Tail bounds in Probability Theory

X : non-negative random variable

Then for any $a > 0$ $\quad Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}$

$\mathbb{E}[X] = \mu \qquad Pr[X \geq t \cdot \mu] \leq \frac{1}{t}$

$\Big\}$ Markov

Chebyshev:

$$Pr[|X-\mu| \geq t] = Pr[(X-\mu)^2 \geq t^2]$$

Apply Markov to random variable $(X-\mu)^2$

$$\mathbb{E}[(X-\mu)^2] = Var(X)$$

$$Pr[|X-\mu| \geq t] = \frac{\mathbb{E}[(X-\mu)^2]}{t^2} = \frac{Var(X)}{t^2}$$

$X = \sum X_i \qquad \mu_i = E[X_i]$

$\mathbb{E}[X] = \sum \mathbb{E}[X_i]$ linearity of expectation

$$Var[X] = \mathbb{E}[(\sum X_i - \sum \mu_i)^2]$$

$$= \mathbb{E}[(\sum (X_i - \mu_i))^2]$$

$$= \mathbb{E}[\sum_{i,j} (X_i - \mu_i)(X_j - \mu_j)]$$

$$= \sum_{i,j} \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)]$$

$i=j \quad \mathbb{E}[(X_i-\mu_i)(X_j-\mu_j)] = \mathbb{E}[(X_i-\mu_i)^2] = Var(X_i)$

$i \neq j \quad$ If i & j are independent

$$\mathbb{E}[(X_i-\mu_i)(X_j-\mu_j)] = \mathbb{E}[(X_i-\mu_i)] \cdot \mathbb{E}[(X_j-\mu_j)]$$

$$= 0$$

If $X_1 \dots X_n$ are pairwise independent,

then $Var[X] = \sum_i Var(X_i)$

$X_1 \ldots X_n$ independent random variables

$\mathbb{E}[X_i] = \mu_i \qquad Var[X_i] = \sigma_i^2$

$$\mathbb{P}\left[\left|\sum X_i - \sum \mu_i\right| \geq a\right] \leq \frac{\sum_{i=1}^{n} \sigma_i^2}{a^2}$$

$X_1 \ldots X_n$ iid expectation μ variance σ^2

$$\mathbb{E}\left[\frac{\sum X_i}{n}\right] = \mu$$

$$Var\left[\frac{\sum X_i}{n}\right] = \frac{\sigma^2}{n}$$

$$P\left[\left|\frac{\sum X_i}{n} - \mu\right| \geq \varepsilon\right] \leq \frac{\sigma^2}{n \cdot \varepsilon^2} \quad \leftarrow$$

$$Pr[X_i = a_i, X_j = a_j] = Pr[X_i = a_i] \cdot Pr[X_j = a_j]$$

$$\forall S \subseteq [n] \quad Pr\left[\bigwedge_{i \in S} (X_i = a_i)\right] = \prod_{i \in S} Pr[X_i = a_i]$$

<u>Chernoff Bounds:</u>

sums of independent 0-1 variables

<u>Thm:</u> $X = \sum_{i=1}^{n} X_i \qquad X_i = 1$ w. prob. p_i, $X_i = 0$ w. prob $1-p_i$

all X_i independent

$$\mu = \mathbb{E}[X] = \sum_{i=1}^{n} p_i$$

(i) Upper tail: $\mathbb{P}(X \geq (1+\delta)\mu) \leq e^{-\frac{\delta^2}{2+\delta}\mu}$ for all $\delta > 0$

(ii) Lower tail: $\mathbb{P}(X \leq (1-\delta)\mu) \leq e^{-\mu\delta^2/2}$ for all $0 < \delta < 1$

Common combination:

$$Pr(|X-\mu| \geq \delta\mu) \leq 2e^{-\mu\delta^2/3} \quad \text{for all } 0 < \delta < 1$$

Fair coin, n tosses, $S_n = \#$ heads

$$\mathbb{P}\left[\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon \right] \leq \frac{1}{4n\varepsilon^2}$$

e.g. $$\mathbb{P}\left[\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \frac{1}{4} \right] \leq \frac{4}{n}$$

Chernoff: $$\mathbb{P}\left(|S_n - n/2| \geq \delta \cdot \left(\frac{n}{2}\right) \right) \leq 2e^{-n\delta^2/6}$$

