COL 758: Advanced Algorithms

Lecture 06

Important Tools

Markov’s Inequality

Chernoff bound

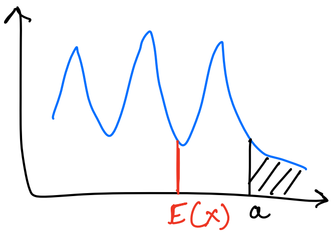
Markov's Inequality

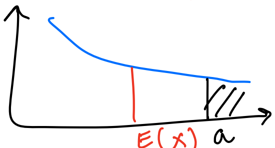
For any non-negative (discrete/continuous) random variable X and any $a > 0$,

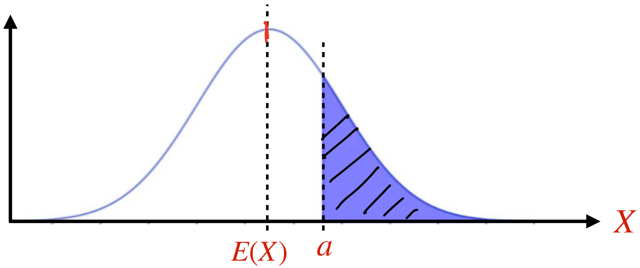
$$Prob(\, X \geq a \,) \leqslant \frac{E(X)}{a}$$

Note:

The inequality is useful only when a is larger than $E(X)$







Markov's Inequality

For any non-negative (discrete/continuous) random variable X and any $a > 0$,

$$Prob(\ X \geq a\ ) \leqslant \frac{E(X)}{a}$$

Proof:

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot Prob(\omega).$$

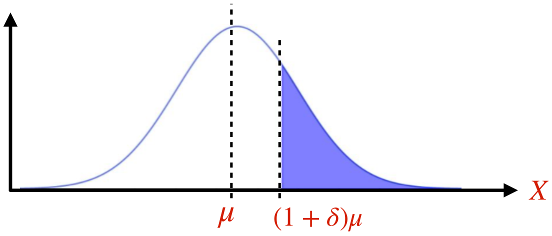
Divide Ω into two parts Ω_1 and Ω_2 such that for

(i) $\omega \in \Omega_1, \quad X(\omega) < a.$

(ii) $\omega \in \Omega_2, \quad X(\omega) \geq a.$

Then, $E(X) \geq \sum_{\omega \in \Omega_2} X(\omega) \cdot Prob(\omega) \geq \sum_{\omega \in \Omega_2} a \cdot Prob(\omega) = a \cdot Prob(X \geq a)$

Chernoff Bound (a)



Let $X_1, X_2, \ldots, X_n$ be n independent Binary random variables with parameters $p_1, p_2, \ldots, p_n$, that is

$$X_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i \end{cases}$$

Let X be the sum of random variables $X_1, X_2, \ldots, X_n$, and let

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

Then, for any $\delta > 0$,

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

Eg. A fair coin is tossed n times. X = No of Heads

Prob$\left(X \geq \frac{n}{2} + 10 \log n\right)$?

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

$\delta = \frac{20 \log n}{n}$

Proof of Chernoff Bound (a)

Let $X_1, X_2, \ldots, X_n$ be n independent Binary random variables with parameters $p_1, p_2, \ldots, p_n$, that is

$$X_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i \end{cases}$$

Let X be the sum of random variables $X_1, X_2, \ldots, X_n$, and let

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

Then, for any $\delta > 0$,

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

Tools

1. Markov's Inequality

2. If Y and Z are independent r.v., then
$E(YZ) = E(Y) \cdot E(Z)$

3. Fact: $e^x \geq 1 + x$

4. At local minima, differentiation is zero

(t is non-ve)

$$\text{Prob}(X \geq (1+\delta)\mu) = \text{Prob}(tX \geq (1+\delta)t\mu) \quad \text{for } t \in \mathbb{R}^+$$

$$= \text{Prob}(e^{tX} \geq e^{t(1+\delta)\mu})$$

$$\leq \frac{E(e^{tX})}{e^{t(1+\delta)\mu}} \qquad (\text{Tool 1 - Markov})$$

			e^{tX_i}
$X_i =$	1	p_i	e^t
	0	$1-p_i$	1

$$E(e^{tX}) = E(e^{tX_1 + tX_2 + \dots + tX_n})$$

$$= \prod_{i=1}^{n} E(e^{tX_i}) \qquad (\text{Tool 2 - } X_1 \dots X_n \text{ are indep.})$$

