Algorithm Design Paradigms

• Exhaustive Search

• **Greedy Algorithms:** Build a solution incrementally piece by piece

• **Divide and Conquer:** Divide into parts, solve each part, combine results

• **Dynamic Programming:** Divide into subtasks, perform subtask by size. Combine smaller subtasks to larger ones

• **Hill-climbing:** Start with a solution, improve it

• **Randomized Algorithms:** Algorithm can make random choices
Randomized Algorithms

• Algorithm can make random decisions
• Why randomized algorithms? Simple and efficient
• Examples: Symmetry-breaking, graph algorithms, quicksort, hashing, load balancing, cryptography, etc
Given \( n \) processors and a resource which they all wish to access, design a symmetry breaking access protocol.

Restriction: Processors can’t communicate.
Simultaneous access blocks the resource.

Need symmetry-breaking!
Contestion Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

Restriction: Processors can’t communicate
Simultaneous access blocks the resource

Protocol: At time t, each processor accesses resource w.p. $p = 1/n$
Contention Resolution

Given $n$ processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate  
Simultaneous access blocks the resource

**Protocol:** At time $t$, each processor accesses resource w.p. $p = \frac{1}{n}$

**Goals:**
- Low contention
- Low waiting time for all processors
Contention Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. $p = \frac{1}{n}$

**Goals:**
- Low contention
- Low waiting time for all processors

$S(i, t) =$ Event that processor i succeeds at time t = At time t, only processor i accesses
Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. $p = 1/n$

**Goals:**
- Low contention
- Low waiting time for all processors

$S(i, t) = \text{Event that processor } i \text{ succeeds at time } t = \text{At time } t, \text{ only processor } i \text{ accesses}$

$$\Pr[S(i, t)] = p \times (1 - p)^{n-1} = \frac{1}{n} \times (1 - 1/n)^{n-1}$$
Contention Resolution

Given $n$ processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time $t$, each processor accesses resource w.p. $p = 1/n$

**Goals:**
- Low contention
- Low waiting time for all processors

**Independent Events A, B:**
$\Pr(A \text{ and } B) = \Pr(A) \times \Pr(B)$

$S(i, t) =$ Event that processor $i$ succeeds at time $t =$ At time $t$, only processor $i$ accesses

$\Pr[S(i, t)] = \frac{p}{n} \times (1 - \frac{p}{n})^{n-1} = \frac{1}{n} \times \left(1 - \frac{1}{n}\right)^{n-1}$
Contention Resolution

Given n processors and a resource which they all wish to access, design a
symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. \( p = \frac{1}{n} \)

\[ \Pr[S(i, t)] = p \times (1 - p)^{n-1} = \left(\frac{1}{n}\right) \times (1 - \frac{1}{n})^{n-1} \]

- \( p \) \( \uparrow \) processor i accesses
- \( (1 - p)^{n-1} \) \( \uparrow \) no other processors access

\( S(i, t) = \) Event that processor i succeeds at time t = At time t, only processor i accesses
Contention Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
   Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. $p = 1/n$

$$Pr[S(i, t)] = p \times (1 - p)^{n-1} = \frac{1}{n} \times (1 - \frac{1}{n})^{n-1}$$

**Fact:** As n increases,
(1) $(1 - 1/n)^n$ grows monotonically up to $1/e$
(2) $(1 - 1/n)^{n-1}$ decreases monotonically down to $1/e$
Contention Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. \( p = \frac{1}{n} \)

\[
\Pr[S(i, t)] = \frac{1}{n} \times (1 - \frac{1}{n})^{n-1} \geq \frac{1}{en}
\]

\[
\Pr[S(i, t)] = \frac{1}{n} \times (1 - \frac{1}{n})^{n-1} \leq \frac{1}{2n}
\]

**Fact:** As n increases,

(1) \((1 - \frac{1}{n})^n\) grows monotonically up to \(\frac{1}{e}\)

(2) \((1 - \frac{1}{n})^{n-1}\) decreases monotonically down to \(\frac{1}{e}\)
Contestion Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. p = 1/n

\[ S(i, t) = \text{Event that processor } i \text{ succeeds at time } t = \text{At time } t, \text{ only processor } i \text{ accesses} \]

\[ \Pr[S(i, t)] = p \times (1 - p)^{n-1} = \frac{1}{n} \times (1 - \frac{1}{n})^{n-1} \]

\[ \geq \frac{1}{e} \]

\[ \leq \frac{1}{2n} \]
Contention Resolution

Given $n$ processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time $t$, each processor accesses resource w.p. $p = 1/n$

