CS648 Project

Aayush Gupta (210020)
Ridin Datta (210840)
Shrey Wadhawan (200948)

Problem Statement

- **Finding the expected length of the smallest interval**

n points are picked randomly uniformly and independently from the $[0,1]$ line segment. This will create $n+1$ intervals. What is the expected length of the smallest interval among these intervals?

Claim

- The expected length of the smallest interval is $O(n^{-2})$

Proofs

- By Randomized Incremental Construction
  - we will be proving that the expected length of smallest interval is $\theta(n^{-2})$.
- By a novel method

Proof 1: Randomized Incremental Construction

As all points are picked uniformly and independently, so we can analyze the problem as gradually picking the points one by one.

We will be analysing the case when we are picking the $(i + 1)^{th}$ point.

Proof 1: Randomized Incremental Construction

- Lemma 1: $\alpha \cdot (i+1) \leq P(\delta_{i+1} < \delta_i | \delta_i = \alpha) \leq 2\alpha \cdot (i+1)$
- Lemma 2: $\mathrm{P}(\delta_{\mathrm{i+1}} < \delta_{\mathrm{i}}) = \frac{2}{\mathrm{i+1}}$

δ_{i} := Length of the smallest interval after dropping i points randomly and uniformly on the $[0,1]$ line

Proof 1: Randomized Incremental Construction

Using Lemma 1:

- $\alpha \cdot (i+1) \leq P(\delta_{i+1} < \delta_i | \delta_i = \alpha) \leq 2\alpha \cdot (i+1)$
- $\Rightarrow \alpha \cdot (i+1) \cdot P(\delta_i = \alpha) \leq P(\delta_{i+1} < \delta_i | \delta_i = \alpha) \cdot P(\delta_i = \alpha) \leq 2\alpha \cdot (i+1) \cdot P(\delta_i = \alpha)$
- $\Rightarrow \int \alpha \cdot (i+1) \cdot P(\delta_i = \alpha) d\alpha \leq \int P(\delta_{i+1} < \delta_i | \delta_i = \alpha) \cdot P(\delta_i = \alpha) d\alpha \leq \int 2\alpha \cdot (i+1) \cdot P(\delta_i = \alpha) d\alpha$
- $\Rightarrow (i+1)E[\delta_i] \leq P(\delta_{i+1} < \delta_i) \leq 2(i+1)E[\delta_i]$
- $\Rightarrow \frac{P(\delta_{i+1}<\delta_i)}{2(i+1)} \leq E[\delta_i] \leq \frac{P(\delta_{i+1}<\delta_i)}{i+1}$

Proof 1: Randomized Incremental Construction

Using Lemma 2:

- $\frac{1}{(i+1)^2} \leq E[\delta_i] \leq \frac{2}{(i+1)^2}$
- Thus, $E[\delta_i] = \theta(i^{-2})$
- $\Rightarrow E[\delta_n] = \theta(n^{-2})$

Proof of Lemma 1

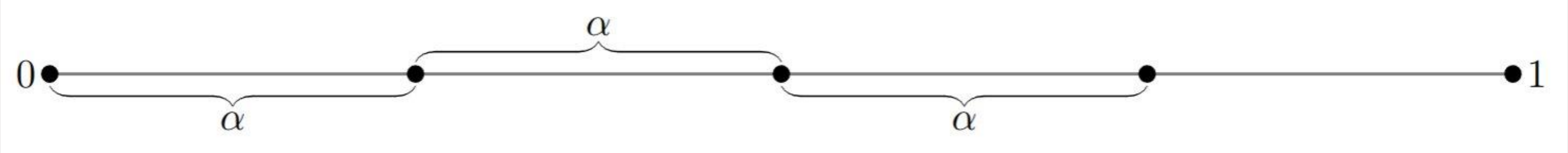
- Let us call the region where dropping the $(i+1)th$ point causes $\delta_{i+1} < \delta_i$ the "favourable" region. If $\delta_i = \alpha$, for maximum possible favorable region all the i neighboring points need to be separated by a distance of 2α. Moreover, the distance between the leftmost point and 0 and the rightmost point and 1 individually need to be greater than 2α . Since the size of the favorable region directly translates to probability of the $(i+1)th$ point falling in the favorable region, from these arguments,

$$P(\delta_{i+1} < \delta_i | \delta_i = \alpha) \leq 2\alpha \cdot (i+1)$$

Proof of Lemma 1

- The situation representing the minimum possible favorable region is the case where all neighboring points are separated by a distance of α including the distance between 0 and the leftmost point and 1 and the rightmost point. Thus,

$$\alpha \cdot (i + 1) \leq P(\delta_{i+1} < \delta_i | \delta_i = \alpha)$$



Proof of Lemma 2

- In the randomized incremental construction framework let us consider that i points have been dropped. The smallest interval is defined by 2 points. Let us call them p and q. If either p or q are chosen as the $(i+1)th$ point they would define the smallest interval. Otherwise, the smallest interval remains the same as δ_i Since the points have been chosen random uniformly, every point has the same probability of being the last dropped point. Hence,

$$\mathrm{P}(\delta_{\mathrm{i}+1} < \delta_{\mathrm{i}}) = \frac{2}{\mathrm{i}+1}$$

