

CSL 857: Randomized Algorithms

Ragesh Jaiswal
CSE, IIT Delhi

Chernoff Bounds

Chernoff Bounds

- Theorem: Let $X_1, \dots, X_n$ be independent, 0/1 random variables, and let $p_i = E[X_i]$ for all $i = 1, \dots, n$. Let $X = X_1 + X_2 + \dots + X_n$ and let $\mu = E[X]$. Let $\delta > 0$ be any real number. Then

$$\Pr[X > (1 + \delta) \cdot \mu] \leq \left[\frac{e^\delta}{(1 + \delta)^{(1+\delta)}} \right]^\mu.$$

- Corollary: Let $X_1, \dots, X_n$ be independent, 0/1 random variables, and let $p_i = E[X_i]$ for all $i = 1, \dots, n$. Let $X = X_1 + X_2 + \dots + X_n$ and let $\mu = E[X]$. Let $0 < \delta \leq 1$ be any real number. Then

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

Chernoff Bounds

- Theorem: Let $X_1, \dots, X_n$ be independent, 0/1 random variables, and let $p_i = E[X_i]$ for all $i = 1, \dots, n$. Let $X = X_1 + X_2 + \dots + X_n$ and let $\mu = E[X]$. Let $\delta > 0$ be any real number. Then

$$\Pr[X > (1 + \delta) \cdot \mu] \leq e^{-f(\delta) \cdot \mu}$$

where $f(\delta) = (1 + \delta) \ln(1 + \delta) - \delta$.

- Claim: $\forall \delta > 0, f(\delta) \geq \frac{\delta^2}{2 + \delta}$.
- So, when $\delta \leq 1$, then the bound is $\exp(-\delta^2 \mu / 3)$.
- The general bound for any positive δ is $\exp\left(-\frac{\delta^2 \mu}{(2 + \delta)}\right)$.

Chernoff Bounds

- Theorem: Let $X_1, \dots, X_n$ be independent, 0/1 random variables, and let $p_i = E[X_i]$ for all $i = 1, \dots, n$. Let $X = X_1 + X_2 + \dots + X_n$ and let $\mu = E[X]$. Let $\delta > 0$ be any real number. Then

$$\Pr[X > (1 + \delta) \cdot \mu] \leq e^{-f(\delta) \cdot \mu}$$

where $f(\delta) = (1 + \delta) \ln(1 + \delta) - \delta$.

- Claim: $\forall \delta > 0, f(\delta) \geq \frac{\delta^2}{2}$.
- The general bound for any positive δ is $\exp\left(-\frac{\delta^2 \mu}{2}\right)$.

Chernoff Bounds: Special Case

- Theorem: Let $X_1, \dots, X_n$ be independent, $\{-1, +1\}$ random variables such that $\Pr[X_i = -1] = \Pr[X_i = +1] = 1/2$ for all $i = 1, \dots, n$. Let $X = X_1 + X_2 + \dots + X_n$. Let $A > 0$ be any real number. Then

$$\Pr[X \geq A] \leq e^{-\frac{A^2}{2n}}.$$

Balls and bins

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: With probability at least $(1 - 1/n)$, all bins have at most $\frac{3 \ln n}{\ln \ln n}$ balls.
- Proof:
 - What is the probability that bin i , has more than k balls?

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: With probability at least $(1 - 1/n)$, all bins have at most $\frac{3 \ln n}{\ln \ln n}$ balls.
- Proof:
 - What is the probability that bin i , has more than k balls?
 - $\leq \binom{n}{k} \cdot \frac{1}{n^k} \leq \left(\frac{e}{k}\right)^k$
 - What is the above bound for $k \geq \frac{3 \ln n}{\ln \ln n}$?

