

# CMPT 210: Probability and Computing

## Lecture 15

---

Sharan Vaswani

March 3, 2023

# Recap

**Expectation**/mean of a random variable  $R$  is denoted by  $\mathbb{E}[R]$  and “summarizes” its distribution.

Formally,  $\mathbb{E}[R] := \sum_{\omega \in \mathcal{S}} \Pr[\omega] R[\omega]$

*Example:* When throwing a standard dice, if  $R$  is the random variable equal to the number on the dice.  $\mathbb{E}[R] = \sum_{i \in \{1,2,\dots,6\}} \frac{1}{6} [i] = \frac{7}{2}$ .

**Alternate definition of expectation:**  $\mathbb{E}[R] = \sum_{x \in \text{Range}(R)} x \Pr[R = x]$ .

This definition does not depend on the sample space.

*Example:* If  $\mathcal{I}_A$  is the indicator random variable for event  $A$ , then

$\mathbb{E}[\mathcal{I}_A] = \Pr[\mathcal{I}_A = 1](1) + \Pr[\mathcal{I}_A = 0](0) = \Pr[A]$ . For  $\mathcal{I}_A$ , the expectation is equal to the probability that event  $A$  happens.

**Linearity of Expectation:** For  $n$  random variables  $R_1, R_2, \dots, R_n$  and constants  $a_1, a_2, \dots, a_n$ ,  
 $\mathbb{E} \left[ \sum_{i=1}^n a_i R_i \right] = \sum_{i=1}^n a_i \mathbb{E}[R_i]$ .

# Recap

If  $R \sim \text{Bernoulli}(p)$ ,  $\mathbb{E}[R] = p$ . *Example:* When tossing a coin, if  $R$  is the random variable equal to 1 if we get a heads.

If  $R \sim \text{Uniform}(\{v_1, \dots, v_n\})$ ,  $\mathbb{E}[R] = \frac{v_1 + v_2 + \dots + v_n}{n}$ . *Example:* When throwing an  $n$ -sided dice with numbers  $v_1, \dots, v_n$ , if  $R$  is the random variable equal to the number.

If  $R \sim \text{Bin}(n, p)$ ,  $\mathbb{E}[R] = np$ . *Example:* When tossing  $n$  independent coins, if  $R$  is the random variable equal to the number of heads.

If  $R \sim \text{Geo}(p)$ ,  $\mathbb{E}[R] = \frac{1}{p}$ . *Example:* When tossing a coin repeatedly, if  $R$  is the random variable equal to the number of tosses required to get the first heads.

## Expectation - Examples - Coupon Collector Problem

**Q:** In a game started by a coffee shop, each time we buy a coffee, we get a coupon. Each coupon has a color (amongst  $n$  different colors) and each time, the color of the coupon is selected uniformly at random from amongst the  $n$  colors. If we collect at least one coupon of each color, we can claim a free coffee. On average, how many coupons should we collect (coffees we should buy) to claim the prize?

Suppose we get the following sequence of coupons:

*blue, green, green, red, blue, orange, blue, orange, gray*

Let us partition this sequence into segments such that a segment ends when we collect a coupon of a new color we did not have before. For this example,

*blue* *green* *green, red* *blue, orange* *blue, orange, gray*  
 $S_1$   $S_2$   $S_3$   $S_4$   $S_5$

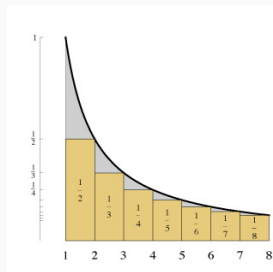
If the number of segments is equal to  $n$ , by definition, we will have collected coupons of the  $n$  different colors. Define  $X_k$  to be the random variable equal to the length of segment  $S_k$  and  $T$  to be the total number of coupons required to have at least one coupon per color.

# Expectation - Examples - Coupon Collector Problem

$T = X_1 + X_2 + \dots + X_n$ . We wish to compute  $\mathbb{E}[T]$ . By linearity of expectation,  $\mathbb{E}[T] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n]$ .

