CS648: Randomized Algorithm

Assignment – I

Ashwani Kumar Gautam(13165)

Solution 1:

Let, The number of primes between 2 to N be = $\pi(n)$

We have matrices

$$A=\begin{pmatrix}1 & 0\\1 & 1\end{pmatrix} \qquad\qquad M[\text{э}]=\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$$

These are all 2×2 order matrices and thus, multiplication of any two matrices will take constant number of addition and multiplication operation.

Thus calculating M[x] for an arbitrary string x of length n would take O(n) time.

We know,

Solution 2:

(a). We are given,

$$X=\sum_{i=1}^{n} X_i \quad , \quad X_i \sim Geom(p_i)$$

$$E(X)=E\left(\sum_{i=1}^{n} X_i\right)=\sum_{i=1}^{n} E(X_i)=2\times n$$

$$p_i=\frac{1}{E(X_i)}=2, \forall\, i\in\{1,2,3,..,n\}$$

Since $X_i's$ are all *Geom(p)* random variable, X_i represent number of trials required to hit the success element i^{th} time.

The random variable *X* represent number of trial required to obtain *n* success.

Consider this,

$$Pr\{X\geq(1+\delta)2n\}$$

In words, this represents the probability that experiment will take more then $(1+\delta)2n$ tosses of a coin to hit *n* success.

Consider a sequence of $(1+\delta)2n$ fair coin tosses, where success element is same as the previous case.

Let Y represent the total number of success obtained, that is, Y follows a Binomial distribution with $p=\frac{1}{2}$ and $(1+\delta)2n$ trials. So,

$$Y\sim Bin\left(\frac{1}{2},(1+\delta)2n\right)$$

Clearly, if we obtain $Y\leq n-1$ success in $(1+\delta)2n$ tosses, then $X>(1+\delta)2n$. Thus,

$$Pr\{X\geq(1+\delta)n\}=Pr\{Y\leq n-1\}$$

$$E(Y)=\frac{1}{2}(1+\delta)2n=(1+\delta)n$$

$$Pr\{Y\leq n-1\}=Pr\{Y<(1-\bar{\delta})E(Y)\} \quad , \quad \bar{\delta}=1-\frac{n-1}{E(Y)}\Rightarrow 0<\bar{\delta}\leq 1$$

Using Chernoff bounds, we get

$$\boldsymbol{Pr\{Y\leq(1-\bar{\delta})E(Y)\}\leq\exp\left(\frac{-E(Y)\bar{\delta}^2}{2}\right)}$$

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

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

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

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

$$\Rightarrow \qquad Pr\{X\geq(1+\delta)2n\}\leq\exp\{-\frac{\delta^2 n}{2(1+\delta)}\}$$

(b).

Consider,

$$Pr\{X \geq (1+\delta)2n\}$$

Using $f(x) = e^{xt}$ and markov inequality,

$$Pr\{X \geq (1+\delta)2n\} = Pr\{e^{xt} \geq e^{t(1+\delta)2n}\}$$

and,

$$Pr\{e^{xt} \geq e^{t(1+\delta)2n}\} \leq \frac{E(e^{xt})}{e^{t(1+\delta)2n}}$$

Lets calculate $E(e^{xt})$.

$$E(e^{xt}) = E\left(e^{t(x_1+x_2+..+x_n)}\right)$$

$$\prod_{i=1}^{\infty} E(e^{tx_i}) \qquad \text{……..(i)}$$

Lets calculate $E(e^{tx_i})$,

$$E(e^{tx_i}) = \sum_{i=1}^{\infty} e^{it} p(1-p)^{i-1}$$

$$\frac{p}{(1-p)} \sum_{i=1}^{\infty} (e^t(1-p))^i$$

$$\frac{p}{(1-p)} \frac{e^t(1-p)}{(1-e^t(1-p))} = \frac{pe^t}{1-pe^t} = \frac{e^t}{2-e^t}$$