$S(i, t) = \text{Event that processor } i \text{ succeeds at time } t = \text{At time } t, \text{ only processor } i \text{ accesses}$

$$\Pr[S(i, t)] = p \times (1 - p)^{n-1} = \frac{1}{n} \times (1 - \frac{1}{n})^{n-1} = \frac{1}{en} \leq \frac{1}{2n}$$

$F(i, t) = \text{Event that processor } i \text{ has not succeeded after } t \text{ steps}$
Contestation Resolution

Given $n$ processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate

Simultaneous access blocks the resource

**Protocol:** At time $t$, each processor accesses resource w.p. $p = 1/n$

$S(i, t) =$ Event that processor $i$ succeeds at time $t = At time $t$, only processor $i$ accesses

$F(i, t) =$ Event that processor $i$ has not succeeded after $t$ steps

$$\Pr[S(i, t)] = p \times (1 - p)^{n-1} = \frac{1}{n} \times (1 - \frac{1}{n})^{n-1} \geq \frac{1}{en} \leq \frac{1}{2n}$$

$$\Pr[F(i, t)] = (1 - S(i, 1))(1 - S(i, 2))...(1 - S(i, t)) \leq (1 - \frac{1}{en})^t$$
Contention Resolution

Given \( n \) processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate

Simultaneous access blocks the resource

**Protocol:** At time \( t \), each processor accesses resource w.p. \( p = 1/n \)

\( S(i, t) = \text{Event that processor } i \text{ succeeds at time } t = \text{At time } t, \text{ only processor } i \text{ accesses} \)

\[
\Pr[S(i, t)] = p \times (1 - p)^{n-1} = \left(\frac{1}{n}\right) \times (1 - \frac{1}{n})^{n-1} \\
\geq \left(\frac{1}{en}\right) \\
\leq \left(\frac{1}{2n}\right)
\]

\( F(i, t) = \text{Event that processor } i \text{ has not succeeded after } t \text{ steps} \)

\[
\Pr[F(i, t)] = (1 - \Pr[S(i, 1)])(1 - \Pr[S(i, 2)])\ldots(1 - \Pr[S(i, t)]) \\
\leq (1 - \frac{1}{en})^t
\]

Picking \( t = en \), \( \Pr[F(i, t)] \leq 1/e \)

Picking \( t = c. \text{en.} \ln n \), \( \Pr[F(i, t)] \leq n^{-c} \)
Contention Resolution

Given $n$ processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time $t$, each processor accesses resource w.p. $p = \frac{1}{n}$

**Properties:**

\[
F(i, t) = \text{Event that processor } i \text{ has not succeeded after } t \text{ steps}
\]

Picking $t = \log n$, $\Pr[F(i, t)] \leq \frac{1}{e}$

Picking $t = c \cdot \log n \cdot \ln n$, $\Pr[F(i, t)] \leq n^{-c}$
Contestation Resolution

Given \( n \) processors and a resource which they all wish to access, design a symmetry breaking access protocol.

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time \( t \), each processor accesses resource w.p. \( p = \frac{1}{n} \)

**Properties:**

\( F(i, t) = \) Event that processor \( i \) has not succeeded after \( t \) steps

Picking \( t = \text{en} \), \( \Pr[F(i, t)] \leq \frac{1}{e} \)

Picking \( t = c \cdot \text{en} \cdot \ln n \), \( \Pr[F(i, t)] \leq n^{-c} \)

**Fact:** W.p. \( 1 - \frac{1}{n} \), all the processors succeed within \( 2\text{en} \cdot \ln n \) rounds
Contestation Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. \( p = \frac{1}{n} \)

**Properties:**

\( F(i, t) = \) Event that processor i has not succeeded after t steps

Picking \( t = e \), \( \Pr[F(i, t)] \leq \frac{1}{e} \)

Picking \( t = c \cdot e \cdot \ln n \), \( \Pr[F(i, t)] \leq n^{-c} \)

