)^t e^{t-m}$$

Note: Here n doesn't appear on the RHS

Pf: let $m_i = EX_i$.

$$P(S_n \ge t) \le e^{-\lambda t}\prod_{i=1}^{n} E e^{\lambda X_i}$$

$$\le e^{-\lambda t}\prod_{i=1}^{n} E\left(X(e^{\lambda}-1)+1\right) \quad (\because$$



$)$

$$= e^{-\lambda t}\prod_{i=1}^{n}\left[1+m_i(e^{\lambda}-1)\right]$$

$$\le e^{-\lambda t}\prod_{i=1}^{n} e^{m_i(e^{\lambda}-1)} = e^{-\lambda t}e^{m(e^{\lambda}-1)} \quad ★$$

$$\le \left(\frac{m}{t}\right)^{t} e^{t-m}$$ by choosing $\lambda = \ln\left(\frac{t}{m}\right)$

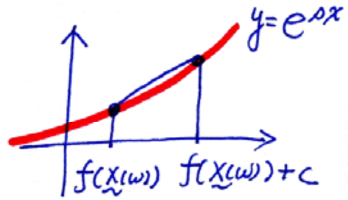
to minimize the bound ★ QED

An extension of Hoeffding's Lemma

Lemma: Suppose for rv V and r.vector $\underset{\sim}{X}$ we have

① $E(V \mid \underset{\sim}{X}) = 0$ a.s. ② $f(\underset{\sim}{X}) \le V \le f(\underset{\sim}{X}) + c$, for some fun. f and constant c.

Then for any $s > 0$, $E(e^{sV} \mid \underset{\sim}{X}) \le e^{\frac{s^2c^2}{8}}$.



pf: Note that

$$e^{sV} \le \frac{V - f(\underset{\sim}{X})}{c}\left(e^{s(f(\underset{\sim}{X})+c)} - e^{sf(\underset{\sim}{X})}\right) + e^{sf(\underset{\sim}{X})}$$

Thus $E(e^{sV} \mid \underset{\sim}{X}) = \frac{-f(\underset{\sim}{X})}{c}\left(e^{s(f(\underset{\sim}{X})+c)} - e^{sf(\underset{\sim}{X})}\right) + e^{sf(\underset{\sim}{X})}$.

The remains is similar to Hoeffding's Lemma.

McDiarmid's Inequality

Thm: Let $X_1, X_2, \cdots, X_n$ be independent rvs. Let f be a function $f: \mathbb{R}^n \to \mathbb{R}$ with a vector $(c_1, c_2, \cdots, c_n)$ s.t. $|f(\underset{\sim}{x}) - f(\underset{\sim}{y})| \le c_i$ for all $\underset{\sim}{x}, \underset{\sim}{y}$ in $\mathbb{R}^n$ that differ only at the ith coordinate, $1 \le i \le n$.

Then for any $t > 0$,

$$P\left(\left| f(X_1, \cdots, X_n) - \mathcal{E} f(X_1, \cdots, X_n) \right| \ge t \right) \le 2 e^{-\frac{2t^2}{\sum_{i=1}^n c_i^2}}$$

pf: Let $\underset{\sim}{X} = (X_1, \cdots, X_n)$, $Z_0 = \mathcal{E} f(\underset{\sim}{X})$, $Z_i = \mathcal{E}(f(\underset{\sim}{X}) \mid X_1, \cdots; X_i)$, $Z_n = f(\underset{\sim}{X})$.

Claim $\mathcal{E}\left(e^{s(Z_k - Z_{k-1})} \mid X_1, \cdots; X_{k-1}\right) \le e^{\frac{s^2 c_k^2}{8}}$, $1 \le k \le n$.

pf: Let $U_k = \sup_u \left\{ \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_{k-1}, u\right) - \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_{k-1}\right) \right\}$

$L_k = \inf_l \left\{ \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_{k-1}, l\right) - \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_{k-1}\right) \right\}$

Note that $U_k - L_k \le \sup_{l,u} \left\{ \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_{k-1}, u\right) - \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_{k-1}, l\right) \right\}$

$\because X_1, \cdots, X_n$ are independent

$\le \sup_{l,u} \left\{ \sum_{y_{k+1}, \cdots, y_n} \left[f(X_1, \cdots, X_{k-1}, u, y_{k+1}, \cdots, y_n) - f(X_1, \cdots, X_{k-1}, l, y_{k+1}, \cdots, y_n) \right] \prod_{j=k+1}^{n} P(X_j = y_j) \right\}$

$\le c_k$ ($\because$ Lipschitz condition)

So $L_k \le Z_k - Z_{k-1} \le U_k \le L_k + c$ and hence the claim is true by the extension of Hoeffding's Lemma, since $\mathcal{E}(Z_k - Z_{k-1} \mid X_1, \cdots, X_{k-1}) = 0$ a.s.

