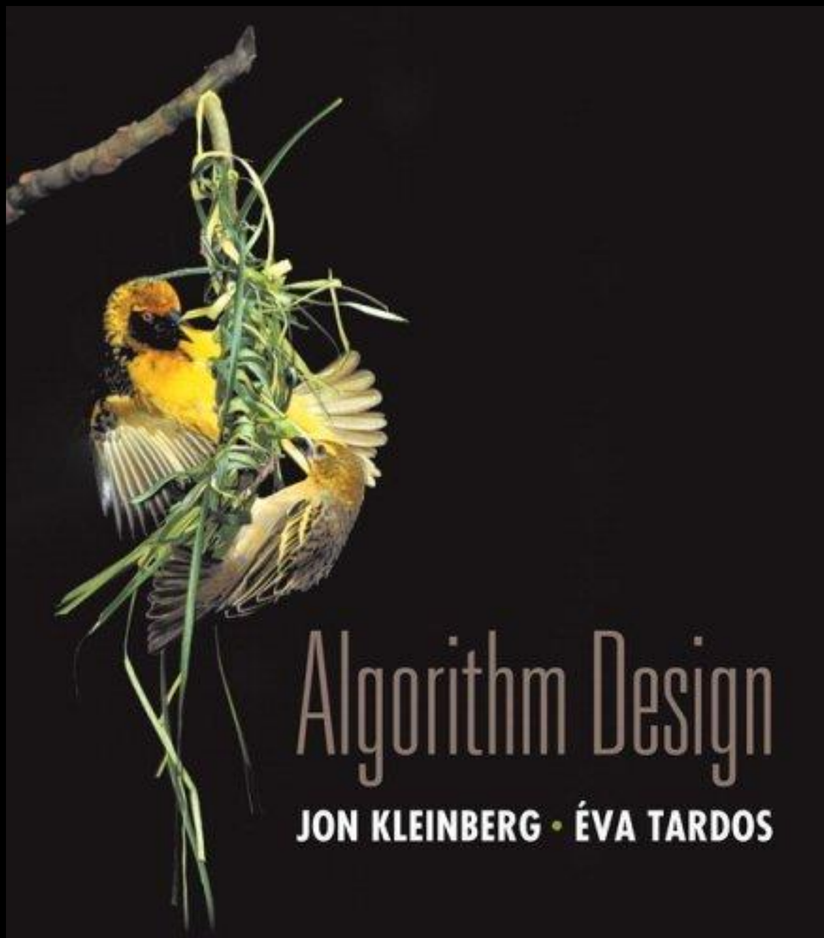


Chapter 13

Randomized Algorithms

Slides by Kevin Wayne.

Desigualdades de Cauda

- Como gerar limites superiores para a probabilidade de uma variável aleatória se afastar da média?

Ferramenta fundamental para caracterizar o tempo de execução e/ou a probabilidade de sucesso de algoritmos aleatorizados

Desigualdade de Markov

Seja **X** uma V.A. que assume somente valores não negativos. Então, para todo t positivo,

$$\Pr[X \geq t] \leq \frac{E[X]}{t}$$

Prova: Considere a variavel 0-1 Y que assume valor 1 se X>=t e 0, caso contrário. Note que $Y \leq X/t$. Logo,

$\Pr[X \geq t] = E[Y]$ e $E[Y] \leq E[X]/t$.

Portanto, $\Pr[X \geq t] \leq E[X]/t$

Desigualdade de Markov

Qual é a probabilidadade obtermos mais de 75 caras ao jogar uma moeda justa 100 vezes ?

X: número de vezes que o resultado é cara
X_i :1 se o resultado da i-ésima tentativa é cara e 0 caso contrário.

$$E[X_i] = 1/2 \qquad E[X] = \sum_{i=1}^{100} E[X_i] = 50$$

Aplicando Markov temos

$$\Pr[X>75] \leq 50/75 = 2/3$$

Variância de uma distribuição

A variância de uma variável aleatória X é definida como

$$Var(X)= E[(X-E[X])^2]$$

- A variância mede o quanto a distribuição "foge" da média.
- O desvio padrão de X, σ_X , é a raiz quadrada da variância

Variância de uma distribuição

Temos que:

$$Var(X)= E[(X-E[X])^2] = E[(X^2 - 2XE[X]+ E[X]^2] =$$

$$E[(X^2 - XE[X]+ E[X]^2] = E[(X^2]- E[X]^2$$

Variância de uma distribuição

Def Duas variáveis aleatórias X e Y são independentes se para qualquer par de reais x,y

$$Pr[X=x, Y=y] = Pr[X=x]Pr[Y=y]$$

Lema: Se X e Y são V.A. independentes, então

$$E[XY]=E[X]E[Y]$$

$$Var(X+Y)=Var(X)+Var(Y)$$

Variância de uma distribuição

Definição. Uma variável binomial X com parâmetros n e p conta o número de caras em uma sequência de n lançamentos aonde a probabilidade cara ocorrer em um lançamento é p

Lema. O valor esperado de uma V.A. binomial X com parâmetros n e p é np.