**Fact:** W.p. \( 1 - \frac{1}{n} \), all the processors succeed within \( 2e \cdot \ln n \) rounds

For \( t = 2e \cdot \ln n \), \( \Pr[F(i,t)] = n^{-2} \) for a fixed i
Contention Resolution

Given $n$ processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time $t$, each processor accesses resource w.p. $p = 1/n$

**Properties:**

$F(i, t) = \text{Event that processor } i \text{ has not succeeded after } t \text{ steps}$

Picking $t = en$, $Pr[F(i, t)] \leq 1/e$

Picking $t = c \cdot en \cdot \ln n$, $Pr[F(i, t)] \leq n^{-c}$

**Fact:** W.p. $1 - 1/n$, all the processors succeed within $2en \cdot \ln n$ rounds

For $t = 2en \cdot \ln n$, $Pr[F(i, t)] = n^{-2}$ for a fixed $i$

**Union Bound:** Any $A, B$

$Pr[A \cup B] \leq Pr[A] + Pr[B]$
Contention Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. $p = 1/n$

**Properties:**

$F(i, t) = $ Event that processor i has not succeeded after t steps

Picking $t = en$, $Pr[F(i, t)] \leq 1/e$

Picking $t = c \cdot en \cdot \ln n$, $Pr[F(i, t)] \leq n^{-c}$

**Fact:** W.p. $1 - 1/n$, all the processors succeed within $2en \cdot \ln n$ rounds

For $t = 2en \cdot \ln n$, $Pr[F(i, t)] = n^{-2}$ for a fixed i

Therefore,

$$Pr[\bigcup_{i=1}^{n} F(i, t)] \leq \sum_{i=1}^{n} Pr[F(i, t)] \leq n \cdot n^{-2} \leq 1/n$$

**Union Bound:** Any A, B

$Pr[A \cup B] \leq Pr[A] + Pr[B]$
Summary: Contention Resolution

Given n processors and a resource which they all wish to access, design a symmetry breaking access protocol

**Restriction:** Processors can’t communicate
   Simultaneous access blocks the resource

**Protocol:** At time t, each processor accesses resource w.p. \( p = \frac{1}{n} \)
\( F(i, t) = \) Event that processor i has not succeeded after t steps

**Facts:** Picking \( t = c \cdot e \cdot n \cdot \ln n \), \( \Pr[F(i, t)] \leq n^{-c} \)
   WiP \( 1 - \frac{1}{n} \), all processors succeed within \( 2en \ln n \) rounds

**Facts we learnt:**

**Independent Events** \( A, B: \) \( \Pr(A \text{ and } B) = \Pr(A) \times \Pr(B) \)

**Union Bound:** For any two events \( A \) and \( B \), \( \Pr[A \cup B] \leq \Pr[A] + \Pr[B] \)
   As n increases,
   (1) \( (1 - \frac{1}{n})^n \) converges monotonically up to \( 1/e \)
   (2) \( (1 - \frac{1}{n})^{n-1} \) converges monotonically down to \( 1/e \)
Randomized Algorithms

- Contention Resolution
- Some Facts about Random Variables
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$
Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Examples:**

1. Let $X = 1$ if a fair coin toss comes up heads, 0 ow. What is $E[X]$?
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

Examples:

1. Let $X = 1$ if a fair coin toss comes up heads, 0 otherwise. What is $E[X]$?

2. We are tossing a coin with head probability $p$, tail probability $1 - p$. Let $X = \#$ independent flips until first head. What is $E[X]$?
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

Examples:

1. Let $X = 1$ if a fair coin toss comes up heads, 0 ow. What is $E[X]$?

2. We are tossing a coin with head probability $p$, tail probability $1 - p$. Let $X = \#$ independent flips until first head. What is $E[X]$?

$$\Pr[X = j] = p \times (1 - p)^{j-1}$$

head on toss $j$ \ first $j-1$ tails

$$E[X] = \sum_{j=1}^{\infty} j \cdot p(1 - p)^{j-1} = \frac{p}{1 - p} \sum_{j=0}^{\infty} j(1 - p)^{j} = \frac{p}{1 - p} \cdot \frac{1 - p}{p^2} = \frac{1}{p}$$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Guessing a card**

Shuffle $n$ cards, then turn them over one by one. Guess what the next card is. How many guesses are correct on expectation?
Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Guessing a card**

Shuffle $n$ cards, then turn them over one by one. Guess what the next card is. How many guesses are correct on expectation?