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: With probability at least $(1 - 1/n)$, all bins have at most $\frac{3 \ln n}{\ln \ln n}$ balls.
- Proof:
 - What is the probability that bin i , has more than k balls?
 - $\leq \binom{n}{k} \cdot \frac{1}{n^k} \leq \left(\frac{e}{k}\right)^k$
 - What is the above bound for $k \geq \frac{3 \ln n}{\ln \ln n}$?
 - $\leq \frac{1}{n^2}$ for sufficiently large values of n .

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: With probability at least $(1 - 1/n)$, all bins have at most $\frac{3 \ln n}{\ln \ln n}$ balls.
- Proof:
 - What is the probability that bin i , has more than k balls?
 - $\leq \binom{n}{k} \cdot \frac{1}{n^k} \leq \left(\frac{e}{k}\right)^k$
 - What is the above bound for $k \geq \frac{3 \ln n}{\ln \ln n}$?
 - $\leq \frac{1}{n^2}$ for sufficiently large values of n .
 - $\Pr \left[\text{no bin has more than } \frac{3 \ln n}{\ln \ln n} \text{ balls} \right] \geq (1 - 1/n)$

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: With probability at least $(1 - 1/n)$, all bins have at most $\frac{3 \ln n}{\ln \ln n}$ balls.
- Question: Is the above bound tight?
 - What is the probability that a bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls?

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.
- Proof:
 - What is the probability that bin i , has at least k balls?

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?

- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.

- Proof:

- What is the probability that bin i , has at least k balls?

- $\geq \binom{n}{k} \cdot \frac{1}{n^k} \cdot \left(1 - \frac{1}{n}\right)^{n-k} \geq \binom{n}{k}^k \cdot \frac{1}{n^k} \cdot \frac{1}{2e} = \frac{1}{2ek^k}$

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?

- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.

- Proof:

- What is the probability that bin i , has at least k balls?

- $\geq \binom{n}{k} \cdot \frac{1}{n^k} \cdot \left(1 - \frac{1}{n}\right)^{n-k} \geq \binom{n}{k}^k \cdot \frac{1}{n^k} \cdot \frac{1}{2e} = \frac{1}{2ek^k}$

- What is the above bound for $k = \frac{\ln n}{3 \ln \ln n}$?

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?

- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.

- Proof:

- What is the probability that bin i , has at least k balls?

- $\geq \binom{n}{k} \cdot \frac{1}{n^k} \cdot \left(1 - \frac{1}{n}\right)^{n-k} \geq \left(\frac{n}{k}\right)^k \cdot \frac{1}{n^k} \cdot \frac{1}{2e} = \frac{1}{2ek^k}$

- What is the above bound for $k = \frac{\ln n}{3 \ln \ln n}$?

- $\geq \Omega(n^{-\frac{1}{3}})$.

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?

- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.

- Proof:

- What is the probability that bin i , has at least k balls?

- $\geq \binom{n}{k} \cdot \frac{1}{n^k} \cdot \left(1 - \frac{1}{n}\right)^{n-k} \geq \left(\frac{n}{k}\right)^k \cdot \frac{1}{n^k} \cdot \frac{1}{e} = \frac{1}{ek^k}$

- What is the above bound for $k = \frac{\ln n}{3 \ln \ln n}$?

- $\geq \Omega(n^{-\frac{1}{3}})$.

- What is the expected number of bins having at least $\frac{\ln n}{3 \ln \ln n}$ balls?

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.
- Proof:
 - What is the probability that bin i , has at least k balls?
 - $\geq \binom{n}{k} \cdot \frac{1}{n^k} \cdot \left(1 - \frac{1}{n}\right)^{n-k} \geq \left(\frac{n}{k}\right)^k \cdot \frac{1}{n^k} \cdot \frac{1}{e} = \frac{1}{ek^k}$
 - What is the above bound for $k = \frac{\ln n}{3 \ln \ln n}$?
 - $\geq \Omega(n^{-\frac{1}{3}})$.
 - What is the expected number of bins having at least $\frac{\ln n}{3 \ln \ln n}$ balls?
 - $\Omega(n^{2/3})$.

