Lecture 13a: Probabilistic Method
Lale Özkahya

Resources:
Kenneth Rosen, “Discrete Mathematics and App.”
http://www.math.caltech.edu/ 2015-16/2term/ma006b
Independent Events

- Two events $A, B \subset \Omega$ are independent if
  \[ \Pr[A \cap B] = \Pr[a] \cdot \Pr[b]. \]

- Example. We flip two fair coins.
  - Let $\omega_{i,j}$ be the elementary event that coin $A$ landed on $i$ and coin $B$ on $j$, where $i, j \in \{h, t\}$. Each of the four events has a probability of 0.25.
  - The event where coin $A$ lands on heads is $a = \{\omega_{h,t}, \omega_{h,h}\}$. For $B$ it is $b = \{\omega_{t,h}, \omega_{h,h}\}$.
  - The events are independent since \( \Pr[a \text{ and } b] = \Pr[\omega_{h,h}] = 0.25 = \Pr[a] \cdot \Pr[b]. \)

(Discrete) Uniform Distribution

- In a uniform distribution we have a set $\Omega$ of elementary events, each occurring with probability \( \frac{1}{|\Omega|} \).
  - For example, when flipping a fair die, we have a uniform distribution over the six possible results.
**Union Bound**

- For any two events $A, B$, we have
  \[\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B].\]
- This immediately implies that
  \[\Pr[A \cup B] \leq \Pr[A] + \Pr[B],\]
  where equality holds iff $A, B$ are **disjoint**.
- **Union bound.** For any finite set of events $A_1, ..., A_k$, we have
  \[\Pr[\bigcup_i A_i] \leq \sum_i \Pr[A_i].\]

**Recall: Ramsey Numbers**

- $R(p, p)$ is the smallest number $n$ such that each blue-red edge coloring of $K_n$ contains a **monochromatic** $K_p$.
- **Theorem.** $R(p, p) > 2^{p/2}$.
  - We proved this in a previous class.
  - Now we provide another proof, using probability.
Probabilistic Proof

• For some $n$, we color the edges of $K_n$.
  ◦ Each edge is independently and uniformly colored either red or blue.
  ◦ For any fixed set $S$ of $p$ vertices, the probability that it forms a monochromatic $K_p$ is $2^{1-(\binom{p}{2})}$.
  ◦ There are $\binom{n}{p}$ possible sets of $p$ vertices. By the union bound, the probability that there is a monochromatic $K_p$ is at most
    \[
    \sum_{S} 2^{1-(\binom{p}{2})} = \binom{n}{p} 2^{1-(\binom{p}{2})}.
    \]

Proof (cont.)

• For some $n$, we color the edges of $K_n$.
  ◦ Each edge is colored blue with probability of 0.5, and otherwise red.
  ◦ The probability for a monochromatic $K_p$ is
    \[
    \leq \binom{n}{p} 2^{1-(\binom{p}{2})}.
    \]
  ◦ If $n \leq 2^{p/2}$, this probability is smaller than 1.
  ◦ In this case, the probability that we do not have any monochromatic $K_p$ is positive, so there exists a coloring of $K_n$ with no such $K_p$. 
Non-Constructive Proofs

- We proved that there exists a coloring of $K_n$ with no monochromatic $K_p$, but we have no idea how to find this coloring.
- Such a proof is called non-constructive.
- The probabilistic method often proves the existence of objects with surprising properties, but we still have no idea how they look like.

A Tournament

- We have $n$ people competing in thumb wrestling.
  - Every pair of contestants compete once.
  - How can we decide who the overall winner is?
- We build a directed graph:
  - A vertex for every participant.
  - An edge between every two vertices, directed from the winner to the loser.
  - An orientation of $K_n$ is called a tournament.
The King of the Tournament

- The winner can be the vertex with the maximum outdegree (the contestant winning the largest number of matches), but it might not be unique.
- A king is a contestant \( x \) such that for every other contestant \( y \) either \( x \rightarrow y \) or there exists \( z \) such that \( x \rightarrow z \rightarrow y \).
- **Theorem.** Every tournament has a king.