Let $X_i = 1$ if guess $i$ is correct, 0 otherwise
### Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]

#### Example: Guessing a card

Shuffle $n$ cards, then turn them over one by one. Guess what the next card is. How many guesses are correct on expectation?

Let $X_i = 1$ if guess $i$ is correct, 0 otherwise

$$\Pr[X_i = 1] = \frac{1}{n - i + 1}$$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Guessing a card**

Shuffle $n$ cards, then turn them over one by one. Guess what the next card is. How many guesses are correct on expectation?

Let $X_i = 1$ if guess $i$ is correct, 0 otherwise

$$\Pr[X_i = 1] = \frac{1}{n - i + 1} \quad \quad E[X_i] = \frac{1}{n - i + 1}$$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]

**Example: Guessing a card**

Shuffle $n$ cards, then turn them over one by one. Guess what the next card is. How many guesses are correct on expectation?

Let $X_i = 1$ if guess $i$ is correct, 0 otherwise

$$\Pr[X_i = 1] = \frac{1}{n - i + 1} \quad E[X_i] = \frac{1}{n - i + 1}$$

Expected # of correct guesses = $E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{i} = \Theta(\log n)$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$ 

**Example: Guessing a card**

Shuffle $n$ cards, then turn them over one by one. Guess what the next card is. How many guesses are correct on expectation?

Let $X_i = 1$ if guess $i$ is correct, 0 otherwise

$$\Pr[X_i = 1] = \frac{1}{n - i + 1}$$

$$E[X_i] = \frac{1}{n - i + 1}$$

Expected # of correct guesses = $E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{i} = \Theta(\log n)$

What if we insert the selected card into the pile randomly and pull another?
**Expectation**

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Coupon Collector’s Problem**

Balls tossed randomly into $n$ bins. How many balls in expectation before each bin has a ball?
**Expectation**

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_ip_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]

**Example: Coupon Collector’s Problem**

Balls tossed randomly into $n$ bins. How many balls in expectation before each bin has a ball?

Let $X_j =$ time spent when there are exactly $j$ non-empty bins (Phase $j$)
**Expectation**

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$ 

**Example: Coupon Collector’s Problem**  
Balls tossed randomly into $n$ bins. How many balls in expectation before each bin has a ball?

Let $X_j = \text{time spent when there are exactly } j \text{ non-empty bins (Phase } j)$  
Let $X = \text{total } \#\text{steps} = X_1 + X_2 + ... + X_{n-1}$
## Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$ 

### Example: Coupon Collector’s Problem

Balls tossed randomly into $n$ bins. How many balls in expectation before each bin has a ball?

Let $X_j =$ time spent when there are exactly $j$ non-empty bins (Phase $j$)

Let $X =$ total steps $= X_1 + X_2 + ... + X_{n-1}$

We move from phase $j$ to $j+1$ when a ball hits one of $n-j$ bins, so w.p. $(n-j) / n$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Coupon Collector’s Problem**

Balls tossed randomly into $n$ bins. How many balls in expectation before each bin has a ball?

Let $X_i =$ time spent when there are exactly $j$ non-empty bins (Phase $j$)

Let $X =$ total #steps $= X_1 + X_2 + ... + X_{n-1}$

We move from phase $j$ to $j+1$ when a ball hits one of $n-j$ bins, so w.p. $(n-j) / n$

Therefore, $E[X_i] = n/(n - j)$ [From previous slide on expected waiting times]
**Expectation**

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]

**Example: Coupon Collector’s Problem**

Balls tossed randomly into $n$ bins. How many balls in expectation before each bin has a ball?

Let $X_i = \text{time spent when there are exactly } j \text{ non-empty bins (Phase } j)$

Let $X = \text{total steps} = X_1 + X_2 + … + X_{n-1}$

We move from phase $j$ to $j+1$ when a ball hits one of $n-j$ bins, so w.p. $(n-j) / n$

Therefore, $E[X_j] = n/(n - j)$  [From previous slide on expected waiting times]

$E[X] = E[X_1 + … + X_{n-1}] = E[X_1] + … + E[X_{n-1}] = n + n/2 + … + n/(n-1) = \Theta(n \log n)$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Birthday Paradox**

$m$ balls tossed randomly into $n$ bins. What is the expected #collisions?
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Birthday Paradox**

$m$ balls tossed randomly into $n$ bins. What is the expected #collisions?