Proof 2: By Geometry

- We will prove the following claim in a novel geometric fashion

The kth largest interval's expected length is equal to

$$\frac{\frac{1}{k} + \frac{1}{k+1} + \cdots + \frac{1}{n+1}}{n+1}$$

Proof 2: By Geometry

- Without loss of generality, assume the [0, 1] segment is broken into segments of length
$s_1 \geq s_2 \geq \cdots \geq s_n \geq s_{n+1}$, in that order

$$\sum_{i=1}^{n+1} s_i = 1$$

Define:

$$x_i := s_i - s_{i+1}, \forall i \in [1, n] \quad \text{with} \quad x_{n+1} = s_{n+1}$$

This implies,

$$s_i = \sum_{k=i}^{n+1} x_k$$

Note that $x_i \geq 0$ for each i

Proof 2: By Geometry

$$\sum_{i=1}^{n+1} i x_i = 1$$

Define: $y_i = i x_i$

$$\implies y_1 + \cdots + y_{n+1} = 1$$

Claim: $\mathbb{E}[y_i] = \frac{1}{n+1}$ for all i

Proof 2: By Geometry

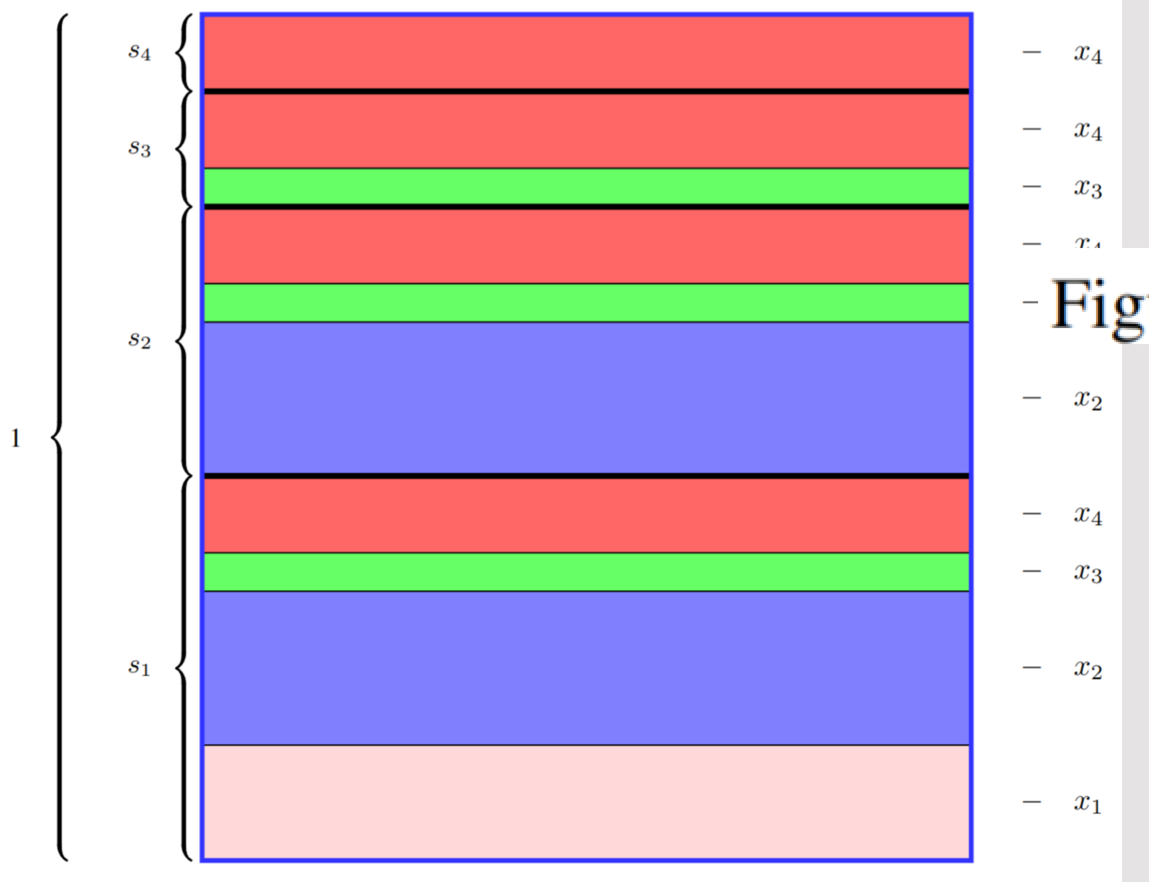


Figure 2

Figure 1

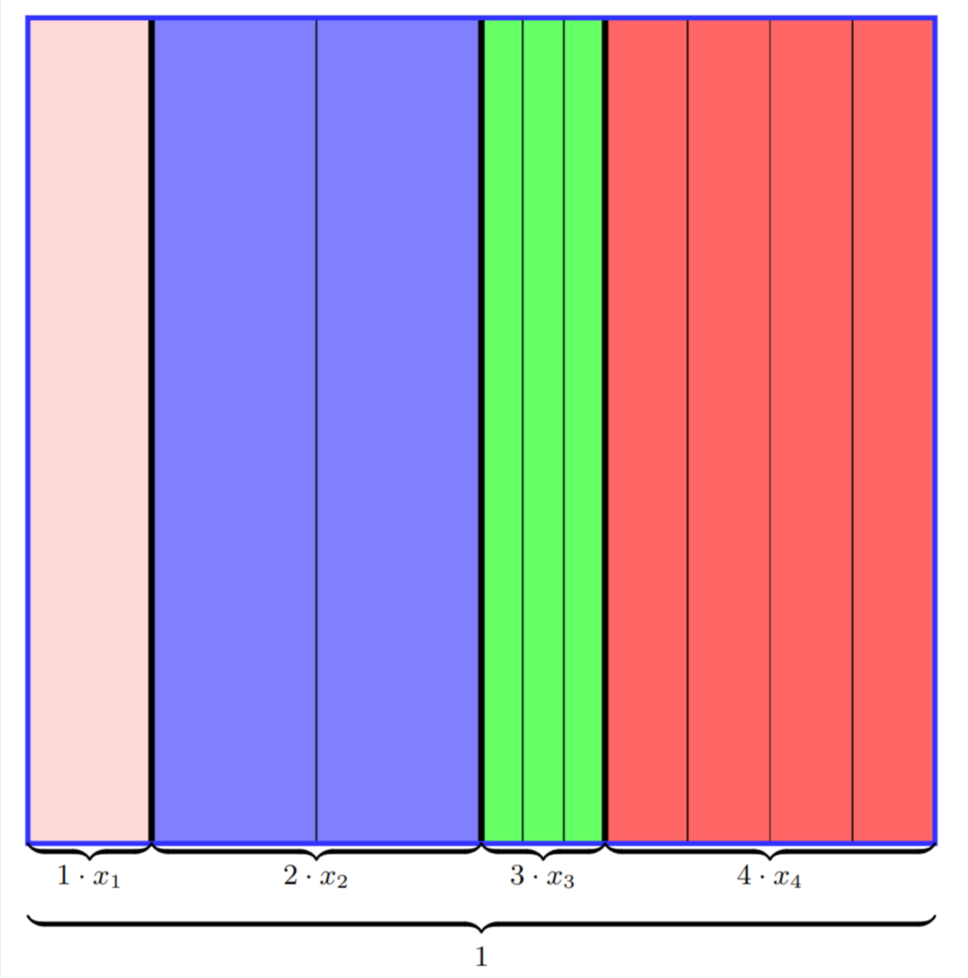


Figure 2

Proof 2: By Geometry

Consider the rearrangement in Figure-2 where each ix_i are separated by darker lines. Once these darker lines are placed, the lighter lines are *uniquely* determined by the dark lines as they equally partition each subrectangle formed by ix_i. In this case, these three random lines can be put anywhere randomly uniformly independently on the side with total length 1. Hence the expected value of each such $ix_i = \frac{1}{4} = \frac{1}{n+1}$.