$\delta = 1/2$

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \frac{1}{4} \right) \leq 2e^{-n/24}$$

$$\delta = \sqrt{\frac{6\ln n}{n}}$$

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \frac{1}{2}\sqrt{\frac{6\ln n}{n}} \right) \leq 2e^{-\ln n} = \frac{2}{n}$$

$$\delta = 2\sqrt{6\ln(n)/n}$$

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \sqrt{\frac{6\ln n}{n}} \right) \leq 2e^{-4\ln n} = \frac{2}{n^4}$$

Proof: $X = \sum X_i$

$$\Pr(X \geq a) = \mathbb{P}(e^{sX} \geq e^{sa}) \qquad s > 0$$

$$\leq \frac{\mathbb{E}[e^{sX}]}{e^{sa}} \qquad \text{(Markov)}$$

$$\Pr(X \leq a) = \mathbb{P}(e^{-sX} \geq e^{-sa})$$

$$\leq \frac{\mathbb{E}[e^{-sX}]}{e^{-sa}}$$

$M_X(s) = \mathbb{E}[e^{sX}]$ moment generating f^n of X

$$M_X(s) = \mathbb{E}\left[1 + sX + \frac{1}{2}s^2X^2 + \frac{1}{3!}s^3X^3 + \dots\right) = \sum_{i=0}^{\infty} \frac{1}{i!} s^i \underbrace{\mathbb{E}[X^i]}_{\text{moments}}$$

Lemma If $X = \sum_{i=1}^{n} X_i$, $X_1 \ldots X_n$ independent

$$M_X(s) = \prod_{i=1}^{n} M_{X_i}(s)$$

Pf:

$$M_X(s) = \mathbb{E}[e^{sX}] = \mathbb{E}[e^{s \sum X_i}]$$
$$= \mathbb{E}\left(\prod_{i=1}^{n} e^{sX_i}\right)$$
$$= \prod_{i=1}^{n} \mathbb{E}[e^{sX_i}] = \prod_{i=1}^{n} M_{X_i}(s)$$

Lemma Y be random variable that is
1 w. prob. p
0 w. prob $1-p$

$$M_Y(s) = \mathbb{E}[e^{sY}] \leq e^{p(e^s-1)}$$

Pf:

$$\mathbb{E}[e^{sY}] = p \cdot e^s + (1-p) \cdot 1$$
$$= 1 + p(e^s - 1)$$
$$\leq e^{p(e^s-1)} \qquad \boxed{1+t \leq e^t}$$

Pf of thm:

$$\mathbb{E}[e^{sX}] = \prod_{i=1}^{n} \mathbb{E}[e^{sX_i}] \leq \prod_{i=1}^{n} e^{p_i(e^s-1)}$$
$$= e^{(e^s-1)\sum_{i=1}^{n} p_i}$$
$$= e^{(e^s-1)\mu}$$

$$\mathbb{P}[X \geq (1+\delta)\mu] \leq \frac{\mathbb{E}[e^{sX}]}{e^{s(1+\delta)\mu}} \leq \frac{e^{(e^s-1)\mu}}{e^{s(1+\delta)\mu}}$$
$$= \left(\frac{e^{e^s-1}}{e^{s(1+\delta)}}\right)^{\mu}$$

Exponent $e^s - 1 - s(1+\delta)$

$$\frac{d}{ds}(\quad) = e^s - (1+\delta) = 0 \Rightarrow s = \ln(1+\delta)$$

Sub. $s = \ln(1+\delta)$

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

Using $\ln(1+x) \geq \frac{x}{1+x/2}$

$$\frac{e^{\delta}}{(1+\delta)^{1+\delta}} \leq e^{-\frac{\delta^2}{2+\delta}}$$

Lower tail: $s = \ln(1-\delta)$

$$\ln(1-\delta) \geq -\delta + \frac{\delta^2}{2}$$

$X_1 \ldots X_n$ independent

Chernoff bounds holds also when $X_1 \ldots X_n$ negatively correlated.

