CS388R: Randomized Algorithms Fall 2015

Lecture 5 — Sep 14, 2015

Prof. Eric Price *Scribe: Chao-Yuan Wu, Jin Zhang*

1 Overview

In the last lecture we introduced coupon collector problem, concentration in equalities, subguassian random variables and their basic properties.

In this lecture we will see these properties are equivalent up to a scaling factor and introduce subexponential and subgamma random variables.

2 Starter Problem

2.1

Question: Suppose x is drawn from an unknown distribution with mean μ and variance σ^2. How many samples do we need to estimate $\mu \pm \epsilon\sigma$ with probability $\geq 1 - \delta$?

One solution: Let $\bar{x} := \frac{\sum x_i}{n}$. Then

$$\mathrm{Var}[\bar{x}] = \frac{\sigma^2}{n} \tag{1}$$

$$\mathbb{E}[\bar{x}] = \mu \tag{2}$$

By Chebyshev inequality,

$$\mathbb{P}(|\bar{x} - \mu| \geq \alpha \frac{\sigma}{\sqrt{n}}) \leq \frac{1}{\alpha^2}.$$

Let $\delta = \frac{1}{\alpha^2}$, and we conclude that

$$n \geq \frac{1}{\delta\epsilon^2}$$

2.2 Is this bound tight?

Let's look at an example. Suppose $x \in \{0, 1\}$ with $\mathbb{P}(x = 1) = p = \frac{1}{10n^2}$. Then we have

$$\sigma = \sqrt{p(1-p)} \approx \sqrt{p} < \frac{1}{3n}$$

If any of the sample is 1, the empirical mean

$$\frac{\sum x_i}{n} \geq \frac{1}{n} \geq \frac{1}{10n^2} + \frac{1}{3n} > \mu + \sigma$$

The probability that any sample is 1 is

$$\mathbb{P}(\text{any sample} = 1) = 1 - (1-p)^n \approx pn = \frac{1}{10n}$$

In this case, $\frac{1}{\delta}$ is tight.

2.3 Can we do better?

How about median? How far can median of x be from mean of x (in terms of σ)? By Chebyshev inequality, we have

$$\mathbb{P}[|x - \mu| \geq \sqrt{2}\sigma] \leq \frac{1}{2}.$$

This means more than half of the samples fall into the $\pm\sqrt{2}\sigma$ range, and so does the median.

How many samples do we need if we use median? Let

$$Y_i = \begin{cases} 1 & if\ |x_i - \mu| \leq 2\sigma \\ 0 & if\ |x_i - \mu| \geq 2\sigma \end{cases}$$

By Chebyshev's inequality we get $\mathbb{P}[Y_i = 1] \geq \frac{3}{4}$. Note that if $\sum Y_i \geq \frac{n}{2}$, $|\text{median}(x_i) - \mu| \leq 2\sigma$.

$$\begin{aligned} \mathbb{P}\left(\sum Y_i \geq \frac{n}{2}\right) &\geq 1 - \mathbb{P}\left(\sum Yi \leq \mathbb{E}[\sum Y_i] - \frac{n}{4}\right) \\ &\geq 1 - e^{-\frac{2(\frac{n}{4})^2}{n}} \\ &= 1 - e^{-\frac{n}{8}} \end{aligned}$$

We thus need only

$$n = 8 \log \frac{1}{\delta}$$

samples to get an estimation within $\mu \pm O(\sigma)$ with probability $1 - \delta$.

Combining mean and median: Suppose we divide the samples into $S_1, S_2, \ldots, S_m$ buckets. For each bucket, we have $\text{mean}(S_i) = \mu$ and standard deviation $\sigma_{S_i} = \frac{\sigma}{\sqrt{|S_i|}}$. Let $|S_i| = \frac{4}{\epsilon^2}$, we get $\sigma_{s_i} = \frac{\epsilon\sigma}{2}$. If we take median of $m = 8 \log \frac{1}{\delta}$ bucket means, according to Section **??**, we get estimation within $\pm\epsilon\sigma$. Overall we need only

$$|S_i| * m = \frac{4}{\epsilon^2} * 8 \log \frac{1}{\delta} = 32 \frac{1}{\epsilon^2} \log \frac{1}{\delta}$$

samples.

3 Subgaussian random variables

Recall in last class we see x is subguassian with parameter σ if any of following holds.

1. $\mathbb{E}[e^{\lambda x}] \leq e^{\frac{\lambda^2 \sigma^2}{2}}$
2. $\mathbb{P}[|x| \geq t] \leq 2e^{-\frac{t^2}{2\sigma^2}}$
3. $\mathbb{E}[|X|^k]^{\frac{1}{k}} \leq \sigma\sqrt{k}$

In fact, these three are equivalent up to scaling σ by a constant.