For $1 \leq i < j \leq m$, let $X_{ij} = 1$ if balls $i$ and $j$ land in the same bin, 0 otherwise.
Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Birthday Paradox**

$m$ balls tossed randomly into $n$ bins. What is the expected #collisions?

For $1 \leq i < j \leq m$, let $X_{ij} = 1$ if balls $i$ and $j$ land in the same bin, $0$ otherwise

$$\Pr[X_{ij} = 1] = \frac{1}{n}, \text{ so } E[X_{ij}] = \frac{1}{n}$$
Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$

**Example: Birthday Paradox**

$m$ balls tossed randomly into $n$ bins. What is the expected #collisions?

For $1 <= i < j <= m$, let $X_{ij} = 1$ if balls $i$ and $j$ land in the same bin, 0 otherwise

$\Pr[ X_{ij} = 1 ] = 1/n$, so $E[X_{ij}] = 1/n$

So, expected number of collisions from tossing $m$ balls $= \sum_{i,j} E[X_{ij}] = \binom{m}{2} \cdot \frac{1}{n} = \frac{m(m-1)}{2n}$

So when $m < \sqrt{2n}$, expected #collisions $< 1$; otherwise, it’s more
Summary: Expectation and Linearity of Expectation

Given discrete random variable $X$, which takes $m$ values $x_i$ w.p. $p_i$, the expectation $E[X]$ is defined as:

$$E[X] = \sum_{i=1}^{m} x_i \cdot \Pr[X = x_i] = \sum_{i=1}^{m} x_i p_i$$

**Linearity of Expectation:** $E[X + Y] = E[X] + E[Y]$
Variance

Given a random variable $X$, its variance $\text{Var}[X]$ is defined as:

$$\text{Var}(X) = E[(X - E[X])^2]$$

Variance of a random variable measures its “spread”
Variance

Given a random variable $X$, its variance $\text{Var}[X]$ is defined as:

$$\text{Var}(X) = E[(X - E[X])^2]$$

**Examples:**

1. Let $X = 1$ if a fair coin toss comes up heads, 0 otherwise. What is $\text{Var}(X)$?
Variance

Given a random variable $X$, its variance $\text{Var}[X]$ is defined as:

$$\text{Var}(X) = E[(X - E[X])^2]$$

**Examples:**

1. Let $X = 1$ if a fair coin toss comes up heads, 0 otherwise. What is $\text{Var}(X)$?
2. Let $X =$ outcome of a fair dice throw. What is $\text{Var}(X)$?
Variance

Given a random variable $X$, its variance $\text{Var}[X]$ is defined as:

$$\text{Var}(X) = E[(X - E[X])^2]$$

**Properties of Variance:**

1. $\text{Var}(X) = E[X^2] - (E[X])^2$
Variance

Given a random variable $X$, its variance $\text{Var}[X]$ is defined as:

$$\text{Var}(X) = E[(X - E[X])^2]$$

**Properties of Variance:**

1. $\text{Var}(X) = E[X^2] - (E[X])^2$
2. If $X$ and $Y$ are independent random variables, then, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
Variance

Given a random variable $X$, its variance $\text{Var}[X]$ is defined as:

$$\text{Var}(X) = E[(X - E[X])^2]$$

**Properties of Variance:**

1. $\text{Var}(X) = E[X^2] - (E[X])^2$
2. If $X$ and $Y$ are independent random variables, then, $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$
3. For any constants $a$ and $b$, $\text{Var}(aX + b) = a^2 \text{Var}(X)$

Variance of a random variable measures its “spread”
Randomized Algorithms

- Contention Resolution
- Some Facts about Random Variables
- Global Minimum Cut Algorithm
**Global Min-Cut**

**Problem:** Given undirected, unweighted graph $G = (V, E)$, find a cut $(A, V - A)$ which has the minimum number of edges across it.

**Example:**

Note: A global min-cut is not the same as an s-t min cut.

How can we find a global min-cut using $n - 1$ max-flows?

We can do better for the global min-cut.