CSE 202: Design and Analysis of Algorithms

Lecture 18

Instructor: Kamalika Chaudhuri

Announcements

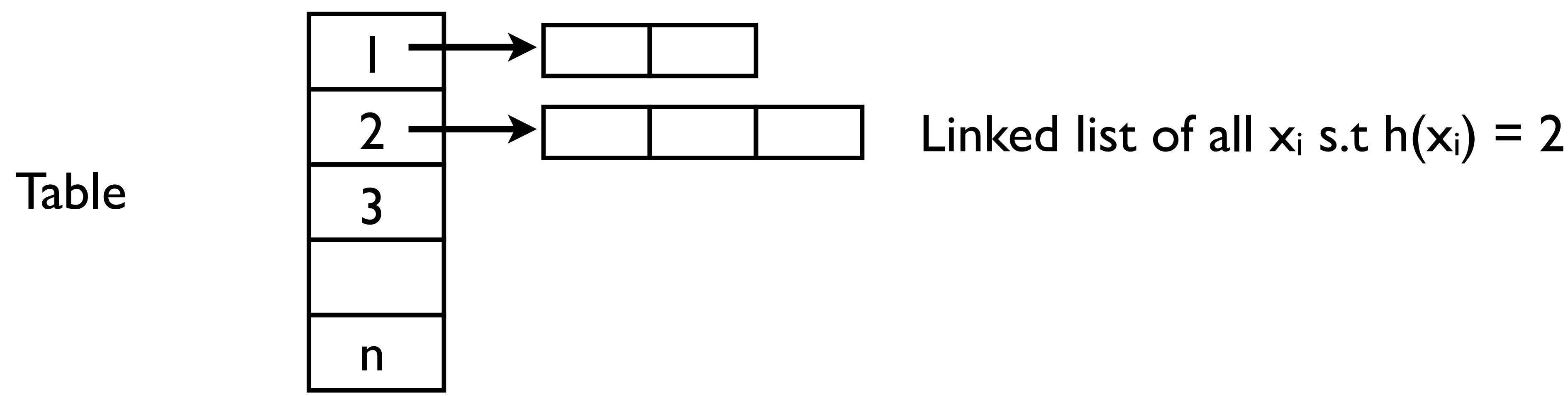
- No TA Office Hours on Thu May 26
- Extra Instructor Office Hours Thu May 26 5-6pm (instead of 2:30-3:30pm), CSE 4110
- No class on Monday May 30 (Memorial Day)
- Extra Instructor Office Hours Tue May 31 5-6pm at CSE 4110
- TA Office Hours on Tue May 31 moved to Wed Jun 1, 11-12, B250A
- Pick up graded Homework 3 after class

Randomized Algorithms

- Contention Resolution
- Some Facts about Random Variables
- Global Minimum Cut Algorithm
- Randomized Selection and Sorting
- Max 3-SAT
- Three Concentration Inequalities
- Hashing and Balls and Bins

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not

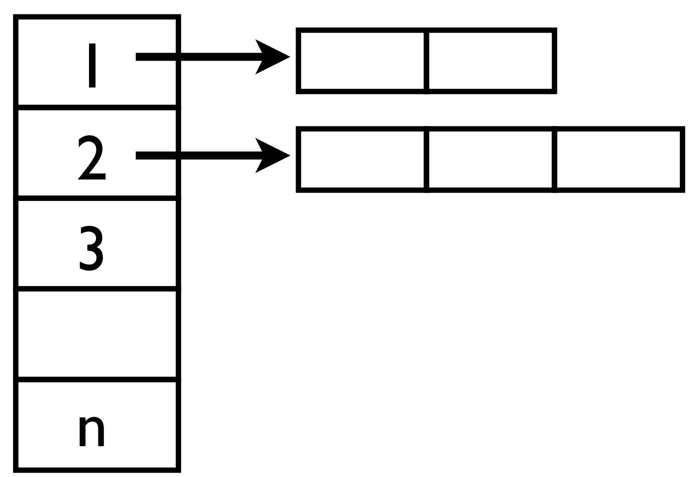


Popular Data Structure: A Hash table

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not

Table



Linked list of all x_i s.t $h(x_i) = 2$

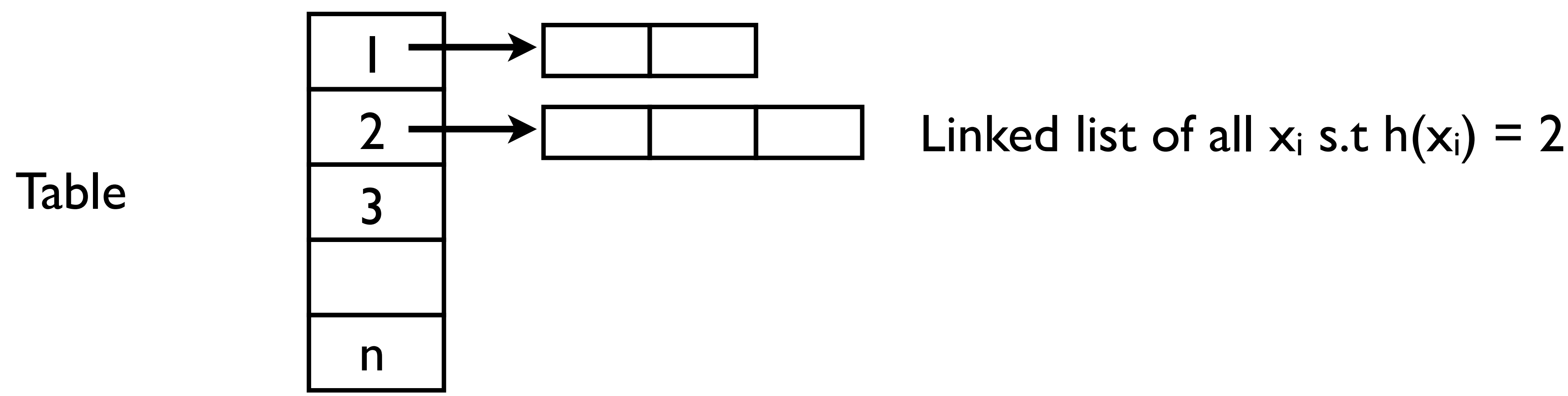
Popular Data Structure: A Hash table

Algorithm:

1. Pick a completely random function $h : \mathcal{U} \rightarrow \{1, \ldots, n\}$
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not



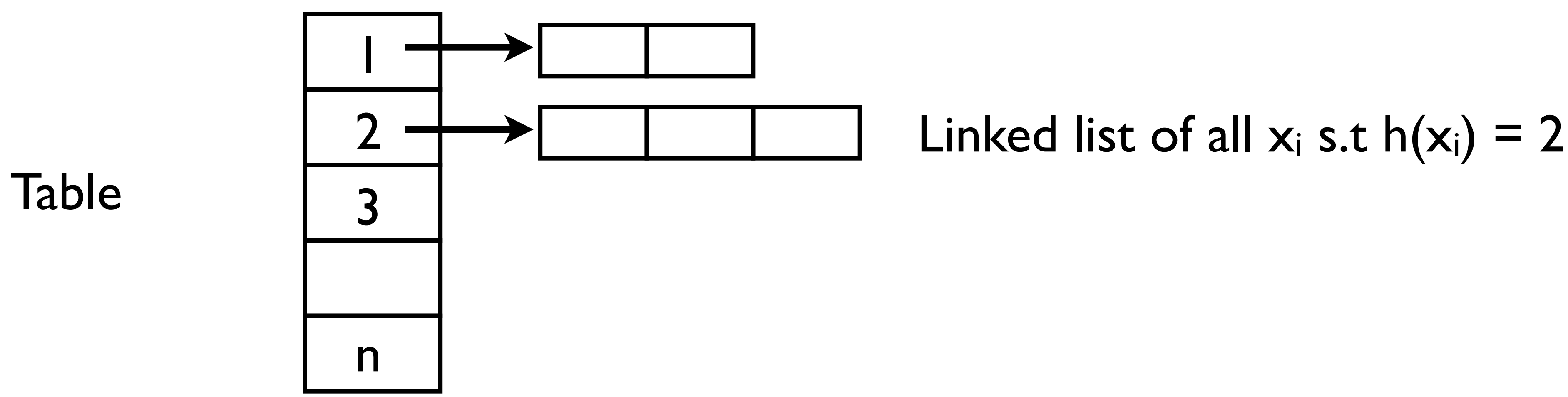
Popular Data Structure: A Hash table

Algorithm:

1. Pick a completely random function $h : \mathcal{U} \rightarrow \{1, \ldots, n\}$
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, look at the linked list at location h(q) of table to see if q is there

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not



Popular Data Structure: A Hash table

Algorithm:

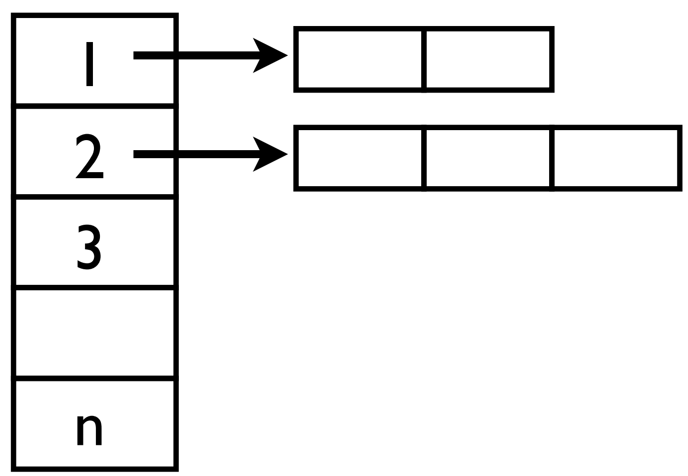
1. Pick a completely random function $h : \mathcal{U} \to \{1, \ldots, n\}$
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, look at the linked list at location h(q) of table to see if q is there

What is the query time of the algorithm?

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not

Table



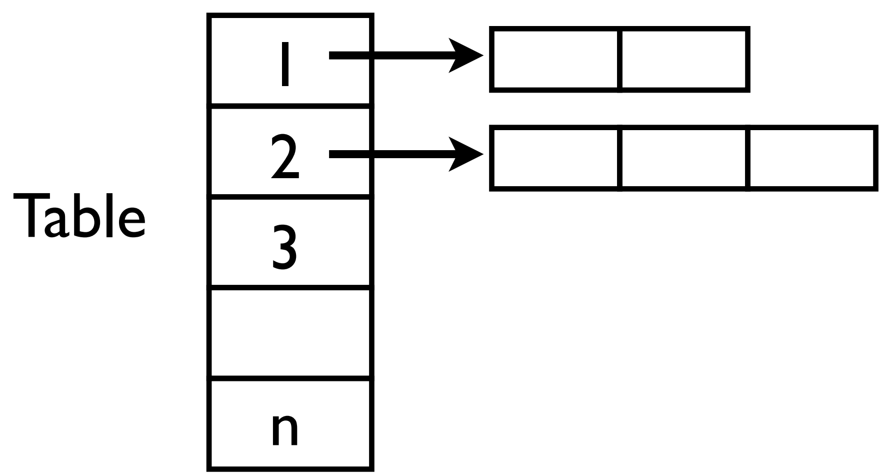
Algorithm:

1. Pick a completely random function h
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, check the linked list at location h(q)

Average Query Time: Suppose q is picked at random s.t it is equally likely to hash to 1, .., n. What is the expected query time?

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not



Algorithm:

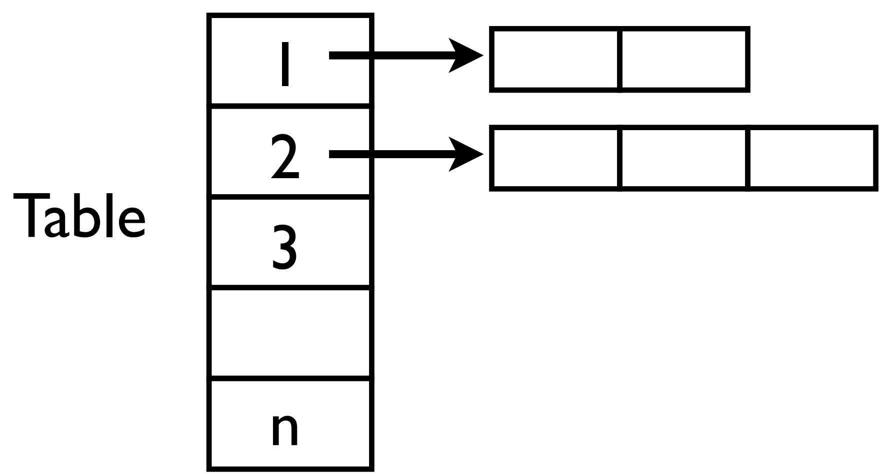
1. Pick a completely random function h
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, check the linked list at location h(q)

Average Query Time: Suppose q is picked at random s.t it is equally likely to hash to 1, .., n. What is the expected query time?

$$\text{Expected Query Time} = \sum_{i=1}^{n} \Pr[q \text{ hashes to location } i] \cdot (\text{length of list at } T[i])]$$

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not



Algorithm:

1. Pick a completely random function h
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, check the linked list at location h(q)

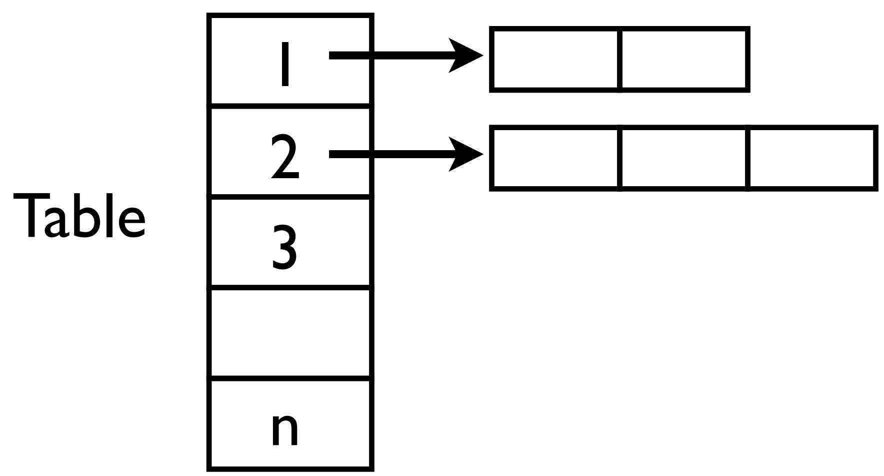
Average Query Time: Suppose q is picked at random s.t it is equally likely to hash to 1, .., n. What is the expected query time?

$$\text{Expected Query Time} = \sum_{i=1}^{n} \Pr[q \text{ hashes to location } i] \cdot (\text{length of list at } T[i])]$$

$$= \frac{1}{n} \sum_{i} (\text{length of list at } T[i])$$

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not



Algorithm:

1. Pick a completely random function h
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, check the linked list at location h(q)

Average Query Time: Suppose q is picked at random s.t it is equally likely to hash to 1, .., n. What is the expected query time?

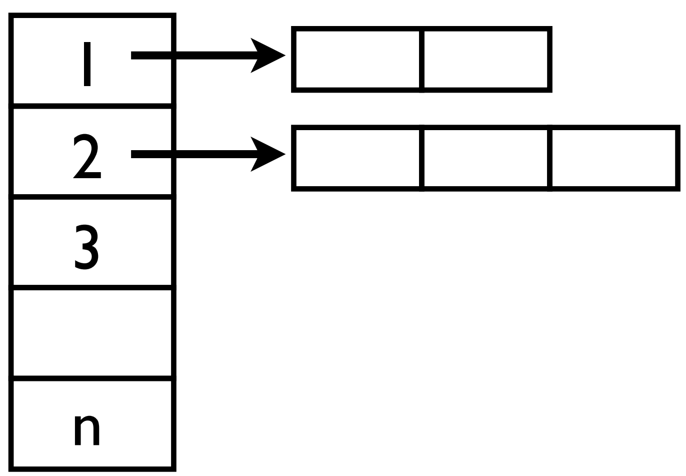
Expected Query Time $= \sum_{i=1}^{n} \Pr[q \text{ hashes to location } i] \cdot (\text{length of list at } T[i])]$

$$= \frac{1}{n} \sum_{i} (\text{length of list at } T[i]) = \frac{1}{n} \cdot n = 1$$

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not

Table



Algorithm:

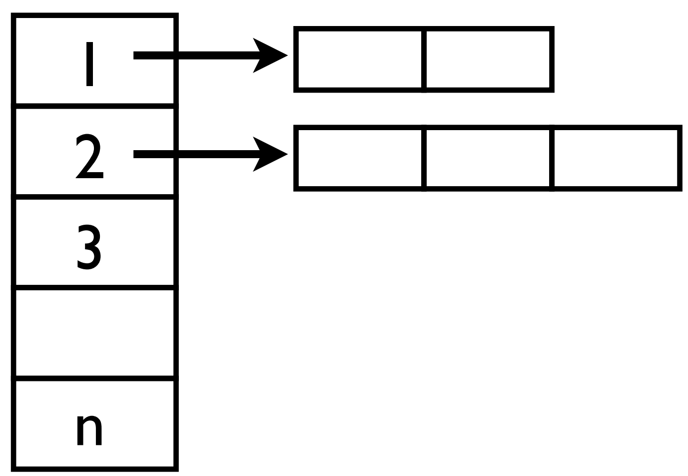
1. Pick a completely random function h
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, check the linked list at location h(q)

Worst Case Query Time: For any q, what is the query time? (with high probability over the choice of hash functions)

Hashing and Balls-n-Bins

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not

Table



Algorithm:

1. Pick a completely random function h
2. Create a table of size n, initialize it to null
3. Store x_i in the linked list at position $h(x_i)$ of table
4. For a query q, check the linked list at location h(q)

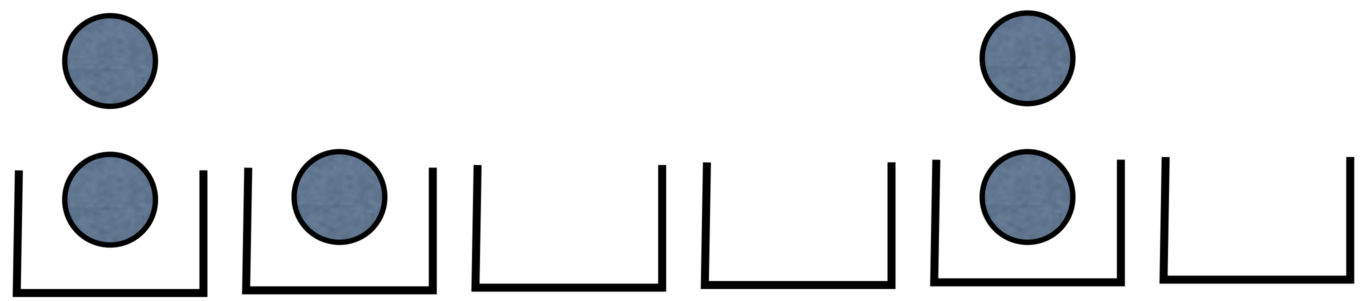
Worst Case Query Time: For any q, what is the query time? (with high probability over the choice of hash functions)

Equivalent to the following Balls and bins Problem:
Suppose we toss n balls u.a.r into n bins. What is the max #balls in a bin with high probability?