QED of claim

Pf (continued)

$$\Pr\{f(\underset{\sim}{X}) - \mathcal{E} f(\underset{\sim}{X}) \geq t\}$$

$$\leq e^{-st}\, \mathcal{E}\, e^{s[f(\underset{\sim}{X}) - \mathcal{E} f(\underset{\sim}{X})]}$$

$$= e^{-st}\, \mathcal{E}\, e^{s \sum_{k=1}^{n} [Z_k - Z_{k-1}]}$$

$$= e^{-st}\, \mathcal{E}\left(\mathcal{E}\left(e^{s \sum_{k=1}^{n} [Z_k - Z_{k-1}]} \,\middle|\, X_1, \cdots, X_{n-1}\right)\right) \quad (\because \text{ Tower property})$$

$$= e^{-st}\, \mathcal{E}\left(e^{s \sum_{k=1}^{n-1} [Z_k - Z_{k-1}]}\, \mathcal{E}\left(e^{s(Z_n - Z_{n-1})} \,\middle|\, X_1, \cdots, X_{n-1}\right)\right)$$

$$\leq e^{-st}\, \mathcal{E}\left(e^{s \sum_{k=1}^{n-1} [Z_k - Z_{k-1}]}\, e^{\frac{s^2 c_n^2}{8}}\right) \quad (\text{by the claim})$$

$$\leq e^{-st} \prod_{k=1}^{n} e^{\frac{s^2 c_k^2}{8}} \quad (\text{by repeating the same argument } n \text{ times})$$

$$= e^{-st + s^2 \sum_{k=1}^{n} \frac{c_k^2}{8}} \leq e^{-\frac{2t^2}{\sum_{k=1}^{n} c_k^2}} \quad \text{by choosing } s = \frac{4t}{\sum_{k=1}^{n} c_k^2}$$

QED

Bin Packing

- The Bin Packing Problem requires finding the minimum number of unit size bins needed to pack a given collection of items with sizes in [0,1]. In our model, we have n items to pack, and the size of items are $X_1,\ldots,X_n$, i.i.d. over [0,1]. *We use a fixed procedure for packing*.

Ref:
Rhee & Talagrand (1987) “Martingale inequalities and NP-complete problem” Math. of Oper. Res., 12, 177-181.

Bin Packing

Discussion: We denote by $f(X_1,\cdots,X_n)$ the numbers of bins needed to pack $X_1, X_2, \cdots, X_n$ using procedure $\mathcal{A}$.

(a) If $\mathcal{A}$ is optimum then f is 1-Lipschitz. McDiarmid's Ineq. says that $P(|f(\underset{\sim}{X}) - \mathcal{E}f(\underset{\sim}{X})| \geq t) \leq 2e^{-\frac{2t^2}{n}}$.

(b) If $\mathcal{A}$ is the procedure Next Fit then f is 2-Lipschitz, and hence $P(|f(\underset{\sim}{X}) - \mathcal{E}f(\underset{\sim}{X})| \geq t) \leq 2e^{-\frac{t^2}{n}}$.

Next Fit Procedure

- Where the bins are filled one at a time and a new bin is started when the current element does not fit in the bin being currently filled.

To see f is 2-Lipschitz:

$(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, \cdots, x_n)$ $\quad$ $(x_1, \cdots, x_{i-1}, x_i^{new}, x_{i+1}, \cdots, x_n)$

Concentration of the Chromatic Number

Thm (Shamir-Spencer) Let $n \geq 2$ and $p \in (0,1)$.

Then we have

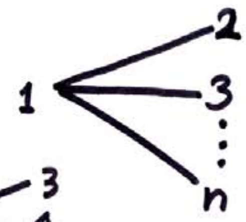
$$P\left(\left|\chi(G_{n,p}) - \mathcal{E}\chi(G_{n,p})\right| \geq t\right) \leq 2e^{-\frac{2t^2}{n-1}}$$

Remark: probability space $\Omega = \{\omega : \omega \text{ is a graph on the vertex set } \{1,2,\cdots,n\}\}$ with probability measure $P(\omega) = p^m (1-p)^{\binom{n}{2}-m}$ where $m = |E(\omega)|$.