Proof

- \( D^+(v) \) – the number of vertices reachable from \( v \) by a path of length \( \leq 2 \).
- Let \( v \) be a vertex that maximizes \( D^+(v) \).
  - Assume for contradiction \( v \) is not a king.
  - Then there exists \( u \) such that \( u \rightarrow v \) and there is no path of length two from \( v \) to \( u \).
  - That is, for every \( w \) such that \( v \rightarrow w \), we also have \( u \rightarrow w \).
  - But this implies that \( D^+(u) \geq D^+(v) + 1 \), contradicting the maximality of \( v \)!
The $S_k$ Property

- We say that a tournament $T$ has the $S_k$ property if for every subset $S$ of $k$ participants, there exists a participant that won against everyone in $S$.
  - Formally, this is an orientation of $K_n$, such that for every subset $S$ of $k$ vertices there exists a vertex $v \in V \setminus S$ with an edge from $v$ to every vertex of $S$.

- Example. A tournament with the $S_1$ property.

Tournaments with the $S_k$ Property

- Theorem. If \( \binom{n}{k} (1 - 2^{-k})^{n-k} < 1 \) then there is a tournament on $n$ vertices with the $S_k$ property.

- Proof.
  - For some $n$ satisfying the above, we randomly orient $K_n = (V, E)$, such that the orientation of every $e \in E$ is chosen uniformly.
  - Consider a subset $S \subset V$ of $k$ vertices. The probability that a given vertex $v \in V \setminus S$ does not beat all of $S$ is $1 - 2^{-k}$. 
Proof (cont.)

- Consider a subset $S \subset V$ of $k$ vertices. The probability that a specific vertex $v \in V \setminus S$ does not beat all of $S$ is $1 - 2^{-k}$.
- $A_S$ – the event of $S$ not being beat by any vertex of $V \setminus S$.
- We have $\Pr[A_S] = \left(1 - 2^{-k}\right)^{n-k}$, since we ask for $n - k$ independent events to hold.
- By the union bound, we have
  \[
  \Pr\left[\bigvee_{S \subset V, \ |S| = k} A_S\right] \leq \sum_{S \subset V, \ |S| = k} \Pr[A_S] = \binom{n}{k} \left(1 - 2^{-k}\right)^{n-k} < 1.
  \]

Completing the Proof

- $A_S$ – the event of $S$ not being beat by any vertex of $V \setminus S$.
- we have
  \[
  \Pr\left[\bigvee_{S \subset V, \ |S| = k} A_S\right] < 1.
  \]
- That is, there is a positive probability that every subset $S \subset V$ of size $k$ is beat by some vertex of $V \setminus S$. So such a tournament exists.
Which NBA Player is Related to Mathematics?

Michael Jordan  Shaquille O'Neal  LeBron James

Random Variables

- A random variable is a function from the set of possible events to \( \mathbb{R} \).
- **Example.** Say that we flip five coins.
  - We can define the random variable \( X \) to be the number of coins that landed on heads.
  - We can define the random variable \( Y \) to be the percentage of heads in the tosses.
  - Notice that \( Y = 20X \).
Indicator Random Variables

- An **indicator random variable** is a random variable $X$ that is either 0 or 1, according to whether some event happens or not.

**Example.** We toss a fair die.
- We can define the six indicator variable $X_1, ..., X_6$ such that $X_i = 1$ iff the result of the roll is $i$.

Expectation

- The **expectation** of a random variable $X$ is

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \Pr[\omega].$$

  - Intuitively, $E[X]$ is the expected value of $X$ in the long-run average value when repeating the experiment $X$ represents.
Expectation Example

- We roll a fair six-sided die.
  - Let $X$ be a random variable that represents the outcome of the roll.

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \Pr[\omega] = \sum_{i \in \{1, \ldots, 6\}} i \cdot \frac{1}{6} = 3.5$$

While a prisoner of war during World War II, J. Kerrich conducted an experiment in which he flipped a coin 10,000 times and kept a record of the outcomes. A portion of the results is given in the table below.

<table>
<thead>
<tr>
<th>Number of Tosses</th>
<th>Number of Heads</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>100</td>
<td>44</td>
</tr>
<tr>
<td>500</td>
<td>255</td>
</tr>
<tr>
<td>1,000</td>
<td>502</td>
</tr>
<tr>
<td>5,000</td>
<td>2,533</td>
</tr>
<tr>
<td>10,000</td>
<td>5,067</td>
</tr>
</tbody>
</table>
Linearity of Expectation

- If $X$ is a random variable, then $5X$ is a random variable with a value five times that of $X$.

