Advanced Probability

Indicator Random Variables

Daniel Baumgartner

Fall 2025

Rutgers Competitive Programming

Table of Contents

Introduction

Definition

Let A be an event. Its indicator function is then

$$\mathbb{1}_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$$

Properties

Let A and B be events.

- $(\mathbb{1}_A)^2 = \mathbb{1}_A$
- $\mathbb{1}_{A^c} = 1 - \mathbb{1}_A$
- $\mathbb{1}_{A \cap B} = \mathbb{1}_A \mathbb{1}_B$
- $\mathbb{1}_{A \cup B} = \mathbb{1}_A + \mathbb{1}_B - \mathbb{1}_A \mathbb{1}_B$

Fundamental bridge between probability and expectation:

$$\mathbb{P}(A) = \mathbb{E}(\mathbb{1}_A)$$

General Strategy

Recall linearity of expectation: $\mathbb{E}(aX + Y) = a\mathbb{E}(X) + \mathbb{E}(Y)$

This holds for dependent random variables too!

The idea behind indicators is to use them as building blocks for the random variables that we are actually interested in.

Exercises

Fixed Points

A fixed point in an array a is an index i satisfying $a_i = i$.

Find the expected number of fixed points in a random permutation.

Fixed Points

- Let $\mathbb{1}_i$ be the indicator for if index i is a fixed point

Fixed Points

- Let $\mathbb{1}_i$ be the indicator for if index i is a fixed point
- We want $\mathbb{1}_1 + \mathbb{1}_2 + \cdots + \mathbb{1}_n$

Fixed Points

- Let $\mathbb{1}_i$ be the indicator for if index i is a fixed point
- We want $\mathbb{1}_1 + \mathbb{1}_2 + \cdots + \mathbb{1}_n$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i \text{ is a fixed point}) = \frac{1}{n}$

Fixed Points

- Let $\mathbb{1}_i$ be the indicator for if index i is a fixed point
- We want $\mathbb{1}_1 + \mathbb{1}_2 + \cdots + \mathbb{1}_n$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i \text{ is a fixed point}) = \frac{1}{n}$
- So, $\mathbb{E}(\sum_{i=1}^{n} \mathbb{1}_i) = \sum_{i=1}^{n} \mathbb{E}(\mathbb{1}_i) = n \cdot \frac{1}{n} = 1$

Inversions

An inversion in an array a is a pair (i, j) satisfying $i < j$ and $a_i > a_j$ (i.e. a pair of elements with the wrong relative order).

Find the expected number of inversions in a random permutation.

Inversions

- Let $\mathbb{1}_{ij}$ be the indicator for if (i, j) forms an inversion

Inversions

- Let $\mathbb{1}_{ij}$ be the indicator for if (i, j) forms an inversion
- We want $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$

Inversions

- Let $\mathbb{1}_{ij}$ be the indicator for if (i, j) forms an inversion
- We want $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}((i, j)$ forms an inversion$) = \frac{1}{2}$ by symmetry

Inversions

- Let $\mathbb{1}_{ij}$ be the indicator for if (i,j) forms an inversion
- We want $\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{1}_{ij}$
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}((i,j)$ forms an inversion$) = \frac{1}{2}$ by symmetry
- So, $\mathbb{E}(\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{1}_{ij}) = \sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{E}(\mathbb{1}_{ij}) = \frac{n(n-1)}{4}$

Local Maxima

A local maximum in an array a occurs at i if $a_i > a_{i-1}$ and $a_i > a_{i+1}$ (for $2 \leq i \leq n-1$; only need $a_1 > a_2$ for $i = 1$ and $a_n > a_{n-1}$ for $i = n$).

Find the expected number of local maxima in a random permutation.

