CMPT 210: Probability and Computing

Lecture 22

Sharan Vaswani

March 30, 2023

Sums of Random Variables

If we know that the r.v X is (i) non-negative and (ii) $\mathbb{E}[X]$, we can use Markov's Theorem to bound the probability of deviation from the mean.

If we know both (i) $\mathbb{E}[X]$ and (ii) $\text{Var}[X]$, we can use Chebyshev's Theorem to bound the probability of deviation.

In many cases the random variable of interest is a sum of r.v's (e.g., for the voter poll application), and we can use the Chernoff bound to obtain tighter bounds on the deviation from the mean.

Chernoff Bound: Let $T_1, T_2, \ldots, T_n$ be mutually independent r.v's such that $0 \leq T_i \leq 1$ for all i. If $T := \sum_{i=1}^{n} T_i$, for all $c \geq 1$ and $\beta(c) := c \ln(c) - c + 1$,

$$\Pr[T \geq c\mathbb{E}[T]] \leq \exp(-\beta(c)\,\mathbb{E}[T])$$

If $T_i \sim \text{Ber}(p)$ and are mutually independent, then $T_i \in \{0, 1\}$ and we can use the Chernoff bound to bound the deviation from the mean for $T \sim \text{Bin}(n, p)$. In general, if $T_i \in [0, 1]$, the Chernoff Bound can be used even if the T_i's have different distributions!

Chernoff Bound – Binomial Distribution

Q: Bound the probability that the number of heads that come up in 1000 independent tosses of a fair coin exceeds the expectation by 20% or more.

Let T_i be the r.v. for the event that coin i comes up heads, and let T denote the total number of heads. Hence, $T = \sum_{i=1}^{1000} T_i$. For all i, $T_i \in \{0, 1\}$ and are mutually independent r.v's. Hence, we can use the Chernoff Bound.

We want to compute the probability that the number of heads is larger than the expectation by 20% meaning that $c = 1.2$ for the Chernoff Bound. Computing $\beta(c) = c\ln(c) - c + 1 \approx 0.0187$. Since the coin is fair, $\mathbb{E}[T] = 1000\,\frac{1}{2} = 500$. Plugging into the Chernoff Bound,

$$\Pr[T \geq c\mathbb{E}[T]] \leq \exp(-\beta(c)\mathbb{E}[T]) \implies \Pr[T \geq 1.2\,\mathbb{E}[T]] \leq \exp(-(0.0187)\,(500)) \approx 0.0000834.$$

Comparing this to using Chebyshev's inequality,

$$\Pr[T \geq c\mathbb{E}[T]] = \Pr[T - \mathbb{E}[T] \geq (c-1)\mathbb{E}[T]] \leq \Pr[|T - \mathbb{E}[T]| \geq (c-1)\,\mathbb{E}[T]]$$

$$\leq \frac{\mathrm{Var}[T]}{(c-1)^2\,(\mathbb{E}[T])^2} = \frac{1000\,\frac{1}{4}}{(1.2-1)^2(500^2)} = \frac{250}{0.2^2\,500^2} = \frac{250}{10000} = 0.025.$$

Chernoff Bound – Lottery Game

Q: Pick-4 is a lottery game in which you pay \$1 to pick a 4-digit number between 0000 and 9999. If your number comes up in a random drawing, then you win \$5,000. Your chance of winning is 1 in 10000. If 10 million people play, then the expected number of winners is 1000. When there are 1000 winners, the lottery keeps \$5 million of the \$10 million paid for tickets. The lottery operator's nightmare is that the number of winners is much greater – especially at the point where more than 2000 win and the lottery must pay out more than it received. What is the probability that will happen? (Assume that the players' picks and the winning number are random, independent and uniform)

Let T_i be an indicator for the event that player i wins. Then $T := \sum_{i=1}^{n} T_i$ is the total number of winners. Using the independence assumptions, we can conclude that T_i are independent, as required by the Chernoff bound.