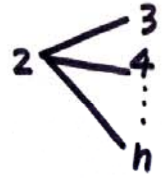
One may consider $G_{n,p}$ as a random element $G_{n,p} : \Omega \to \Omega$ s.t. $G_{n,p}(\omega) = \omega$.

Sometimes we consider $G_{n,p}$ as a prob. space.

pf: Let $\underset{\sim}{X_1} = (X_{12}, X_{13}, X_{14}, \cdots, X_{1n})$

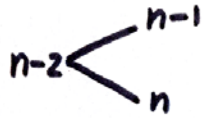


$\underset{\sim}{X_2} = (X_{23}, X_{24}, \cdots\cdots, X_{2n})$



$\vdots$

$\underset{\sim}{X_{n-2}} = (X_{n-2,n-1}, X_{n-2,n})$



$\underset{\sim}{X_{n-1}} = X_{n-1,n}$ $\qquad n-1 \text{ ——— } n$

We have $\chi(G_{n,p}) = f(\underset{\sim}{X_1}, \underset{\sim}{X_2}, \cdots, \underset{\sim}{X_{n-1}})$ for some function f.

Note that $f: S_1 \times S_2 \times \cdots \times S_{n-1} \longrightarrow \mathbb{R}$, where $S_i = \{0,1\}^{n-i}$, and

f is a 1-Lipschitz function.

By McDiarmid's Inequality, we arrive at

$$P\left(|\chi(G_{n,p}) - E\chi(G_{n,p})| \geq t \right) \leq 2e^{-\frac{2t^2}{n-1}}$$

QED

A Remark on Shamir-Spencer's Thm.

- Since $\mathbb{P}(|\chi(G_{n,p})-\mathbb{E}\chi(G_{n,p})| \geq t\sqrt{n-1}) \leq 2e^{-2t^2}$, so the chromatic number is almost always concentrated on about $O(\sqrt{n})$ values.
- Note that we have no clue to what the value of $\mathbb{E}\chi(G_{n,p})$ is.

A technique lemma...

Thm Lemma 7.3.4 (Alon & Spencer) Let α, c be fixed and $\alpha > \frac{5}{6}$. let $p = n^{-\alpha}$. Then almost always every $c\sqrt{n\log n}$ vertices of $G_{n,p}$ induce a 3-colorable subgraph.

pf: $\mathcal{P}(\forall$ subgraph H of $G_{n,p}$ with $v_H \leq c\sqrt{n\log n}$ having $\chi(H) \leq 3)$

$$= 1 - \mathcal{P}\left(\bigcup_{t=4}^{c\sqrt{n\log n}} \{\exists\, H \subseteq G_{n,p} \text{ with } v_H = t \text{ having } \chi(H) > 3\}\right)$$

$$\star \leq \sum_{t=4}^{c\sqrt{n\log n}} \mathcal{P}(\exists \text{ minimal } H \subseteq G_{n,p} \text{ with } v_H = t \text{ having } \chi(H) > 3)$$

i.e. $\chi(H-v) \leq 3$ for $\forall v \in V(H)$

$$\leq \sum_{t=4}^{c\sqrt{n\log n}} \mathcal{P}(\exists\, H \subseteq G_{n,p} \text{ with } v_H = t \text{ having } d_H(x) \geq 3 \text{ for all } x \in V(H))$$

$$\leq \sum_{t=4}^{c\sqrt{n\log n}} \binom{n}{t} \mathcal{P}(H \cong G_{n,p}[\{1,2,\cdots,t\}] \text{ has } e_H \geq \tfrac{3}{2}t)$$

Pf (continued)

$$\leq \sum_{t=4}^{c\sqrt{n\log n}} \binom{n}{t} \binom{\binom{t}{2}}{\frac{3}{2}t} p^{\frac{3}{2}t}$$

$$\leq \sum_{t=4}^{c\sqrt{n\log n}} \left(\frac{en}{t}\right)^t \left(\frac{e\binom{t}{2}}{\frac{3}{2}t}\right)^{\frac{3}{2}t} p^{\frac{3}{2}t} \quad \left(\because \binom{n}{k} \leq \left(\frac{en}{k}\right)^k\right)$$

$$\leq \sum_{t=4}^{c\sqrt{n\log n}} \left(\frac{en}{t} \frac{t^{\frac{3}{2}} e^{\frac{3}{2}}}{3^{\frac{3}{2}}} n^{-\frac{3}{2}\alpha}\right)^t \quad (\because p = n^{-\alpha})$$