Local Maxima

- Let $\mathbb{1}_i$ be the indicator for if index i is a local maximum

Local Maxima

- Let $\mathbb{1}_i$ be the indicator for if index i is a local maximum
- We want $\sum_{i=1}^{n} \mathbb{1}_i$

Local Maxima

- Let $\mathbb{1}_i$ be the indicator for if index i is a local maximum
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- If $i = 1$ or $i = n$, then $\mathbb{E}(\mathbb{1}_i) = \frac{1}{2}$ like before

Local Maxima

- Let $\mathbb{1}_i$ be the indicator for if index i is a local maximum
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- If $i = 1$ or $i = n$, then $\mathbb{E}(\mathbb{1}_i) = \frac{1}{2}$ like before
- Otherwise, $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(a_i > a_{i-1}$ and $a_i > a_{i+1}) = \frac{1}{3}$

Local Maxima

- Let $\mathbb{1}_i$ be the indicator for if index i is a local maximum
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- If $i = 1$ or $i = n$, then $\mathbb{E}(\mathbb{1}_i) = \frac{1}{2}$ like before
- Otherwise, $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(a_i > a_{i-1} \text{ and } a_i > a_{i+1}) = \frac{1}{3}$
- So, $\mathbb{E}(\sum_{i=1}^{n} \mathbb{1}_i) = \sum_{i=1}^{n} \mathbb{E}(\mathbb{1}_i) = 2 \cdot \frac{1}{2} + (n-2) \cdot \frac{1}{3} = \frac{n+1}{3}$

Records

Consider the following process: we iterate over an array and maintain the "max seen so far", or record.

Find the expected number of times this value is updated for a random permutation.

Records

- Let $\mathbb{1}_i$ be the indicator for if index i is a record

Records

- Let $\mathbb{1}_i$ be the indicator for if index i is a record
- We want $\sum_{i=1}^{n} \mathbb{1}_i$

Records

- Let $\mathbb{1}_i$ be the indicator for if index i is a record
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i \text{ is a record}) = \frac{1}{i}$

Records

- Let $\mathbb{1}_i$ be the indicator for if index i is a record
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i \text{ is a record}) = \frac{1}{i}$
- So, $\mathbb{E}(\sum_{i=1}^{n} \mathbb{1}_i) = \sum_{i=1}^{n} \mathbb{E}(\mathbb{1}_i) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = O(\log n)$

Balls in Boxes

Randomly, k distinguishable balls are placed into n distinguishable boxes (all possibilities are equally likely). Find the expected number of empty boxes.

Balls in Boxes

- Let $\mathbb{1}_i$ be the indicator for if the *i*-th box is empty

Balls in Boxes

- Let $\mathbb{1}_i$ be the indicator for if the i-th box is empty
- We want $\sum_{i=1}^{n} \mathbb{1}_i$

Balls in Boxes

- Let $\mathbb{1}_i$ be the indicator for if the i-th box is empty
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th box is empty}) = (1 - \frac{1}{n})^k$

Balls in Boxes

- Let $\mathbb{1}_i$ be the indicator for if the i-th box is empty
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th box is empty}) = (1 - \frac{1}{n})^k$
- So, $\mathbb{E}(\sum_{i=1}^{n} \mathbb{1}_i) = \sum_{i=1}^{n} \mathbb{E}(\mathbb{1}_i) = n(1 - \frac{1}{n})^k$

Coin Tosses I

In a sequence of n independent fair coin tosses, find the expected number of occurrences of the pattern HTH (consecutively).

Note that overlap is allowed. For example, HTHTH has two instances of the pattern.

Coin Tosses I

- Let $\mathbb{1}_i$ be the indicator for if the pattern appears starting at the i-th toss

Coin Tosses I

- Let $\mathbb{1}_i$ be the indicator for if the pattern appears starting at the i-th toss
- We want $\sum_{i=1}^{n-2} \mathbb{1}_i$

Coin Tosses I

- Let $\mathbb{1}_i$ be the indicator for if the pattern appears starting at the i-th toss
- We want $\sum_{i=1}^{n-2} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(\text{pattern is seen starting at the } i\text{-th toss}) = \frac{1}{8}$

Coin Tosses I

- Let $\mathbb{1}_i$ be the indicator for if the pattern appears starting at the i-th toss
- We want $\sum_{i=1}^{n-2} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(\text{pattern is seen starting at the } i\text{-th toss}) = \frac{1}{8}$
- So, $\mathbb{E}(\sum_{i=1}^{n-2} \mathbb{1}_i) = \sum_{i=1}^{n-2} \mathbb{E}(\mathbb{1}_i) = \frac{n-2}{8}$

Coin Tosses II

A fair coin is flipped n times. The sequence of outcomes can be divided into runs (blocks of H's or blocks of T's). Find the expected number of runs.