With high probability (w.h.p) = With probability 1 - 1/poly(n)

Balls and Bins, again

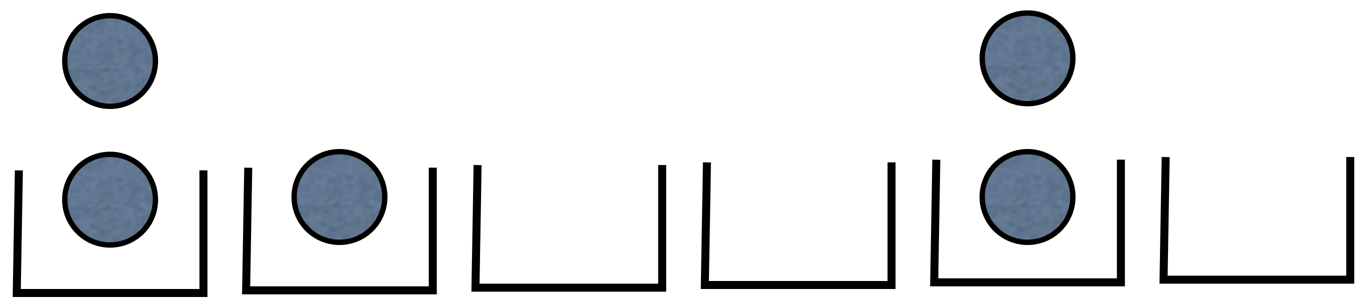
Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Some Facts:

Balls and Bins, again

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



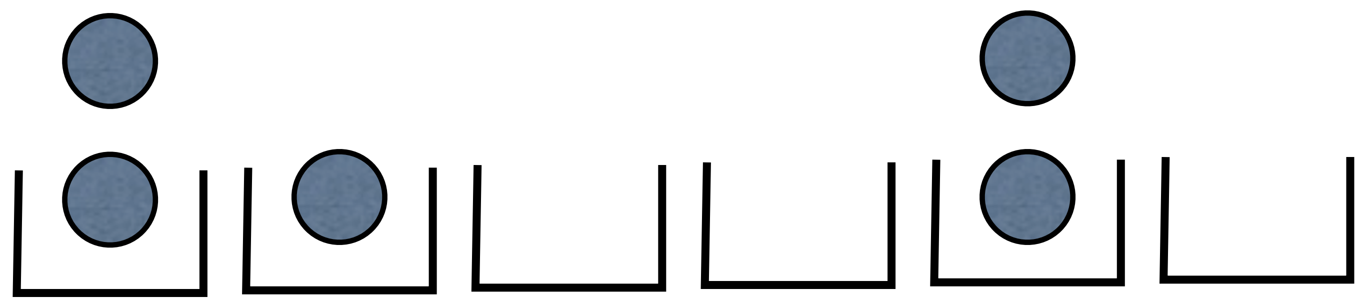
Some Facts:

1. The expected load of each bin is 1

2. What is the probability that each bin has load 1?

Balls and Bins, again

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Some Facts:

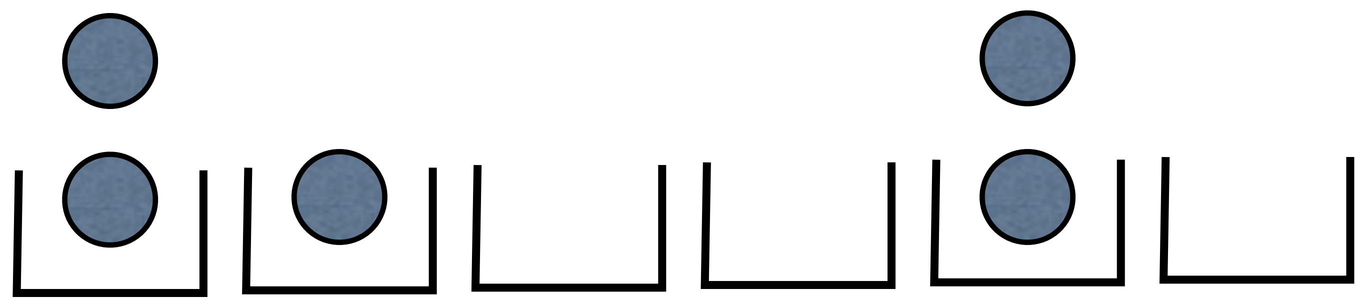
1. The expected load of each bin is 1

2. What is the probability that each bin has load 1?

$$\text{Probability} = \frac{\#\text{ permutations}}{\#\text{ ways of tossing n balls to n bins}} = \frac{n!}{n^n}$$

Balls and Bins, again

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Some Facts:

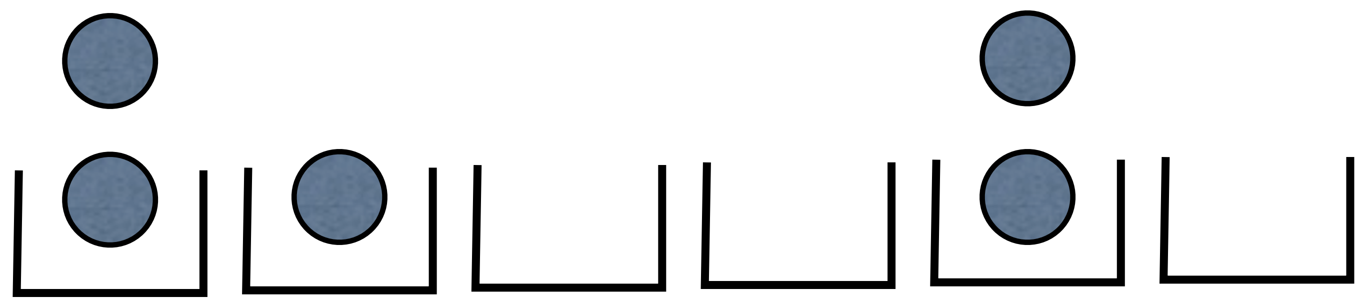
1. The expected load of each bin is 1
2. What is the probability that each bin has load 1?

$$\text{Probability} = \frac{\#\text{ permutations}}{\#\text{ ways of tossing n balls to n bins}} = \frac{n!}{n^n}$$

3. What is the expected #empty bins?

Balls and Bins, again

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Some Facts:

1. The expected load of each bin is 1

2. What is the probability that each bin has load 1?

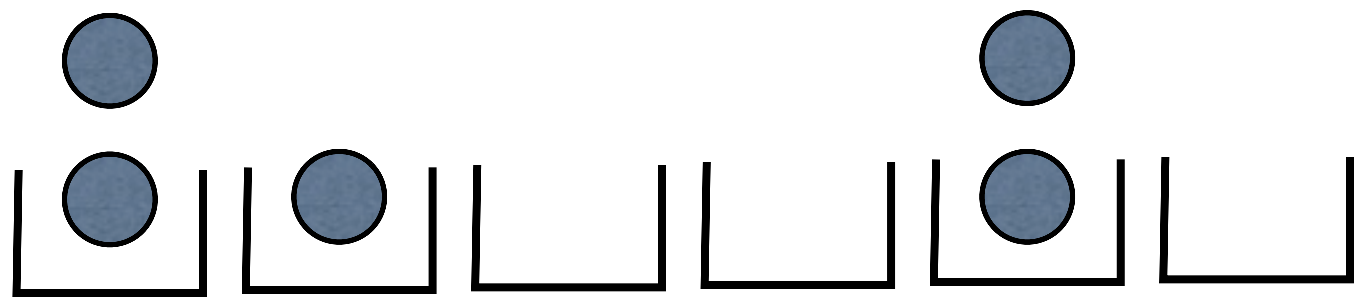
$$\text{Probability} = \frac{\#\text{ permutations}}{\#\text{ ways of tossing n balls to n bins}} = \frac{n!}{n^n}$$

3. What is the expected #empty bins?

$$\Pr[\text{Bin } i \text{ is empty}] = \left(1 - \frac{1}{n}\right)^n$$

Balls and Bins, again

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Some Facts:

1. The expected load of each bin is 1

2. What is the probability that each bin has load 1?

$$\text{Probability} = \frac{\#\text{ permutations}}{\#\text{ ways of tossing n balls to n bins}} = \frac{n!}{n^n}$$

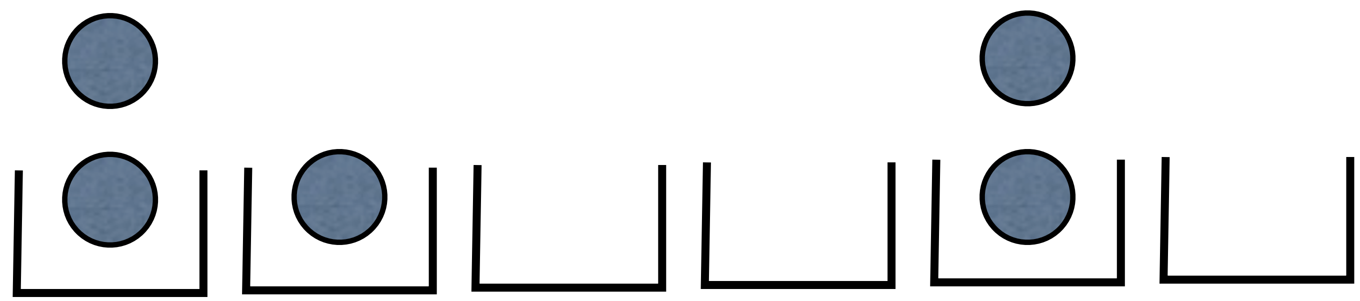
3. What is the expected #empty bins?

$$\Pr[\text{Bin } i \text{ is empty}] = \left(1 - \frac{1}{n}\right)^n$$

$$E[\#\text{ empty bins}] = n\left(1 - \frac{1}{n}\right)^n = \Theta(n)$$

Balls and Bins, again

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Some Facts:

1. The expected load of each bin is 1

2. What is the probability that each bin has load 1?

$$\text{Probability} = \frac{\#\text{ permutations}}{\#\text{ ways of tossing n balls to n bins}} = \frac{n!}{n^n}$$

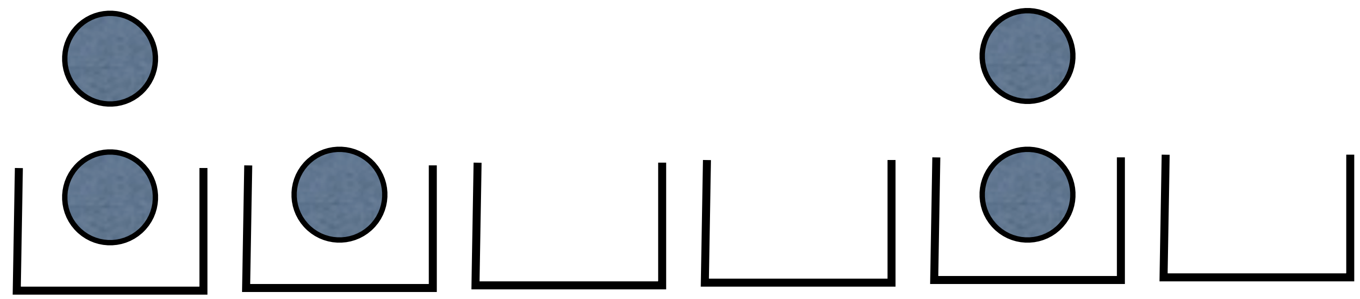
3. What is the expected #empty bins?

$$\Pr[\text{Bin } i \text{ is empty}] = \left(1 - \frac{1}{n}\right)^n$$

$$E[\#\text{ empty bins}] = n\left(1 - \frac{1}{n}\right)^n = \Theta(n)$$ ($(1-1/n)^n$ lies between 1/4 and 1/e for n>=2)

Balls and Bins

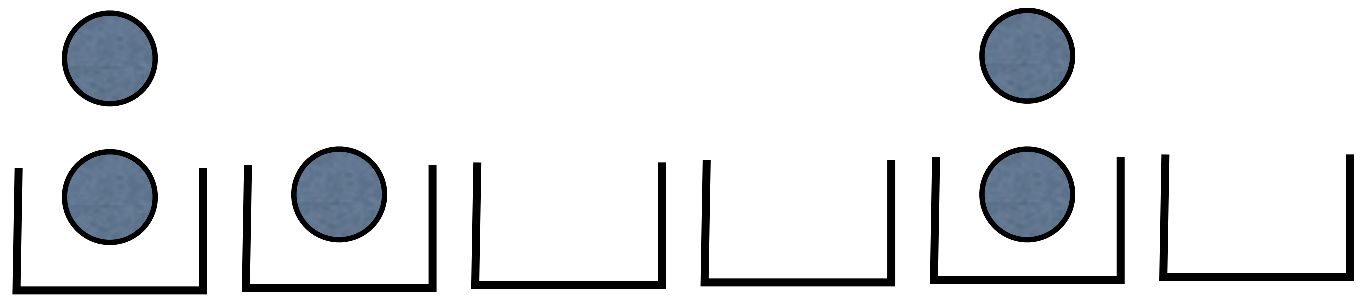
Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?

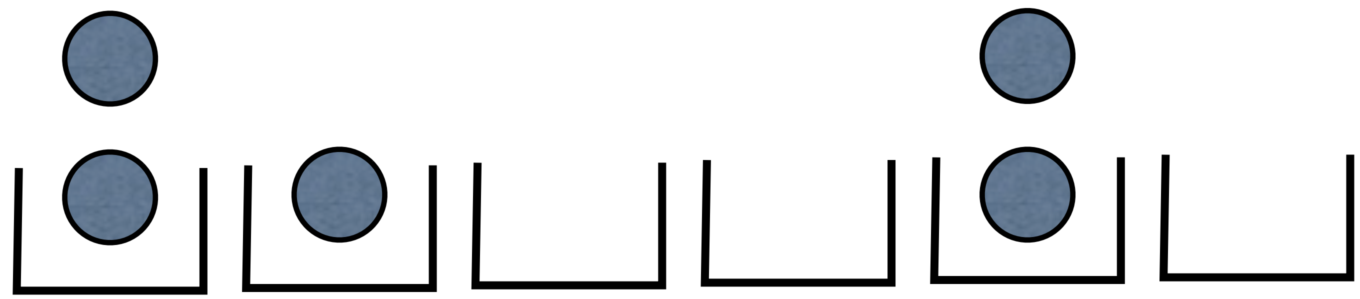


Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t} \frac{1}{n^t}$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t}$$

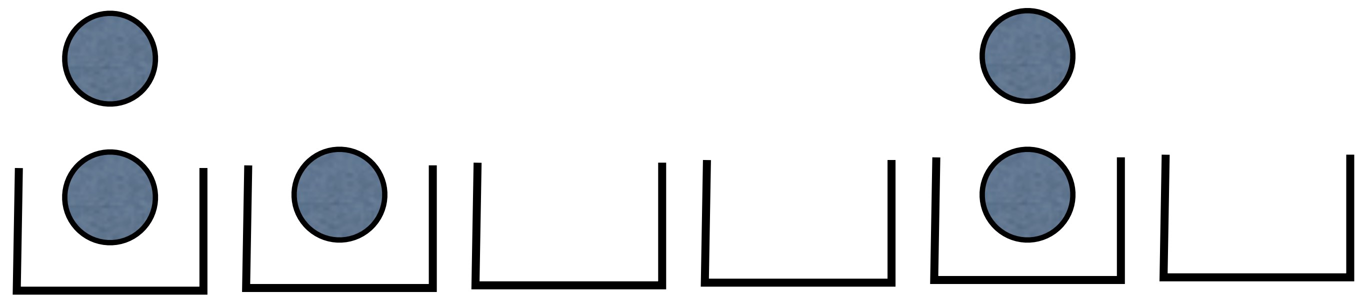
From **Fact**

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \quad \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \quad \leq \left(\frac{e}{t}\right)^t$$

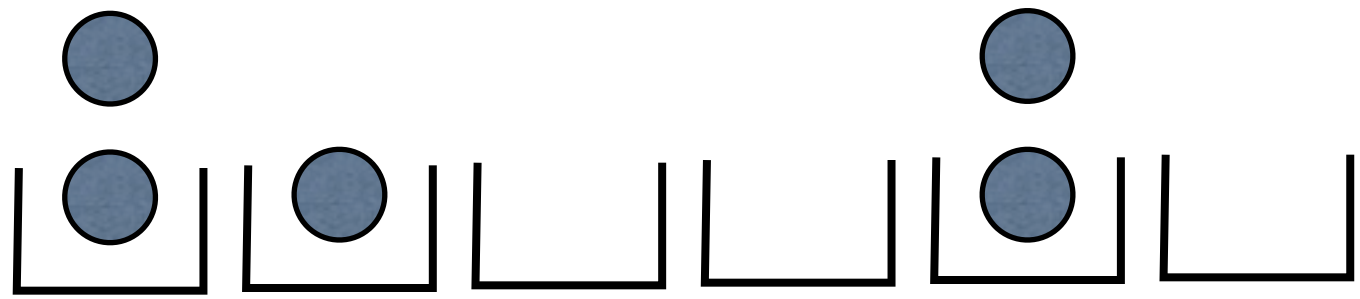
From **Fact**

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \leq \left(\frac{e}{t}\right)^t \leq \frac{1}{n^2}$$