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?
- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.
- Proof:
 - Let X_i denote the probability that bin i , has at least $\frac{\ln n}{3 \ln \ln n}$ balls and let $X = X_1 + \dots + X_n$.
 - We know that $\forall i, E[X_i] \geq \Omega(n^{-\frac{1}{3}})$ and $E[X] = \Omega(n^{2/3})$.
 - What is $\Pr[X = 0]$?

Balls and Bins

- Problem: n balls are thrown into n bins randomly. What is the maximum load that a bin can have?

- Claim: The max loaded bin has $\Omega\left(\frac{\ln n}{\ln \ln n}\right)$ balls with probability at least $(1 - 1/n^{1/3})$.

- Proof:

- Let X_i denote the probability that bin i , has at least $\frac{\ln n}{3 \ln \ln n}$ balls and let $X = X_1 + \dots + X_n$.

- We know that $\forall i, E[X_i] \geq \Omega(n^{-\frac{1}{3}})$ and $E[X] = \Omega(n^{2/3})$.

- What is $\Pr[X = 0]$?

- $$\Pr[X = 0] \leq \Pr[|X - E[X]| \geq E[X]] \leq \frac{\text{Var}[X]}{E[X]^2} =$$

$$\frac{\sum_i \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}(X_i, X_j)}{E[X]^2} = O(n^{-\frac{1}{3}}) \text{ (by Chebychev and } \text{Cov}(X_i, X_j) < 1)$$

Balls and Bins

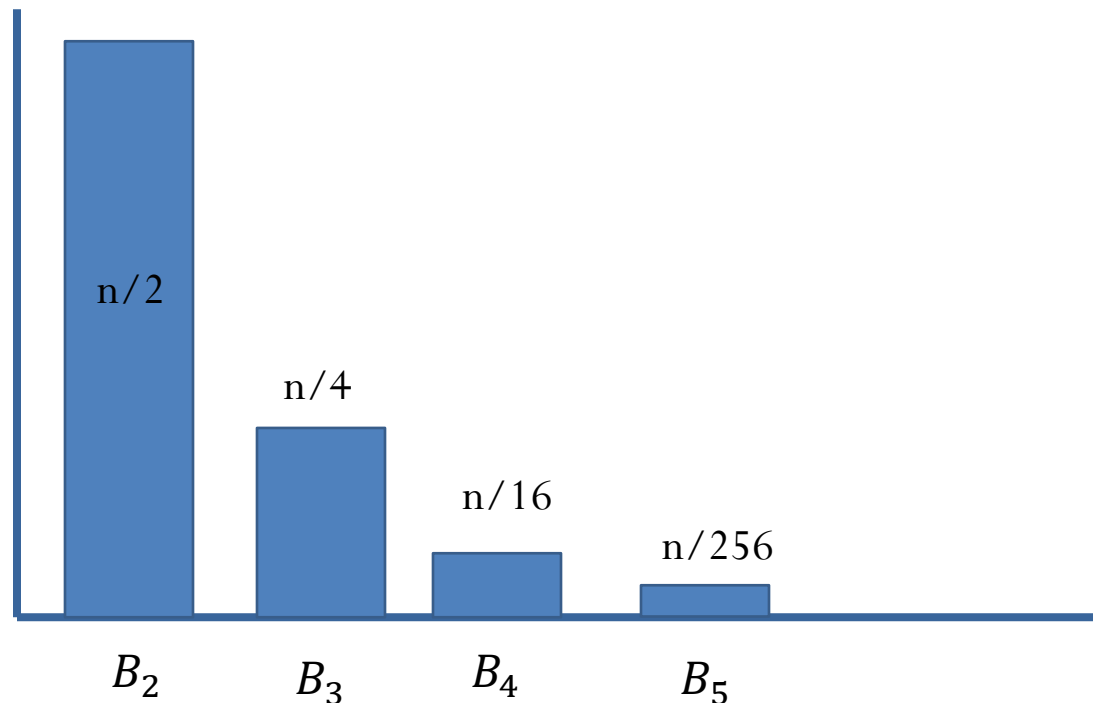
“Power of two choices”