Let us calculate  $\mathbb{E}[X_k]$ . If we are on segment  $k$ , we have seen  $k - 1$  colors before. Hence, the probability of seeing a new (one that we have not seen before) colored coupon in  $S_k$  is  $\frac{n-(k-1)}{n}$ .  $X_k \sim \text{Geo}\left(\frac{n-(k-1)}{n}\right)$ , and we know that  $\mathbb{E}[X_k] = \frac{n}{n-k+1}$ .

$$\begin{aligned}\mathbb{E}[T] &= \sum_{k=1}^n \frac{n}{n-k+1} = n \left[ \frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{1} \right] \\ &\leq n \left[ 1 + \int_1^n \frac{dx}{x} \right] = n [1 + \ln(n)]\end{aligned}$$

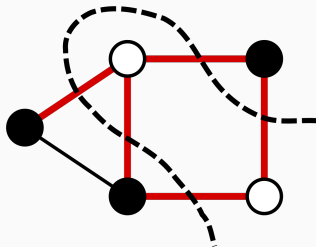


We also know that  $\mathbb{E}[T] \geq n \ln(n+1)$ . Hence,  $\mathbb{E}[T] = O(n \ln(n))$ , meaning that we need to buy  $O(n \ln(n))$  coffees to collect coupons of  $n$  colors and get a free coffee.

Questions?

# Max Cut

Given a graph  $G = (\mathcal{V}, \mathcal{E})$ , partition the graph's vertices into two complementary sets  $\mathcal{S}$  and  $\mathcal{T}$ , such that the number of edges between the set  $\mathcal{S}$  and the set  $\mathcal{T}$  is as large as possible.



Max Cut has applications to VLSI circuit design.

Equivalently, find a set  $\mathcal{U} \subseteq \mathcal{V}$  of vertices that solve the following

$$\max_{\mathcal{U} \subseteq \mathcal{V}} |\delta(\mathcal{U})| \text{ where } \delta(\mathcal{U}) := \{(u, v) \in \mathcal{E} \mid u \in \mathcal{U} \text{ and } v \notin \mathcal{U}\}$$

Here,  $\delta(\mathcal{U})$  is referred to as the “cut” corresponding to the set  $\mathcal{U}$ .

# Max Cut

- Max Cut is NP-hard (Karp, 1972), meaning that there is no polynomial (in  $|\mathcal{E}|$ ) time algorithm that solves Max Cut exactly.
- We want to find an approximate solution  $\mathcal{U}$  such that, if  $OPT$  is the size of the optimal cut, then,  $|\delta(\mathcal{U})| \geq \alpha OPT$  where  $\alpha \in (0, 1)$  is the multiplicative approximation factor.
- Randomized algorithm that guarantees an approximate solution with  $\alpha = \frac{1}{2}$  with probability close to 1 (Erdos, 1967).
- Algorithm with  $\alpha = 0.878$ . (Goemans and Williamson, 1995).
- Under some technical conditions, no efficient algorithm has  $\alpha > 0.878$  (Khot et al, 2004).

We will use Erdos' randomized algorithm and first prove the result in expectation. We wish to prove that for  $\mathcal{U}$  returned by Erdos' algorithm,

$$\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} OPT$$

**Algorithm:** Select  $\mathcal{U}$  to be a random subset of  $\mathcal{V}$  i.e. for each vertex  $v$ , choose  $v$  to be in the set  $\mathcal{U}$  independently with probability  $\frac{1}{2}$  (do not even look at the edges!).



**Claim:** For Erdos' algorithm,  $\mathbb{E}[|\delta(\mathcal{U})|] \geq \frac{1}{2} \text{OPT}$ .

**Proof:** For each edge  $(u, v) \in \mathcal{E}$ , let  $X_{u,v}$  be the indicator random variable equal to 1 iff the event  $E_{u,v} = \{(u, v) \in \delta(\mathcal{U})\}$  happens.

$$\mathbb{E}[|\delta(\mathcal{U})|] = \mathbb{E}\left[\sum_{(u,v) \in \mathcal{E}} X_{u,v}\right] = \sum_{(u,v) \in \mathcal{E}} \mathbb{E}[X_{u,v}] = \sum_{(u,v) \in \mathcal{E}} \Pr[E_{u,v}]$$

(Linearity of expectation, and Expectation of indicator r.v's.)

$$\begin{aligned}\Pr[E_{u,v}] &= \Pr[(u, v) \in \delta(\mathcal{U})] = \Pr[(u \in \mathcal{U} \cap v \notin \mathcal{U}) \cup (u \notin \mathcal{U} \cap v \in \mathcal{U})] \\ &= \Pr[(u \in \mathcal{U} \cap v \notin \mathcal{U})] + \Pr[(u \notin \mathcal{U} \cap v \in \mathcal{U})] \quad (\text{Union rule for mutually exclusive events})\end{aligned}$$

$$\Pr[E_{u,v}] = \Pr[u \in \mathcal{U}] \Pr[v \notin \mathcal{U}] + \Pr[u \notin \mathcal{U}] \Pr[v \in \mathcal{U}] = \frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} = \frac{1}{2}.$$

(Independent events)

$$\implies \mathbb{E}[|\delta(\mathcal{U})|] = \sum_{(u,v) \in \mathcal{E}} \Pr[E_{u,v}] = \frac{|\mathcal{E}|}{2} \geq \frac{\text{OPT}}{2}.$$

Questions?