In the last step, in the denominator $(1-p)$ is replaced with *p* because we are given *p* as $\frac{1}{2}$

Putting this value in equation (i), we have

$$E(e^{xt}) = \prod_{i=1}^{\infty} \frac{e^t}{2-e^t}$$

$$= \left(\frac{2e^t}{2-e^t}\right)^n$$

=> $$Pr\{e^{xt} \geq e^{t(1+\delta)2n}\} = \frac{\left(\frac{e^t}{2-e^t}\right)^n}{e^{t(1+\delta)2n}} \quad \text{…... (1)}$$

$$f(x) = \frac{\left(\frac{e^t}{2-e^t}\right)^n}{e^{t(1+\delta)2n}}$$

Differentiating the above equation w.r.t *t* and setting it 0 we get the point where the function reaches its extremes.

For $t = \ln\left(1 + \frac{\delta}{1+\delta}\right)$

$$f'(x) = 0$$

$$Pr\{X > (1+\delta)2n\} = \frac{\left(\frac{e^t}{2-e^t}\right)^n}{e^{t(1+\delta)2n}}$$

$$<= \left(\left(1 - \frac{\delta}{1+\delta}\right)\left(1 + \frac{\delta}{1+\delta}\right)^{1+2\delta}\right)^{-n}$$

C)

$$\left(\left(1 - \frac{\delta}{1+\delta}\right)\left(1 + \frac{\delta}{1+\delta}\right)^{1+2\delta}\right)^{-n} < \exp\left(-э + (1-э)\left(\frac{\delta^2}{1+\delta} + \delta\right)\right)^{-n} \quad \text{for some } э > 0$$

$$<= \exp(-n((1-э)\delta^2 - э))$$

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

Thus, bound in part (b) is tighter than the one obtained in (a).

Question 2:

Total number of Tosses = n

Let the number of Heads be = n_h

Naturally we would estimate the biasness of the coin by the term $\frac{n_h}{n}$.

Now, we need to show

$$Pr\{|p-\bar{p}|>эp\}<\delta \qquad \forall\, э, \delta, a \in (0,1) \quad \text{and} \quad p \geq a \qquad \ldots (1)$$

эδ

Consider this,

$$Pr\{|p-\bar{p}|>эp\}=Pr\{p-\bar{p}>эp\}+Pr\{\bar{p}-p>эp\}$$

$$= Pr\{np-n\bar{p}>npэ\}+Pr\{n\bar{p}-np>npэ\}$$

$$= Pr\{n\bar{p}<np-npэ\}+Pr\{n\bar{p}>npэ+np\}$$

$$= Pr\{n\bar{p}<(1-э)np\}+Pr\{n\bar{p}>(1+э)np\}$$

$$< e^{-\frac{npэ^2}{2}} + e^{-\frac{npэ^2}{4}}$$

$$< e^{-\frac{apэ^2}{2}} + e^{-\frac{apэ^2}{4}}$$

So to prove our claim, we need to find a suitable value of N such that above equation is bounded by δ.

$$e^{-\frac{apэ^2}{2}}+e^{-\frac{apэ^2}{4}}<\delta$$

$$< e^{-\frac{na\text{э}^2}{4}} < \frac{\delta}{2} \qquad \left(\text{as } e^{-\frac{na\text{э}^2}{2}} > e^{-\frac{na\text{э}^2}{4}}\right)$$

$$\frac{na\text{э}^2}{4} > \ln\left(\frac{2}{\delta}\right)$$

$$n > \frac{4\ln\left(\frac{2}{\delta}\right)}{a\text{э}^2}$$

Thus for all n satisfying above equation we can guarantee the equation (1) give above.

Formulas Used :

1) $Pr\{X > (1+\delta)\mu\} < e^{-\frac{\delta^2 \mu}{4}}$, $1 < 1+\delta < 2e$

2) $Pr\{X < (1-\delta)\mu\} < e^{-\frac{\delta^2 \mu}{2}}$, $\delta > 0$