Power of two choices

- When n balls are thrown into n bins randomly, we know that the maximum loaded bin has $\theta \left(\frac{\log n}{\log \log n} \right)$ balls w.h.p.
- Main idea: *For each ball pick two bins uniformly at random and then throw the ball into the bin with lesser load.*
- Theorem(main result): The maximum load w.r.t. the above randomized procedure is $\frac{\ln \ln n}{\ln 2} + O(1)$ w.h.p.

Power of two choices

- The height of a ball is defined to be the place of the ball in its bin where the balls are ordered as per time of arrival.
- Let B_i denote the number of bins with load at least i .
- Main idea: We will argue that w.h.p. B_i 's behave as follows:



Power of two choices

- Let $\alpha_4 = n/4$, and for all $i \geq 5$, $\alpha_{i+1} = \frac{2(\alpha_i)^2}{n}$.
- Let E_i denote the event that $B_i \leq \alpha_i$.
- Claim 1: $\Pr[E_4] = 1$.
- Claim 2: $\Pr[\neg E_{i+1} \mid E_i] \leq \Pr \left[X \left(\left(\frac{\alpha_i}{n} \right)^2 \right) > \alpha_{i+1} \right]$.

Here $X(\beta)$ is the sum of n independent random 0/1 variables each with expectation β .

- Proof: Suppose α_i bins have been marked before starting ball allocation. A ball will have height at least $(i + 1)$ if both its choices are marked. The total number of such balls gives an upper bound on B_{i+1} . The number of such balls is denoted by random variable $X(\cdot)$.

Power of two choices

- Let $\alpha_4 = n/4$, and for all $i \geq 5$, $\alpha_{i+1} = \frac{2(\alpha_i)^2}{n}$.
- Let E_i denote the event that $B_i \leq \alpha_i$.
- Claim 1: $\Pr[E_4] = 1$.
- Claim 2: $\Pr[\neg E_{i+1} \mid E_i] \leq \Pr\left[X\left(\left(\frac{\alpha_i}{n}\right)^2\right) > \alpha_{i+1}\right]$.

Here $X(\beta)$ is the sum of n independent random 0/1 variables each with expectation β .

- Let $j = \min\{i: (\alpha_i)^2 < 6n \ln n\}$.
- Claim 3: $j = \frac{\ln \ln n}{\ln 2} + O(1)$.
- Claim 4: $\forall i, 4 \leq i \leq j, \Pr[\neg E_i] \leq \frac{1}{n}$.

Power of two choices

- Let $\alpha_4 = n/4$, and for all $i \geq 5$, $\alpha_{i+1} = \frac{2(\alpha_i)^2}{n}$.
- Let E_i denote the event that $B_i \leq \alpha_i$.
- Claim 1: $\Pr[E_4] = 1$.
- Claim 2: $\Pr[\neg E_{i+1} \mid E_i] \leq \Pr\left[X\left(\left(\frac{\alpha_i}{n}\right)^2\right) > \alpha_{i+1}\right]$. Here $X(\beta)$ is the sum of n independent random 0/1 variables each with expectation β .
- Let $j = \min\{i: (\alpha_i)^2 < 6n \ln n\}$.
- Claim 3: $j = \frac{\ln \ln n}{\ln 2} + O(1)$.
- Claim 4: $\forall i, 4 \leq i \leq j, \Pr[\neg E_i] \leq \frac{1}{n}$.
- Claim 5: $\Pr[B_{j+1} > 12 \ln n] \leq \frac{2}{n}$.
- Claim 6: $\Pr[B_{j+2} > 1] \leq O\left(\frac{(\ln n)^2}{n}\right)$.

End