From **Fact**

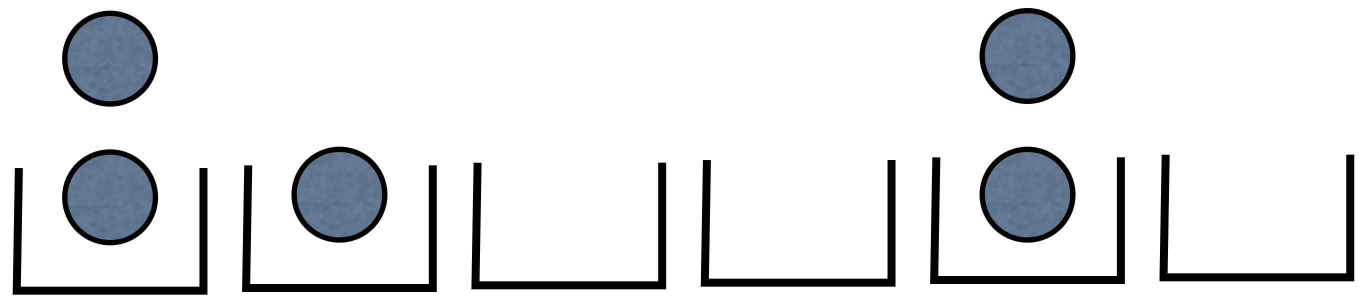
Would like this for whp condition

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \quad \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \quad \leq \left(\frac{e}{t}\right)^t \quad \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

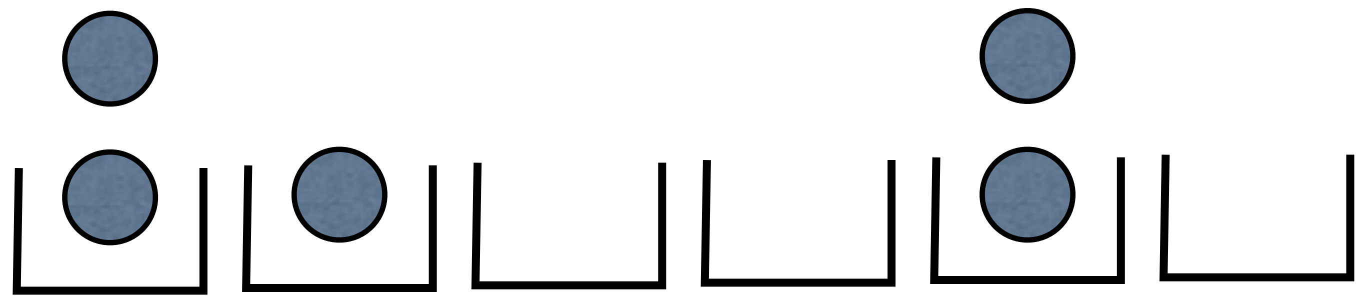
Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

$$\text{Let } t = \frac{c \log n}{\log \log n} \text{ for constant } c$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \quad \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \quad \leq \left(\frac{e}{t}\right)^t \quad \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

Fact: If n >= k

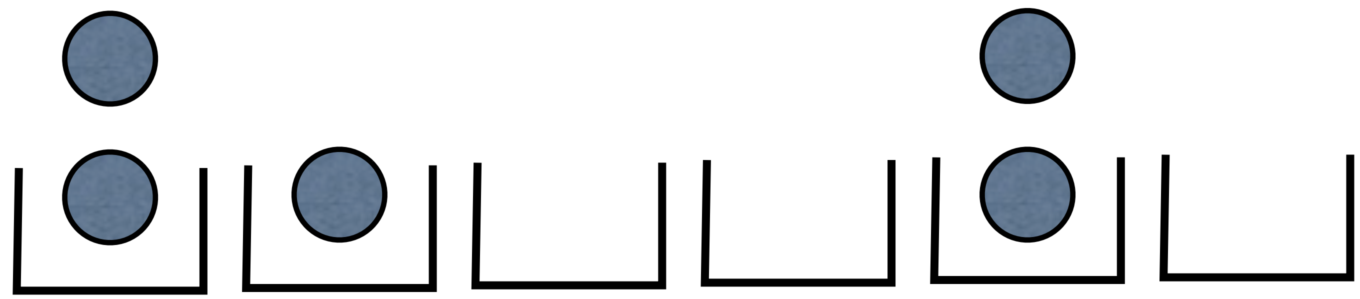
$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

$$\text{Let } t = \frac{c \log n}{\log \log n} \text{ for constant } c$$

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

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \quad \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \quad \leq \left(\frac{e}{t}\right)^t \quad \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

Fact: If n >= k

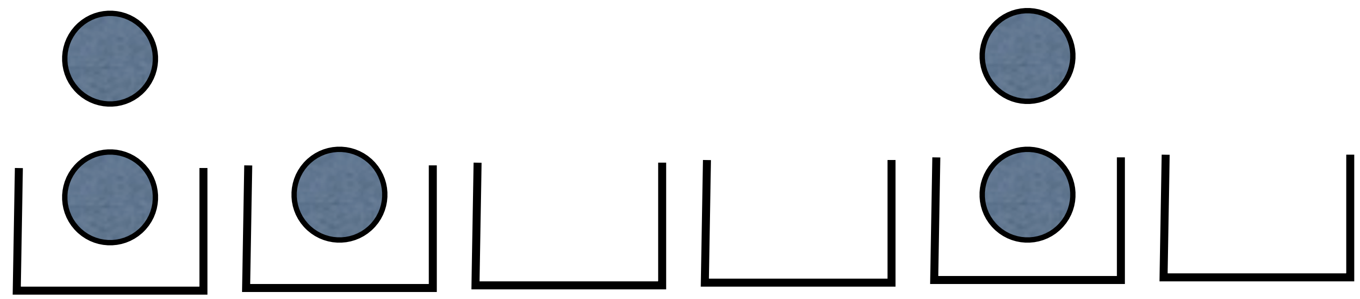
$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

$$\text{Let } t = \frac{c\log n}{\log\log n} \text{ for constant } c$$

$$\log\left(\frac{t}{e}\right)^t = t\log t - t \quad = \frac{c\log n}{\log\log n}\cdot(\log c + \log\log n - \log\log\log n)$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \leq \left(\frac{e}{t}\right)^t \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

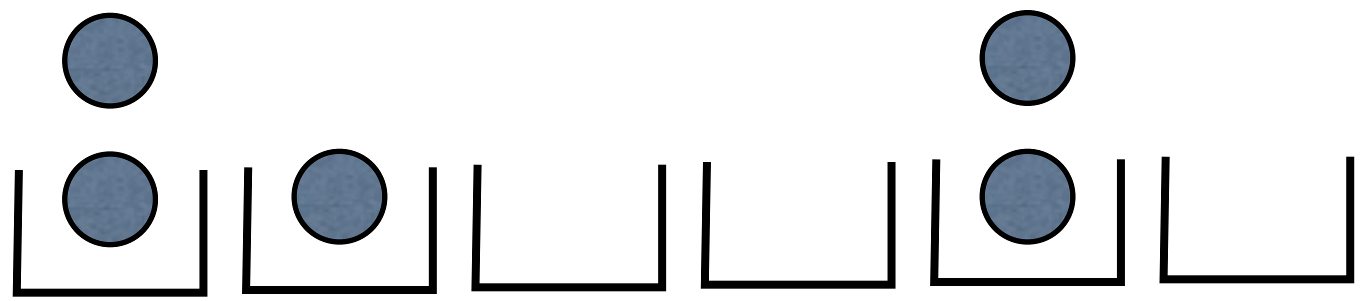
$$\text{Let } t = \frac{c\log n}{\log\log n} \text{ for constant } c$$

$$\log\left(\frac{t}{e}\right)^t = t\log t - t = \frac{c\log n}{\log\log n}\cdot \boxed{(\log c + \log\log n - \log\log\log n)}$$

For large n, this is $\geq \frac{1}{2}\log\log n$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \leq \left(\frac{e}{t}\right)^t \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

$$\text{Let } t = \frac{c\log n}{\log\log n} \text{ for constant } c$$

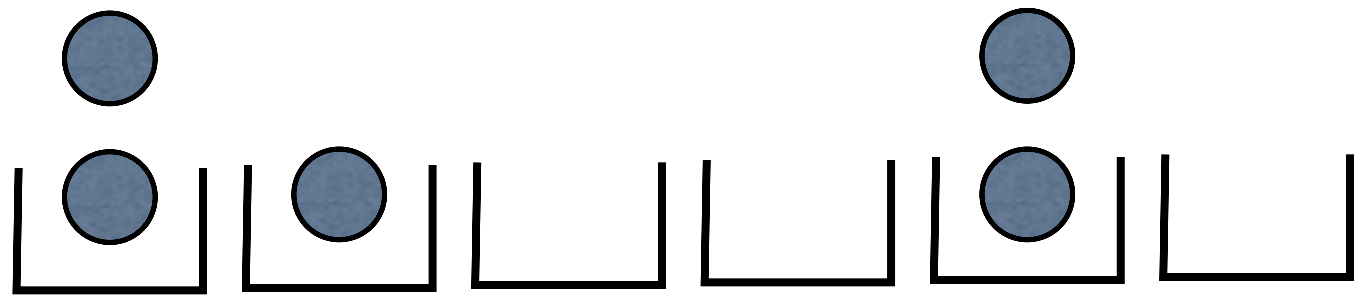
$$\log\left(\frac{t}{e}\right)^t = t\log t - t = \frac{c\log n}{\log\log n}\cdot\boxed{(\log c + \log\log n - \log\log\log n)}$$

$$\geq \frac{c}{2}\log n$$

For large n, this is $\geq \frac{1}{2}\log\log n$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \leq \left(\frac{e}{t}\right)^t \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

$$\text{Let } t = \frac{c\log n}{\log\log n} \text{ for constant } c$$

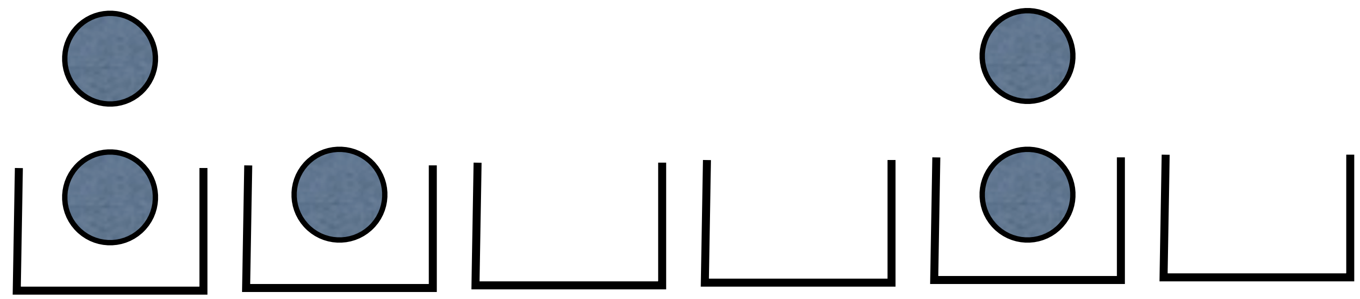
$$\log\left(\frac{t}{e}\right)^t = t\log t - t = \frac{c\log n}{\log\log n}\cdot\boxed{(\log c + \log\log n - \log\log\log n)}$$

$$\geq \frac{c}{2}\log n \geq 2\log n, \text{for } c \geq 4$$

For large n, this is $\geq \frac{1}{2}\log\log n$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \leq \left(\frac{e}{t}\right)^t \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

$$\text{Let } t = \frac{c\log n}{\log\log n} \text{ for constant } c$$

$$\log\left(\frac{t}{e}\right)^t = t\log t - t = \frac{c\log n}{\log\log n} \cdot \boxed{(\log c + \log\log n - \log\log\log n)}$$

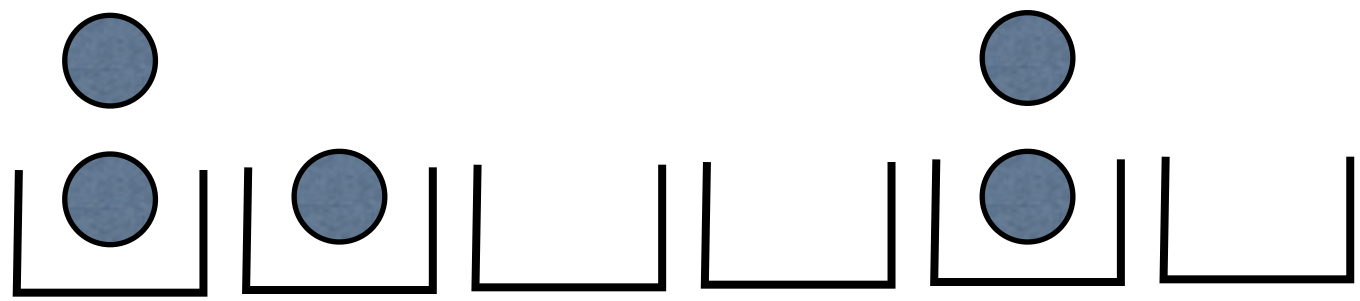
For large n, this is $\geq \frac{1}{2}\log\log n$

$$\geq \frac{c}{2}\log n \quad \geq 2\log n, \text{for } c \geq 4$$

Therefore, w.p. $1/n^2$, there are at least t balls in Bin i. What is Pr(All bins have <= t balls) ?

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \leq \binom{n}{t}\frac{1}{n^t} \leq \left(\frac{ne}{t}\right)^t \frac{1}{n^t} \leq \left(\frac{e}{t}\right)^t \leq \frac{1}{n^2}$$

From **Fact**

Would like this for whp condition

Fact: If n >= k

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$$

$$\text{Let } t = \frac{c\log n}{\log\log n} \text{ for constant } c$$

$$\log\left(\frac{t}{e}\right)^t = t\log t - t = \frac{c\log n}{\log\log n}\cdot \boxed{(\log c + \log\log n - \log\log\log n)}$$

$$\geq \frac{c}{2}\log n \quad \geq 2\log n, \text{for } c \geq 4$$

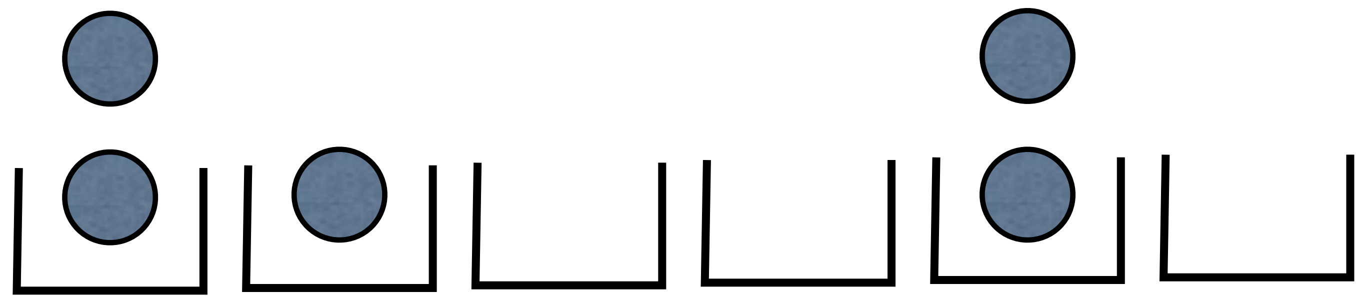
For large n, this is $\geq \frac{1}{2}\log\log n$

Therefore, w.p. $1/n^2$, there are at least t balls in Bin i. What is Pr(All bins have <= t balls) ?

Applying Union Bound, Pr(All bins have <=t balls) >= 1 - 1/n

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?

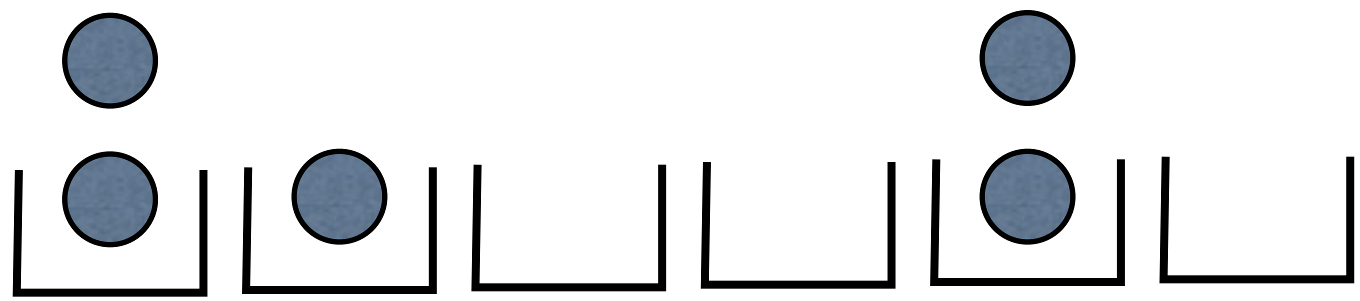


Fact: W.p. $1-1/n$, the maximum load of each bin is at most $O(\log n/\log \log n)$

Fact: The max loaded bin has $(\log n/3\log \log n)$ balls with probability at least 1 - const.$/n^{(1/3)}$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



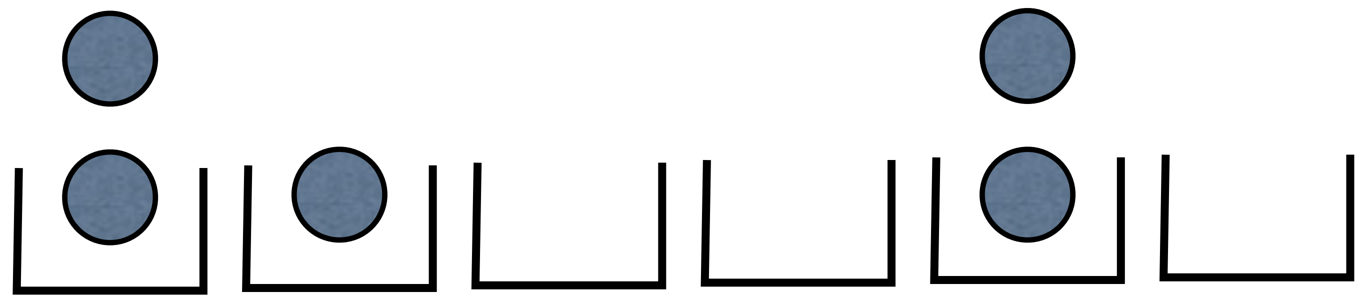
Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

