EE 382V: Parallel Algorithms **Summer 2017**

Lecture 16: July 1

Lecturer: Vijay Garg *Scribe: John Martinez*

16.1 Introduction

This section will involve the notion of estimating probability (bound calculation) by using the following methods: `Markov, Chebyshev` and `Chernoff`. The following notes are based on *Probability and Computation* by Mitzenmacher and Upfal, pages 44-49 & 64.

16.2 Flipping coin example and other preliminaries

Given a coin x_i, where 0 is tails and heads is 1, the probability $\mathbb{P}$ that once flipped the coin will be heads can be denoted:

$$\mathbb{P}[x_i = 1] = \frac{1}{2}, \quad x_i = \begin{cases} 0 \\ 1 \end{cases}$$

This type of experiment, where only two outcomes are possible, each with the same probability, is called a `Bernoulli (binomial) trial`. The expected value $\mathbb{E}$ of this trial using coin x_i, also called the average, is:

$$\mathbb{E}[x_i] = \frac{1}{2}$$

If you flip it n times then we obtain a random variable X:

$$X = \sum_{i=1}^{n} x_i$$

16.2.1 Linearity of Expectation aka Golden Rule

The expected value of a sum of random variables is equal to the sum of the individual expectation. Hence, given random variables x_i and not assuming any relationship among them, then:

$$\forall x_i \mid X = x_1 + x_2 + \ldots + x_n \Longleftrightarrow \mathbb{E}[X] = \mathbb{E}[x_1] + \mathbb{E}[x_2] + \ldots + \mathbb{E}[x_n]$$

Using the coin example, and since $\sum \mathbb{E} = \mathbb{E} \sum$, then:

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[x_i] = \frac{1}{2} \sum_{i=1}^{n} 1 = \frac{n}{2}$$

16.2.2 Variance & Covariance

`Variace` can be informally thought of as the measure of how far a set of random numbers are spread out from their mean. Formally, given that X is a random variable, $\mathbb{E}[X]$ is a constant, $\mathbb{E}[\mathbb{E}[X]] = \mathbb{E}[X]$, and $\mathbb{E}[cX] = c\,\mathbb{E}[X]$, then :

$$\begin{aligned} Var(X) &= \mathbb{E}[\,(X - \mathbb{E}[X])^2\,] \\ &= \mathbb{E}[\,X^2 + (\mathbb{E}[X])^2 - 2X\,\mathbb{E}[X]\,] \\ &= \mathbb{E}[X^2] + \mathbb{E}[X]^2 - 2\,\mathbb{E}[X]\,\mathbb{E}[X] \\ &= \mathbb{E}[X^2] + \mathbb{E}[X]^2 - 2\,\mathbb{E}[X]^2 \\ &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \end{aligned}$$

For two random variables X and Y:

$$\begin{aligned} Var(X+Y) &= \mathbb{E}[\,(X + Y - \mathbb{E}[X+Y])^2\,] \\ &= \mathbb{E}[\,(X + Y - \mathbb{E}[X] - \mathbb{E}[Y])^2\,] \\ &= \mathbb{E}[\,(X - \mathbb{E}[X])^2 + (Y - \mathbb{E}[Y])^2 + 2(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])\,] \\ &= Var(X) + Var(Y) + 2\,cov(X,Y) \end{aligned}$$

A property of a joint probability distribution, `covariance` is a measure of the joint variability of two random variables (like X and Y above). Formally, it is defined as $cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y]$.

If X and Y are *independent* then $cov(X,Y) = 0$ and $\mathbb{P}[X = x, Y = y] = \mathbb{P}[X = x] \cdot \mathbb{P}[Y = y]$. Moreover:

$$\begin{aligned} \mathbb{E}[X,Y] &= \sum_{x,y} xy\,\mathbb{P}[X = x, Y = y] \\ &= \sum_{x,y} xy\,\mathbb{P}[X = x]\,\mathbb{P}[Y = y] \\ &= \sum_{x,y} x\,\mathbb{P}[X = x] \cdot y\,\mathbb{P}[Y = y] \\ &= \mathbb{E}[X] \cdot \mathbb{E}[Y] \end{aligned}$$

Using the coin example above, let $p = \mathbb{P}[x = 1] = \frac{1}{2}$ and $q = \mathbb{P}[x = 0] = 1 - p = \frac{1}{2}$ then:

$$\mathbb{E}[X] = 1 \cdot p + 0 \cdot q = p = \frac{1}{2}$$

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2 = p(1-p) = p \cdot q = \frac{1}{4}$$

