13.9 Chernoff Bounds

Recall that the random variables X and Y are independent if the events $X=i$ and $Y=j$ are independent, that is $p(X=i \wedge Y=j) = p(X=i)\cdot p(Y=j)$

Consider a collection $X_1, X_2, \ldots, X_n$ of independent 0-1 valued (indicator) random variables.

Then with $X = \sum_{i=1}^{n} X_i$ we have

$$E(X) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} p_i$$

when $p_i = p(X_i = 1)$

Intuition: If $X_1, \ldots, X_n$ are independent then fluctuations are likely to cancel out so that X should stay close to $E(X)$

Our goal: derive bounds on

$$p(X > E(X)) \text{ and } p(X < E(X))$$

Called Chernoff bounds after their inventor.

(13.42) Let $X_1, X_2, \dots, X_n$ be independent 0-1 random variables

let $X = \sum X_i$ and let $\mu \geq E(X)$

Then $\forall \delta > 0$ we have $p[X > (1+\delta)\mu] < \left[\frac{e^{\delta}}{(1+\delta)^{1+\delta}}\right]^{\mu}$

proof: We use a sequence of transformations

(1) $\forall t > 0 \quad p[X > (1+\delta)\mu] = p[e^{tX} > e^{t(1+\delta)\mu}]$

as e^{ty} is monotone increasing with y

(2) By Markov's inequality we have for every non-negative random variable Y and positive number γ

$p[Y > \gamma] \leq \frac{E(Y)}{\gamma}$ so $\gamma\, p[Y > \gamma] \leq E(Y)$ (*)

Combining (1) and (*) we get

(3) $p[X > (1+\delta)\mu] = p[e^{tX} > e^{t(1+\delta)\mu}] \leq e^{-t(1+\delta)\mu} E[e^{tX}]$

minus forgotten in Video

So we need to bound $E[e^{tX}]$

$$E(e^{tX}) = E(e^{t\sum x_i}) = E(e^{\sum t x_i}) = E\left(\prod_{i=1}^{n} e^{tX_i}\right) = \prod_{i=1}^{n} E(e^{tX_i})$$

Here the last equality follows from the fact that $X_1, X_2 \dots X_n$ are independent

Recall that Y, Z independent $\Rightarrow E(Y \cdot Z) = E(Y) \cdot E(Z)$

$$E(e^{tX_i}) = p_i \cdot e^t + (1-p_i)\cdot e^{t\cdot 0} = p_i e^t + (1-p_i) = 1+p_i(e^t-1)$$

so $E(e^{tX_i}) \leq e^{p_i(e^t-1)}$ as $1+x \leq e^x$ when $x \geq 0$

and we get

$$E(e^{tX}) = \prod_{i=1}^{n} E(e^{tX_i}) \leq \prod_{i=1}^{n} e^{p_i(e^t-1)}$$

E forgotten in video

$$= e^{\sum p_i(e^t-1)}$$

$$= e^{(e^t-1)\sum p_i}$$

$$\leq e^{(e^t-1)\mu} \quad \text{as } \sum p_i = E(X) \leq \mu$$

Inserting this in $P[X > (1+\delta)\mu] \leq e^{-t(1+\delta)\mu} \cdot E(e^{tX})$

we get

$$P[X > (1+\delta)\mu] \leq e^{-t(1+\delta)\mu} \cdot e^{(e^t-1)\mu}$$

This holds for all $t > 0$ so taking $t = \ln(1+\delta)$ we get

$$P[X > (1+\delta)\mu] \leq e^{-\ln(1+\delta)\cdot(1+\delta)\mu} \cdot e^{(e^{\ln(1+\delta)}-1)\mu}$$

$$= (1+\delta)^{-(1+\delta)\mu} \cdot e^{(1+\delta-1)\mu}$$

$$= \left[\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right]^\mu$$

$\square$

Similarly one can show

13.43 Let $X_1, X_2, \ldots, X_n$ be independent 0-1 variables

$X = \sum_{i=1}^{n} X_i$ and let $\mu \geq E(X)$

Then $\forall \delta$ with $0 < \delta < 1$ we have

$$\boxed{P[X < (1-\delta)\mu] < e^{-\frac{1}{2}\mu\delta^2}}$$

Easier formulas to use

$$P(X > (1+\delta)\mu) \leq e^{-\frac{\delta^2}{3}\mu} \quad \text{when } 0 < \delta$$

$$P(X < (1-\delta)\mu) \leq e^{-\frac{\delta^2}{2}\mu} \quad \text{when } 0 < \delta < 1$$

Example of application of Chernoff bounds

X = # heads in n flips of a fair coin

We have seen that

$$E(X) = \frac{n}{2} \quad \text{and} \quad V(X) = \frac{n}{4}$$

We want to bound the probability that

$$|X - \tfrac{n}{2}| \geq \tfrac{n}{4} \qquad \left(\text{so } X \leq \tfrac{n}{4} \text{ or } X \geq \tfrac{3n}{4}\right)$$

By Chebyshev:

$$p\left[|X - \tfrac{n}{2}| \geq \tfrac{n}{4}\right] \leq \frac{V(X)}{\left(\frac{n}{4}\right)^2} = \frac{\frac{n}{4}}{\left(\frac{n}{4}\right)^2} = \frac{4}{n}$$

By Chernoff:

$$p\left(X - \tfrac{n}{2} \geq \tfrac{n}{4}\right) = p\left(X \geq \left(1 + \tfrac{1}{2}\right)\tfrac{n}{2}\right)$$
$$\leq e^{-\left(\frac{1}{2}\right)^2 \cdot \frac{1}{3} \cdot \frac{n}{2}} = e^{-\frac{n}{24}}$$

$$p\left(X - \tfrac{n}{2} \leq \tfrac{n}{4}\right) = p\left(X \leq \left(1 - \tfrac{1}{2}\right) \cdot \tfrac{n}{2}\right)$$
$$\leq e^{-\left(\frac{1}{2}\right)^2 \cdot \frac{1}{2} \cdot \frac{n}{2}} = e^{-\frac{n}{16}}$$

So $p\left[|X - \tfrac{n}{2}| \geq \tfrac{n}{4}\right] \leq e^{-\frac{n}{24}} + e^{-\frac{n}{16}} \leq 2 \cdot e^{-\frac{n}{24}}$

Chebyshev $p\left[|X-\frac{n}{2}| \geq \frac{n}{4}\right] \leq \frac{4}{n}$

Chernoff $p\left[|X-\frac{n}{2}| \geq \frac{n}{4}\right] \leq 2 \cdot e^{-\frac{n}{24}}$

n	24	240	2400
Chebyshev	1/6	1/60	1/600
Chernoff	0,73	$9 \cdot 10^{-5}$	$7{,}4 \cdot 10^{-44}$

New calculation:

set $\delta = \sqrt{\frac{6 \ln n}{n}}$ then $\frac{n}{2} \cdot \delta = \frac{1}{2}\sqrt{6 n \ln n}$

then by Chernoff bound

$$p\left[|X-\frac{n}{2}| \geq \frac{1}{2}\sqrt{6 n \ln n}\right) \leq 2 \cdot e^{-\frac{1}{3} \cdot \frac{n}{2} \cdot \frac{6 \ln n}{n}}$$

$$= \frac{2}{n}$$

so very unlikely with deviation larger than $\sqrt{\frac{6 \ln n}{n}}$ from $\frac{n}{2}$