CS 574: Randomized Algorithms

Lecture 12. Chernoff Bounds: Proofs

October 1, 2015

Large Deviations

- Let X_i be i.i..d, with $E[X_i] = p_i$ and $Var[X_i] = \sigma_i^2$ (all identical). Let $X = \sum X_i$, $\mu = np_i$ and $\sigma^2 = n\sigma_i^2$.

Large Deviations

- Let X_i be i.i..d, with $E[X_i] = p_i$ and $Var[X_i] = \sigma_i^2$ (all identical). Let $X = \sum X_i$, $\mu = np_i$ and $\sigma^2 = n\sigma_i^2$.
- Central Limit Theorem says that as $n \to \infty$, $\frac{X-\mu}{\sigma} \to \mathcal{N}(0, 1)$.

Large Deviations

- Let X_i be i.i..d, with $E[X_i] = p_i$ and $Var[X_i] = \sigma_i^2$ (all identical). Let $X = \sum X_i$, $\mu = np_i$ and $\sigma^2 = n\sigma_i^2$.
- Central Limit Theorem says that as $n \to \infty$, $\frac{X-\mu}{\sigma} \to \mathcal{N}(0,1)$.
- Namely,

$$Pr[|X - \mu| > \beta\sigma] \to \frac{1}{\sqrt{2\pi}} \int_{\beta}^{\infty} e^{-t^2/2} dt \approx \frac{1}{\sqrt{2\pi}\beta} e^{-\beta^2/2}$$

.

Large Deviations

- Let X_i be i.i..d, with $E[X_i] = p_i$ and $Var[X_i] = \sigma_i^2$ (all identical). Let $X = \sum X_i$, $\mu = np_i$ and $\sigma^2 = n\sigma_i^2$.
- Central Limit Theorem says that as $n \to \infty$, $\frac{X-\mu}{\sigma} \to \mathcal{N}(0,1)$.
- Namely,

$$Pr[|X - \mu| > \beta\sigma] \to \frac{1}{\sqrt{2\pi}} \int_{\beta}^{\infty} e^{-t^2/2} dt \approx \frac{1}{\sqrt{2\pi}\beta} e^{-\beta^2/2}$$

.

- This result is only asymptotic, and says nothing about when n finite.

Large Deviations

- Let X_i be i.i..d, with $E[X_i] = p_i$ and $Var[X_i] = \sigma_i^2$ (all identical). Let $X = \sum X_i$, $\mu = np_i$ and $\sigma^2 = n\sigma_i^2$.
- Central Limit Theorem says that as $n \to \infty$, $\frac{X-\mu}{\sigma} \to \mathcal{N}(0,1)$.
- Namely,

$$Pr[|X-\mu| > \beta\sigma] \to \frac{1}{\sqrt{2\pi}} \int_{\beta}^{\infty} e^{-t^2/2} dt \approx \frac{1}{\sqrt{2\pi}\beta} e^{-\beta^2/2}$$

.

- This result is only asymptotic, and says nothing about when n finite.
- It applies only to deviations in the order of standard deviation.

In order to prove exponential bounds, we need to use independence in a strong way. $E[X_1 X_2 \cdots X_n] = E[X_1]E[X_2] \cdots E[X_n]$.

In order to prove exponential bounds, we need to use independence in a strong way. $E[X_1 X_2 \cdots X_n] = E[X_1]E[X_2] \cdots E[X_n]$.

Theorem

(Chernoff Bound) Let X_i be independent bernoulli r.v.s (not necessarily identical), with $E[X_i] = p_i$. Let $x = \sum_{i=1}^{n} X_i$ and $\mu = \sum_{i=1}^{n} p_i$ then for every $\beta \geq 1$ we have

$$Pr[X \geq \beta\mu] \leq \left(\frac{e^{\beta-1}}{\beta^{\beta}}\right)^{\mu}$$

and

$$Pr[X \leq \frac{\mu}{\beta}] \leq \left(\frac{e^{1/\beta-1}}{\beta^{\beta}}\right)^{\mu}$$

Method of Laplace Transforms

- Fix real number x, let $t > 0$ be a parameter t.b.d. We need an expression involving all products of the variables.

Method of Laplace Transforms

- Fix real number x, let $t > 0$ be a parameter t.b.d. We need an expression involving all products of the variables.
- Use Taylor expansion

$$e^{tx} = \sum_{k=0}^{\infty} \frac{t^k x^k}{k!}$$

Method of Laplace Transforms

- Fix real number x, let $t > 0$ be a parameter t.b.d. We need an expression involving all products of the variables.
- Use Taylor expansion

$$e^{tx} = \sum_{k=0}^{\infty} \frac{t^k x^k}{k!}$$

- Bulk of the sum concentrated around $k \approx t|x|$. Choosing appropriate value for t, we can get the sum to concentrate around $t|x|$.

