Tail bounds / concentration inequalities

tell us how close r.v.'s are to their expectation

Markov Inequality

bounds right tail

very weak, but no assumptions (except ≥ 0)

Let X be a r.v. that assumes only nonnegative values

Then for all $a > 0$

$$\Pr(X \geq a) \leq \frac{E(X)}{a}$$

Proof:

$$\Pr(X \geq a) = \int_a^\infty f(x)\,dx \leq \int_a^\infty \frac{x}{a} f(x)\,dx \leq \frac{E(X)}{a}$$

$$Var(X) = E\left[(X - E(X))^2\right] = E(X^2) - [E(X)]^2$$

Chebychev Bound

$$\forall\, a > 0 \qquad Pr\left(|X - E(X)| \geq a\right) \leq \frac{Var(X)}{a^2}$$

Proof:

$$Pr\left(|X - E(X)| \geq a\right) = Pr\Big(\underbrace{(X - E(X))^2}_{Y} \geq a^2\Big)$$

$$\leq \frac{E(Y)}{a^2} = \frac{Var(X)}{a^2}$$

Markov Inequality

Chernoff - Hoeffding Bounds

for tails of sums of indep r.v's.

proof technique can be applied in many settings

Version 1.0

Let $X_1, \ldots, X_n$ be indep r.v's s.t. $X_i \in [0,1]$

Let $X = \sum_{i=1}^{n} X_i$ and $\mu = E(X)$ [Therefore, if $E(X_i) = p_i \Rightarrow \mu = \sum_{i=1}^{n} p_i$]

Chernoff-Hoeffding Upper Tail

$\forall \delta > 0$

$$\Pr(X \geq (1+\delta)\mu) \leq e^{-\mu[(1+\delta)\ln(1+\delta) - \delta]} \qquad (1)$$

Weaker but useful versions:

$\forall \delta \in [0,1]$ $$\Pr(X \geq (1+\delta)\mu) \leq e^{-\frac{\mu\delta^2}{3}} \qquad (2)$$

$\forall \delta \geq 1$ $$\Pr(X \geq (1+\delta)\mu) \leq e^{-\frac{\mu\delta}{3}} \qquad (3)$$

Remark: Suffices to have $\mu \geq E(X)$

Chernoff - Hoeffding Lower Tail

$\forall \delta \in [0,1]$ $\quad Pr(X \leq (1-\delta)\mu) \leq e^{-\frac{\mu\delta^2}{2}}$ $\quad (4)$

Remark: Suffices to have $\mu \leq E(X)$

Proofs:

Key property of indep r.v.'s : $E(YZ) = E(Y)E(Z)$

but we have sums.

Idea: convert sums to products using exponential fn.

Define $Y_i = e^{tX_i}$ (t to be determined)

$$Y = e^{tX} = e^{t\Sigma X_i} = \prod_{i=1}^{n} e^{tX_i} = \prod Y_i$$

X_i's indep $\Rightarrow$ Y_i's indep

$$\Rightarrow E(Y) = \prod_{i=1}^{n} E(Y_i)$$

Why is this useful?

$$\Pr(X \geq a) = \Pr(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}} = \frac{E(Y)}{e^{ta}} = \frac{\prod E(Y_i)}{e^{ta}}$$

$\forall t > 0$

Markov Inequality

Choose t to get best possible upper bound.

$$\Rightarrow \quad \boxed{\Pr(X \geq a) \leq \min_{t>0} \frac{E(e^{tX})}{e^{ta}}}$$

Similarly, $\forall t < 0$

$$\Pr(X \leq a) = \Pr(e^{tX} \geq e^{ta}) \leq \frac{E(e^{tX})}{e^{ta}}$$

$$\Rightarrow \quad \boxed{\Pr(X \leq a) \leq \min_{t<0} \frac{E(e^{tX})}{e^{ta}}}$$

$$\boxed{\text{Claim: } E(e^{tX_i}) \le 1 + p_i(e^t - 1) \le e^{(e^t-1)p_i}}$$

$1+x \le e^x \quad \forall x$

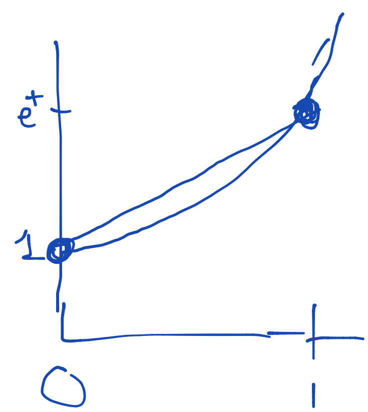
Proof:

Suppose $X_i = \begin{cases} 1 & \text{w.p. } p_i \\ 0 & \text{o.w.} \end{cases}$

$$E(e^{tX_i}) = e^t p_i + (1-p_i)$$

For $X_i \in [0,1]$, observe e^{tx} is convex fn of x [2^{nd} deriv > 0]

$$\Rightarrow e^{tx} \le 1 + (e^t - 1)x \qquad x \in [0,1]$$



$$\Rightarrow E(e^{tX_i}) \le E(1 + (e^t-1)X_i)$$

$$= 1 + (e^t - 1)E(X_i)$$

$$= 1 + (e^t - 1)p_i$$

Putting it together

$$\boxed{\Pr(X \geq (1+\delta)\mu) \leq \frac{E(e^{tX})}{e^{t(1+\delta)\mu}} \leq \frac{\prod_{i=1}^{n} e^{(e^t-1)p_i}}{e^{t(1+\delta)\mu}} = e^{[(e^t-1)\Sigma p_i - t(1+\delta)\mu]}}$$