- **Lemma.** Let $X_1, X_2, \ldots, X_k$ be a collection set of random variables over the same discrete probability. Let $c_1, \ldots, c_k$ be constants. Then

$$E[c_1X_1 + c_2X_2 + \cdots + c_kX_k] = \sum_{i=1}^{k} c_iE[X_i].$$

Fixed Elements in Permutations

- Let $\sigma$ be a uniformly chosen permutation of $\{1, 2, \ldots, n\}$.
  - For $1 \leq i \leq n$, let $X_i$ be an indicator variable that is 1 if $i$ is fixed by $\sigma$.
  - $E[X_i] = \Pr[\sigma(i) = i] = \frac{(n-1)!}{n!} = \frac{1}{n}$.
  - Let $X$ be the number of fixed elements in $\sigma$.
  - We have $X = X_1 + \cdots + X_n$.
  - By linearity of expectation

$$E[X] = \sum_i E[X_i] = n \cdot \frac{1}{n} = 1.$$
Hamiltonian Paths

- Given a directed graph $G = (V, E)$, a **Hamiltonian path** is a path that visits every vertex of $V$ exactly once.
  - Major problem in theoretical computer science: Does there exist a polynomial-time algorithm for finding whether a Hamiltonian path exists in a given graph.

Hamiltonian Paths in Tournaments

- **Theorem.** There exists a tournament $T$ with $n$ players that contains at least $n! \cdot 2^{-(n-1)}$ Hamiltonian paths.
Proof

- We uniformly choose an orientation of the edges of $K_n$ to obtain a tournament $T$.
  - There is a bijection between the possible Hamiltonian paths and the permutations of $\{1, 2, \ldots, n\}$. Every possible path defines a unique permutation, according to the order in which it visits the vertices.
  - For a permutation $\sigma$, let $X_\sigma$ be an indicator variable that is 1 if the path corresponding to $\sigma$ exists in $T$.
  - We have $E[X_\sigma] = \Pr[X_\sigma = 1] = 2^{-n+1}$.

- We uniformly choose an orientation of the edges of $K_n$ to obtain a tournament $T$.
  - For a permutation $\sigma$, let $X_\sigma$ be an indicator variable that is 1 if the path corresponding to $\sigma$ exists in $T$.
  - We have $E[X_\sigma] = \Pr[X_\sigma = 1] = 2^{-n+1}$.
  - Let $X$ be a random variable of the number of Hamiltonian paths in $T$. Then $X = \sum_\sigma X_\sigma$.
    \[
    E[X] = E\left[\sum_\sigma X_\sigma\right] = \sum_\sigma E[X_\sigma] = n! \cdot 2^{-n+1}.
    \]
  - Since this is the expected number of paths in a uniformly chosen tournament, there must be at least as many paths.
Reminder: Indicator Random Variables

• An **indicator random variable** is a random variable $X$ that is either 0 or 1, according to whether some event happens or not.

• **Example.** We toss a die.
  - We can define the six indicator variable $X_1, \ldots, X_6$ such that $X_i = 1$ iff the result of the roll is $i$.

Reminder: Expectation

• The **expectation** of a random variable $X$ is

$$E[X] = \sum_{\omega \in \Omega} X(\omega) \Pr[\omega].$$

- Intuitively, $E[X]$ is the expected value of $X$ in the long-run average value of repetitions of the experiment it represents.
Reminder: Expectation Example

• We roll a fair six-sided die.
  ◦ Let $X$ be a random variable that represents the outcome of the roll.
  
  $E[X] = \sum_{\omega \in \Omega} X(\omega) \Pr[\omega] = \sum_{i \in \{1, \ldots, 6\}} i \cdot \frac{1}{6} = 3.5$

Reminder: Linearity of Expectation

• If $X$ is a random variable, then $5X$ is a random variable with a value five times that of $X$
• Lemma. Let $X_1, X_2, \ldots, X_k$ be a collection set of random variables over the same discrete probability. Let $c_1, \ldots, c_k$ be constants. Then
  
  $E[c_1X_1 + c_2X_2 + \cdots + c_kX_k] = \sum_{i=1}^{k} c_i E[X_i]$. 
Independent Sets

• Consider a graph $G = (V, E)$. An independent set in $G$ is a subset $V' \subseteq V$ such that there is no edge between any two vertices of $V'$.

• Finding a maximum independent set in a graph is a major problem in theoretical computer science.
  ◦ No polynomial-time algorithm is known.