3.1 Proof (1) ⇒ (2)

$$\begin{aligned} Pr[X \geq t] &= \mathbb{P}[e^{\lambda x} \geq e^{\lambda t}] \\ &\leq \frac{E[e^{\lambda x}]}{e^{\lambda t}} \\ &\leq e^{\frac{\lambda^2\sigma^2}{2} - \lambda t} \\ &= e^{\frac{-t^2}{2\sigma^2}} \end{aligned}$$

The inequality is by Markov's inequality, and the last equality is obtained by choosing $\lambda = \frac{t}{\sigma^2}$.

3.2 Proof (2) ⇒ (3)

Let $z = |x|^k$ and assume $\sigma = \frac{1}{\sqrt{2}}$,

$$\mathbb{E}(z) = \int_0^\infty \mathbb{P}(z \geq \mu) d\mu, \forall z \geq 0. \tag{3}$$

$$\begin{aligned} \mathbb{P}\left(|x|^k \geq t^k\right) &= \mathbb{P}\left(|x| \geq t\right) \qquad (4) \\ &\leq 2e^{-t^2} \qquad (5) \end{aligned}$$

$$\mathbb{E}(|x|^k) \leq \int_0^\infty 2e^{-t^2} kt^{k-1} dt \tag{6}$$

Equation **??** is by Chernoff bound, and Equation **??** is obtained by combining Equation **??** and Equation **??**. After some integration by parts, we can get

$$\begin{aligned} \mathbb{E}(|x|^k) &\leq 2(\frac{k}{2})! \\ &\leq 2(\frac{k}{2})^{\frac{k}{2}} \end{aligned}$$

3.2.1 Proof (3) $\Rightarrow$ (1)

By Taylor expansion,

$$\begin{aligned} e^{\lambda x} &= 1 + \lambda x + \frac{\lambda^2 x^2}{2} + \dots \\ \mathbb{E}[e^{\lambda x}] &\le 1 + \sum_{k=2}^{\infty} \frac{\lambda^k (ck)^{\frac{k}{2}}}{k!} \\ &\le 1 + \sum_{k=2}^{\infty} (\frac{\lambda^2 cke^2}{k^2})^{\frac{k}{2}} \\ &\le 1 + \sum_{k=2}^{\infty} (\frac{\lambda^2 e}{k})^{\frac{k}{2}} \end{aligned}$$

Here we use the fact $k! \ge \left(\frac{k}{e}\right)^k$

for $\sigma = O(1)$

$$\begin{aligned} e^{\lambda^2\sigma^2} &= 1 + \sum_{k=1}^{\infty} \frac{(\lambda^2\sigma^2)^k}{k!} \\ &\ge 1 + \sum_{k=1}^{\infty} (\frac{\lambda^2\sigma^2}{k})^k \\ &= 1 + \sum_{k \in 2\mathbb{N}} (\frac{2\lambda^2\sigma^2}{k})^{\frac{k}{2}} \end{aligned}$$

for $\sigma = \Omega(1)$ we have all even terms

$$\begin{aligned} (\frac{\sqrt{e}\lambda}{\sqrt{k}})^k &\le (\frac{\sqrt{e}\lambda}{\sqrt{k}})^{k-1} + (\frac{\sqrt{e}\lambda}{\sqrt{k}})^{k+1} \\ &\le (\frac{\sqrt{e}\lambda}{\sqrt{k-1}})^{k-1} + (\frac{\sqrt{e}\lambda}{\sqrt{k+1}})^{k+1} (\sqrt{\frac{k+1}{k}})^{k+1} \end{aligned}$$

3.3 Subexponential random variables

Random variable x is subexponential with parameter σ if one of the following holds:

1. $E[e^{\lambda x}] \le e^{\frac{\lambda^2\sigma^2}{2}}, \forall |\lambda| \le \frac{1}{\sigma}$
2. $Pr[|x| \ge t] \le 2e^{-\frac{t}{2\sigma}}$
3. $E[|X|^k]^{\frac{1}{k}} \le \sigma k$

3.4 Subgamma random variables

Here we state some of the properties of subgamma random variables without proof. Random variable x is subgamma(σ, B) if

$$\mathbb{E}[e^{\lambda x}] \leq e^{\frac{\lambda^2 \sigma^2}{2}},\ \forall |\lambda| \leq B$$

or equivalently,

$$\mathbb{P}[|x| \geq t] \leq \begin{cases} 2e^{-\frac{t^2}{\sigma^2}} & \text{, if } \frac{|t|}{\sigma^2} \leq B \\ e^{\frac{1}{2}B^2\sigma^2 - Bt} \leq e^{-\frac{Bt}{2}} & \text{, otherwise} \end{cases} \tag{7}$$

Combining the two cases:

$$\mathbb{P}[|x| \geq t] \leq \max(2e^{-\frac{t^2}{\sigma^2}}, 2e^{-\frac{Bt}{2}})$$

Suppose x_i are subgamma(σ_i, β_i). Then $\sum x_i$ is subgamma$(\sqrt{\sum \sigma_i^2}, \min(B_i))$.

References

[MR] Rajeev Motwani, Prabhakar Raghavan Randomized Algorithms. *Cambridge University Press*, 0-521-47465-5, 1995.