$$\leq \sum_{t=4}^{c\sqrt{n\log n}} \left(c_1 n^{1-\frac{3}{2}\alpha} t^{\frac{1}{2}}\right)^t$$

$$\leq \sum_{t=4}^{c\sqrt{n\log n}} \left(c_2 n^{1-\frac{3}{2}\alpha} n^{\frac{1}{4}} (\log n)^{\frac{1}{4}}\right)^t$$

$$= \sum_{t=4}^{c\sqrt{n\log n}} \left(c_2 n^{-\varepsilon} (\log n)^{\frac{1}{4}}\right)^t \quad \left(\because \frac{5}{4} - \frac{3}{2}\alpha < 0 \Leftrightarrow \frac{5}{6} < \alpha\right)$$

$$\leq \left(c_2 n^{-\varepsilon} (\log n)^{\frac{1}{4}}\right)^4 \frac{1}{1 - c_2 n^{-\varepsilon} (\log n)^{\frac{1}{4}}} = o(1)$$

QED

Four-Value Concentration

Thm (see Alon & Spencer Thm 7.3.3) Let $\alpha > \frac{5}{6}$ be fixed, and let $p = n^{-\alpha}$ (i.e. p is not too large)

Then for any n, $\exists$ $u = u(\alpha, n)$ such that

$$\chi(G_{n,p}) \in \{u, u+1, u+2, u+3\} \text{ almost surely}$$

i.e. $$\Pr\{\chi(G_{n,p}) \notin \{u, u+1, u+2, u+3\}\} \to 0 \text{ as } n \to \infty$$

pf: Let $u = u(n,\alpha)$ be the smallest integer s.t. $P(\chi(G_{n,p}) \leq u) > \frac{1}{n}$.

Claim A: $P(\chi(G_{n,p}) \geq u) \geq 1 - \frac{1}{n}$

pf: The choice of $u \Rightarrow P(\chi(G_{n,p}) \leq u-1) \leq \frac{1}{n} \Rightarrow P(\chi(G_{n,p}) \geq u) \geq 1 - \frac{1}{n}$ ∎A

Let X be the minimum number of vertices whose deletion makes $G_{n,p}$ u-colorable.

Then $X = f(\underset{\sim}{X_1}, \underset{\sim}{X_2}, \cdots, \underset{\sim}{X_{n-1}})$ for some function f, where $\underset{\sim}{X_1}, \underset{\sim}{X_2}, \cdots, \underset{\sim}{X_{n-1}}$ were defined in the proof of Shamir-Spencer's Thm. Note that f is 1-Lipschitz.

Claim B: $\sqrt{2(n-1)\log n} > \mathcal{E}X$

pf:
$$\begin{aligned}
\frac{1}{n} < P(\chi(G_{n,p}) \leq u) &= P(X=0) \\
&= P(X \leq \mathcal{E}X - \mathcal{E}X) \\
&= P(f(\underset{\sim}{X_1}, \cdots, \underset{\sim}{X_{n-1}}) \leq \mathcal{E}f(\underset{\sim}{X_1}, \cdots, \underset{\sim}{X_{n-1}}) - \mathcal{E}X) \\
&\leq \exp\left(-\frac{2(\mathcal{E}X)^2}{n-1}\right) \quad \text{by McDiarmid's Inequality}
\end{aligned}$$

Therefore $\sqrt{\frac{1}{2}(n-1)\log n} > \mathcal{E}X$. ∎B

Pf (Continued)

Claim C: $P\left(X < 2\sqrt{2(n-1)\log n}\right) \ge 1 - \frac{1}{n}$

Pf: $\text{LHS} = 1 - P\left(X \ge 2\sqrt{2(n-1)\log n}\right)$

$$\ge 1 - P\left(X \ge EX + \sqrt{2(n-1)\log n}\right)$$

$$\ge 1 - \exp\left(-\frac{2(n-1)\log n}{2(n-1)}\right) \quad \text{by McDiarmid's Ineq. again!}$$

$$= 1 - \frac{1}{n}. \qquad \blacksquare_C$$

Let $A = \{\chi(G_{n,p}) \ge u\}$ and $B = \{X < 2\sqrt{2(n-1)\log n}\}$ be two events.

