Chernoff Bounds

They're tools from modern probability theory that can be used for the analysis of randomized algorithms

They're more powerful version of Markov's lemma

Phenomenon of "concentration of measure":

Toss a coin
- one time ⟶ outcome is unpredictable
- 1000 times ⟶ outcome is sharply predictable!

Application: $T(n)$ guaranteed to be concentrated around some value

In many cases the study of $\Pr\left(T(n) > c \cdot f(n)\right)$ can be rephrased as the study of the distribution of some sum of random variables

Indicator random variable : 0-1 random variable

$$Y = \begin{cases} 1 & \text{if trial is a success} \\ 0 & \text{otherwise} \end{cases}$$

In general

$$X = \sum_{i=1}^{n} X_i \qquad X_i \text{ indicator random variables}$$

we'll usually have that X_i's are <u>independent</u>

$$Pr(X_i = 1) = p_i$$

$$E[X] = E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i] =$$

$$= \sum_{i=1}^{n} p_i \overset{\text{notation}}{=} \mu$$

We want to analyze the probability that X deviates from $E[X]$

$$Pr\left(X > (1+\delta)\mu\right) \underset{\text{Markov}}{\leq} \frac{E[X]}{(1+\delta)\mu} = \frac{\mu}{(1+\delta)\mu} = \boxed{\frac{1}{1+\delta}}$$

usually not a very good bound

A more powerful probabilistic tool:

Chernoff bound: let $X_1, X_2, \dots, X_n$ independent indicator random variables where $E[X_i] = p_i$, $0 < p_i < 1$
Let $X = \sum_{i=1}^{n} X_i$ and $\mu = E[X]$. Then $\forall \delta > 0$

$$\Pr\left(X > (1+\delta)\mu\right) < \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

Example: coin tossing

n coin flips $\longrightarrow X_1, X_2, \dots, X_n$

$$\Pr(X_i = \underset{\text{head}}{\underset{\shortparallel}{1}}) = \frac{1}{2} \quad \forall i$$

$$X = \sum_{i=1}^{n} X_i = n^\circ \text{ of heads}$$

$$E[X] = \frac{n}{2}$$

Possible question: what's the prob. of getting more than $\frac{3}{4}n$ heads?

Let's apply:

1) Markov

$$\Pr\Big(X > \underbrace{\left(1+\tfrac{1}{2}\right)\mu}_{=3/4\,n}\Big) \leq \frac{\mu}{\left(1+\frac{1}{2}\right)\mu} = \underbrace{\frac{2}{3}}_{\text{Constant}}$$

2) Chernoff

$$\Pr\Big(X > \underbrace{\left(1+\tfrac{1}{2}\right)\mu}_{3/4\,n}\Big) < \left(\frac{e^{\frac{1}{2}}}{\left(3/2\right)^{3/2}}\right)^{\frac{n}{2}} < \underbrace{(0.95)^{n}}_{\text{exponential!}}$$

Variants of Chernoff bounds

1) $Pr\left(X < (1-\delta)\mu\right) < e^{-\frac{\mu\delta^2}{2}}$ $\quad 0<\delta\leq 1$

2) $Pr\left(X > (1+\delta)\mu\right) < e^{-\frac{\mu\delta^2}{2}}$ $\quad 0<\delta\leq 2e-1$