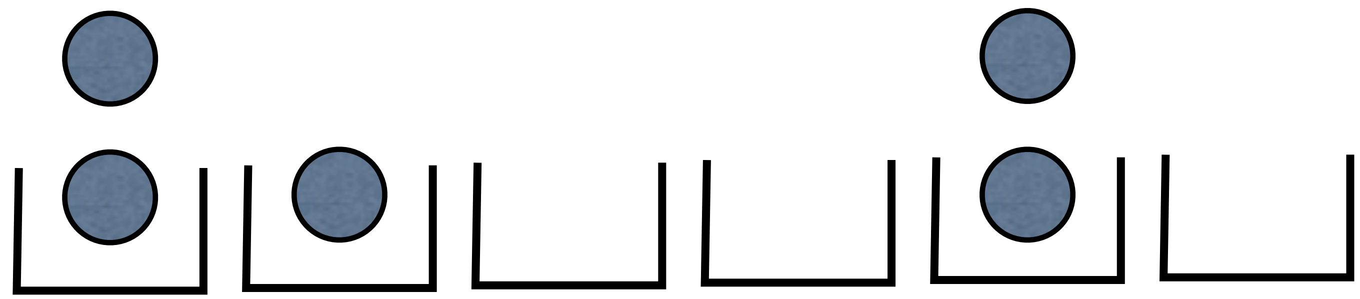
Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t} \frac{1}{n^t} \left(1 - \frac{1}{n}\right)^{n-t}$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

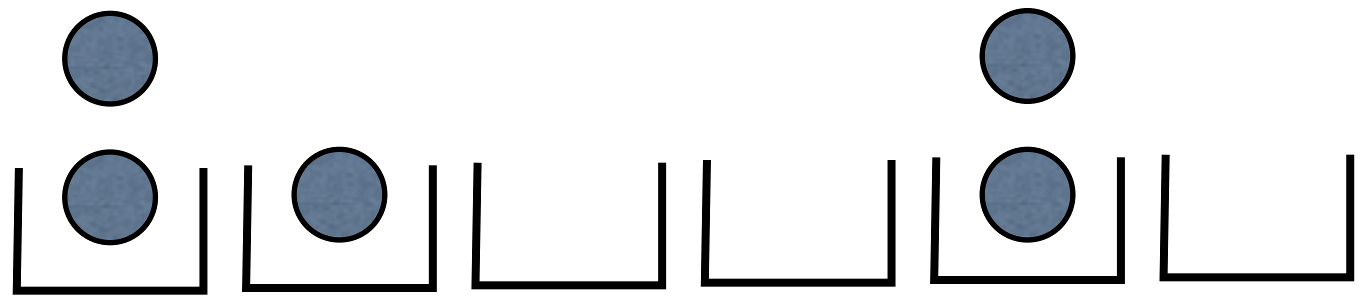
Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1}$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

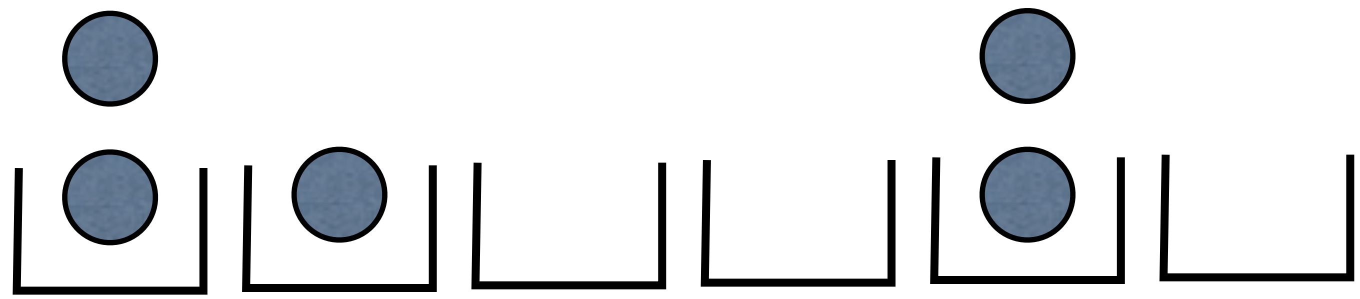
Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./n$^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1} \quad \geq \frac{1}{et^t}$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

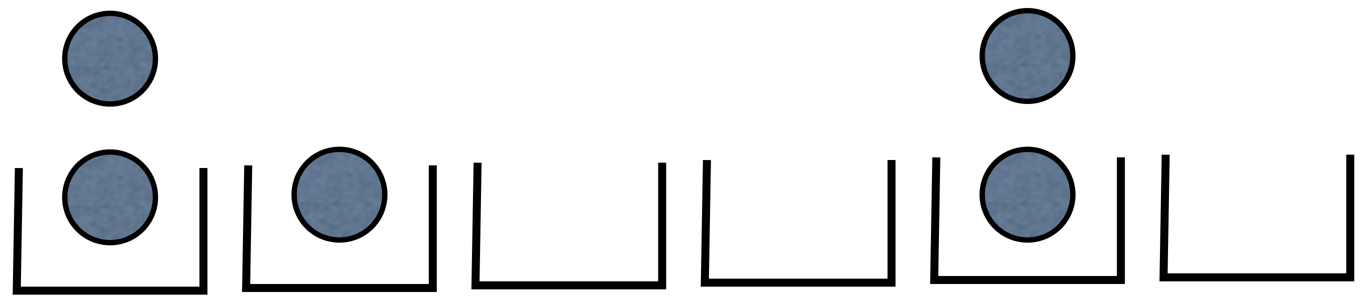
Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1} \quad \boxed{\geq \frac{1}{et^t}}$$

At least $1/en^{1/3}$
for t = log n/3 log log n

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1} \quad \boxed{\geq \frac{1}{et^t}}$$

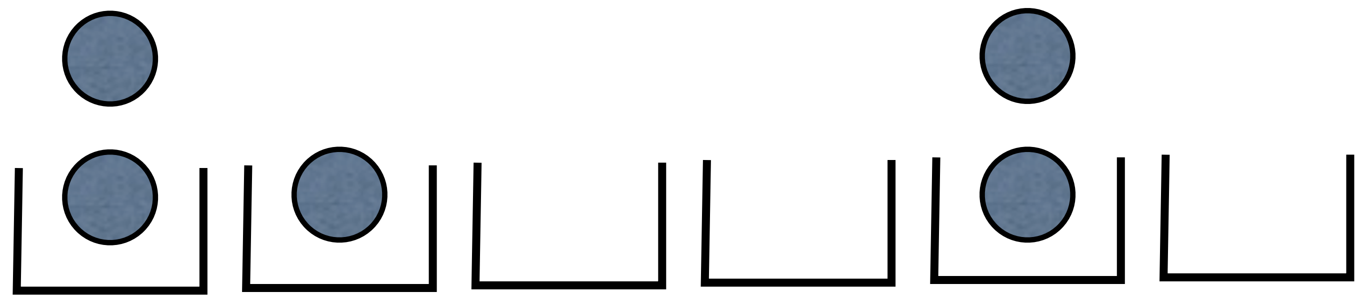
At least $1/en^{1/3}$
for t = log n/3 log log n

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise

$Y = Y_1 + Y_2 + .. + Y_n$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1} \boxed{\geq \frac{1}{et^t}}$$

At least $1/en^{1/3}$
for t = log n/3 log log n

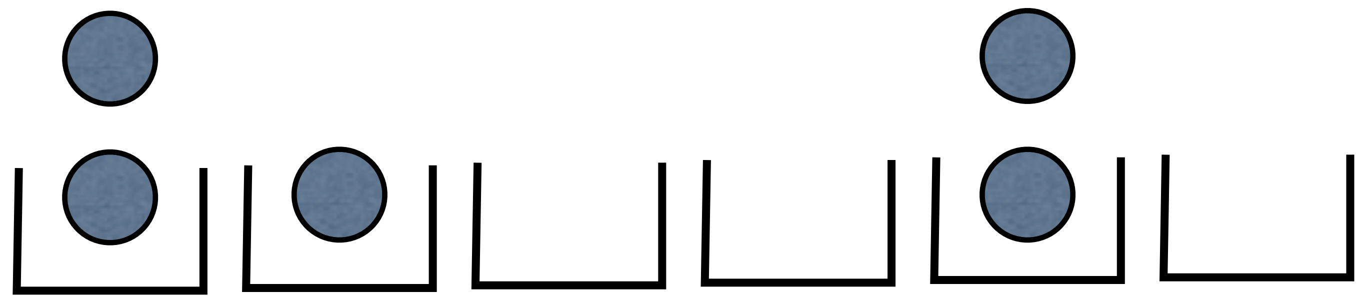
Let Y_i = 1 if bin i has load t or more,
= 0 otherwise

$Y = Y_1 + Y_2 + .. + Y_n$

$\Pr(Y_i = 1) >= 1/en^{1/3}$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1} \quad \boxed{\geq \frac{1}{et^t}}$$

At least $1/en^{1/3}$ for t = log n/3 log log n

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise

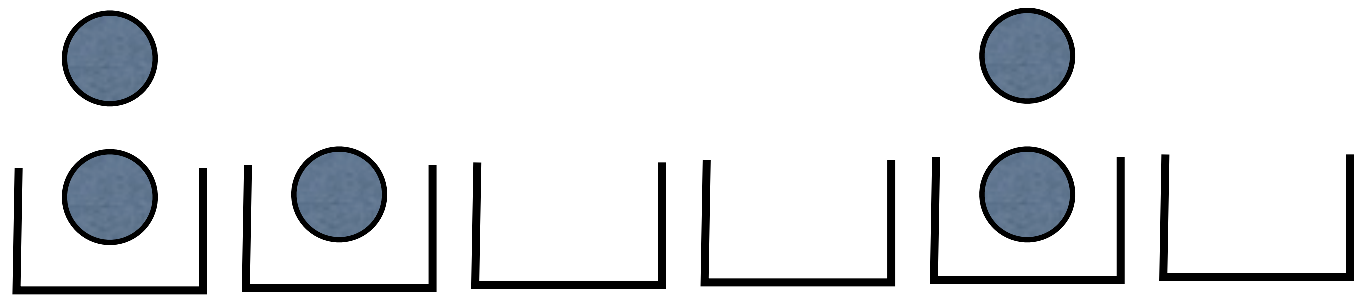
$Y = Y_1 + Y_2 + .. + Y_n$

$\Pr(Y_i = 1) >= 1/en^{1/3}$

$E(Y) >= n^{2/3}/e$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1} \quad \boxed{\geq \frac{1}{et^t}}$$

At least $1/en^{1/3}$ for t = log n/3 log log n

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise

$Y = Y_1 + Y_2 + .. + Y_n$

$\Pr(Y_i = 1) >= 1/en^{1/3}$

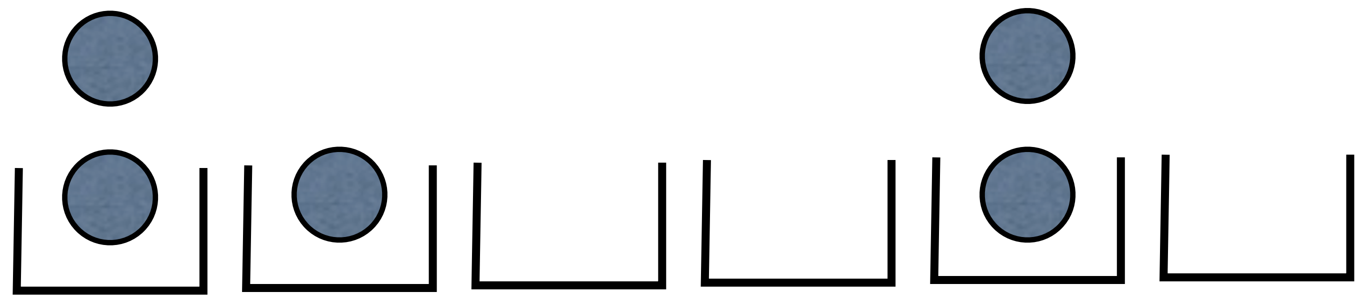
$E(Y) >= n^{2/3}/e$

Pr(Y = 0) = Pr(No bin has load t or more) <= Pr(|Y - E[Y]| >= E[Y])

Which concentration bound to use?

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let X_i = #balls in bin i

$$\Pr(X_i \geq t) \geq \binom{n}{t}\frac{1}{n^t}\left(1-\frac{1}{n}\right)^{n-t} \geq \left(\frac{n}{t}\right)^t \cdot \frac{1}{n^t} \cdot e^{-1} \boxed{\geq \frac{1}{et^t}}$$

At least $1/en^{1/3}$
for t = log n/3 log log n

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise
$Y = Y_1 + Y_2 + .. + Y_n$

$\Pr(Y_i = 1) >= 1/en^{1/3}$

$E(Y) >= n^{2/3}/e$

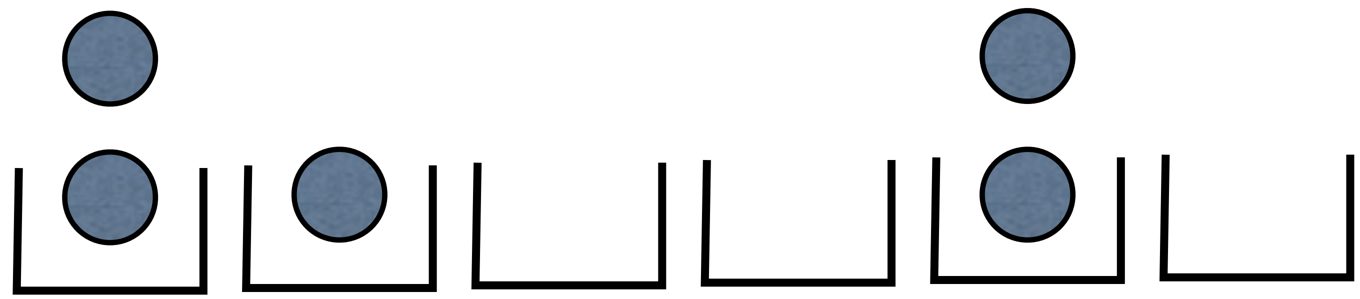
Pr(Y = 0) = Pr(No bin has load t or more) <= Pr(|Y - E[Y]| >= E[Y])

Which concentration bound to use?

Using Chebyshev, Pr(|Y - E[Y]| >= E[Y]) <= $\mathrm{Var}(Y)/E(Y)^2$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. $1-1/n$, the maximum load of each bin is at most $O(\log n/\log\log n)$

Fact: The max loaded bin has $(\log n/3\log\log n)$ balls with probability at least 1 - const.$/n^{(1/3)}$

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise

$Y = Y_1 + Y_2 + .. + Y_n$

$\Pr(Y_i = 1) >= 1/en^{1/3}$

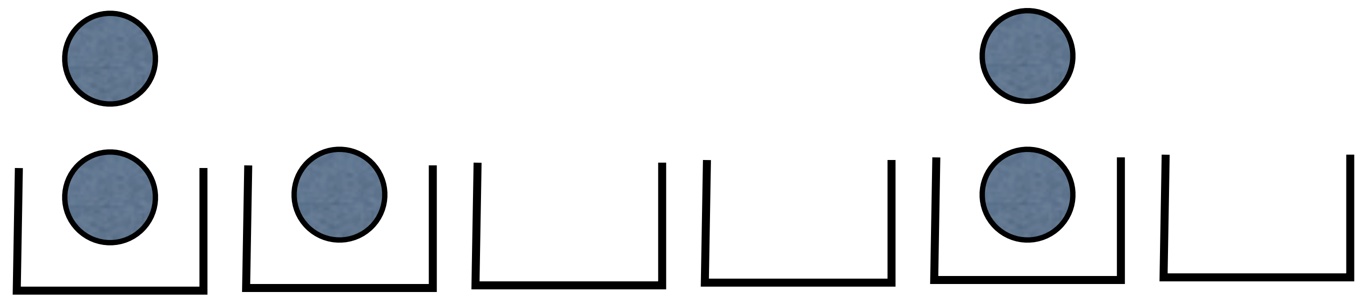
$E(Y) >= n^{2/3}/e$

$\Pr(Y = 0) = \Pr(\text{No bin has load} >= t) <= \Pr(|Y - E[Y]| >= E[Y]) <= \mathrm{Var}(Y)/E(Y)^2$

Chebyshev

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. $1-1/n$, the maximum load of each bin is at most $O(\log n/\log\log n)$

Fact: The max loaded bin has $(\log n/3\log\log n)$ balls with probability at least 1 - const./$n^{(1/3)}$

Let $Y_i = 1$ if bin i has load t or more,
$= 0$ otherwise
$Y = Y_1 + Y_2 + .. + Y_n$

$Pr(Y_i = 1) >= 1/en^{1/3}$

$E(Y) >= n^{2/3}/e$

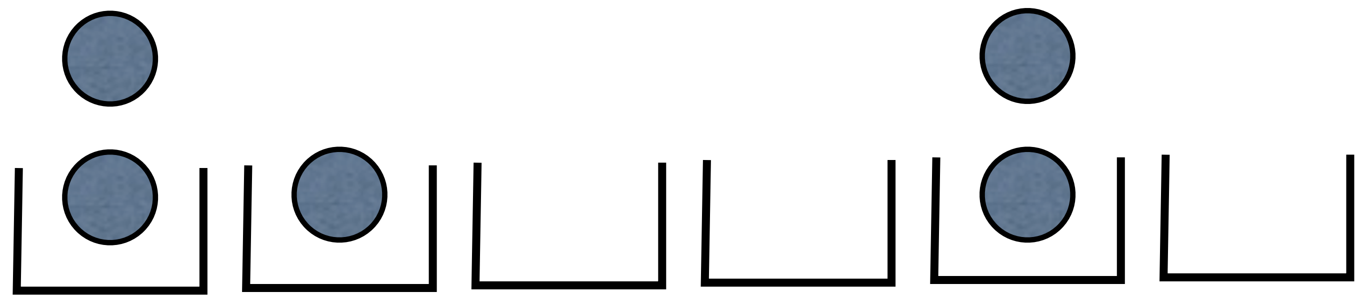
$Pr(Y = 0) = Pr(\text{No bin has load} >= t) <= Pr(|Y - E[Y]| >= E[Y]) <= Var(Y)/E(Y)^2$

Chebyshev

$Var[Y] = Var[(Y_1 + .. + Y_n)^2] =$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. $1-1/n$, the maximum load of each bin is at most $O(\log n/\log\log n)$

Fact: The max loaded bin has $(\log n/3\log\log n)$ balls with probability at least 1 - const.$/n^{(1/3)}$