Teorema. A variância de uma V.A. binomial X com parâmetros n e p é np(1-p).

Desigualdade de Chebyschev

Seja X uma V.A. com desvio padrão σ_X. Então, para todo t>0,

$$\Pr[|X - E[X]| \geq t \cdot \sigma_X] \leq \frac{1}{t^2}$$

Prova:

- $\Pr[|X - E[X]| \geq t \cdot \sigma_X] = \Pr[(X-E[X])^2 \geq t^2 \cdot \sigma_X^2]$
- Se $Y = (X - E[X])^2$, por Markov temos,

$\Pr[(X-E[X])^2 \geq t^2 \, Var(X)] \leq Var(X)/(t^2 Var(X)) = 1/t^2$

Quicksort: Expected Number of Comparisons

Theorem. Expected # of comparisons is 2nln n.
Pf.

Theorem. [Knuth 1973] Stddev of number of comparisons is ~ 0.65N.
↑

Ex. If n = 1 million, the probability that randomized quicksort takes less than 4n ln n comparisons is at least 99.4%.

Chebyshev's inequality. $\Pr[|X - \mu| \geq k\delta] \leq 1 / k^2$.

The result is established by setting k= 2 ln(n)/0.65

Desigualdade de Chebyschev

Qual é a probabilidadade obtermos mais de 75 caras ao jogar uma moeda justa 100 vezes ?

X: número de vezes que o resultado é cara
X_i :1 se o resultado da i-ésima tentativa é cara e 0, caso contrário

$$E[X] = \sum_{i=1}^{100} X_i = 50 \text{ e } Var(X) = \sum_{i=1}^{100} Var(X_i) = 25$$

Aplicando Chebyschev temos

$\Pr[|X-50| \geq 5x5] \leq 1/25 \rightarrow \leq 4\%$

$\Pr[\, X \geq 75 \,] \leq 2\%$ (simetria)

Coupon Collector

- X: number of rounds until all coupons have been collected.
- From previous class: $E[X] = nH_n$
- Markov inequality ➔ $Pr[X \geq 2H_n] \leq 1/2$

Coupon Collector

- X_i: number of rounds between the appereance of the (i-1)-th coupon and the i-th coupon

- X_i and X_j are independent for i<>j

- Variance of X

$$Var[X] = Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i]$$

- X_i is geometric random variable with

Coupon Collector

Lemma 3.8: If Z is a geometric random variable with parameter p then

$$Var[Z] = (1-p)/p^2 \leq 1/p^2$$

Proof

Page 51

Coupon Collector

- We have that

$$Var[X] = \sum_{i=1}^{n} Var[X_i] \le \sum_{i=1}^{n} \left(\frac{n}{n-i+1}\right)^2 \le \frac{\pi^2 n^2}{6}$$

By Chebyshev's

$$\Pr(|X - nH_n| \ge nH_n) \le \frac{\pi^2 n^2}{6(n^2 H_n{}^2)} = O\left(\frac{1}{ln^2\, n}\right)$$

Coupon Collector

- E(i,k): event that we do not obtain the i-th coupon after k rounds

$$\Pr\left[E[i,k]\right] = \left(1 - \frac{1}{n}\right)^k$$

- E[k]: event that we do not obtain some coupon after k rounds

$$\Pr\left[\bigcup_{k=1\ldots n} E[i,k]\right] \leq n\left(1 - \frac{1}{n}\right)^k$$

- For k=2nln n we have probability at most 1/n

Median Problem

Input: S = {a1,a2,…….,an}

Output: median m of S

Median Problem
Deterministic Algorithms

Deterministic Algorithms

- Sorting Approach with a time complexity
  - $T(n) = n \log n$
- Medians of Medians:
  - Linear time but large constant

Randomized Algorithm: Approach

- Sample to find elements u and d in S such that
  (a) $u<m<d$ and
  (b) $T=\{x \in S \mid u \leq x \leq d\}$ has $O(n / \log n)$ elements