We choose $t > 0$ to minimize $f(t) = e^{[(e^t-1) - t(1+\delta)]\mu}$

$$f'(t) = [e^t - (1+\delta)]\, e^{[(e^t-1) - t(1+\delta)]\mu}$$

$$f'(t) = 0 \text{ when } e^t = 1+\delta \equiv \boxed{t = \ln(1+\delta)}$$

$$f''(t) = [e^t - (1+\delta)]^2 e^{[\ldots]} + e^t e^{[\ldots]} \geq 0$$

$$\Pr(X \geq (1+\delta)\mu) \leq e^{[(e^t-1)\Sigma p_i - t(1+\delta)\mu]}$$

with $t = \ln(1+\delta)$

$$\Pr(X \geq (1+\delta)\mu) \leq e^{[e^{\ln(1+\delta)} - 1 - (1+\delta)\ln(1+\delta)]\mu}$$

$$= e^{[1+\delta - 1 - (1+\delta)\ln(1+\delta)]\mu}$$

$$= e^{-\mu[(1+\delta)\ln(1+\delta) - \delta]}$$

$\square$

The Union Bound

Let $E_1, \dots, E_k$ be any collection of events in prob space

Then $\Pr(E_1 \vee E_2 \vee \dots E_k) \leq \sum_{i=1}^{k} \Pr(E_i)$

$\boxed{\text{Application \#1 : Balls in Bins}}$

n bins. Throw n balls, one at a time, independently at random into a bin (Prob $\rightarrow i^{th}$ bin $= \frac{1}{n}$)

Let B_i = # balls in i^{th} bin at end.

Theorem: $\Pr\left(\max_i B_i \geq e \frac{\ln n}{\ln \ln n}\right) \leq \frac{1}{n^c}$ for some $c > 0$

Proof: Outline: prove that prob that a particular bin so unlikely to exceed this bound that a union bound gives result.

Consider bin 1

$$\text{Let } X_i := \begin{cases} 1 & i^{th} \text{ ball} \rightarrow \text{bin } 1 \\ 0 & \text{o.w.} \end{cases}$$

$$\Rightarrow B_1 = \sum_{i=1}^{n} X_i$$

$$E(X_i) = \frac{1}{n} \Rightarrow E(B_1) = 1$$

By Chernoff bound (1)

$$\Pr(\underbrace{B_1 \geq (1+\delta)E(B_1)}_{1+\delta}) \leq e^{-E(B_1)[(1+\delta)\ln(1+\delta)-\delta]}$$

Set $\delta = \boxed{\frac{e \ln n}{\ln \ln n}} - 1$

call this C

$$\Pr(B_1 \geq C) \leq e^{-C\ln C + C - 1} \leq e^{-C(\ln C - 1)}$$

Claim: $\ln C = 1 + \ln\ln n - \ln\ln\ln n$

$$\geq 1 + \frac{\ln\ln n}{2}$$

$\frac{\ln\ln n}{2} \geq \ln\ln\ln n \quad \equiv \quad e^v \geq 2v \quad [v = \ln\ln\ln n]$

$e^v = 1 + v + \frac{v^2}{2} + \ldots \qquad$ trivial for $v \geq 1$

$$\Rightarrow \Pr(B_1 \geq C) \leq e^{-C(\ln C - 1)}$$

$$\leq e^{-C(1 + \frac{\ln\ln n}{2} - 1)}$$

$$\leq e^{-C\frac{\ln\ln n}{2}} = e^{-\frac{e \ln n}{2}} \qquad C = \frac{e \ln n}{\ln\ln n}$$

$$= n^{-\frac{e}{2}} \leq n^{-1.35}$$

$$\Pr(\max_i B_i \geq C) = \Pr(E_1 \vee \dots \vee E_n) \leq \sum_{i=1}^{n} \Pr(E_i) \leq \sum_{i=1}^{n} n^{-1.35}$$

E_j event that $B_j \geq C$

$$\leq n^{-0.35}$$

Application #2: Congestion Minimization (randomized rounding)

Input: directed graph $G=(V,E)$

Set of pairs (s_i, t_i) $\quad i=1,..,k$

Output: path P_i from $s_i \rightsquigarrow t_i$ $\quad \forall i=1..k$

such that $\underbrace{\text{congestion}}_{\text{max \# of paths that intersect any edge}}$ minimized

NP-hard

We will develop an approx alg using randomized rounding

Approach:

1. Solve fractional version of problem (multicommodity flow)

Vars $f_i(e)$: flow from $s_i \rightsquigarrow t_i$ on edge e
$\forall e \in E$
and C

$$\min C$$

$$\sum_{\substack{e \text{ s.t.} \\ e=(u\to v) \text{ for} \\ \text{some } u}} f_i(e) = \sum_{\substack{e \text{ s.t.} \\ e=(v\to w) \text{ for} \\ \text{some } w}} f_i(e) \quad \forall v \neq s_i, t_i$$

conservation of flow

$$\sum_{\substack{e \text{ s.t.} \\ e=(s_i\to v) \\ \text{for some } v}} f_i(e) = 1$$

route 1 unit of flow from s_i to t_i

$$\sum_i f_i(e) \le C \quad \forall e$$

congestion bound

$$f_i(e) \ge 0 \quad \forall i, e$$

(2) Solve LP $\longrightarrow f_i^*(e)$, C^*

Observation: $C^* \le$ OPT congestion for integer version

(3) Decompose flows f^* into paths

$$\mathcal{P}_{s_i,t_i} = \{ P \mid P \text{ is a path from } s_i \text{ to } t_i \}$$

for each i, find set $\mathcal{P}_i^* \subseteq \mathcal{P}_{s_i,t_i}$ and $f_P^i \ \forall P \in \mathcal{P}_i^*$

s.t. $|\mathcal{P}_i^*|$ polynomial

Consider all edges with $f_i^*(e) > 0$

$$\sum_{\substack{P \in \mathcal{P}_i^* \\ \text{s.t. } e \in P}} f_P^i = f_i^*(e)$$

Find path P in graph from s_i to t_i

set $f_P^i = \min_{e \in P} f_i^*(e)$

$$\sum_{P \in \mathcal{P}_i^*} f_P^i = 1$$

remove f_P^i units of flow from all edges in P

repeat.

4. $\forall\, 1 \le i \le k$, pick path $P \in \mathcal{P}_i^*$ with prob f_P^i

call this path P_i

Output $P_1, \ldots, P_k$

Theorem $\Pr\left(\text{any edge has congestion} \ge \boxed{\frac{6 \log n}{\log\log n}}\, C^*\right) \le \frac{1}{n}$

α

Proof:

Fix $e = (u \to v)$

Let $X_i(e) = \begin{cases} 1 & \text{if } e \in P_i \\ 0 & \text{otherwise} \end{cases}$

$$E(X_i) = \sum_{\substack{P \in \mathcal{P}_i \text{ s.t. } e \in P}} f_P^i = f_i^*(e)$$

Let $X(e) = \sum_{i=1}^{k} X_i(e)$ $\Rightarrow E(X(e)) = \sum_i f_i^*(e) \leq C^*$

use Chernoff bound (1) with C^* upper bound on μ

$$\Pr(X(e) \geq (1+\delta)C^*) \leq e^{-C^*[(1+\delta)\ln(1+\delta) - \delta]}$$

Take $(1+\delta)C^* = \alpha C^*$ $\quad 1-\alpha = -\delta$

$$\Rightarrow \Pr(X(e) \geq \alpha C^*) \leq e^{-C^*[\alpha \ln\alpha + 1 - \alpha]}$$

$C^* \geq 1$

$$\leq e^{-\alpha(\ln\alpha - 1)}$$

$\alpha = \frac{6 \ln n}{\ln\ln n}$

$$\leq e^{-\left(\frac{6}{2}\right)\ln n}$$

$\frac{6 \ln n}{\ln\ln n}\underbrace{\left[\ln 6 + \ln\ln n - \ln\ln\ln n - 1\right]}_{\geq \frac{\ln\ln n}{2}}$ as before

$$= \frac{1}{n^3}$$

$$\Pr(\text{any edge has congestion} \geq \alpha C^*) \leq \sum_e \Pr(X(e) \geq \alpha C^*)$$

union bound

$$\leq n^2 \cdot \frac{1}{n^3} = \frac{1}{n}$$

$\square$

Optimal!

$\exists$ graphs with $\frac{OPT}{C^*} = \Omega\left(\frac{\log n}{\log\log n}\right)$ (integrality gap)

so can't do better with such an approach

Hardness (directed graphs)

Every poly time alg has $\Omega\left(\frac{\log n}{\log\log n}\right)$ approx ratio

assuming $NP \not\subseteq BPTIME\left(n^{O(\log\log n)}\right)$

If $C^* \geq c \ln n$ then w.h.p. congestion on every edge $\leq C^* + \sqrt{c C^* \ln n}$

$$\Pr\left(X(e) \geq C^* + \sqrt{c C^* \ln n}\right) \leq e^{-C^* \delta^2/3} = e^{-\frac{c \ln n}{3}} = n^{-\frac{c}{3}}$$

$$(1+\delta)C^* \Rightarrow \delta = \sqrt{\frac{c \ln n}{C^*}}$$

Take $c = 9$

$\Pr(\text{edge } e \text{ bad}) \leq \frac{1}{n^3}$

$\Rightarrow \Pr(\exists \text{ bad edge}) \leq \frac{1}{n}$

Observe that

$$C^* + \sqrt{c\,C^* \ln n} = O(C^*) \qquad \text{if } C^* = \Omega(\ln n)$$