$$= \prod_{i=1}^{n} (e^t p_i + 1 - p_i)$$

$$= \prod_{i=1}^{n} (1 + (e^t - 1)p_i) \leq \prod_{i=1}^{n} e^{[e^t - 1]p_i} \qquad (\text{Tool 3} \quad e^x \geq 1+x)$$

$$\leq e^{[e^t-1]\sum p_i} = e^{[e^t-1]\mu}$$

$$\text{Prob}(X \geq (1+\delta)\mu) = \text{Prob}(tX \geq (1+\delta)t\mu) \quad \text{for } t \in \mathbb{R}^{+}$$

(t is non-ve)

$$= \text{Prob}\left(e^{tX} \geq e^{t(1+\delta)\mu}\right)$$

$$\leq \frac{E(e^{tX})}{e^{t(1+\delta)\mu}}$$

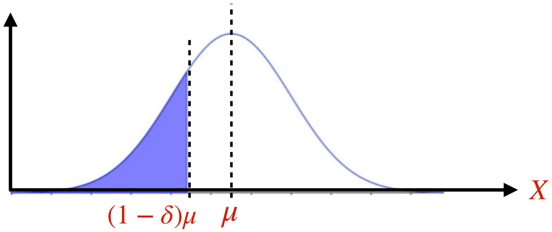
$$\leq \frac{e^{[e^t - 1]\mu}}{e^{t(1+\delta)\mu}}$$

(see last page)

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

$$\text{Prob}(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

Chernoff Bound (b)



Let $X_1, X_2, \ldots, X_n$ be n independent Binary random variables with parameters $p_1, p_2, \ldots, p_n$, that is

$$X_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i \end{cases}$$

Let X be the sum of random variables $X_1, X_2, \ldots, X_n$, and let

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

Then, for any $\delta > 0$,

$$P(X \leq (1-\delta)\mu) \;\leq\; \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}} \right)^{\mu}$$

Homework

Hint $\text{Prob}(X \leq (1-\delta)\mu) = \text{Prob}(tX \geq (1-\delta)t\mu) = \text{Prob}(e^{tX} \geq e^{(1-\delta)t\mu})$

for $t < 0$

Chernoff Bound (More usable forms)

Let $X_1, X_2, \ldots, X_n$ be n independent Binary random variables with parameters $p_1, p_2, \ldots, p_n$, that is

$$X_i = \begin{cases} 1 & \text{with probability } p_i \\ 0 & \text{with probability } 1 - p_i \end{cases}$$

Let X be the sum of random variables $X_1, X_2, \ldots, X_n$, and let $\mu = E(X) = \sum_{i=1}^{n} p_i$.

Then, for any $\delta > 0$,

$$\star \quad P(X \geq (1+\delta)\mu) \;\leq\; e^{-\mu\delta^2/(2+\delta)}$$

$$\star \quad P(X \leq (1-\delta)\mu) \;\leq\; e^{-\mu\delta^2/2}$$

Ques: Suppose $A = \{a_1, a_2 \dots a_n\}$ is a set, and $J \subseteq A$ is random set of size $k = 16 \log n$ picked uniformly randomly with replacement.

(a) Show that $E(|J \cap \{a_1 \dots a_{\frac{n}{4}}\}|) = 4 \log n$.

(b) Obtain a bound on $\text{Prob}(|J \cap \{a_1 \dots a_{\frac{n}{4}}\}| \geq 8 \log n)$?

(a) For $i \in [1, 16 \log n]$,

$$Y_i = \begin{cases} 1 & \text{if } i^{th} \text{ element of } J \text{ lies in } \{a_1, a_2 \dots a_{\frac{n}{4}}\} \\ 0 & \text{o/w} \end{cases}$$

$E(Y_i) = 1/4$

Note $Y = Y_1 + Y_2 + \dots + Y_{16 \log n}$ is size of set $|J \cap \{a_1 \dots a_{\frac{n}{4}}\}|$ and $E(Y) = 4 \log n$.

(b)

Prove that $Y_1\ Y_2 \dots Y_{16 \log n}$ are independent.

Use this in Chernoff bound to obtain bound on

$$\text{Prob}(Y \geq \underbrace{8 \log n}_{\substack{\text{|||} \\ 2E(Y)}}) \leq \_\_\_\_\_\_$$

Homework