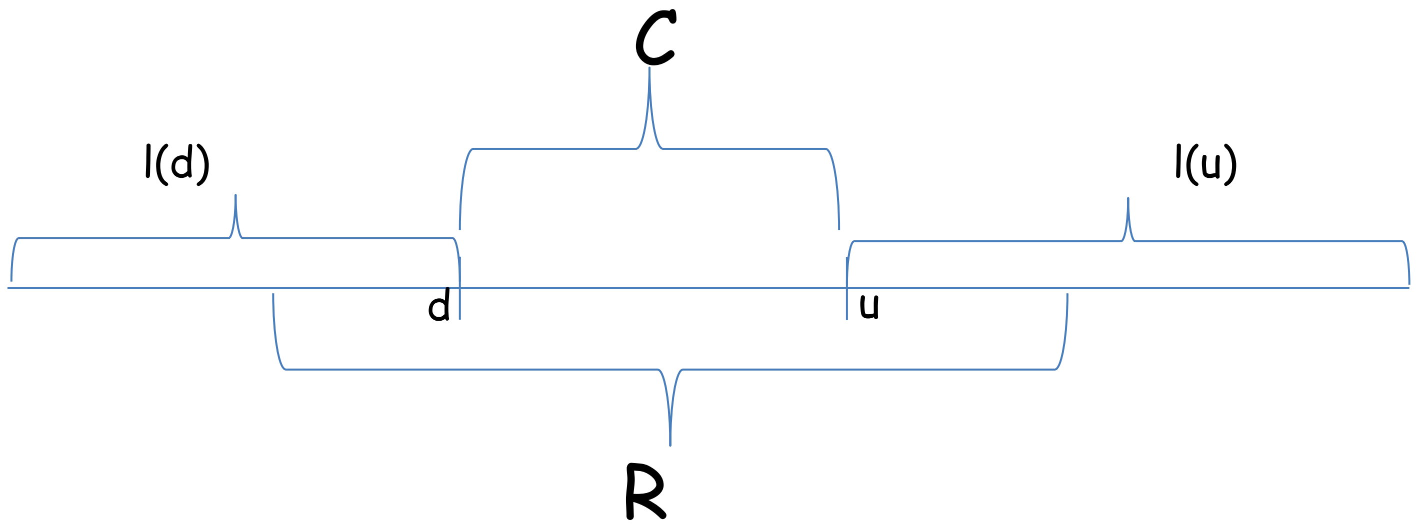
- Find m by sorting T and counting the number of elements in S smaller than u.

Randomized Algorithm for median

1. Sample with replacement a multiset R with $O(n^{3/4})$ elements
2. Sort R in sublinear time
3. u ← $((1/2)n^{3/4} - n^{1/2})$-th smallest element of R
4. d ← $((1/2)n^{3/4} + n^{1/2})$-th smallest element of R
5. Find C={elements of S between d and u};
   l_d ← number elements in S smaller than d and
   l_u ← number elements in S larger than u

6. If $l_d > n/2$ or $l_u > n/2$ FAIL (median is not in C)
7. If $|C| <= 4n^{3/4}$ Sort C, otherwise FAIL (set C is large)
8. Output the $(n/2 - l_d)$-th smallest element of C.

Randomized Algorithm for median



- If C is "small" and the median is between d and u we succeed.

Randomized Algorithm for median

Analysis

E1: $Y(1)=|\{r \in R \mid r \leq m\}| < (1/2)n^{3/4} - n^{1/2}$

E2: $Y(2)=|\{r \in R \mid r>=m\}| < 1/2n^{3/4} - n^{1/2}$

E3: $|C| > 4n^{3/4}$

Lemma. The algorithm fails iff at least one the of the events E1, E2, E3 happens

Proof Sketch.