# Conditional Expectation

Similar to probabilities, expectations can be conditioned on some event.

For random variable  $R$ , the expected value of  $R$  conditioned on an event  $A$  is given by:

$$\mathbb{E}[R|A] = \sum_{x \in \text{Range}(R)} x \Pr[R = x|A]$$

**Q:** If we throw a standard dice and define  $R$  to be the random variable equal to the number that comes up, what is the expected value of  $R$  given that the number is at most 4?

Let  $A$  be the event that the number is at most 4.

$$\Pr[R = 1|A] = \frac{\Pr[(R=1) \cap A]}{\Pr[A]} = \frac{\Pr[R=1]}{\Pr[A]} = \frac{1/6}{4/6} = 1/4.$$

$$\Pr[R = 2|A] = \Pr[R = 3|A] = \Pr[R = 4|A] = \frac{1}{4} \text{ and } \Pr[R = 5|A] = \Pr[R = 6|A] = 0.$$

$$\mathbb{E}[R|A] = \sum_{x \in \{1,2,3,4\}} x \Pr[R = x|A] = \frac{1}{4}[1 + 2 + 3 + 4] = \frac{5}{2}.$$

**Q:** What is the expected value of  $R$  given that the number is at least 4?

# Law of Total Expectation

If  $R$  is a random variable  $\mathcal{S} \rightarrow V$  and events  $A_1, A_2, \dots, A_n$  form a partition of the sample space i.e. for all  $i, j$ ,  $A_i \cap A_j = \emptyset$  and  $A_1 \cup A_2 \cup \dots \cup A_n = \mathcal{S}$ , then,

$$\mathbb{E}[R] = \sum_i \mathbb{E}[R|A_i] \Pr[A_i].$$

*Proof:*

$$\begin{aligned} \mathbb{E}[R] &= \sum_{x \in \text{Range}(R)} x \Pr[R = x] = \sum_{x \in \text{Range}(R)} x \sum_i \Pr[R = x|A_i] \Pr[A_i] \\ &\quad \text{(Law of total probability)} \\ &= \sum_i \Pr[A_i] \sum_{x \in \text{Range}(R)} x \Pr[R = x|A_i] \\ \implies \mathbb{E}[R] &= \sum_i \Pr[A_i] \mathbb{E}[R|A_i]. \end{aligned}$$

## Conditional Expectation - Examples

**Q:** Suppose that 49.6% of the people in the world are male and the rest female. If the expected height of a randomly chosen male is 5 feet 11 inches, while the expected height of a randomly chosen female is 5 feet 5 inches, what is the expected height of a randomly chosen person?

Define  $H$  to be the random variable equal to the height (in feet) of a randomly chosen person. Define  $M$  to be the event that the person is male and  $F$  the event that the person is female.

We wish to compute  $\mathbb{E}[H]$  and we know that  $\mathbb{E}[H|M] = 5 + \frac{11}{12}$  and  $\mathbb{E}[H|F] = 5 + \frac{5}{12}$ .

$\Pr[M] = 0.496$  and  $\Pr[F] = 1 - 0.496 = 0.504$ .

Hence,  $\mathbb{E}[H] = \mathbb{E}[H|M] \Pr[M] + \mathbb{E}[H|F] \Pr[F] = \frac{71}{12}(0.496) + \frac{65}{12}(0.504)$ .

Questions?

# Randomized Quick Select

Given an array  $A$  of  $n$  distinct numbers, return the  $k^{\text{th}}$  smallest element in  $A$  for  $k \in [1, n]$ .

---

**Algorithm** Randomized Quick Select

---

```
1: function QuickSelect( $A, k$ )
2:   If  $\text{Length}(A) = 1$ , return  $A[1]$ .
3:   Select  $p \in A$  uniformly at random.
4:   Construct sets  $\text{Left} := \{x \in A \mid x < p\}$  and  $\text{Right} := \{x \in A \mid x > p\}$ .
5:    $r = |\text{Left}| + 1$  {Element  $p$  is the  $r^{\text{th}}$  smallest element in  $A$ .}
6:   if  $k = r$  then
7:     return  $p$ 
8:   else if  $k < r$  then
9:     QuickSelect( $\text{Left}, k$ )
10:  else
11:    QuickSelect( $\text{Right}, k - r$ )
12:  end if
```

---

## Randomized QuickSelect

If  $A = \{2, 7, 0, 1, 3\}$  and we wish to find the  $2^{nd}$  smallest element meaning that  $k = 2$ .

According to the algorithm,  $p \sim \text{Uniform}(A)$ . Say  $p = 3$ .