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise

$Y = Y_1 + Y_2 + .. + Y_n$

$\Pr(Y_i = 1) >= 1/en^{1/3}$

$E(Y) >= n^{2/3}/e$

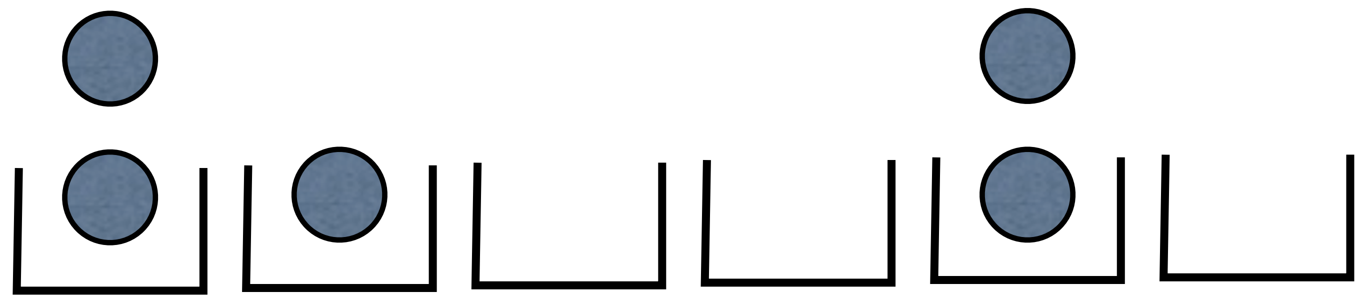
$\Pr(Y = 0) = \Pr(\text{No bin has load} >= t) <= \Pr(|Y - E[Y]| >= E[Y]) <= Var(Y)/E(Y)^2$

Chebyshev

$$Var[Y] = Var[(Y_1 + .. + Y_n)^2] = \sum_i Var(Y_i) + 2\sum_{i\neq j}(E[Y_iY_j] - E[Y_i]E[Y_j])$$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise
$Y = Y_1 + Y_2 + .. + Y_n$

$Pr(Y_i = 1) >= 1/en^{1/3}$

$E(Y) >= n^{2/3}/e$

Pr(Y = 0) = Pr(No bin has load >= t) <= Pr(|Y - E[Y]| >= E[Y]) <= $Var(Y)/E(Y)^2$

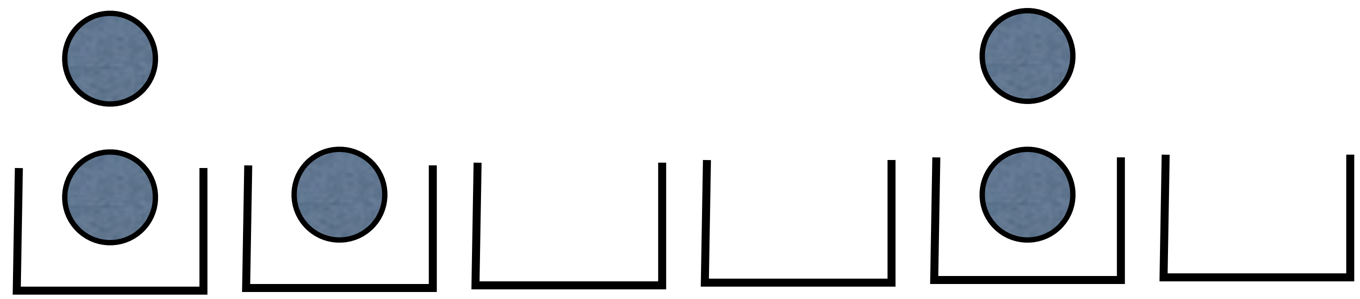
Chebyshev

$Var[Y] = Var[(Y_1 + .. + Y_n)^2] = \sum_i Var(Y_i) + 2\sum_{i \neq j}(E[Y_i Y_j] - E[Y_i]E[Y_j])$

Now if i is not j, Y_i and Y_j are negatively correlated, which means that $E[Y_i Y_j] < E[Y_i]E[Y_j]$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise
Y = Y_1 + Y_2 + .. + Y_n

Pr(Y_i = 1) >= 1/e$n^{1/3}$

E(Y) >= $n^{2/3}$ /e

Pr(Y = 0) = Pr(No bin has load >= t) <= Pr(|Y - E[Y]| >= E[Y]) <= Var(Y)/E(Y)^2

Chebyshev

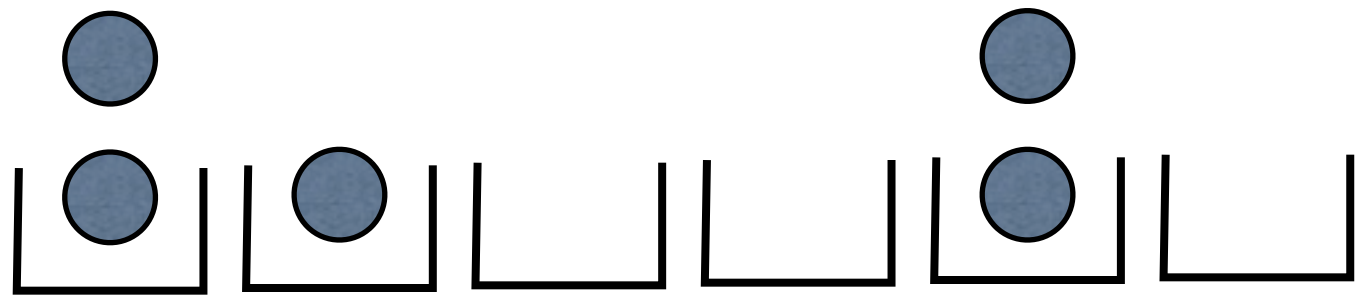
Var[Y] = Var[(Y_1 + .. + Y_n)^2] = $\sum_i Var(Y_i) + 2\sum_{i\neq j}(E[Y_iY_j] - E[Y_i]E[Y_j])$

Now if i is not j, Y_i and Y_j are negatively correlated, which means that $E[Y_iY_j] < E[Y_i]E[Y_j]$

Thus, $Var(Y) \leq \sum_{i=1}^{n} Var(Y_i) \leq n \cdot 1$

Balls and Bins

Suppose we toss n balls u.a.r into n bins. What is the max load of a bin with high probability?



Fact: W.p. 1-1/n, the maximum load of each bin is at most O(log n/log log n)

Fact: The max loaded bin has (log n/3log log n) balls with probability at least 1 - const./$n^{(1/3)}$

Let Y_i = 1 if bin i has load t or more,
= 0 otherwise
$Y = Y_1 + Y_2 + .. + Y_n$

$Pr(Y_i = 1) >= 1/en^{1/3}$

$E(Y) >= n^{2/3} /e$

$Pr(Y = 0) = Pr(\text{No bin has load} >= t) <= Pr(|Y - E[Y]| >= E[Y]) <= Var(Y)/E(Y)^2$

Chebyshev

$$Var[Y] = Var[(Y_1 + .. + Y_n)^2] = \sum_i Var(Y_i) + 2\sum_{i \neq j}(E[Y_iY_j] - E[Y_i]E[Y_j])$$

Now if i is not j, Y_i and Y_j are negatively correlated, which means that $E[Y_iY_j] < E[Y_i]E[Y_j]$

Thus, $Var(Y) \leq \sum_{i=1}^{n} Var(Y_i) \leq n \cdot 1$

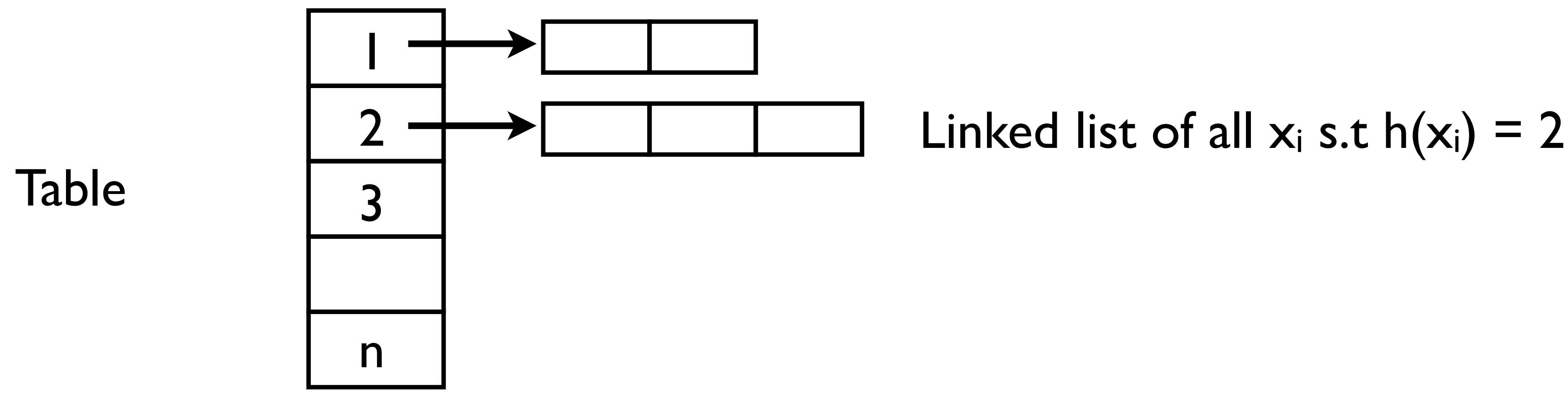
$$\Pr(Y = 0) \leq \frac{Var(Y)}{E(Y)^2} \leq \frac{ne^2}{n^{4/3}} \leq \frac{e^2}{n^{1/3}}$$

Randomized Algorithms

- Contention Resolution
- Some Facts about Random Variables
- Global Minimum Cut Algorithm
- Randomized Selection and Sorting
- Max 3-SAT
- Three Concentration Inequalities
- Hashing and Balls and Bins
  - The Power of Two Choices

The Power of Two Choices

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not

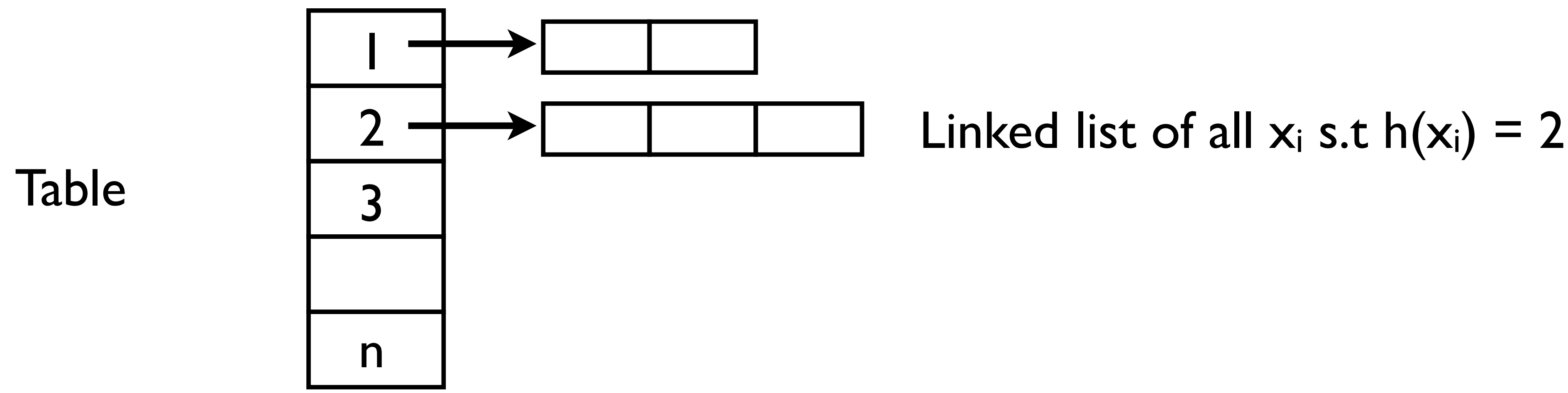


Algorithm:

1. Pick **two** completely random functions $h_1 : \mathcal{U} \to \{1, \ldots, n\},$ and $h_2 : \mathcal{U} \to \{1, \ldots, n\}$
2. Create a table of size n, initialize it to null
3. Store x_i at linked list at position $h_1(x_i)$ or $h_2(x_i)$, whichever is shorter
4. For a query q, look at the linked list at location $h_1(q)$ and $h_2(q)$ of table to see if q is there

The Power of Two Choices

Problem: Given a large set S of elements $x_1, .., x_n$, store them using O(n) space s.t it is easy to determine whether a query item q is in S or not



Algorithm:

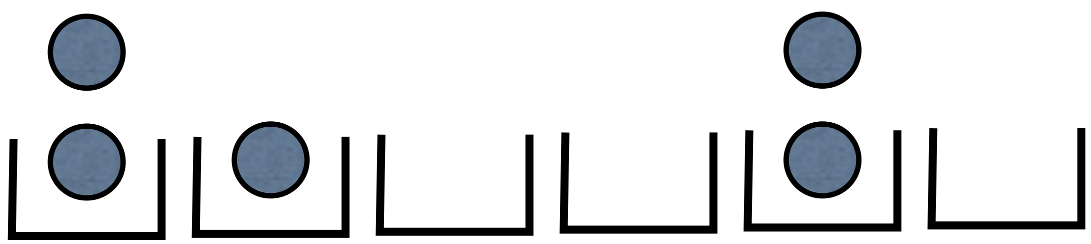
1. Pick **two** completely random functions $h_1 : \mathcal{U} \to \{1, \ldots, n\},$ and $h_2 : \mathcal{U} \to \{1, \ldots, n\}$
2. Create a table of size n, initialize it to null
3. Store x_i at linked list at position $h_1(x_i)$ or $h_2(x_i)$, whichever is shorter
4. For a query q, look at the linked list at location $h_1(q)$ and $h_2(q)$ of table to see if q is there

Equivalent to the following Balls and Bins Problem: Toss n balls into n bins. For each ball, pick two bins u.a.r and put the ball into the lighter of the two bins.

What is the worst case query time?

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?

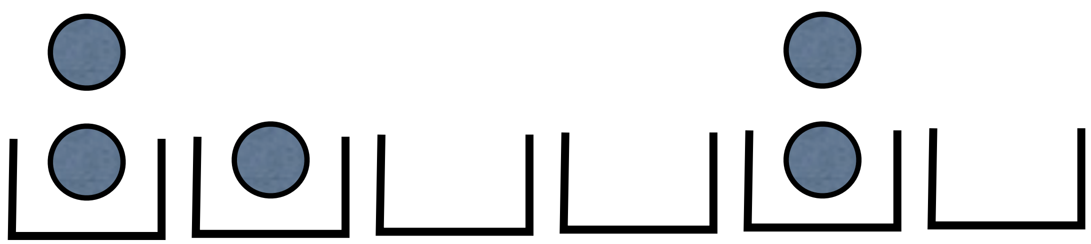


Fact: Max load of a bin = O(log log n), w.p. 1-1/n

We will prove this for the rest of class

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

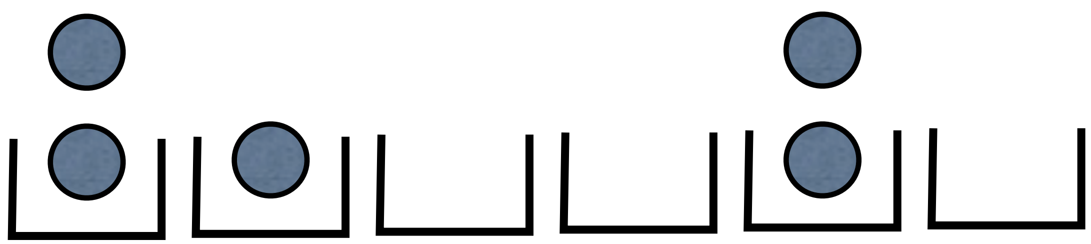
We will prove this for the rest of class

Main Idea (Not a formal Proof):

If a_i = fraction of bins with i or more balls, then, fraction of bins with i+1 or more balls is:

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

We will prove this for the rest of class

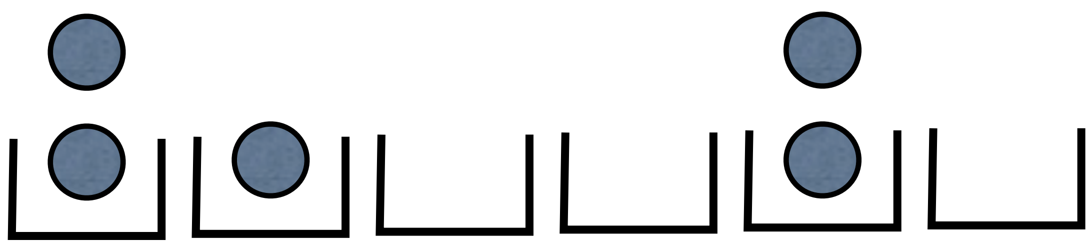
Main Idea (Not a formal Proof):

If a_i = fraction of bins with i or more balls, then, fraction of bins with i+1 or more balls is:

$a_{i+1} \approxeq a_i^2$ (To put a ball in a bin with i or more balls, need to pick two such bins out of n)

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

We will prove this for the rest of class

Main Idea (Not a formal Proof):

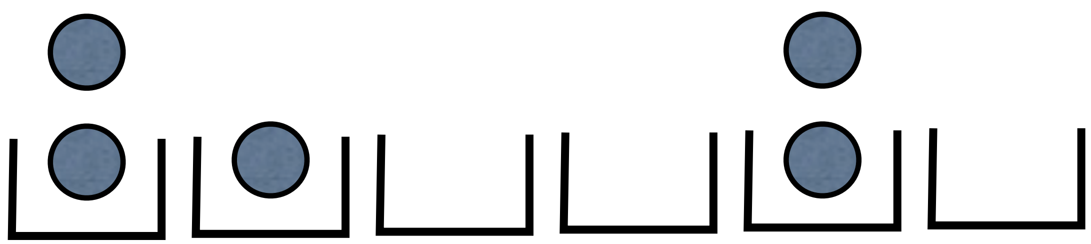
If a_i = fraction of bins with i or more balls, then, fraction of bins with i+1 or more balls is:

$a_{i+1} \approxeq a_i^2$ (To put a ball in a bin with i or more balls, need to pick two such bins out of n)

$a_2 \leq \frac{1}{2}$ $\quad a_i \approxeq \frac{1}{2^{2^{i-1}}} \quad < \frac{1}{n^2}$ for $i = O(\log \log n)$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

We will prove this for the rest of class

Main Idea (Not a formal Proof):

If a_i = fraction of bins with i or more balls, then, fraction of bins with i+1 or more balls is:

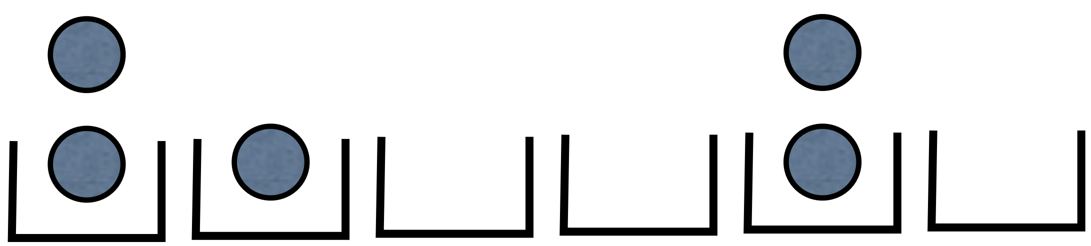
$a_{i+1} \approxeq a_i^2$ (To put a ball in a bin with i or more balls, need to pick two such bins out of n)

$a_2 \leq \frac{1}{2}$ $\quad a_i \approxeq \frac{1}{2^{2^{i-1}}} \quad < \frac{1}{n^2}$ for $i = O(\log \log n)$

Now we prove this formally. But what makes the proof hard?