For example, HHHTTHTTTH has 5 runs:

HHH	TT	H	TTT	H

Coin Tosses II

- The number of runs is equal to the number of tosses where the outcome changes from the previous value (plus one)

Coin Tosses II

- The number of runs is equal to the number of tosses where the outcome changes from the previous value (plus one)
- Let $\mathbb{1}_i$ be the indicator for if this happens on the i-th toss

Coin Tosses II

- The number of runs is equal to the number of tosses where the outcome changes from the previous value (plus one)
- Let $\mathbb{1}_i$ be the indicator for if this happens on the i-th toss
- We want $1 + \sum_{i=2}^{n} \mathbb{1}_i$

Coin Tosses II

- The number of runs is equal to the number of tosses where the outcome changes from the previous value (plus one)
- Let $\mathbb{1}_i$ be the indicator for if this happens on the i-th toss
- We want $1 + \sum_{i=2}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(\text{outcome changes on the } i\text{-th toss}) = \frac{1}{2}$

Coin Tosses II

- The number of runs is equal to the number of tosses where the outcome changes from the previous value (plus one)
- Let $\mathbb{1}_i$ be the indicator for if this happens on the i-th toss
- We want $1 + \sum_{i=2}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(\text{outcome changes on the } i\text{-th toss}) = \frac{1}{2}$
- So, $\mathbb{E}(1 + \sum_{i=2}^{n} \mathbb{1}_i) = 1 + \sum_{i=2}^{n} \mathbb{E}(\mathbb{1}_i) = 1 + (n-1) \cdot \frac{1}{2} = \frac{n+1}{2}$

Prizes

There are n prizes, with values $\$1, \$2, \cdots, \$n$. You randomly choose k prizes, without replacement.

Find the expected total value of the prizes you get.

Prizes

- Let $\mathbb{1}_i$ be the indicator for selecting the *i*-th prize

Prizes

- Let $\mathbb{1}_i$ be the indicator for selecting the *i*-th prize
- We want $\sum_{i=1}^{n} i \cdot \mathbb{1}_i$

Prizes

- Let $\mathbb{1}_i$ be the indicator for selecting the i-th prize
- We want $\sum_{i=1}^{n} i \cdot \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th prize is selected}) = \frac{k}{n}$

Prizes

- Let $\mathbb{1}_i$ be the indicator for selecting the *i*-th prize
- We want $\sum_{i=1}^{n} i \cdot \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th prize is selected}) = \frac{k}{n}$
- So, $\mathbb{E}(\sum_{i=1}^{n} i \cdot \mathbb{1}_i) = \sum_{i=1}^{n} i \cdot \mathbb{E}(\mathbb{1}_i) = \frac{n(n+1)}{2} \cdot \frac{k}{n} = \frac{k(n+1)}{2}$

Problems

Game on Tree

You are given a rooted tree with n nodes and play a game on it.

On each step, choose a node uniformly at random (among all remaining nodes) and remove its subtree. The game ends when there are no nodes left.

Find the expected number of steps in the game.

Constraints

$1 \leq n \leq 10^5$

Game on Tree - Solution

- The length of the game is equal to the number of selected nodes

Game on Tree - Solution

- The length of the game is equal to the number of selected nodes
- Let $\mathbb{1}_v$ be the indicator for if node v is chosen in the game

Game on Tree - Solution

- The length of the game is equal to the number of selected nodes
- Let $\mathbb{1}_v$ be the indicator for if node v is chosen in the game
- We want $\sum_{i=1}^{n} \mathbb{1}_i$

Game on Tree - Solution

- The length of the game is equal to the number of selected nodes
- Let $\mathbb{1}_v$ be the indicator for if node v is chosen in the game
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_v) = \mathbb{P}(v \text{ is chosen in the game}) = \frac{1}{d_v}$ where d_v is the depth of v, because we must choose v before any of its ancestors

Game on Tree - Solution

- The length of the game is equal to the number of selected nodes
- Let $\mathbb{1}_v$ be the indicator for if node v is chosen in the game
- We want $\sum_{i=1}^{n} \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_v) = \mathbb{P}(v \text{ is chosen in the game}) = \frac{1}{d_v}$ where d_v is the depth of v, because we must choose v before any of its ancestors
- So, $\mathbb{E}(\sum_{i=1}^{n} \mathbb{1}_i) = \sum_{i=1}^{n} \mathbb{E}(\mathbb{1}_i) = \sum_{i=1}^{n} \frac{1}{d_i}$

Inversion Expectation

You are given a permutation of size n. However, the numbers on some positions are replaced by -1.