Then after step 1,  $\text{Left} = \{2, 0, 1\}$  and  $\text{Right} = \{7\}$ .  $r := |\text{Left}| + 1 = 3 + 1 = 4$ . Since  $r > k$ , we recurse on the left-hand side by calling the algorithm on  $\{2, 0, 1\}$  with  $k = 2$ .

$p \sim \text{Uniform}(\{2, 0, 1\})$ . Say  $p = 1$ . After step 2,  $\text{Left} = \{0\}$  and  $\text{Right} = \{2\}$ .

$r := |\text{Left}| + 1 = 1 + 1 = 2$ . Since  $r = k$ , we terminate the recursion and return  $p = 1$  as the second-smallest element in  $A$ .

Q: Run the algorithm if  $p = 0$  in the first step?

Q: Run the algorithm if  $p = 1$  in the first step?



## Randomized Quick Select – Analysis

**Alternate way:** Sort the elements in  $A$  and return the  $k^{\text{th}}$  element in the sorted list. Uses  $O(n \log(n))$  comparisons.

**Q:** Can Randomized Quick Select do better – what is the maximum number of comparisons required by Randomized Quick Select in the worst-case?

In the worst case, Randomized Quick Select is worse than the naive strategy of sorting and returning the  $k^{\text{th}}$  element. What about the average (over the pivot selection) case?

**Claim:** For any array  $A$  with  $n$  distinct elements, and for any  $k \in [n]$ , Randomized Quick Select performs fewer than  $8n$  comparisons in expectation.

In order to prove this claim, we will need to prove the following lemma.

## Randomized Quick Select – Analysis

**Lemma:** The child sub-problem's array (either Left or Right) after the partitioning (in Line 4 of the algorithm) has expected size smaller than  $\frac{7n}{8}$ .

*Proof:* Let us define a “good” event  $\mathcal{E}$  that the randomly chosen pivot splits the array roughly in half. Formally, if  $n$  is the length of the array, then  $\mathcal{E}$  is the event that  $r \in (\frac{n}{4}, \frac{3n}{4}]$  (for simplicity, let us assume that  $n$  is divisible by 4.) Since  $r$  is chosen randomly,  $\Pr[\mathcal{E}] = \frac{3n/4 - n/4}{n} = \frac{1}{2}$ .

Recall that  $|\text{Left}| = r - 1$  and  $|\text{Right}| = n - r$ . Hence if event  $\mathcal{E}$  happens, then  $|\text{Left}| < \frac{3n}{4}$  and  $|\text{Right}| < \frac{3n}{4}$ . Hence,  $|\text{Child}| < \frac{3n}{4}$ . If event  $\mathcal{E}$  does not happen, in the worst-case,  $|\text{Child}| < n$ . By using the law of total expectation,

$$\begin{aligned}\mathbb{E}[|\text{Child}|] &= \mathbb{E}[|\text{Child}| \mid \mathcal{E}] \Pr[\mathcal{E}] + \mathbb{E}[|\text{Child}| \mid \mathcal{E}^c] \Pr[\mathcal{E}^c] \\ &< \frac{3n}{4} \frac{1}{2} + (n) \frac{1}{2} = \frac{7n}{8}.\end{aligned}$$

Hence on average, the size of the child sub-problem is smaller than  $\frac{7n}{8}$ , proving the lemma.

## Randomized Quick Select – Analysis

In order to upper-bound the total number of comparisons, we use the Lemma with an induction on  $n$ . Recall that we need to prove that Randomized Quick Select requires fewer than  $8n$  comparisons in expectation.

**Base case:** If  $n = 1$ , then we require  $0 < 8$  comparisons. Hence the base case is satisfied.

**Inductive Step:** Assume that for  $m < n$ ,

$\mathbb{E}[\text{Total number of comparisons for size } m \text{ array}] < 8m$ .

$$\begin{aligned} & \mathbb{E}[\text{Total number of comparisons for size } n \text{ array}] \\ &= \mathbb{E}[(n - 1) + \text{Total number of comparisons in child sub-problem}] \\ &= (n - 1) + \mathbb{E}[\text{Total number of comparisons in child sub-problem}] \quad (\text{Linearity of expectation}) \\ &< (n - 1) + 8 \mathbb{E}[|\text{Child}|] \quad (\text{Induction hypothesis}) \\ &< (n - 1) + 8 \frac{7n}{8} < 8n. \quad (\text{Lemma}) \end{aligned}$$

Hence, for any  $k \in [n]$ , on average, Randomized Quick Select requires fewer than  $8n$  comparisons, even though it might require  $O(n^2)$  comparisons in the worst-case.

Questions?