Method of Laplace Transforms

- Fix real number x, let $t > 0$ be a parameter t.b.d. We need an expression involving all products of the variables.
- Use Taylor expansion

$$e^{tx} = \sum_{k=0}^{\infty} \frac{t^k x^k}{k!}$$

- Bulk of the sum concentrated around $k \approx t|x|$. Choosing appropriate value for t, we can get the sum to concentrate around $t|x|$.
- We can apply this with $x = X = \sum X_i$ and get

$$E[e^{tX}] = \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{i_1,\cdots,i_k} E[X_{i_1}] \cdots E[X_{i_k}]$$

Method of Laplace Transforms

- Fix real number x, let $t > 0$ be a parameter t.b.d. We need an expression involving all products of the variables.
- Use Taylor expansion

$$e^{tx} = \sum_{k=0}^{\infty} \frac{t^k x^k}{k!}$$

- Bulk of the sum concentrated around $k \approx t|x|$. Choosing appropriate value for t, we can get the sum to concentrate around $t|x|$.
- We can apply this with $x = X = \sum X_i$ and get

$$E[e^{tX}] = \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{i_1,\cdots,i_k} E[X_{i_1}] \cdots E[X_{i_k}]$$

- We can work just with the exponential and avoid messy right hand side.

Proof of Chernoff

- Use method of Laplace Transofms

Proof of Chernoff

- Use method of Laplace Transofms
- Apply Markov.

Proof of Chernoff

- Use method of Laplace Transofms
- Apply Markov.
- Use the independence of X_i to take product of expectations

Proof of Chernoff

- Use method of Laplace Transofms
- Apply Markov.
- Use the independence of X_i to take product of expectations
- Plug in specific information about the distributions.

Proof of Chernoff

- Use method of Laplace Transofms
- Apply Markov.
- Use the independence of X_i to take product of expectations
- Plug in specific information about the distributions.
- Use concavity to get a bound.

Proof of Chernoff

- Use method of Laplace Transofms
- Apply Markov.
- Use the independence of X_i to take product of expectations
- Plug in specific information about the distributions.
- Use concavity to get a bound.
- Optimize with respect to t.

Other Forms of Chernoff Bounds

(Please prove them at home)

Theorem

*$Pr[X \geq \mu + \lambda] \leq e^{-nH_p(p+\frac{\lambda}{n})}$ for $0 < \lambda < n - \mu$ where $p = \frac{\mu}{n}$ and $Pr[X \leq \mu - \lambda] \leq e^{-nH_{1-p}(1-p+\frac{\lambda}{n})}$ for $0 < \lambda < \mu$.
Here $H_p(x) = x \ln(\frac{x}{p}) + (1-x)\ln(\frac{1-x}{1-p})$ is the relative entropy of x with respect to p.*

Other Forms of Chernoff Bounds

(Please prove them at home)

Theorem

*$Pr[X \geq \mu + \lambda] \leq e^{-nH_p(p+\frac{\lambda}{n})}$ for $0 < \lambda < n - \mu$ where $p = \frac{\mu}{n}$ and $Pr[X \leq \mu - \lambda] \leq e^{-nH_{1-p}(1-p+\frac{\lambda}{n})}$ for $0 < \lambda < \mu$.
Here $H_p(x) = x\ln(\frac{x}{p}) + (1-x)\ln(\frac{1-x}{1-p})$ is the relative entropy of x with respect to p.*

Corollary

$Pr[X \leq \mu - \lambda] \leq e^{-\frac{2\lambda^2}{n}}$ and $Pr[X \geq \mu + \lambda] \leq e^{-\frac{2\lambda^2}{n}}$

Other Forms of Chernoff Bounds

(Please prove them at home)

Theorem

$Pr[X \geq \mu + \lambda] \leq e^{-nH_p(p+\frac{\lambda}{n})}$ for $0 < \lambda < n - \mu$ where $p = \frac{\mu}{n}$ and $Pr[X \leq \mu - \lambda] \leq e^{-nH_{1-p}(1-p+\frac{\lambda}{n})}$ for $0 < \lambda < \mu$.
Here $H_p(x) = x\ln(\frac{x}{p}) + (1-x)\ln(\frac{1-x}{1-p})$ is the relative entropy of x with respect to p.

Corollary

$Pr[X \leq \mu - \lambda] \leq e^{-\frac{2\lambda^2}{n}}$ and $Pr[X \geq \mu + \lambda] \leq e^{-\frac{2\lambda^2}{n}}$

Corollary

$Pr[X \leq (1-\delta)\mu] \leq e^{-\frac{\delta^2\mu}{2}}$ and $Pr[X \geq (1+\delta)\mu] \leq e^{-\frac{\delta^2\mu}{3}}$

Some more Applications

- Coin Flips (fair, biased).

Some more Applications

- Coin Flips (fair, biased).
- Random Routing.