Warm Up

• What are the sizes of the maximum independent sets in:

  2

  4
Large Independent Sets

- **Theorem.** A graph $G = (V, E)$ has an independent set of size at least
  \[ \sum_{v \in V} \frac{1}{1 + \deg v}. \]

- **Proof.** We uniformly choose an ordering for the vertices of $V = \{v_1, \ldots, v_n\}$.
  - The set of vertices that appear before all of their neighbors is an independent set.

We uniformly choose an ordering for the vertices of $V = \{v_1, \ldots, v_n\}$.
- The set $S$ of vertices that appear before all of their neighbors is an independent set.
- $X_i$ - indicator that is 1 if $v_i \in S$.
  \[ E[X_i] = \Pr[X_i = 1] = \frac{1}{1 + \deg v_i}. \]
- $X$ – the random variable of the size of $S$.
  \[ E[X] = E \left[ \sum_{i=1}^{n} X_i \right] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{1}{1 + \deg v_i}. \]
- There must exist an ordering for which $|S|$ has at least this value.
Sum-free Sets

- Consider a set $A$ of positive integers. We say that $A$ is sum-free if for every $x, y \in A$, we have that $x + y \not\in A$ (including the case where $x = y$).

- What are large sum-free subsets of $S = \{1, 2, \ldots, N\}$?
  - We can take all of the odd numbers in $S$.
  - We can take all of the numbers in $S$ of size larger than $N/2$.

Large Sum-Free Sets Always Exist

- **Theorem.** For any set of positive integers $A$, there is a sum-free subset $B \subseteq A$ of size $|B| \geq \frac{1}{3} |A|$.

- **Proof.** Consider a prime $p$ such that $p > a$ for every $a \in A$.
  - From now on, calculations are mod $p$.
  - Notice that if $B$ is sum-free mod $p$, it is also sum-free under standard addition.
  - Thus, it suffices to find a large set that is sum-free mod $p$. 
Proof (cont.)

- The calculations are \( \text{mod } p \).
- The set \( S = \{[p/3], ..., [2p/3]\} \) is sum-free and \( |S| \geq (p - 1)/3 \).
- We uniformly choose \( x \in \{1, 2, ..., p - 1\} \) and set \( A_x = \{a \in A \mid ax \in S\} \).
- Consider \( b, c \in A_x \). Since \( bx, cx \in S \), we have that \((b + c)x = bx + cx \notin S \). Thus, \((b + c) \notin A_x \), and \( A_x \) is sum-free.
- \( X_a \) – indicator that is 1 if \( a \in A_x \).

\[
E[|A_x|] = E \left[ \sum_{a \in A} X_a \right] = \sum_{a \in A} E[X_a] = \sum_{a \in A} \Pr[X_a = 1].
\]

- The set \( S = \{[p/3], ..., [2p/3]\} \) is sum-free and \( |S| \geq (p - 1)/3 \).
- We uniformly choose \( x \in \{1, 2, ..., p - 1\} \) and set \( A_x = \{a \in A \mid ax \in S\} \). \( A_x \) is sum-free.
- \( X_a \) – indicator that is 1 if \( a \in A_x \).
- \( E[|A_x|] = \sum_{a \in A} \Pr[X_a = 1] \).
- Recall: If \( x \equiv x' \text{ mod } p \) then \( ax \equiv ax' \text{ mod } p \).
- We thus have \( \Pr[X_a = 1] = |S|/(p - 1) \).

\[
E[|A_x|] = \sum_{a \in A} \Pr[X_a = 1] = \frac{|A||S|}{p - 1} \geq |A| \frac{1}{3}.
\]
- Thus, there exists an \( x \) for which \( |A_x| \geq |A| \frac{1}{3} \).
Which Super Villain is a Mathematician?

Austin Powers’
Dr. Evil

Spiderman’s
Dr. Octopus

Sherlock Holmes’
Professor Moriarty

More Ramsey Numbers

• In the previous class, we used a basic probabilistic argument to prove
  \( R(p, p) > 2^{p/2} \).

• Theorem. For any integer \( n > 0 \), we have
  \[ R(p, p) > n - \binom{n}{p} 2^{1 - \frac{p}{2}}. \]
Proof

• Consider a random red-blue coloring of $K_n = (V, E)$. The color of each edge is chosen uniformly and independently.

• For every subset $S \subset V$ of size $p$, we denote by $X_S$ the indicator that $S$ induces a monochromatic $K_p$. We set $X = \sum_{|S|=p} X_S$.