A valid permutation is a replacement of the unknown -1's so that the resulting sequence is a permutation of size n.

The given sequence was turned into a valid permutation randomly with an equal probability of getting each valid permutation.

Find the expected number of inversions in the resulting permutation.

Constraints

$1 \leq n \leq 10^5$

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Four types of inversions:

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Four types of inversions:
  (1) $a_i = a_j = -1$

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Four types of inversions:
  (1) $a_i = a_j = -1$
  (2) $a_i, a_j \neq -1$

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Four types of inversions:
  (1) $a_i = a_j = -1$
  (2) $a_i, a_j \neq -1$
  (3) $a_i = -1,\ a_j \neq -1$

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Four types of inversions:
  (1) $a_i = a_j = -1$
  (2) $a_i, a_j \neq -1$
  (3) $a_i = -1,\ a_j \neq -1$
  (4) $a_i \neq -1,\ a_j = -1$

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Four types of inversions:
  (1) $a_i = a_j = -1$
  (2) $a_i, a_j \neq -1$
  (3) $a_i = -1$, $a_j \neq -1$
  (4) $a_i \neq -1$, $a_j = -1$
- Recall an earlier exercise to compute the contribution of (1)

Inversion Expectation - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Four types of inversions:
  (1) $a_i = a_j = -1$
  (2) $a_i, a_j \neq -1$
  (3) $a_i = -1,\ a_j \neq -1$
  (4) $a_i \neq -1,\ a_j = -1$
- Recall an earlier exercise to compute the contribution of (1)
- The contribution of (2) can be found with ordered sets

Inversion Expectation - Solution

- Considering (3), fix indices i and j

Inversion Expectation - Solution

- Considering (3), fix indices i and j
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}(p_i > a_j)$ where p is the final valid permutation

Inversion Expectation - Solution

- Considering (3), fix indices i and j
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}(p_i > a_j)$ where p is the final valid permutation
- This is equivalent to $\frac{\text{number of missing numbers greater than } a_j}{\text{total number of missing numbers}}$

Inversion Expectation - Solution

- Considering (3), fix indices i and j
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}(p_i > a_j)$ where p is the final valid permutation
- This is equivalent to $\frac{\text{number of missing numbers greater than } a_j}{\text{total number of missing numbers}}$
- Compute this using ordered sets!

Inversion Expectation - Solution

- Considering (3), fix indices i and j
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}(p_i > a_j)$ where p is the final valid permutation
- This is equivalent to $\frac{\text{number of missing numbers greater than } a_j}{\text{total number of missing numbers}}$
- Compute this using ordered sets!
- Multiply by the count of how many -1's there are before a_j

Inversion Expectation - Solution

- Considering (3), fix indices i and j
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}(p_i > a_j)$ where p is the final valid permutation
- This is equivalent to $\frac{\text{number of missing numbers greater than } a_j}{\text{total number of missing numbers}}$
- Compute this using ordered sets!
- Multiply by the count of how many -1's there are before a_j
- (4) is similar, except use suffix counts instead of prefix counts

Inversions After Shuffle

You are given a permutation of size n.

Consider the following operation: pick a random segment and shuffle its elements. All $\frac{n(n+1)}{2}$ segments are equiprobable.

Find the expected number of inversions after this operation is applied exactly once.