- Line 7 FAIL iff E3 happens
- Line 6 FAILS iff E1 or E2 happens

Randomized Algorithm for median

Lemma 3.11 Pr(E1)=Pr(E2) $\leq$ 1/ (4 $n^{1/4}$)

Proof.

$$X_i = \begin{cases} 1, if\ sample\ i\ is \leq median\ m \\ 0, \qquad otherwise \end{cases}$$

We have $P[X_i=1] = 1/2 + 1/2n$

Event E1 is equivalent to

$$Y = \sum_{i=1}^{n^{3/4}} X_i \geq \frac{n^{\frac{3}{4}}}{2} - n^{\frac{1}{2}}$$

Randomized Algorithm for median

Lemma 3.11 Pr(E1)=Pr(E2) $\leq$ 1/ (4 $n^{1/4}$)

We use Chebyshev to give an upper bound on

$$\Pr\left[Y = \sum_{i=1}^{n^{3/4}} X_i \geq \frac{n^{\frac{3}{4}}}{2} - n^{\frac{1}{2}}\right]$$

Sampling without replacement implies that variable X's are independent

Y is a binomial random variable with parameters $n^{3/4}$ and (1/2 + 1/2n)

Randomized Algorithm for median

Lemma 3.11 $Pr(E1)=Pr(E2) \leq 1/(4\ n^{1/4})$

Y is a binomial random variable with parameters $n^{3/4}$ and $(1/2 + 1/2n)$

$$E[Y] = n^{3/4}/2 + n^{3/4}/2n$$

$$VAR[Y] = n^{3/4}(1/2 + 1/2n)(1/2 - 1/2n) < 1/4\ n^{3/4}$$

Applying Chebyshev we get that

$$Pr[E1]=Pr[E2] \leq 1/(4\ n^{1/4})$$

Randomized Algorithm for median

Lemma 3.12 $Pr(E3) \leq 1/(2\, n^{1/4})$

Proof. If event E3 occurs than one of the following events also occurs

Event A : at least $2n^{3/4}$ elements of C are larger than m

Event B : at least $2n^{3/4}$ elements of C are smaller than m

Thus,

$$Pr[E3] <= Pr[A] + Pr[B]$$

Randomized Algorithm for median

Lemma 3.12 Pr(E3) $\leq 1/(2\ n^{1/4})$

Proof.

If Event A occurs => at least $2n^{3/4}$ elements of C are larger than m

=> Rank of u in the list S is at least $1/2n + 2n^{3/4}$

$\Rightarrow$ Set R has at least $(1/2)\ n^{3/4} - (1/2)\ n^{1/2}$ samples among the $1/2n - 2n^{3/4}$ largest elements of S

We can bound the probability of the occurrence of the last event.

Randomized Algorithm for median

Lemma 3.12 $\Pr(E3) \leq 1/(2\, n^{1/4})$

$$X_i = \begin{cases} 1, \textit{if the ith sample i is among the } \frac{n}{2} - 2n^{\frac{3}{4}} \textit{ largest of } S \\ 0, \qquad \textit{otherwise} \end{cases}$$

Let

$$X = \sum_{i=1}^{n^{3/4}} X_i$$

We have to give an upper bound on

$$\Pr\left[X > \frac{n}{2} - 2n^{\frac{3}{4}}\right]$$

Randomized Algorithm for median

X is a binomial random variable with parameters $n^{3/4}$ and $\frac{1}{2} - 2n^{\frac{-1}{4}}$. Thus,

E[X] =

VAR[X] =

Applying Chebyshev

Pr[B]=Pr[A]<=$Pr[X > \frac{n}{2} - 2n^{\frac{3}{4}}] \leq$ 1/ (4 $n^{1/4}$)

Pr[E3]<=Pr[B]+Pr[A] $\leq$ 1/ (2 $n^{1/4}$)

Randomized Algorithm for median

The probabiliity that randomized median fails is at most

$$Pr[E1]+Pr[E2]+Pr[E3] \leq 1/n^{1/4}$$

Chernoff Bounds (above mean)

Theorem. Suppose $X_1, \ldots, X_n$ are independent 0-1 random variables. Let $X = X_1 + \ldots + X_n$. Then for any $\mu \geq E[X]$ and for any $\delta > 0$, we have

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

↑ sum of independent 0-1 random variables is tightly centered on the mean

Pf. We apply a number of simple transformations.

- For any $t > 0$,

$$\Pr[X > (1+\delta)\mu] = \Pr\left[e^{tX} > e^{t(1+\delta)\mu}\right] \leq e^{-t(1+\delta)\mu} \cdot E[e^{tX}]$$

↑ $f(x) = e^{tx}$ is monotone in x

↑ Markov's inequality: $\Pr[X > a] \leq E[X] / a$

- Now $E[e^{tX}] = E[e^{t\sum_i X_i}] = \prod_i E[e^{tX_i}]$

↑ definition of X

↑ independence

Chernoff Bounds (above mean)

Pf. (cont)