Then $P(A \cap B) = 1 - P(\bar{A} \cup \bar{B}) \ge 1 - P(\bar{A}) - P(\bar{B})$

$$\ge 1 - \frac{1}{n} - \frac{1}{n} = 1 - \frac{2}{n} \to 1 \text{ as } n \to \infty$$

The thm. follows from the technique lemma we proved. **QED**

Kim's Lemma

Thm (J.H. Kim 1995) Suppose that

- $X_1, X_2, \cdots, X_n$ are independent rvs s.t. $X_i \sim B(1, p_i)$ $\forall i$.
- $f : \{0,1\}^n \longrightarrow \mathbb{R}$ and g is a convex function.

Then for any i, $1 \leq i \leq n$, we have

$q_i = 1 - p_i$

$$\mathcal{E}\left(g(V_i) \mid X_1, \cdots, X_{i-1}\right) \leq \mathcal{E}\left(p_i\, g(q_i r_i) + q_i\, g(-p_i r_i) \mid X_1, \cdots, X_{i-1}\right)$$

where $r_i = f(X_1, \cdots, X_{i-1}, 1, X_{i+1}, \cdots, X_n) - f(X_1, \cdots, X_{i-1}, 0, X_{i+1}, \cdots, X_n)$

$$V_i = \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_i\right) - \mathcal{E}\left(f(\underset{\sim}{X}) \mid X_1, \cdots, X_{i-1}\right)$$

$\underset{\sim}{X} = (X_1, \cdots, X_n)$

An extension of Devroye's Inequality

Thm If we have

(1) $X_1, X_2, \cdots, X_n$ are independent. $X_i \sim B(1, p_i)$ $\forall i$

(2) $f : \{0,1\}^n \longrightarrow \mathbb{R}$ having a vector $(c_1, c_2, \cdots, c_n)$ s.t.

$|f(\underset{\sim}{a}) - f(\underset{\sim}{b})| \leq c_i$ for all $\underset{\sim}{a}, \underset{\sim}{b}$ that differ only in the ith coordinate.

Then

$$\mathrm{Var}[f(X_1, X_2, \cdots, X_n)] \leq \sum_{i=1}^{n} c_i^2 p_i (1-p_i)$$

pf: Let $f = f(X_1, X_2, \dots, X_n)$ and $V_i = \mathcal{E}(f \mid \underbrace{X_1, \dots, X_i}_{\underset{\sim}{X}_i}) - \mathcal{E}(f \mid \underbrace{X_1, \dots, X_{i-1}}_{\underset{\sim}{X}_{i-1}})$ $2 \le i \le n$

$V_1 \overset{\text{def}}{=} \mathcal{E}(f \mid X_1) - \mathcal{E}(f \mid \mathcal{F}_0)$

Then $\mathrm{Var}\, f = \mathcal{E}(f - \mathcal{E}f)^2$

$$= \mathcal{E}\Big(\big(\textstyle\sum_{i=1}^n V_i\big)^2 \Big)$$

$\because$ $\mathcal{E}(f \mid \mathcal{F}_0) = \mathcal{E}f$ where $\mathcal{F}_0 = \{\phi, \Omega\}$

$\mathcal{E}(f \mid \underset{\sim}{X}_n) = f$

$$= \mathcal{E}\Big(\textstyle\sum_{i=1}^n V_i^2 + 2 \sum_{1 \le i < j \le n} V_i V_j \Big)$$

$$= \mathcal{E}\big(\textstyle\sum_{i=1}^n V_i^2 \big)$$

$\because$ $\mathcal{E} V_i V_j = \mathcal{E}\big(\mathcal{E}(V_i V_j \mid X_1, \dots, X_{j-1})\big)$ $1 \le i < j \le n$

$= \mathcal{E}\big(V_i\, \mathcal{E}(V_j \mid X_1, \dots, X_{j-1})\big) = \mathcal{E}(V_i \cdot 0) = 0$

$$= \textstyle\sum_{i=1}^n \mathcal{E}\big(\mathcal{E}(V_i^2 \mid \underset{\sim}{X}_{i-1})\big)$$

here we define $\mathcal{E}(V_1^2 \mid \underset{\sim}{X}_0) \overset{\text{def}}{=} \mathcal{E}(V_1^2 \mid \mathcal{F}_0)$

$$\le \textstyle\sum_{i=1}^n \mathcal{E}\Big(\mathcal{E}\big(p_i (q_i r_i)^2 + q_i (p_i r_i)^2 \mid \underset{\sim}{X}_{i-1}\big) \Big)$$

Applying Kim's Lemma with $g(x) = x^2$.

$$= \textstyle\sum_{i=1}^n p_i q_i r_i^2$$

$$\le \textstyle\sum_{i=1}^n p_i q_i c_i^2$$

$\because$ $|r_i| \le c_i$ Lipschitz property

QED