Constraints

$1 \leq n \leq 10^5$

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Decompose further by conditioning if i and j are shuffled

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Decompose further by conditioning if i and j are shuffled
- When they are, the contribution totals to $\sum_{l=1}^{n} \frac{n-l+1}{\frac{n(n+1)}{2}} \cdot \frac{l(l-1)}{4}$

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Decompose further by conditioning if i and j are shuffled
- When they are, the contribution totals to $\sum_{l=1}^{n} \frac{n-l+1}{\frac{n(n+1)}{2}} \cdot \frac{l(l-1)}{4}$
- Hereafter, $\mathbb{1}_{ij}$ will also indicate i or j (or possibly both) not being in the subsegment

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Decompose further by conditioning if i and j are shuffled
- When they are, the contribution totals to $\sum_{l=1}^{n} \frac{n-l+1}{\frac{n(n+1)}{2}} \cdot \frac{l(l-1)}{4}$
- Hereafter, $\mathbb{1}_{ij}$ will also indicate i or j (or possibly both) not being in the subsegment
- Observe that relative order remains unchanged in these cases, so i and j must form an inversion in the original permutation

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Decompose further by conditioning if i and j are shuffled
- When they are, the contribution totals to $\sum_{l=1}^{n} \frac{n-l+1}{\frac{n(n+1)}{2}} \cdot \frac{l(l-1)}{4}$
- Hereafter, $\mathbb{1}_{ij}$ will also indicate i or j (or possibly both) not being in the subsegment
- Observe that relative order remains unchanged in these cases, so i and j must form an inversion in the original permutation
- Fix such an inversion

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Decompose further by conditioning if i and j are shuffled
- When they are, the contribution totals to $\sum_{l=1}^{n}\frac{n-l+1}{\frac{n(n+1)}{2}}\cdot\frac{l(l-1)}{4}$
- Hereafter, $\mathbb{1}_{ij}$ will also indicate i or j (or possibly both) not being in the subsegment
- Observe that relative order remains unchanged in these cases, so i and j must form an inversion in the original permutation
- Fix such an inversion
- $\mathbb{E}(\mathbb{1}_{ij})=\mathbb{P}(i \text{ or } j \text{ are not in the subsegment})=1-\frac{i\cdot(n-j+1)}{\frac{n(n+1)}{2}}$

Inversions After Shuffle - Solution

- Consider $\sum_{i=1}^{n} \sum_{j=i+1}^{n} \mathbb{1}_{ij}$ where $\mathbb{1}_{ij}$ is the indicator for if indices i and j form an inversion
- Decompose further by conditioning if i and j are shuffled
- When they are, the contribution totals to $\sum_{l=1}^{n} \frac{n-l+1}{\frac{n(n+1)}{2}} \cdot \frac{l(l-1)}{4}$
- Hereafter, $\mathbb{1}_{ij}$ will also indicate i or j (or possibly both) not being in the subsegment
- Observe that relative order remains unchanged in these cases, so i and j must form an inversion in the original permutation
- Fix such an inversion
- $\mathbb{E}(\mathbb{1}_{ij}) = \mathbb{P}(i \text{ or } j \text{ are not in the subsegment}) = 1 - \frac{i \cdot (n-j+1)}{\frac{n(n+1)}{2}}$
- This yields an $O(n^2)$ solution, but we need to do better

Inversions After Shuffle - Solution

- Inspiration: Given an array a, count the number of inversions

Inversions After Shuffle - Solution

- Inspiration: Given an array a, count the number of inversions
- Process elements in increasing order while maintaining a segment tree (1 if the i-th element has been visited, 0 otherwise)

Inversions After Shuffle - Solution

- Inspiration: Given an array a, count the number of inversions
- Process elements in increasing order while maintaining a segment tree (1 if the i-th element has been visited, 0 otherwise)
- When a_i is processed, querying the suffix corresponds to the number of inversions that have a_i as the larger value

Inversions After Shuffle - Solution

- Inspiration: Given an array a, count the number of inversions
- Process elements in increasing order while maintaining a segment tree (1 if the i-th element has been visited, 0 otherwise)
- When a_i is processed, querying the suffix corresponds to the number of inversions that have a_i as the larger value
- Mimic this idea by maintaining two segment trees, one for each term of $\mathbb{E}(\mathbb{1}_{ij}) = 1 - \frac{i \cdot (n-j+1)}{\frac{n(n+1)}{2}}$

Inversions After Shuffle - Solution

- Inspiration: Given an array a, count the number of inversions
- Process elements in increasing order while maintaining a segment tree (1 if the i-th element has been visited, 0 otherwise)
- When a_i is processed, querying the suffix corresponds to the number of inversions that have a_i as the larger value
- Mimic this idea by maintaining two segment trees, one for each term of $\mathbb{E}(\mathbb{1}_{ij}) = 1 - \frac{i \cdot (n-j+1)}{\frac{n(n+1)}{2}}$
    1. The first is identical to inversion counting (for 1)