We wish to compute $\Pr[T \geq 2000] = \Pr[T \geq 2\mathbb{E}[T]]$. Hence $c = 2$ and $\beta(c) \approx 0.386$. By the Chernoff bound,

$$\Pr[T \geq 2\mathbb{E}[T]] \leq \exp(-\beta(c)\mathbb{E}[T]) = \exp(-(0.386)\,1000) < \exp(-386) \approx 10^{-168}$$

Questions?

Chernoff Bound – Proof

Chernoff Bound: Let $T_1, T_2, \ldots, T_n$ be mutually independent r.v's such that $0 \le T_i \le 1$ for all i. If $T := \sum_{i=1}^{n} T_i$, for all $c \ge 1$ and $\beta(c) := c\ln(c) - c + 1$,

$$\Pr[T \ge c\mathbb{E}[T]] \le \exp(-\beta(c)\,\mathbb{E}[T])$$

Proof: We want to compute $\Pr[T \ge c\mathbb{E}[T]] = \Pr[f(T) \ge f(c\mathbb{E}[T])]$ where f is a one-one monotonically non-decreasing function. For $c \ge 1$, choosing $f(T) = c^T$ and using Markov's Theorem,

$$\Pr[T \ge c\mathbb{E}[T]] = \Pr[c^T \ge c^{c\mathbb{E}[T]}] \le \frac{\mathbb{E}[c^T]}{c^{c\mathbb{E}[T]}}$$

$$\le \frac{\exp((c-1)\,\mathbb{E}[T])}{c^{c\mathbb{E}[T]}} \qquad \text{(To prove next: } \mathbb{E}[c^T] \le \exp((c-1)\,\mathbb{E}[T]))$$

$$= \frac{\exp((c-1)\,\mathbb{E}[T])}{\exp(\ln(c^{c\mathbb{E}[T]}))} = \frac{\exp((c-1)\,\mathbb{E}[T])}{\exp(c\mathbb{E}[T]\,\ln(c))} = \exp\left(-(c\ln(c) - c + 1)\,\mathbb{E}[T]\right)$$

$$\implies \Pr[T \ge c\mathbb{E}[T]] \le \exp\left(-\beta(c)\,\mathbb{E}[T]\right)$$

The proof would be done if we prove that $\mathbb{E}[c^T] \le \exp((c-1)\,\mathbb{E}[T])$ and we do this next.

Chernoff Bound – Proof

Claim: $\mathbb{E}[c^T] \le \exp((c-1)\,\mathbb{E}[T])$

$$\mathbb{E}[c^T] = \mathbb{E}[c^{\sum_{i=1}^{n} T_i}] = \mathbb{E}\left[\prod_{i=1}^{n} c^{T_i}\right] = \prod_{i=1}^{n} \mathbb{E}[c^{T_i}]$$

(Expectation of product of mutually independent r.v's is equal to the product of the expectation.)

$$\le \prod_{i=1}^{n} \exp((c-1)\,\mathbb{E}[T_i]) \qquad \text{(To prove next: } \mathbb{E}[c^{T_i}] \le \exp((c-1)\,\mathbb{E}[T_i])\text{)}$$

$$= \exp\left((c-1)\sum_{i=1}^{n} \mathbb{E}[T_i]\right) = \exp\left((c-1)\,\mathbb{E}\left[\sum_{i=1}^{n} T_i\right]\right)$$

(Linearity of Expectation)

$$\implies \mathbb{E}[c^T] \le \exp((c-1)\mathbb{E}[T])$$

The proof would be done if we prove that $\mathbb{E}[c^{T_i}] \le \exp((c-1)\,\mathbb{E}[T_i])$ and we do this next.