To show the reverse, Figure 2 $\implies$ Figure 1, follow the steps:

1. On the side with length 1, randomly draw n vertical lines. Call the $n+1$ regions so formed $y_1, y_2, \ldots, y_{n+1}$.
2. Divide y_i into i equal parts, such that $x_i = y_i/i$ is the length of one part of y_i.
3. Now, there are i copies of x_i. Rearrange them into groups such that $([x_1, x_2, \ldots, x_{n+1}]; [x_2, x_3, \ldots, x_{n+1}]; \ldots; [x_n, x_{n+1}]; [x_{n+1}]) = s_1; s_2; \ldots; s_n; s_{n+1})$.

Proof 2 By Geometry

Since the division of y_i into subparts of length x_i in Step 3 takes place after Step 1. So, step 3 does not affect the expected value of y_i in step 1. As the n lines were drawn randomly uniformly, we get $\mathbb{E}[y_i] = \mathbb{E}[y_j]$, $\forall i \neq j$.

There is a clear bijection

$$\sum_{i=1}^{n+1} y_i = 1 \implies \mathbb{E}[y_i] = \frac{1}{n+1}$$

Proof 2: By Geometry

$$\implies \mathbb{E}[x_i] = \frac{\mathbb{E}[y_i]}{i} = \frac{1}{i(n+1)}$$

By linearity of expectation,

$$\implies E[s_i] = E[x_i] + \cdots + E[x_{n+1}] = \frac{1}{n+1}\left(\frac{1}{i} + \cdots + \frac{1}{n+1}\right)$$

QED