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



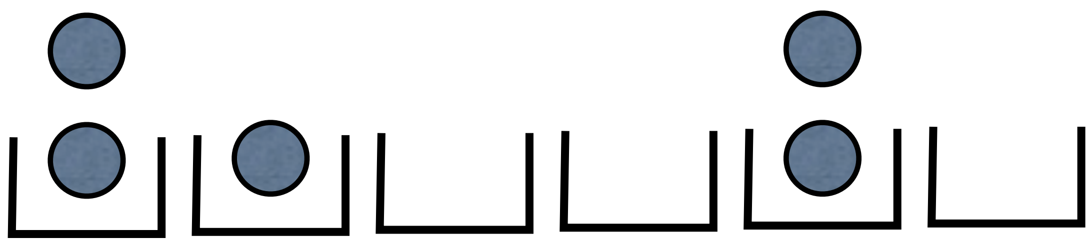
Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

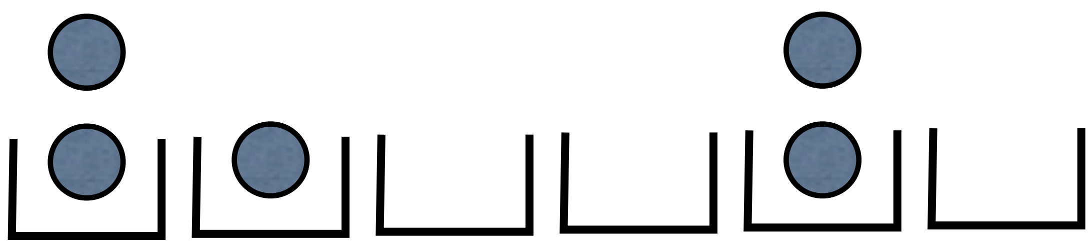
Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

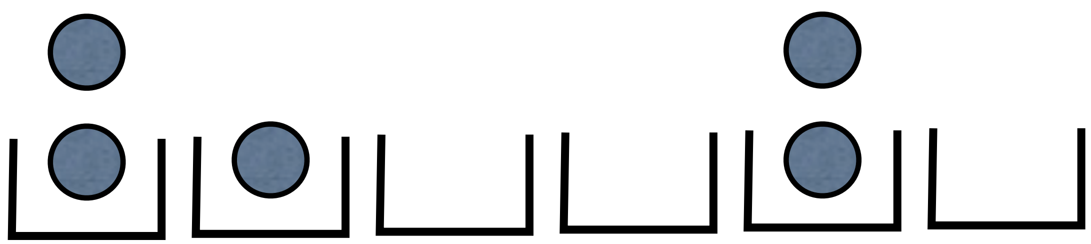
If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p$, then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n, p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1|X_1, \ldots, X_{i-1}) \leq p$, then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n, p) \geq k)$

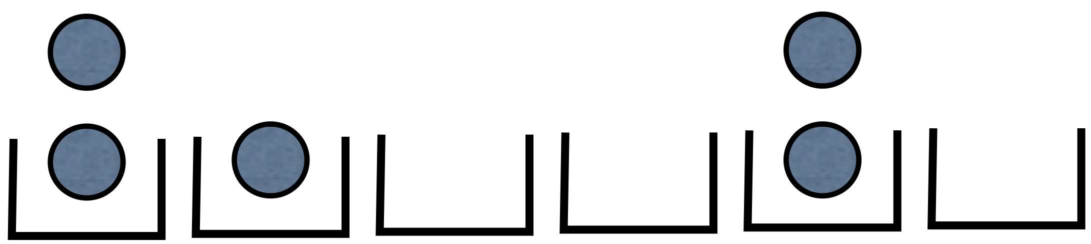
Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t|N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

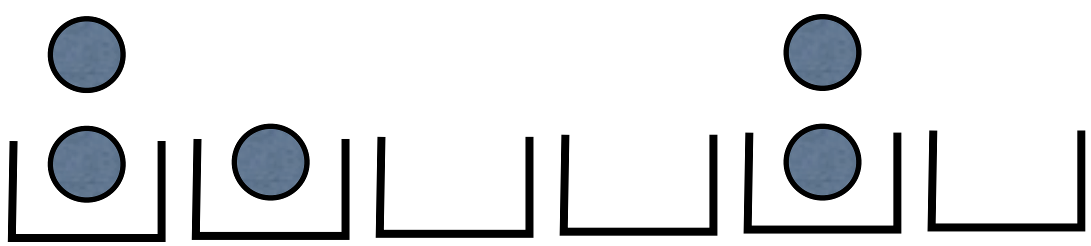
Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

Define: Y_j = 1 if (ball j lands in bin with current load >= i+1) and (N_i at time j-1 <= an)

= 0 otherwise

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

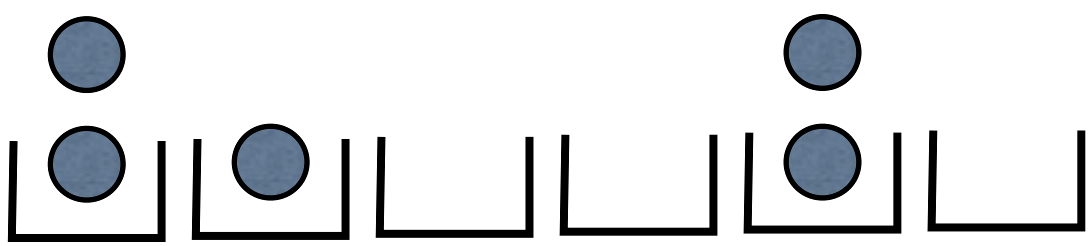
Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

Define: Y_j = 1 if $\underbrace{\boxed{\text{(ball j lands in bin with current load >= i+1)}}}_{\text{event A}}$ and $\underbrace{\boxed{(N_i \text{ at time j-1 <= an})}}_{\text{event B}}$

= 0 otherwise

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

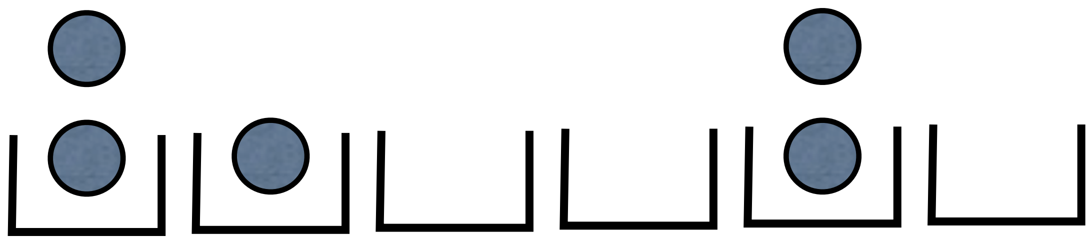
Define: Y_j = 1 if $\underbrace{\boxed{\text{(ball j lands in bin with current load >= i+1)}}}_{\text{event A}}$ and $\underbrace{\boxed{(N_i \text{ at time j-1 <= an})}}_{\text{event B}}$

= 0 otherwise

Y_j depends on $X_1..X_{j-1}$.

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

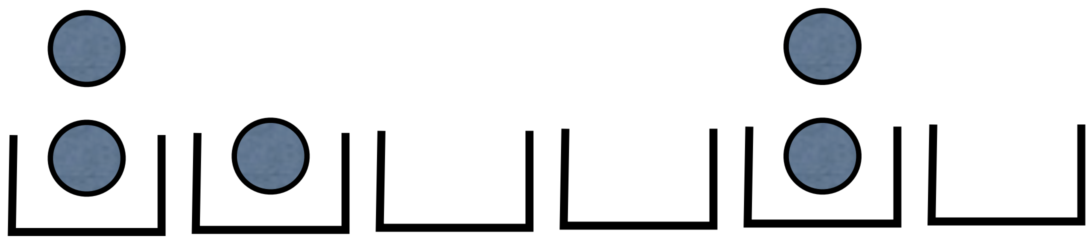
Define: Y_j = 1 if $\underbrace{\boxed{\text{(ball j lands in bin with current load >= i+1)}}}_{\text{event A}}$ and $\underbrace{\boxed{(N_i \text{ at time j-1} \leq an)}}_{\text{event B}}$

= 0 otherwise

Y_j depends on $X_1..X_{j-1}$. $\Pr(Y_j = 1 | X_1,..X_{j-1}) = \Pr(A, B)$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p$, then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

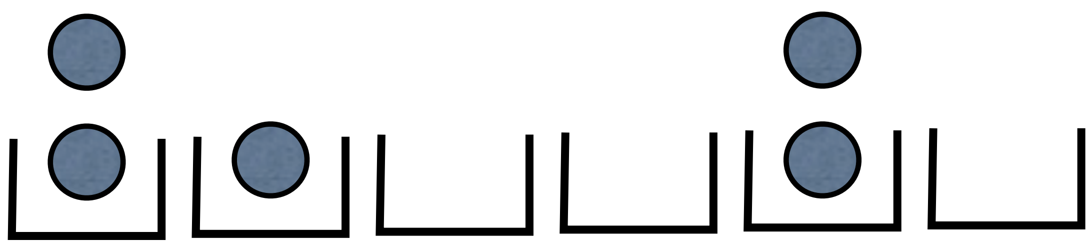
Define: Y_j = 1 if $\underbrace{\boxed{\text{(ball j lands in bin with current load >= i+1)}}}_{\text{event A}}$ and $\underbrace{\boxed{(N_i \text{ at time j-1 <= an})}}_{\text{event B}}$

= 0 otherwise

Y_j depends on $X_1..X_{j-1}$. $\Pr(Y_j = 1 | X_1,..X_{j-1}) = \Pr(A, B) = \Pr(A|B) \Pr(B)$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p$, then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

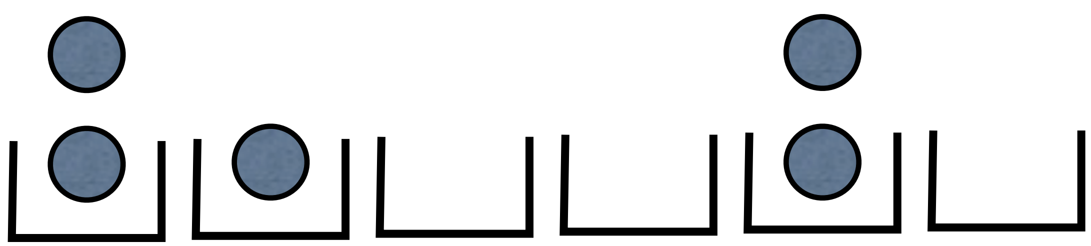
Define: Y_j = 1 if $\underbrace{\boxed{\text{(ball j lands in bin with current load >= i+1)}}}_{\text{event A}}$ and $\underbrace{\boxed{(N_i \text{ at time j-1 <= an})}}_{\text{event B}}$

= 0 otherwise

Y_j depends on $X_1..X_{j-1}$. $\Pr(Y_j = 1 | X_1,..X_{j-1}) = \Pr(A, B) = \Pr(A|B) \Pr(B) <= \Pr(A|B) <= a^2$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

Define: Y_j = 1 if $\underbrace{\boxed{\text{(ball j lands in bin with current load >= i+1)}}}_{\text{event A}}$ and $\underbrace{\boxed{(N_i \text{ at time j-1 <= an})}}_{\text{event B}}$

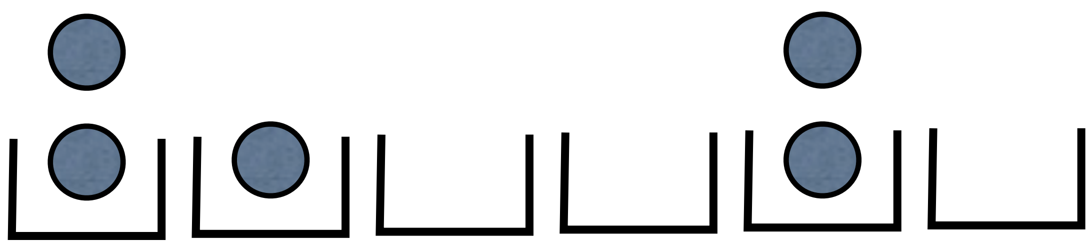
= 0 otherwise

Y_j depends on $X_1..X_{j-1}$. $\Pr(Y_j = 1 | X_1,..X_{j-1}) = \Pr(A, B) = \Pr(A|B)\Pr(B) \leq \Pr(A|B) \leq a^2$

$$\Pr(N_{i+1} > t | N_i \leq an) = \frac{\Pr(N_{i+1} > t, N_i \leq an)}{\Pr(N_i \leq an)}$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1 | X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n,p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

Define: Y_j = 1 if (ball j lands in bin with current load >= i+1) and (N_i at time j-1 <= an)

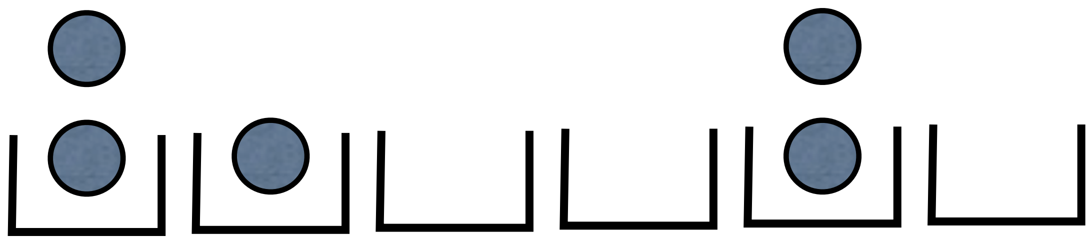
= 0 otherwise — event A — event B

Y_j depends on $X_1..X_{j-1}$. $\Pr(Y_j = 1 | X_1,..X_{j-1}) = \Pr(A, B) = \Pr(A|B)\Pr(B) \leq \Pr(A|B) \leq a^2$

$$\Pr(N_{i+1} > t | N_i \leq an) = \frac{\Pr(N_{i+1} > t, N_i \leq an)}{\Pr(N_i \leq an)} = \frac{\Pr(\sum_{j=1}^{n} Y_j > t)}{\Pr(N_i \leq an)}$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = $O(\log \log n)$, w.p. $1-1/n$

Fact: Let $X_1..X_n$ be independent random variables(rv), $Y_1..Y_n$ be rv s.t Y_i depends on $X_1..X_{i-1}$

If $\Pr(Y_i = 1|X_1, \ldots, X_{i-1}) \leq p,$ then $\Pr(\sum_{i=1}^{n} Y_i \geq k) \leq \Pr(Bin(n, p) \geq k)$

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Proof: For any j, let X_j = k if ball picks in bin k in a toss. Are X_js independent?