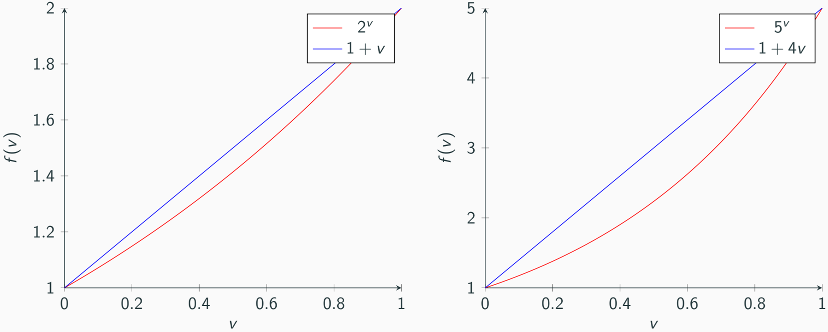
Chernoff Bound – Proof

Claim: $\mathbb{E}[c^{T_i}] \leq \exp((c-1)\,\mathbb{E}[T_i])$

$$\mathbb{E}[c^{T_i}] = \sum_{v \in \text{Range}(T_i)} c^v \Pr[T_i = v] \leq \sum_{v \in \text{Range}(T_i)} (1 + (c-1)v) \Pr[T_i = v]$$

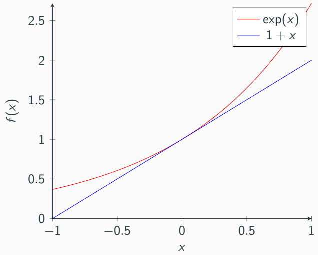
(Since $T_i \in [0,1]$ and $c^v \leq 1 + (c-1)v$ for all $v \in [0,1]$.)

For $c = 2$ and $c = 5$,



Chernoff Bound – Proof

$$\mathbb{E}[c^{T_i}] \le \sum_{v \in \text{Range}(T_i)} (1 + (c-1)v) \Pr[T_i = v]$$
$$= \sum_{v \in \text{Range}(T_i)} \Pr[T_i = v] + (c-1) \sum_{v \in \text{Range}(T_i)} v \Pr[T_i = v]$$
$$= 1 + (c-1)\,\mathbb{E}[T_i] \le \exp((c-1)\mathbb{E}[T_i]) \qquad \text{(Since } 1 + x \le \exp(x) \text{ for all } x\text{)}$$
$$\implies \mathbb{E}[c^{T_i}] \le \exp((c-1)\,\mathbb{E}[T_i])$$



Hence we have proved the Chernoff Bound!

Comparing the Bounds

For r.v's $T_1, T_2, \ldots T_n$, if $T_i \in \{0,1\}$ and $\Pr[T_i = 1] = p_i$. Define $T := \sum_{i=1}^n T_i$. By linearity of expectation, $\mathbb{E}[T] = \sum_{i=1}^n p_i$. For $c \geq 1$,

Markov's Theorem: $\Pr[T \geq c\mathbb{E}[T]] \leq \frac{1}{c}$. Does not require T_i's to be independent.

Chebyshev's Theorem:

$$\Pr[T - \mathbb{E}[T] \geq x] \leq \Pr[|T - \mathbb{E}[T]| \geq x] \leq \frac{\text{Var}[T]}{x^2}$$

$$\implies \Pr[T - \mathbb{E}[T] \geq (c-1)\mathbb{E}[T]] \leq \frac{\text{Var}[T]}{(c-1)^2\,(\mathbb{E}[T])^2)} \qquad (x = (c-1)\,\mathbb{E}[T])$$

If the T_i's are pairwise independent, by linearity of variance, $\text{Var}[T] = \sum_{i=1}^n p_i\,(1-p_i)$. Hence, $\Pr[T \geq c\mathbb{E}[T]] \leq \frac{\sum_{i=1}^n p_i\,(1-p_i)}{(c-1)^2\left(\sum_{i=1}^n p_i\right)^2}$. If for all i, $p_i = 1/2$, then, $\Pr[T \geq c\mathbb{E}[T]] \leq \frac{1}{(c-1)^2}$.

Chernoff Bound: If T_i' are mutually independent, then,
$\Pr[T \geq c\mathbb{E}[T]] \leq \exp(-\beta(c)\,\mathbb{E}[T]) = \exp\left(-(c\ln(c) - c + 1)\,\left(\sum_{i=1}^n p_i\right)\right)$. If for all i, $p_i = 1/2$, $\Pr[T \geq c\mathbb{E}[T]] \leq \exp\left(-\frac{n\,(c\ln(c)-c+1)}{2}\right)$.

Questions?