$$\forall S \subseteq [n] \quad \mathbb{E}\left[\prod_{i \in S} X_i\right] \leq \prod_{i \in S} \mathbb{E}[X_i]$$

$$\mathbb{E}[e^{sX}] = \prod_{i=1}^{n} \mathbb{E}[e^{sX_i}]$$

negatively correlated: $\mathbb{E}[e^{sX}] \leq \prod_{i=1}^{n} \mathbb{E}[e^{sX_i}]$

$\mathbb{E}[X^k]$

Let $Y_1 \ldots Y_n$ be independent 0-1 random variables

$\mathbb{E}[Y_i] = \mathbb{E}[X_i]$

$Y = \sum Y_i$

$$\mathbb{E}[X^k] = \mathbb{E}[(X_1 + X_2 \ldots X_n)^k]$$

$$= \sum_{\alpha} \mathbb{E}[X_1^{\alpha_1} X_2^{\alpha_2} \ldots X_n^{\alpha_n}]$$

$$\alpha_i \geq 0 \quad \sum \alpha_i = k$$

$$\mathbb{E}[X^k] \leq \sum_\alpha \mathbb{E}[X_1^{\alpha_1}]\,\mathbb{E}[X_2^{\alpha_2}] \dots \mathbb{E}[X_n^{\alpha_n}]$$

$$= \sum_\alpha \mathbb{E}[Y_1^{\alpha_1}]\,\mathbb{E}[Y_2^{\alpha_2}] \dots \mathbb{E}[Y_n^{\alpha_n}]$$

$$= \sum_\alpha \mathbb{E}[Y_1^{\alpha_1} Y_2^{\alpha_2} \dots Y_n^{\alpha_n}] = \mathbb{E}[Y^k]$$

$$e^{tX} = 1 + tX + \frac{t^2X^2}{2} \dots$$

$$\mathbb{E}[e^{tX}] \leq \mathbb{E}[e^{tY}] = \prod_{i=1}^n \mathbb{E}[e^{tY_i}] = \prod_{i=1}^n \mathbb{E}[e^{tX_i}]$$

$$\mathbb{E}[e^{tX}] \leq \prod_{i=1}^n \mathbb{E}[e^{tX_i}]$$

<u>Appl^n</u>: $S_1 \dots S_N \subseteq [n]$ $\quad |S_i| = \frac{n}{10}$

$$|S_i \cap S_j| < \frac{n}{20}$$

$N = 2^{cn}$ many sets.

<u>Approach 1</u>: Sample elements independently w. prob. $\frac{1}{10}$

$|S_a \cap S_b| = \sum X_i = X \quad X_i = 1$ if $i \in S_a \cap S_b$

$\mathbb{E}[|S_a \cap S_b| = \frac{n}{100}]$

does not exceed $\frac{n}{20}$ w.h.p.

$$\mu = \frac{n}{100} \qquad (1+\delta) = 5, \quad (1+\delta)\frac{n}{100} = \frac{n}{20}$$

$$\Pr[X > (1+\delta)\mu] \leq \left(\frac{e^\delta}{(1+\delta)^{(1+\delta)}}\right)^\mu \leq e^{-c_1 n}$$

$$\delta = 4, \quad \mu = \frac{n}{100}$$

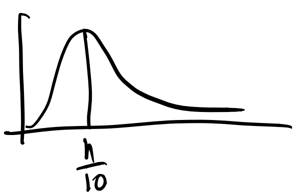
Sample $e^{c_2 n}$ such sets. $\quad c_2 < \frac{c_1}{2}$

$\#$ pairs $\leq e^{2c_2 n} < e^{c_1 n}$

$$\Pr[\exists \text{ pair } S_a, S_b \text{ st. } |S_a \cap S_b| > \tfrac{n}{20}] \leq (\#\text{pairs}) \cdot e^{-c_1 n}$$
$$\leq e^{2c_2 n} \cdot e^{-c_1 n}$$
$$\ll 1$$

$E_1 \ldots E_m$

$$\Pr[\cup E_i] \leq \sum \Pr[E_i]$$



$\approx \frac{1}{\sqrt{n}} > \frac{1}{n}$

$$\frac{e^{c_2 n}}{2n}$$

Approach 2: Sample exactly $\frac{n}{10}$ elements.

$X_1 \; X_2 \; \ldots \; X_{n/10}$

not independent random variables.

Can show: Negatively correlated!