Define: Y_j = 1 if (ball j lands in bin with current load >= i+1) and (N_i at time j-1 <= an)

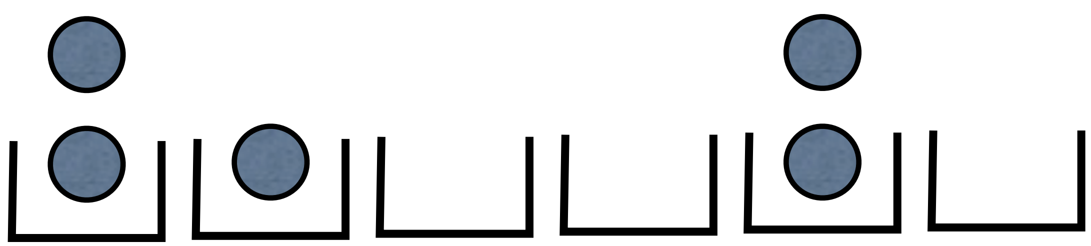
= 0 otherwise — event A — event B

Y_j depends on $X_1..X_{j-1}$. $\Pr(Y_j = 1 | X_1,..X_{j-1}) = \Pr(A, B) = \Pr(A|B) \Pr(B) <= \Pr(A|B) <= a^2$

$$\Pr(N_{i+1} > t | N_i \leq an) = \frac{\Pr(N_{i+1} > t, N_i \leq an)}{\Pr(N_i \leq an)} = \frac{\Pr(\sum_{j=1}^{n} Y_j > t)}{\Pr(N_i \leq an)} \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$$ **from Fact**

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

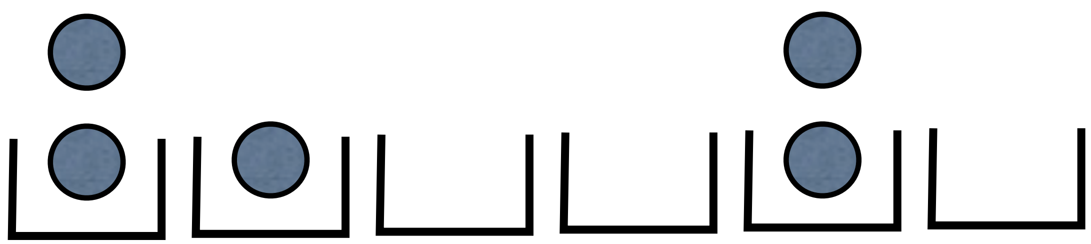
Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

Chernoff Bounds:
Let X_1, .., X_n be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

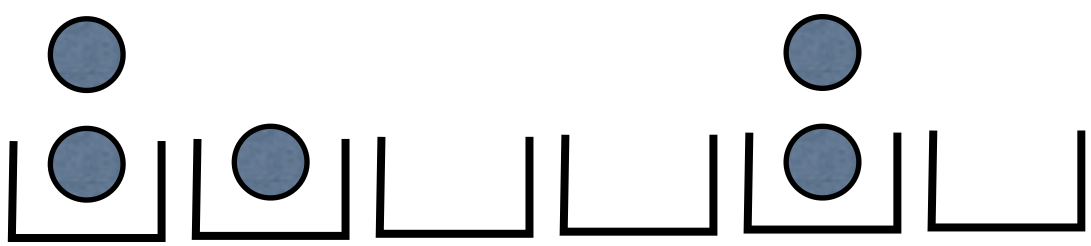
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}}\right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n/n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n/n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$

Chernoff Bounds:

Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

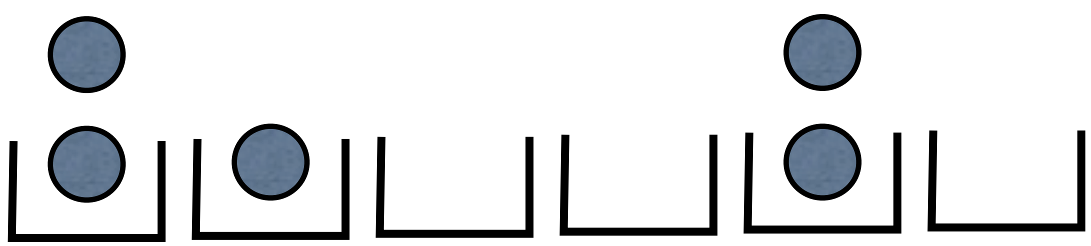
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}}\right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

Chernoff Bounds:
Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

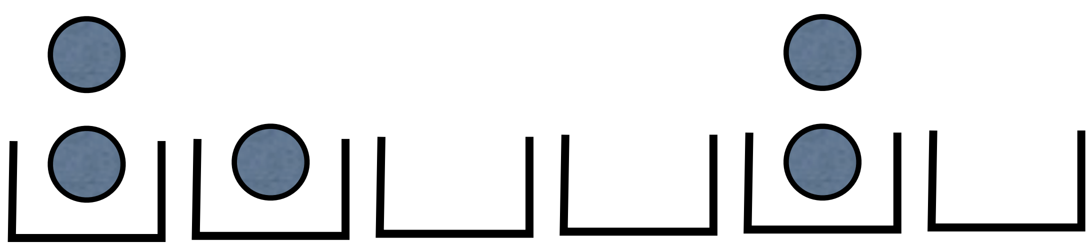
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}}\right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

$\Pr(Bin(n, a^2) > 2na^2) \leq (e/4)^{na^2}$

Chernoff Bounds:
Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

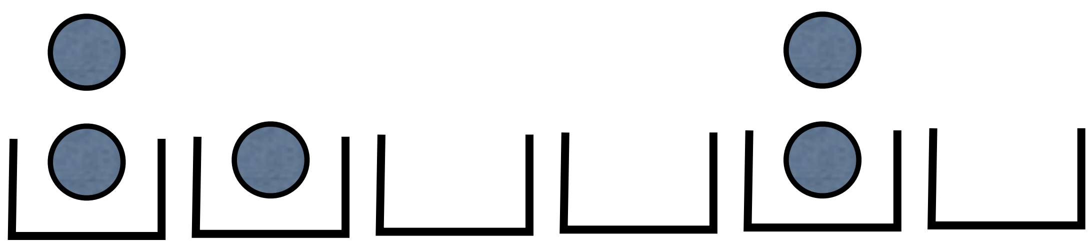
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}} \right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

$$\Pr(Bin(n, a^2) > 2na^2) \leq (e/4)^{na^2} \leq (e/4)^{6 \ln n} \leq 1/n^2$$

Chernoff Bounds:
Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

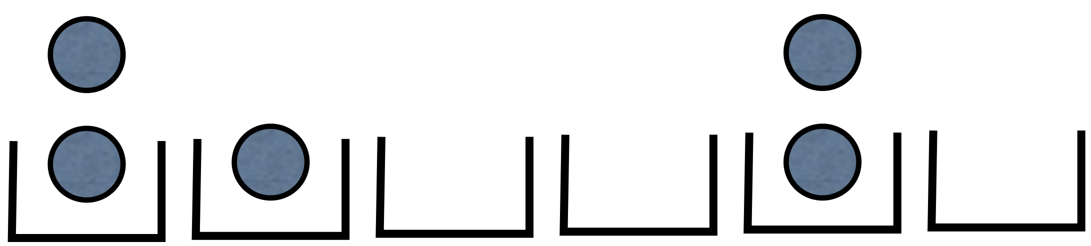
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}} \right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

$$\Pr(Bin(n, a^2) > 2na^2) \leq (e/4)^{na^2} \leq (e/4)^{6 \ln n} \leq 1/n^2$$

For the second part, $E[X] < 6 \ln n$

Chernoff Bounds:

Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

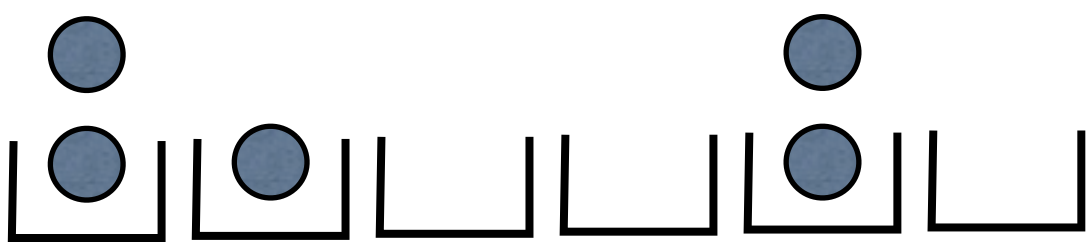
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}} \right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

$\Pr(Bin(n, a^2) > 2na^2) \leq (e/4)^{na^2} \leq (e/4)^{6 \ln n} \leq 1/n^2$

For the second part, $E[X] < 6 \ln n$ so $t > 10 \ln n / E[X]$

Chernoff Bounds:

Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

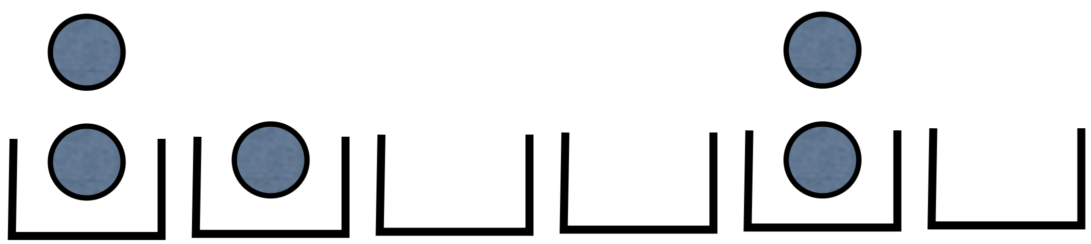
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}} \right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

$$\Pr(Bin(n, a^2) > 2na^2) \leq (e/4)^{na^2} \leq (e/4)^{6 \ln n} \leq 1/n^2$$

For the second part, $E[X] < 6 \ln n$ so $t > 10 \ln n / E[X]$

$$\Pr(Bin(n, a^2) > 16 \ln n) \leq \exp(-(10 \ln n / E[X])^2 \cdot E[X]/4)$$

Chernoff Bounds:

Let X_1, .., X_n be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

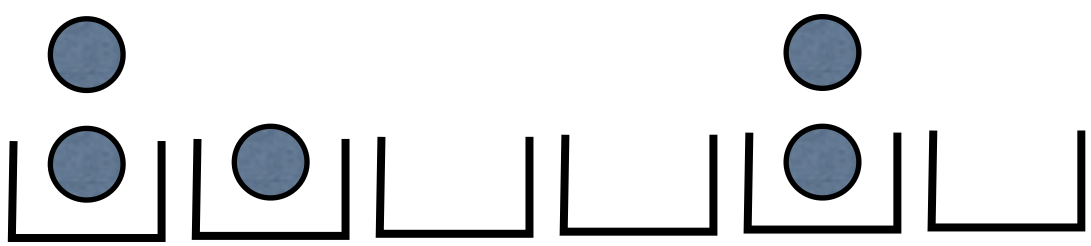
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}}\right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

$$\Pr(Bin(n, a^2) > 2na^2) \leq (e/4)^{na^2} \leq (e/4)^{6 \ln n} \leq 1/n^2$$

For the second part, $E[X] < 6 \ln n$ so $t > 10 \ln n / E[X]$

$$\Pr(Bin(n, a^2) > 16 \ln n) \leq \exp(-(10 \ln n / E[X])^2 \cdot E[X]/4)$$
$$\leq \exp(-25 \ln^2 n / E[X])$$

Chernoff Bounds:
Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

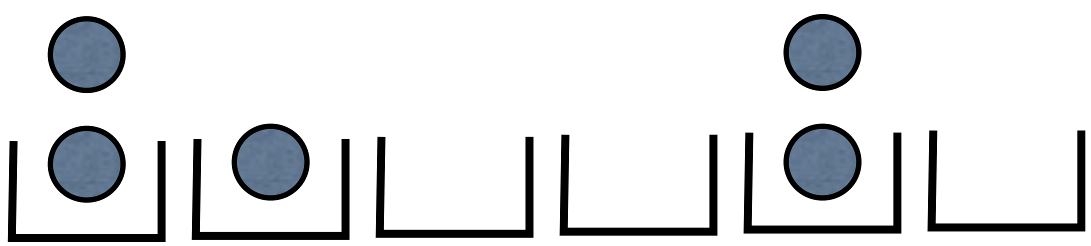
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}}\right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n/n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n/n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Proof: Use Chernoff Bounds. What is X and what is X_i ?

For the first part, $E[X] = na^2$ t = 1

$$\Pr(Bin(n, a^2) > 2na^2) \leq (e/4)^{na^2} \leq (e/4)^{6 \ln n} \leq 1/n^2$$

For the second part, $E[X] < 6 \ln n$ so $t > 10 \ln n/E[X]$

$$\Pr(Bin(n, a^2) > 16 \ln n) \leq \exp(-(10 \ln n/E[X])^2 \cdot E[X]/4)$$
$$\leq \exp(-25 \ln^2 n/E[X]) \leq e^{-4 \ln n} \leq n^{-2}$$

Chernoff Bounds:

Let $X_1, .., X_n$ be independent 0/1 rvs, $X = X_1 + .. + X_n$. For t>0,

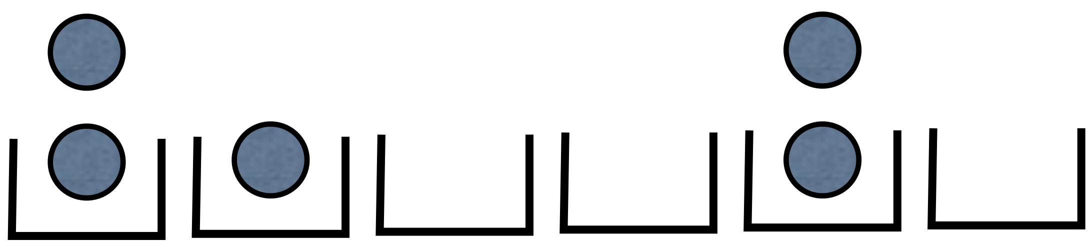
$$\Pr(X \geq (1+t)E[X]) \leq \left(\frac{e^t}{(1+t)^{(1+t)}}\right)^{E[X]}$$

For t < 2e - 1,

$$\Pr(X \geq (1+t)E[X]) \leq \exp(-t^2 E[X]/4)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

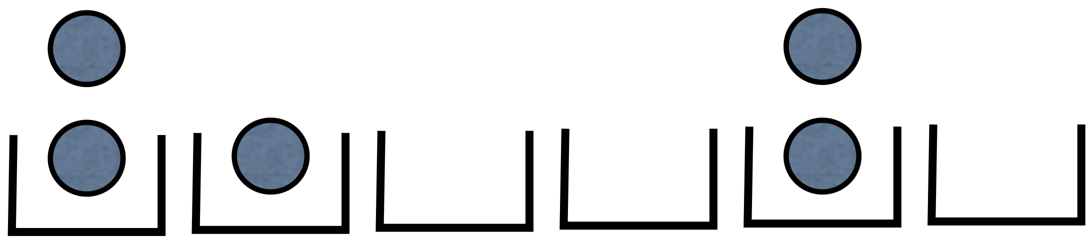
Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

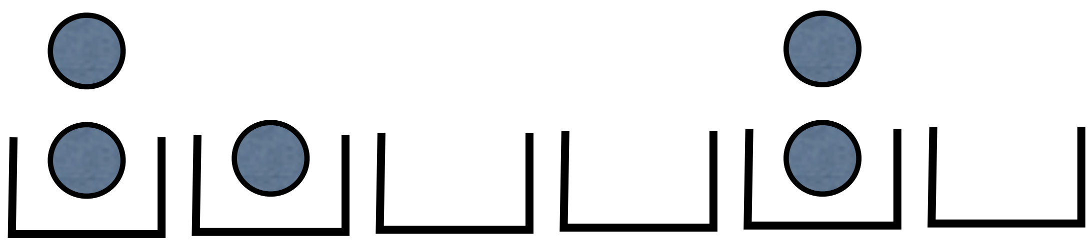
Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i n$)

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

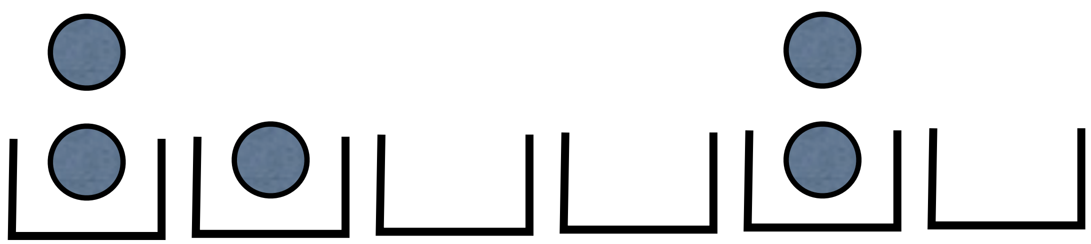
If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

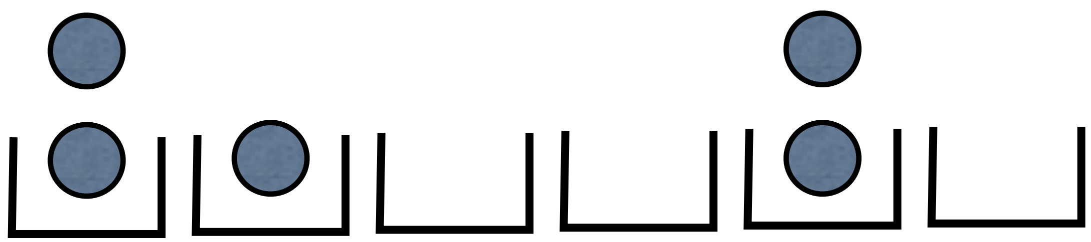
Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event (N_i <= a_i n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Proof: By Induction. Base Case: i = 4. Pr(not(E_4)) = 0, as at most n/4 bins can have >=4 balls

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event (N_i <= a_i n)

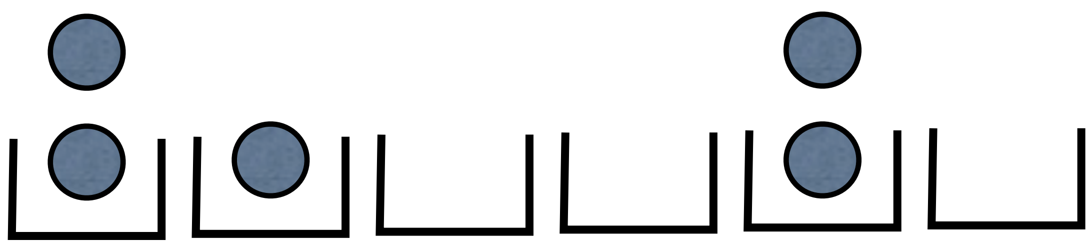
Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Proof: By Induction. Base Case: i = 4. Pr(not(E_4)) = 0, as at most n/4 bins can have >=4 balls

Inductive Case: $\Pr(not(E_{i+1})) \leq \Pr(not(E_{i+1}) | E_i) \Pr(E_i) + \Pr(not(E_i))$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i \le a_i n$)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

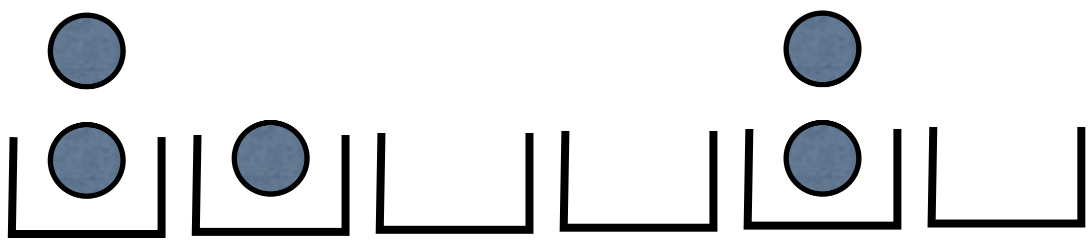
Proof: By Induction. Base Case: i = 4. Pr(not(E_4)) = 0, as at most n/4 bins can have >=4 balls

Inductive Case: $\Pr(not(E_{i+1})) \leq \boxed{\Pr(not(E_{i+1})|E_i) \Pr(E_i)} + \Pr(not(E_i))$

<= $1/n^2$ from previous two lemmas

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event (N_i <= a_i n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

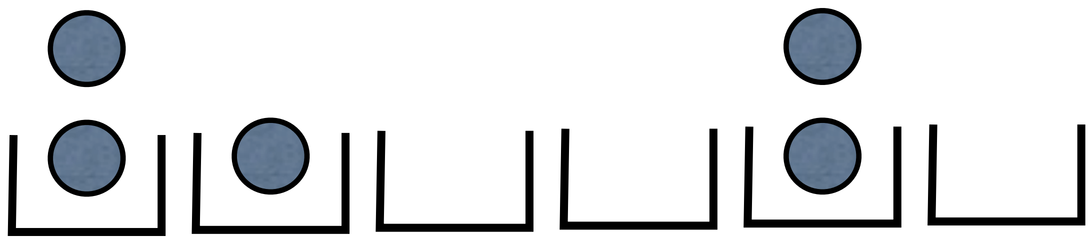
Proof: By Induction. Base Case: i = 4. Pr(not(E_4)) = 0, as at most n/4 bins can have >=4 balls