• $E[X_S] = \Pr[X_S = 1] = 2^{1-\binom{p}{2}}$

• By linearity of expectation, we have

$$E[X] = \sum_{|S|=p} E[X_S] = \binom{n}{p} 2^{1-\binom{p}{2}}.$$

Completing the Proof

• We proved that the in a random red-blue coloring of $K_n$ the expected number of monochromatic copies of $K_p$ is

$$m = \binom{n}{p} 2^{1-\binom{p}{2}}.$$

• There exist a coloring with at most $m$ monochromatic $K_p$’s.

• By removing a vertex from each of these copies, we obtain a coloring of $K_{n-m}$ with no monochromatic $K_p$. 
Recap

- How we used the probabilistic method:
  - Our first applications were simply about making random choices and showing that we obtain some **property with non-zero probability**.
  - We moved to more involved proofs, where we use **linearity of expectation** to talk about the “expected” result.
  - In the previous proof, we used a **two step method** – first we randomly choose an object, and then we alter it. This method is called **the alternation method**.

Transmission Towers

- **Problem.** A company wants to establish **transmission towers** in its large compound.
  - Each tower must be on top of a building and each building must be covered by at least one tower.
  - We are given the pairs of buildings such that a tower on one covers the other.
  - We wish to minimize the number of towers.
Building a graph

- We build a graph $G = (V, E)$.
  - A vertex for every building.
  - An edge between every pair of buildings that can cover each other.
  - We need to find the minimum subset of vertices $V' \subseteq V$ such that every vertex of $V$ has at least one vertex of $V'$ as a neighbor.

Dominating Sets

- Consider a graph $G = (V, E)$. A dominating set of $G$ is a subset $V' \subseteq V$ such that every vertex of $V$ has at least one neighbor in $V'$.
- It is not known whether there exists a polynomial-time algorithm for finding a minimum dominating set in a graph.
Warm Up

- **Problem.** Let \( G = (V, E) \) be a graph with maximum degree \( k \). Give a lower bound for the size of any dominating set of \( G \).

- **Answer.**
  - Every vertex covers itself and at most \( m \) other vertices, so any dominating set is of size at least
    \[
    |V| / (k + 1).
    \]

The Case of a Minimum Degree

- **Theorem.** Let \( G = (V, E) \) be a graph with minimum degree \( k \). Then there exists a dominating set of size at most \( n \cdot \frac{1 + \lg k}{k + 1} \).

- **Proof.** We consider a random subset \( S \subset V \) by independently taking each vertex of \( V \) with probability \( p = \frac{\lg k}{k + 1} \).

  - Let \( T \subset V \setminus S \) be the vertices that have no neighbors in \( S \).
    - \( S \cup T \) is a dominating set.
Proof (cont.)

- $S \subset V$ – a random subset formed by independently taking each vertex of $V$ with probability $p = \frac{\lg (k+1)}{k+1}$.
- $T \subset V \setminus S$ – the vertices with no neighbors in $S$.

$S \cup T$ is a dominating set.

- $E[|S|] = \sum_{v \in V} \Pr[v \in S] = \sum_v p = p|V|$.
- A vertex is in $T$ if it is not in $S$ and none of its neighbors are in $S$. The probability for this is at most $(1-p)^{k+1}$.

- $E[|T|] = \sum_{v \in V} \Pr[v \in T] \leq \sum_v (1-p)^{k+1} = (1-p)^{k+1}|V|$.

Completing the Proof

- $p = \frac{\lg (k+1)}{k+1}$.
- Famous inequality. $1 - p \leq e^{-p}$ for any positive $p$.
- Thus, $(1 - p)^{k+1} \leq e^{-p(k+1)} = \frac{1}{k+1}$.
- We proved
  \[
  E[|S| + |T|] = E[|S|] + E[|T|] 
  \leq (p + (1-p)^{k+1})|V| = \frac{\lg (k+1) + 1}{k+1}|V|.
  \]
- There must exist a dominating set of size.
How to Choose the Probability?

- In the previous problem, we knew to choose \( p = \frac{\log(k+1)}{k+1} \). But how?
  - When you solve a question and the choice is not uniform, first mark the probability as \( p \).
  - At the end of the analysis you will obtain some expression with \( p \) in it. Choose the value of \( p \) that optimizes the expression.

The End: Professor Moriarty

- Professor Moriarty is a mathematician.
  - “At the age of twenty-one he wrote a treatise upon the binomial theorem”.
  - So a combinatorist?!
  - Dr. Octopus is a nuclear physicist.
  - Dr. Evil is a medical doctor.