Inversions After Shuffle - Solution

- Inspiration: Given an array a, count the number of inversions
- Process elements in increasing order while maintaining a segment tree (1 if the i-th element has been visited, 0 otherwise)
- When a_i is processed, querying the suffix corresponds to the number of inversions that have a_i as the larger value
- Mimic this idea by maintaining two segment trees, one for each term of $\mathbb{E}(\mathbb{1}_{ij}) = 1 - \frac{i \cdot (n-j+1)}{\frac{n(n+1)}{2}}$
  1. The first is identical to inversion counting (for 1)
  2. The second queries suffixes as well, but stores $n - j + 1$ as the j-th value; multiply queries by their index i to get the contribution itself

Expected Damage

There are n monsters and m shields. You are given each monster's strength d_i and each shield's durability a_j and defence rating b_j.

When you fight a monster with strength d while having a shield with current durability a and defence b, three outcomes are possible:

- If $a = 0$, then you receive d damage
- If $a > 0$ and $d \geq b$, then you receive no damage but the durability of the shield decreases by 1
- If $a > 0$ and $d < b$, then nothing happens

For each shield, find the expected amount of damage you will receive if you take this shield and fight the monsters in a random order.

Constraints

$1 \leq n, m \leq 10^5$

$1 \leq a_j, b_j, d_i \leq 10^9$

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$
- A monster can only deal damage after the shield "breaks"

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$
- A monster can only deal damage after the shield "breaks"
- That is, at least a monsters with strength at least b must have already hit the shield

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$
- A monster can only deal damage after the shield "breaks"
- That is, at least a monsters with strength at least b must have already hit the shield
- Let k be the number of "strong" monsters with strength at least b

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$
- A monster can only deal damage after the shield "breaks"
- That is, at least a monsters with strength at least b must have already hit the shield
- Let k be the number of "strong" monsters with strength at least b
- Two cases:
    1. $d_i < b$
    2. $d_i \geq b$

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$
- A monster can only deal damage after the shield "breaks"
- That is, at least a monsters with strength at least b must have already hit the shield
- Let k be the number of "strong" monsters with strength at least b
- Two cases:
  1. $d_i < b$ yields $\frac{k+1-a}{k+1}$ by considering the subsequence consisting of strong monsters and the i-th monster
  2. $d_i \geq b$ similarly yields $\frac{k-a}{k}$
- This yields an $O(nm)$ solution, but we need to do better

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$
- A monster can only deal damage after the shield “breaks”
- That is, at least a monsters with strength at least b must have already hit the shield
- Let k be the number of “strong” monsters with strength at least b
- Two cases:
  1. $d_i < b$ yields $\frac{k+1-a}{k+1}$ by considering the subsequence consisting of strong monsters and the i-th monster
  2. $d_i \geq b$ similarly yields $\frac{k-a}{k}$
- This yields an $O(nm)$ solution, but we need to do better
- Observe that these values are constant with respect to d_i

Expected Damage - Solution

- Fix the j-th shield and consider $\sum_{i=1}^{n} d_i \cdot \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if the i-th monster inflicts damage
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(i\text{-th monster inflicts damage})$
- A monster can only deal damage after the shield "breaks"
- That is, at least a monsters with strength at least b must have already hit the shield
- Let k be the number of "strong" monsters with strength at least b
- Two cases:
  1. $d_i < b$ yields $\frac{k+1-a}{k+1}$ by considering the subsequence consisting of strong monsters and the i-th monster
  2. $d_i \geq b$ similarly yields $\frac{k-a}{k}$
- This yields an $O(nm)$ solution, but we need to do better
- Observe that these values are constant with respect to d_i
- So, we can improve by using prefix sums and binary search

Expected Square Beauty

Define the beauty of an array x as the minimal size of a partition of x into subsegments such that all elements in each subsegment are equal. For example, $B([3,3,6,1,6,6,6]) = 4$ using $[3,3|6|1|6,6,6]$.

You are given n intervals $M_i = [l_i, r_i]$. We construct a random array x by choosing x_i to be a uniformly random integer in M_i.

Find the expected value of $B(x)^2$.