16.2.3 More on Bernoulli trials

$$x_i = \begin{cases} 0, \ \mathbb{P} = p \\ 1, \ \mathbb{P} = q = 1 - p \end{cases}$$

$$X = \sum_{i=1}^{n} x_i$$

$$\mathbb{E}[X] = \mathbb{E}[\sum_{i}^{n} x_i] = \sum_{i}^{n} \mathbb{E}[x_i] = \sum_{i}^{n} p = n \cdot p$$

$$Var(X) = Var(\sum_{i}^{n} x_i) = \sum_{i}^{n} Var(x_i) = \sum_{i}^{n} p \cdot q = n \cdot p \cdot q$$

A `binomial distribution` gives the probability distribution of obtaining exactly m successes out of n Bernoulli trials. In other words, the probability that a random variable X with binomial distribution $B(n,p)$ is equal to the value m, where $m = 0, 1, \ldots, n$, is given by

$$\begin{aligned} \mathbb{P}[X = m] &= \binom{n}{m} p^m \cdot q^{n-m} \\ &= \frac{n!}{(n-m)!\, m!}\, p^m \cdot q^{n-m} \end{aligned}$$

16.3 Markov's Inequality

This inequality relates probabilities to expectations by giving an upper bound for the probability that a non-negative function of a random variable is greater than or equal to some positive constant. Thus, given a random variable $x \geq 0$:

$$\forall a > 0 \mid \mathbb{P}[x \geq a] \leq \frac{\mathbb{E}[x]}{a} \;\equiv\; \mathbb{P}[x \geq k\,\mathbb{E}[x]] \leq \frac{1}{k}$$

To prove it, let $I = \begin{cases} 0, \; x \geq a \\ 1, \; otherwise \end{cases}$, hence $I \leq \frac{x}{a}, \forall x \geq 0, \forall a > 0$. It follows:

$$\mathbb{E}[I] = 1 \cdot \mathbb{P}[x \geq a] + 0 \cdot (1 - \mathbb{P}[x \geq a]) = \mathbb{P}[x \geq a]$$

$$\mathbb{E}[I] \leq \mathbb{E}[\frac{x}{a}] \quad \textit{which is the same as} \quad \frac{\mathbb{E}[x]}{a}$$

Using the coin example, the probability ?????? can be expressed $\mathbb{P}[x \geq \frac{3}{4}n] = \mathbb{P}[x \geq \frac{3}{2}\frac{n}{2}] \leq \frac{2}{3}$

16.4 Chebyshev's Inequality

This inequality has great utility because it can be applied to any probability distribution in which the mean and variance are defined.

$$\mathbb{P}[\; |x - \mathbb{E}[x]| \geq a \;] = \mathbb{P}[\; (x - \mathbb{E}[x])^2 \geq a^2 \;)] \leq \frac{Var(x)}{a^2}$$

Using the previous coin example, the bound for $\mathbb{P}[x \geq \frac{3}{4}n]$ can be calculated as follows:

$$\mathbb{P}[\; |x - \mathbb{E}[x]| \geq (\frac{3}{4}n - \mathbb{E}[x]) \;] = \mathbb{P}[\; |x - \mathbb{E}[x]| \geq \frac{n}{4} \;] \leq \frac{Var(x)}{(\frac{n}{4})^2} = \frac{\frac{n}{4}}{(\frac{n}{4})^2} = \frac{4}{n}$$

16.5 Chernoff Bounds

This method gives exponentially decreasing bounds on tail distributions of sums of *independent* random variables. Formally, Let $x_1, \ldots, x_n$ be independent variables, where $\mathbb{P}[x_i = 1] = p_i$ and:

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

Then the bounds are defined as:

1. $\forall \delta > 0 \ , \ \mathbb{P}[x \geq (1+\delta) \cdot \mu] \leq \left(\frac{e^{\delta}}{(1+\delta)^{1+\delta}} \right)^{\mu}$
2. $\forall \delta \in (0,1] \ , \ \mathbb{P}[x \geq (1+\delta) \cdot \mu] \leq e^{\frac{-\mu\delta^2}{3}}$
3. $\forall R \geq 6\mu \ , \ \mathbb{P}[x \geq R] \leq 2^{-R}$