Prof. Sofya Raskhodnikova
TF Gavin Brown

Discussion 6 – Wednesday, February 26, 2020

Problems

1. (**Distributing Jobs**) Suppose that we have n jobs to distribute among m processors. For simplicity, we assume that m divides n. A job takes 1 step with probability p and $k > 1$ steps with probability $1-p$. Use Chernoff bounds to determine upper and lower bounds (that hold with high probability) on when all jobs will be completed if we randomly assign exactly n/m jobs to each processor.

   Play around with the different bounds and parameters and see how the results change.

For machine j: Let X_{ij} be indicator for job i on machine j took 1 step. fast job.

Total for machine j: $X_j = \sum_{i=1}^{n/m} X_{ij}$ ← apply Chernoff bounds.

$$\mathbb{E}[X_j] = \sum_{i=1}^{n/m} \mathbb{E}[X_{ij}] = \frac{n}{m} p$$

Time for machine j: $T_j = 1 \cdot X_j + k(\underbrace{\tfrac{n}{m} - X_j}_{\#\text{ slow job}}) = \frac{kn}{m} - X_j(k-1)$

$$\Pr[X_j \le (1-\delta)\mu]$$

$$\le e^{-\mu\delta^2/2} = e^{-\frac{n\delta^2}{2m}p}$$

$$\Pr\left[\bigcup_{j=1}^{m} (X_j \le (1-\delta)\mu)\right] \le \sum_{j=1}^{m} \Pr[X_j \le (1-\delta)\mu] \le m e^{-\frac{n\delta^2}{2m}p}$$

2. (**Sums of Geometric Random Variables**) Consider a collection $X_1, \ldots, X_n$ of n independent and geometrically distributed random variables with mean 2. Let $X = \sum_{i=1}^{n} X_i$ and $\delta > 0$.

(a) Derive a bound on $\Pr[X \geq (1+\delta)(2n)]$ by applying the Chernoff bound to a sequence of $(1+\delta)(2n)$ fair coin tosses.

$\Pr[X \geq (1+\delta)(2n)] = \Pr[Y \leq n] = \Pr[Y \leq \frac{1}{(1+\delta)}(1+\delta)n] = \Pr[Y \leq (1-\underbrace{(\frac{\delta}{1+\delta})}_{\delta'})\mu]$

Let Y be the number of heads in the sequence.

$E[Y] = (1+\delta)\cdot 2n \cdot \frac{1}{2} = (1+\delta)n = \mu$

$\leq \exp\{-\frac{\mu}{2}(\delta')^2\}$

$= \exp\{-\frac{n\delta^2}{2(1+\delta)}\}$

(b) General idea: to drive specialized Chernoff bounds -

① Calculate $E[e^{tX}] = E[(e^{tX_i})^n]$

② Apply Markov to probability

③ Apply a "good" t to minimize upper bound

(b) Directly derive a Chernoff bound on $\Pr[X \geq (1+\delta)(2n)]$ using the moment generating function for geometric random variables.

$X_i \sim geom(\frac{1}{2})$ R.V.

$E[e^{tX_i}] = \sum_{j=1}^{\infty} \Pr[X_i = j] e^{tj}$

$= \sum_{j=1}^{\infty} (\frac{1}{2})^j (e^t)^j$

$= -1 + \sum_{j=0}^{\infty} (\frac{e^t}{2})^j$

$= \frac{1}{1-\frac{e^t}{2}} - 1 = \frac{\frac{e^t}{2}}{1-\frac{e^t}{2}} = \frac{e^t}{2-e^t}$

$\Pr[X \geq (1+\delta)(2n)]$

$= \Pr[e^{tX} \geq e^{t(1+\delta)2n}]$

$\leq \frac{E[e^{tX}]}{e^{t(1+\delta)2n}} = \left(\frac{e^t}{2-e^t} \frac{1}{e^{t(1+\delta)2}}\right)^n$

minimized at $t = \ln\left(\frac{1+2\delta}{1+\delta}\right)$