Inductive Case: $\Pr(not(E_{i+1})) \leq \Pr(not(E_{i+1}) | E_i) \Pr(E_i) + \Pr(not(E_i))$

<= $1/n^2$ from previous two lemmas   <= i/n^2

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

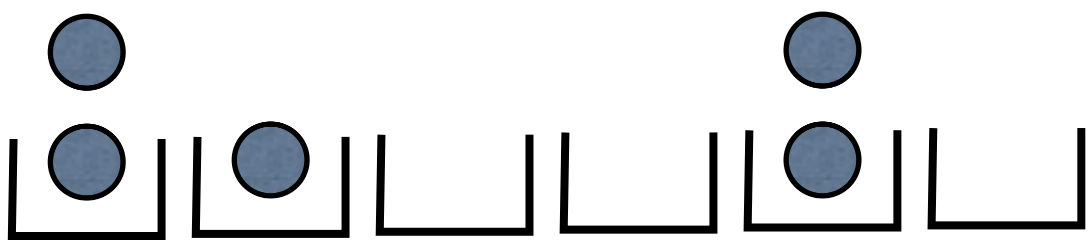
Proof: By Induction. Base Case: i = 4. Pr(not(E_4)) = 0, as at most n/4 bins can have >=4 balls

Inductive Case: $\Pr(not(E_{i+1})) \leq \boxed{\Pr(not(E_{i+1}) | E_i) \Pr(E_i)} + \boxed{\Pr(not(E_i))} \leq (i+1)/n^2$

<= $1/n^2$ from previous two lemmas <= i/n^2

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

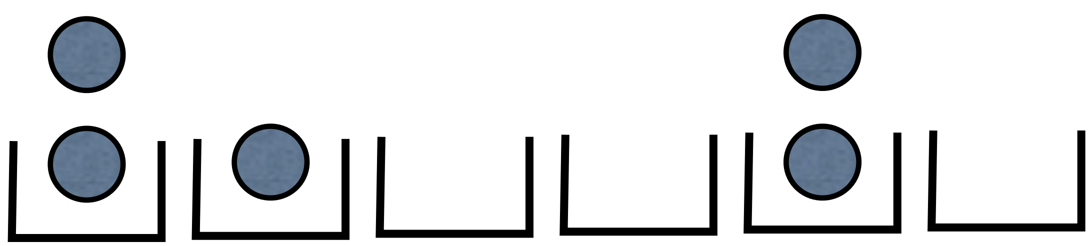
If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i n$)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

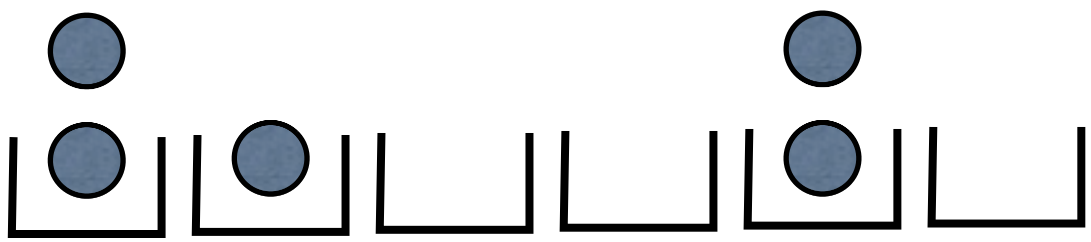
Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i \leq a_i n$)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

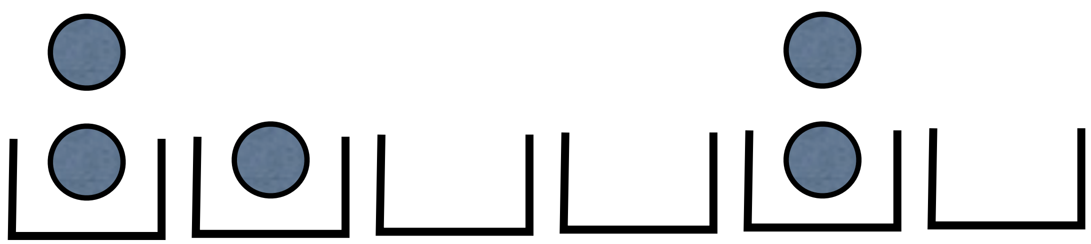
Define: Sequence of a_i values: a_4 = 1/4 $a_{i+1} = 2a_i^2$ E_i = good event ($N_i \leq a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

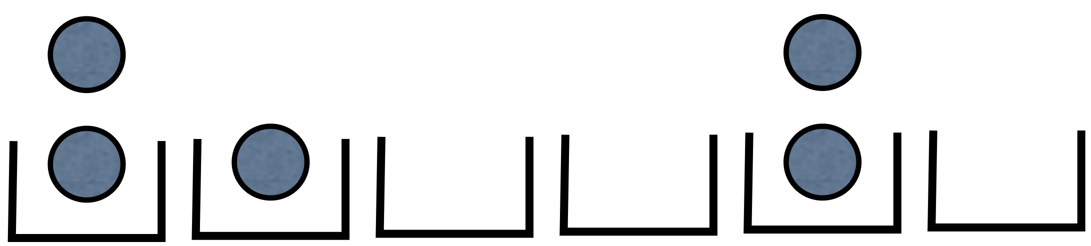
Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

But N_{I^*+1} or N_{I^*+2} may still be large, so we need to bound them

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i n$)

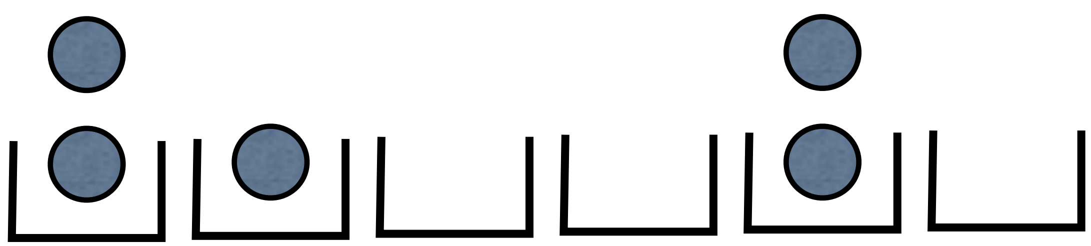
Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

$\Pr(N_{I^*+1} > 16 \ln n) \leq \Pr(N_{I^*+1} > 16 \ln n | E_{I^*}) \Pr(E_{I^*}) + \Pr(not(E_{I^*}))$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $\quad a_{i+1} = 2a_i^2$ $\quad E_i$ = good event ($N_i <= a_i\, n$)

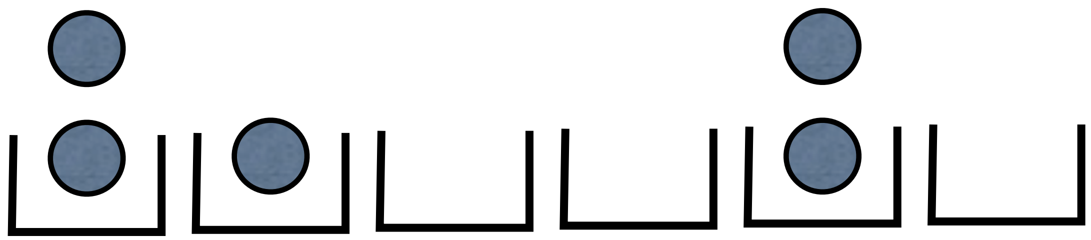
Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

$$\begin{aligned}\Pr(N_{I^*+1} > 16 \ln n) &\leq \Pr(N_{I^*+1} > 16 \ln n | E_{I^*}) \Pr(E_{I^*}) + \Pr(not(E_{I^*})) \\ &\leq \Pr(Bin(n, a_{I^*+1}^2) > 16 \ln n | E_{I^*}) \Pr(E_{I^*}) + \Pr(not(E_{I^*}))\end{aligned}$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i\, n$)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

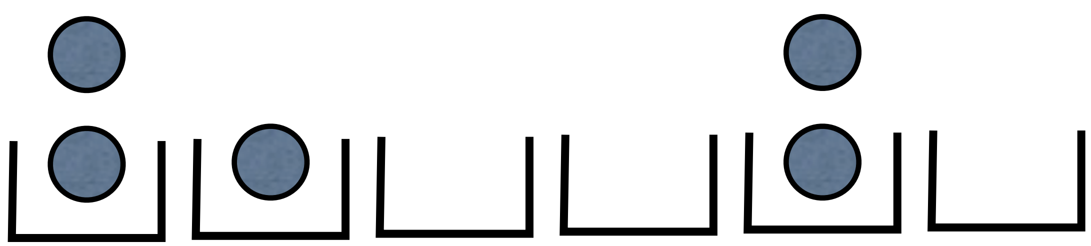
Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

$$\Pr(N_{I^*+1} > 16 \ln n) \leq \Pr(N_{I^*+1} > 16 \ln n | E_{I^*}) \Pr(E_{I^*}) + \Pr(not(E_{I^*}))$$

$$\leq \boxed{\Pr(Bin(n, a_{I^*+1}^2) > 16 \ln n | E_{I^*})} \Pr(E_{I^*}) + \boxed{\Pr(not(E_{I^*}))}$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n/n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n/n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n/n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

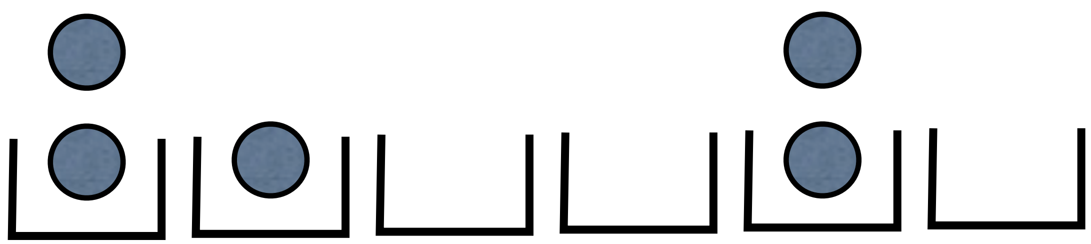
Let $I^* = \min_i \{a_i^2 \leq 6 \ln n/n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

$$\Pr(N_{I^*+1} > 16 \ln n) \leq \Pr(N_{I^*+1} > 16 \ln n | E_{I^*}) \Pr(E_{I^*}) + \Pr(not(E_{I^*}))$$

$$\leq \boxed{\Pr(Bin(n, a_{I^*+1}^2) > 16 \ln n | E_{I^*})} \Pr(E_{I^*}) + \boxed{\Pr(not(E_{I^*}))} \quad \leq (n+1)/n^2$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

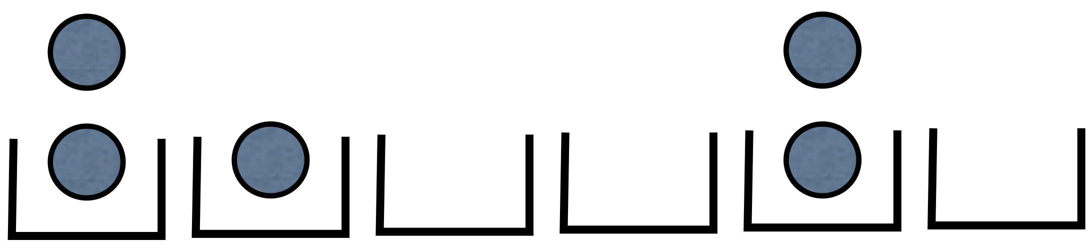
Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

$\Pr(N_{I^*+1} > 16 \ln n) \leq 2/n$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

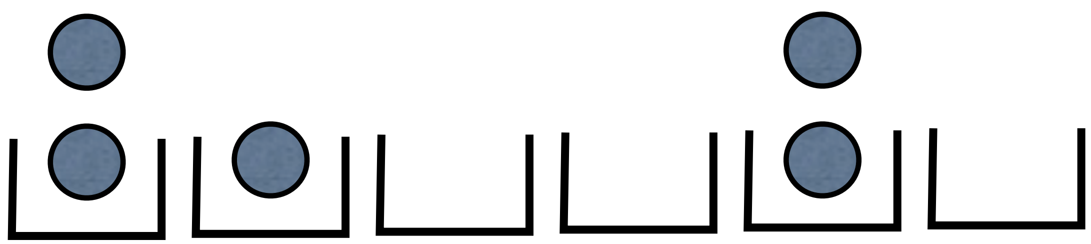
Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

$\Pr(N_{I^*+1} > 16 \ln n) \leq 2/n$

$\Pr(N_{I^*+2} \geq 1) \leq \Pr(Bin(n, \frac{(16 \ln n)^2}{n}) \geq 1) + \Pr(N_{I^*+1} > 16 \ln n)$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: a_4 = 1/4 $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

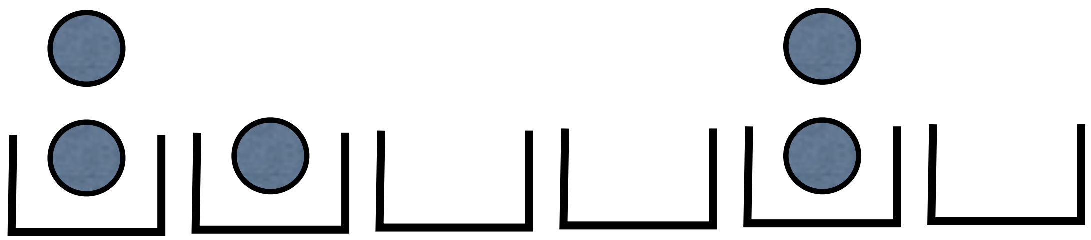
$\Pr(N_{I^*+1} > 16 \ln n) \leq 2/n$

$$\Pr(N_{I^*+2} \geq 1) \leq \Pr(Bin(n, \frac{(16 \ln n)^2}{n}) \geq 1) + \Pr(N_{I^*+1} > 16 \ln n)$$

$$\leq E(Bin(n, \frac{(16 \ln n)^2}{n})) + 2/n$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \dfrac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

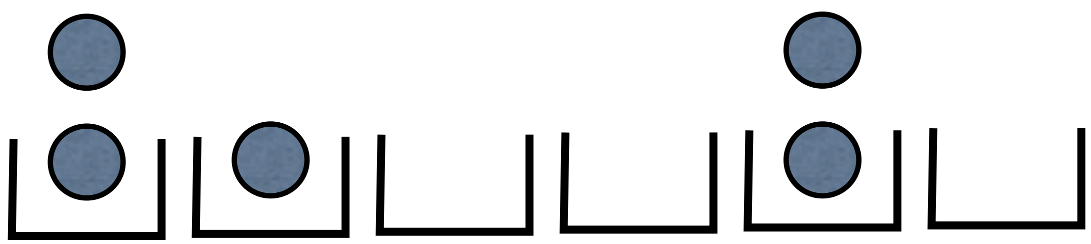
$\Pr(N_{I^*+1} > 16 \ln n) \leq 2/n$

$$\Pr(N_{I^*+2} \geq 1) \leq \Pr(Bin(n, \frac{(16 \ln n)^2}{n}) \geq 1) + \Pr(N_{I^*+1} > 16 \ln n)$$

$$\leq E(Bin(n, \frac{(16 \ln n)^2}{n})) + 2/n \quad \leq O(\log^2 n / n)$$

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: $a_4 = 1/4$ $a_{i+1} = 2a_i^2$ E_i = good event (N_i <= a_i n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

$\Pr(N_{I^*+1} > 16 \ln n) \leq 2/n$

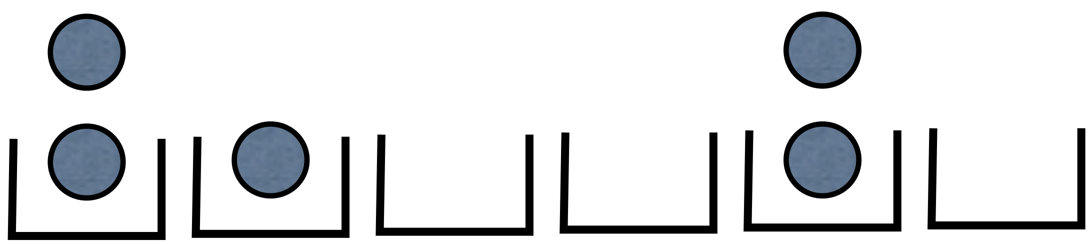
$$\Pr(N_{I^*+2} \geq 1) \leq \Pr(Bin(n, \frac{(16 \ln n)^2}{n}) \geq 1) + \Pr(N_{I^*+1} > 16 \ln n)$$

$$\leq E(Bin(n, \frac{(16 \ln n)^2}{n})) + 2/n \quad \leq O(\log^2 n / n)$$

Lemma 4: Wp (1 - $\log^2 n/n$), there are no bins with load I* + 2 or more

Power of Two Choices

Toss n balls into n bins. For each ball, we pick two bins u.a.r and put it in the lighter of the two. What is the maximum load of a bin with high probability?



Fact: Max load of a bin = O(log log n), w.p. 1-1/n

Define: N_i = #bins with load i or more

Lemma 1: For any a, $\Pr(N_{i+1} > t | N_i \leq an) \leq \frac{\Pr(Bin(n, a^2) > t)}{\Pr(N_i \leq an)}$

Lemma 2: If $a^2 \geq 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 2na^2) \leq 1/n^2$

If $a^2 < 6 \ln n / n$ then $\Pr(Bin(n, a^2) > 16 \ln n) \leq 1/n^2$

Define: Sequence of a_i values: a_4 = 1/4 $a_{i+1} = 2a_i^2$ E_i = good event ($N_i <= a_i$ n)

Lemma 3: If $a_i^2 \geq 6 \ln n / n$ then $\Pr(not(E_{i+1})) \leq (i+1)/n^2$

Let $I^* = \min_i \{a_i^2 \leq 6 \ln n / n\}$ Then, $\Pr[not(E_{I^*})] \leq I^*/n^2 \leq 1/n$

Lemma 4: Wp (1 - $\log^2 n/n$), no bins with load I^* + 2 or more

Can be shown that $a_i \approx \frac{1}{2^{2^i}}$ from which $I^* = O(\log \log n)$