- Let $p_i = \Pr[X_i = 1]$. Then,

$$E[e^{tX_i}] = p_i e^t + (1-p_i)e^0 = 1 + p_i(e^t - 1) \le e^{p_i(e^t-1)}$$

↑ for any $\alpha \ge 0$, $1+\alpha \le e^{\alpha}$

- Combining everything:

$$\Pr[X > (1+\delta)\mu] \le e^{-t(1+\delta)\mu} \prod_i E[e^{tX_i}] \le e^{-t(1+\delta)\mu} \prod_i e^{p_i(e^t-1)} \le e^{-t(1+\delta)\mu}\, e^{\mu(e^t-1)}$$

↑ previous slide; ↑ inequality above; ↑ $\Sigma_i\, p_i = E[X] \le \mu$

- Finally, choose $t = \ln(1 + \delta)$. ▪

Chernoff Bounds (below mean)

Theorem. Suppose $X_1, \ldots, X_n$ are independent 0-1 random variables. Let $X = X_1 + \ldots + X_n$. Then for any $\mu \le E[X]$ and for any $0 < \delta < 1$, we have

$$\Pr[X < (1-\delta)\mu] \; < \; e^{-\delta^2 \mu / 2}$$

Pf idea. Similar.

Remark. Not quite symmetric since only makes sense to consider $\delta < 1$.

Chernoff Bounds

Qual é a probabilidadade obtermos mais de 75 caras ao jogar uma moeda justa 100 vezes ?

X: número de vezes que o resultado é cara
X_i :1 se o resultado da i-ésima tentativa é cara e 0 caso contrário

$$E[X] = \sum_{i=1}^{100} X_i = 50 \text{ e } Var(X) = \sum_{i=1}^{100} Var(X_i) = 25$$

Aplicando Chernoff temos

Pr[|X-50|>(1+0.5) x50]<=0.007 ➔ <= 0.7%

13.10 Load Balancing

Load Balancing

Load balancing. System in which m jobs arrive in a stream and need to be processed immediately on n identical processors. Find an assignment that balances the workload across processors.

Centralized controller. Assign jobs in round-robin manner. Each processor receives at most $\lceil m/n \rceil$ jobs.

Decentralized controller. Assign jobs to processors uniformly at random. How likely is it that some processor is assigned "too many" jobs?

Analysis.

- Let X_i = number of jobs assigned to processor i.
- Let Y_{ij} = 1 if job j assigned to processor i, and 0 otherwise.
- We have $E[Y_{ij}] = 1/n$
- Thus, $X_i = \sum_j Y_{ij}$, and $\mu = E[X_i] = 1$.
- Applying Chernoff bounds with $\delta = c - 1$ yields $\Pr[X_i > c] < \frac{e^{c-1}}{c^c}$
- Let $\gamma(n)$ be number x such that $x^x = n$, and choose $c = e\,\gamma(n)$.

$$\Pr[X_i > c] < \frac{e^{c-1}}{c^c} < \left(\frac{e}{c}\right)^c = \left(\frac{1}{\gamma(n)}\right)^{e\gamma(n)} < \left(\frac{1}{\gamma(n)}\right)^{2\gamma(n)} = \frac{1}{n^2}$$

- Union bound $\Rightarrow$ with probability $\geq 1 - 1/n$ no processor receives more than $e\,\gamma(n) = \Theta(\log n / \log\log n)$ jobs.

Fact: this bound is asymptotically tight: with high probability, some processor receives $\Theta(\log n / \log\log n)$

Load Balancing: Many Jobs

Theorem. Suppose the number of jobs $m = 16n \ln n$. Then on average, each of the n processors handles $\mu = 16 \ln n$ jobs. With high probability every processor will have between half and twice the average load.

Pf.

- Let X_i , Y_{ij} be as before.
- Applying Chernoff bounds with $\delta = 1$ yields

$$\Pr[X_i > 2\mu] < \left(\frac{e}{4}\right)^{16 n \ln n} < \left(\frac{1}{e}\right)^{\ln n} = \frac{1}{n^2} \qquad \Pr[X_i < \tfrac{1}{2}\mu] < e^{-\frac{1}{2}\left(\frac{1}{2}\right)^2 (16 n \ln n)} = \frac{1}{n^2}$$

- Probability of a processor has load larger than 2μ or smaller than $1/2\ \mu$ is at most $2/n^2$.
- Union bound $\Rightarrow$ Probability of some processor has load larger than 2μ or smaller than $1/2\ \mu$ is at most $2/n$
- every processor has load between half and twice the average with probability $\geq 1 - 2/n$. ▪

13.11 Packet Routing