Constraints

$1 \leq n \leq 10^5$

$1 \leq l_i \leq r_i \leq 10^9$

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$
- Expand $B(x)^2$ and first consider contribution of $1 \cdot \mathbb{1}_i = \mathbb{1}_i \cdot \mathbb{1}_i = \mathbb{1}_i$

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$
- Expand $B(x)^2$ and first consider contribution of $1 \cdot \mathbb{1}_i = \mathbb{1}_i \cdot \mathbb{1}_i = \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(x_i \neq x_{i-1}) = 1 - \mathbb{P}(x_i = x_{i-1}) = 1 - \frac{\left|M_{i-1} \cap M_i\right|}{|M_{i-1}| \cdot |M_i|}$

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$
- Expand $B(x)^2$ and first consider contribution of $1 \cdot \mathbb{1}_i = \mathbb{1}_i \cdot \mathbb{1}_i = \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(x_i \neq x_{i-1}) = 1 - \mathbb{P}(x_i = x_{i-1}) = 1 - \frac{\left|M_{i-1} \cap M_i\right|}{|M_{i-1}| \cdot |M_i|}$
- Examine $\mathbb{1}_i \cdot \mathbb{1}_j$ next and WLOG take $i < j$

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$
- Expand $B(x)^2$ and first consider contribution of $1 \cdot \mathbb{1}_i = \mathbb{1}_i \cdot \mathbb{1}_i = \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(x_i \neq x_{i-1}) = 1 - \mathbb{P}(x_i = x_{i-1}) = 1 - \frac{\left|M_{i-1} \cap M_i\right|}{|M_{i-1}| \cdot |M_i|}$
- Examine $\mathbb{1}_i \cdot \mathbb{1}_j$ next and WLOG take $i < j$
- Independence is good because we can take $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_j)$

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$
- Expand $B(x)^2$ and first consider contribution of $1 \cdot \mathbb{1}_i = \mathbb{1}_i \cdot \mathbb{1}_i = \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(x_i \neq x_{i-1}) = 1 - \mathbb{P}(x_i = x_{i-1}) = 1 - \frac{\left|M_{i-1} \cap M_i\right|}{|M_{i-1}| \cdot |M_i|}$
- Examine $\mathbb{1}_i \cdot \mathbb{1}_j$ next and WLOG take $i < j$
- Independence is good because we can take $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_j)$
- Fortunately, $\mathbb{1}_i$ and $\mathbb{1}_j$ are dependent only when $j = i + 1$

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$
- Expand $B(x)^2$ and first consider contribution of $1 \cdot \mathbb{1}_i = \mathbb{1}_i \cdot \mathbb{1}_i = \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(x_i \neq x_{i-1}) = 1 - \mathbb{P}(x_i = x_{i-1}) = 1 - \frac{\left|M_{i-1} \cap M_i\right|}{|M_{i-1}| \cdot |M_i|}$
- Examine $\mathbb{1}_i \cdot \mathbb{1}_j$ next and WLOG take $i < j$
- Independence is good because we can take $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_j)$
- Fortunately, $\mathbb{1}_i$ and $\mathbb{1}_j$ are dependent only when $j = i + 1$
- $\mathbb{E}(\mathbb{1}_i \cdot \mathbb{1}_{i+1}) = \mathbb{P}(x_{i-1} \neq x_i$ and $x_i \neq x_{i+1})$ can be found with the Principle of Inclusion-Exclusion

Expected Square Beauty - Solution

- $B(x) = 1 + \sum_{i=2}^{n} \mathbb{1}_i$ where $\mathbb{1}_i$ is the indicator for if $x_i \neq x_{i-1}$
- Expand $B(x)^2$ and first consider contribution of $1 \cdot \mathbb{1}_i = \mathbb{1}_i \cdot \mathbb{1}_i = \mathbb{1}_i$
- $\mathbb{E}(\mathbb{1}_i) = \mathbb{P}(x_i \neq x_{i-1}) = 1 - \mathbb{P}(x_i = x_{i-1}) = 1 - \frac{\left|M_{i-1} \cap M_i\right|}{|M_{i-1}| \cdot |M_i|}$
- Examine $\mathbb{1}_i \cdot \mathbb{1}_j$ next and WLOG take $i < j$
- Independence is good because we can take $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_j)$
- Fortunately, $\mathbb{1}_i$ and $\mathbb{1}_j$ are dependent only when $j = i + 1$
- $\mathbb{E}(\mathbb{1}_i \cdot \mathbb{1}_{i+1}) = \mathbb{P}(x_{i-1} \neq x_i$ and $x_i \neq x_{i+1})$ can be found with the Principle of Inclusion-Exclusion
- This yields an $O(n^2)$ solution, but we need to do better

Expected Square Beauty - Solution

- Inspiration: Given an array a, compute $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j$ in $O(n)$

Expected Square Beauty - Solution

- Inspiration: Given an array a, compute $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j$ in $O(n)$
- $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j = \frac{1}{2} \cdot \left((\sum_{i=1}^{n} a_i)^2 - \sum_{i=1}^{n} {a_i}^2 \right)$

Expected Square Beauty - Solution

- Inspiration: Given an array a, compute $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j$ in $O(n)$
- $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j = \frac{1}{2} \cdot \left((\sum_{i=1}^{n} a_i)^2 - \sum_{i=1}^{n} {a_i}^2 \right)$
- We mimic this idea by squaring $s = 1 + \sum_{i=2}^{n} \mathbb{E}(\mathbb{1}_i)$

Expected Square Beauty - Solution

- Inspiration: Given an array a, compute $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j$ in $O(n)$
- $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j = \frac{1}{2} \cdot \left((\sum_{i=1}^{n} a_i)^2 - \sum_{i=1}^{n} {a_i}^2 \right)$
- We mimic this idea by squaring $s = 1 + \sum_{i=2}^{n} \mathbb{E}(\mathbb{1}_i)$
- Two corrections need to be made:

Expected Square Beauty - Solution

- Inspiration: Given an array a, compute $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j$ in $O(n)$
- $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j = \frac{1}{2} \cdot \left((\sum_{i=1}^{n} a_i)^2 - \sum_{i=1}^{n} {a_i}^2 \right)$
- We mimic this idea by squaring $s = 1 + \sum_{i=2}^{n} \mathbb{E}(\mathbb{1}_i)$
- Two corrections need to be made:
    1. Instead of $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_i)$, we should add $\mathbb{E}(\mathbb{1}_i)$

Expected Square Beauty - Solution

- Inspiration: Given an array a, compute $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j$ in $O(n)$
- $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j = \frac{1}{2} \cdot \left((\sum_{i=1}^{n} a_i)^2 - \sum_{i=1}^{n} {a_i}^2 \right)$
- We mimic this idea by squaring $s = 1 + \sum_{i=2}^{n} \mathbb{E}(\mathbb{1}_i)$
- Two corrections need to be made:
  1. Instead of $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_i)$, we should add $\mathbb{E}(\mathbb{1}_i)$
  2. Instead of $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_{i+1})$, we should add $\mathbb{E}(\mathbb{1}_i \cdot \mathbb{1}_{i+1})$

Expected Square Beauty - Solution

- Inspiration: Given an array a, compute $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j$ in $O(n)$
- $\sum_{i=1}^{n} \sum_{j=i+1}^{n} a_i \cdot a_j = \frac{1}{2} \cdot \left((\sum_{i=1}^{n} a_i)^2 - \sum_{i=1}^{n} {a_i}^2 \right)$
- We mimic this idea by squaring $s = 1 + \sum_{i=2}^{n} \mathbb{E}(\mathbb{1}_i)$
- Two corrections need to be made:
  1. Instead of $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_i)$, we should add $\mathbb{E}(\mathbb{1}_i)$
  2. Instead of $\mathbb{E}(\mathbb{1}_i) \cdot \mathbb{E}(\mathbb{1}_{i+1})$, we should add $\mathbb{E}(\mathbb{1}_i \cdot \mathbb{1}_{i+1})$
- All of this can be done in $O(n)$

Resources

References

- Introduction to Probability (Blitzstein, Hwang)

Problems

From Lecture	**280 C**	Game on Tree
	1096 F	Inversion Expectation
	749 E	Inversions After Shuffle
	1418 E	Expected Damage
	1187 F	Expected Square Beauty
More Practice	**2020 E**	Expected Power
	1983 E	I Love Balls
	1854 C	Expected Destruction

Thank you!