Combinatorial Mathematics

Mong-Jen Kao (高孟駿)

Monday 18:30 – 21:20

Outline

- Probabilistic Counting – The Framework
  - Ex1. Tournaments
  - Ex2. Universal Sets
  - Ex4. 2-Colorable Families
  - Ex3. Covering by Bipartite Cliques
  - Some Useful Tools & Bounds

Probabilistic Method

– The Framework (in this lecture)

To prove that an object with certain properties exists.

2-Coloring for Set Families

- Let $\mathcal{F}$ be a family of subsets for some finite ground set N, and let

$$g : N \longrightarrow \{R, B\}$$

  be a coloring of the elements in N into *red* or *blue*.

  - A set $A \in \mathcal{F}$ is *monochromatic*, if $g(x) = g(y)$ for all $x, y \in A$, i.e., all the elements in A are colored the same.

  - g is said to be a *valid 2-coloring* for F, if ***none** of the sets* in $\mathcal{F}$ is *monochromatic*.

(Scenario 1) Proving that an ***Object of Interest*** Exists

- Suppose that A is a set of objects we are interested in.
- To prove that $A \neq \emptyset$, i.e., there exists an $x \in A$,
  - One way is to define a probability distribution over some $B \supseteq A$ and show that

$$\Pr_{x \leftarrow B}[x \in A] > 0,$$

i.e., if we sample an element x from B,
then with nonzero probability, the element x is in A.

(Scenario 2) Proving that ***a Good Object*** Exists

- Suppose that A is a set of objects we are interested in and $f : A \longrightarrow \mathbb{R}$ is a weight function of the objects in A.
- To prove that there exists $x \in A$ with $f(x) \geq t$ for some given t,
  - One way is to define a probability distribution over A and show that

$$E_A[f] \geq t,$$

i.e., the expectation of f is at least t.

Ex1. Tournaments

Ex1. Tournaments

It has no self-loop.

- A tournament is a directed graph $G = (V, E)$ such that
  - $(v, v) \notin E$, for all $v \in V$, and
  - For any $u, v \in V$,
    exactly one of $(u, v) \in E$ or $(v, u) \in E$ holds.

There is exactly one directed edge between every pair of vertices.

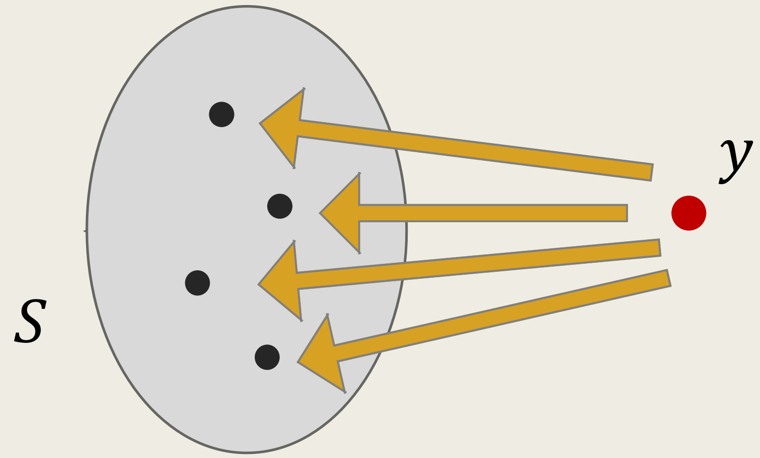
- Intuitively, a tournament graph represents the result of the match between all pair of players.

Ex1. Tournaments

- We say that a tournament $G = (V, E)$ has the ***property*** $\boldsymbol{P_k}$,

  if *for every subset* $S \subseteq V$ *of* k *players*, there exists a player $y \notin S$ that beats all the players in S, i.e., $(y, v) \in E$ for all $v \in S$.

  – P_k implies P_ℓ for all $\ell \leq k$.



What does this mean?

Theorem 1 (Erdös 1963a).

For any $k \geq 2$, if $n \geq k^2 \cdot 2^{k+1}$,

then there is a tournament of n players that has the property P_k.

- Consider a *random tournament* of the n *players*, where

  the *direction of the edges* are *determined by a fair coin*.

- For any subset S of k players,
  let A_S denote the event that there exists no $y \notin S$ that beats all $v \in S$.

- For any subset S of k players,
  let A_S denote the event that there exists no $y \notin S$ that beats all $v \in S$.

  - For any $y \notin S$,

    $$\Pr[\ y \text{ beats all of } v \in S\ ] = 2^{-k}\ .$$

    $$\Pr[\ y \text{ does not beat all of } v \in S\ ] = 1 - 2^{-k}\ .$$

  - There are $n - k$ other vertices that can beat all $v \in S$.
    Hence

    $$\Pr[A_S] = \left(1 - 2^{-k}\right)^{n-k}\ .$$

- For any subset S of k players,
  let A_S denote the event that there exists no $y \notin S$ that beats all $v \in S$.
  - $\Pr[A_S] = \left(1 - 2^{-k}\right)^{n-k}$.

- By the union bound,

$$\Pr[\text{ Some } S \text{ is not dominated by some player }]$$

$$= \Pr[\cup A_S] \le \sum_{S,|S|=k} \Pr[A_S] = \binom{n}{k} \cdot \left(1 - 2^{-k}\right)^{n-k}$$

$$< \frac{n^k}{k!} \cdot e^{-\frac{n-k}{2^k}} \le n^k \cdot e^{-\frac{n}{2^k}},$$

which is less than 1 when $n \ge k^2 \cdot 2^{k+1}$.

Refer to the jamboard for details.

- For any subset S of k players,
  let A_S denote the event that there exists no $y \notin S$ that beats all $v \in S$.
  - $\Pr[A_S] = \left(1 - 2^{-k}\right)^{n-k}$.

- By the union bound,

$$\Pr[\text{ Some } S \text{ is not dominated by some player }] < 1$$

  when $n \geq k^2 \cdot 2^{k+1}$.

- So, when $n \geq k^2 \cdot 2^{k+1}$,

$$\Pr[\text{ All } S \text{ is dominated by some player}] > 0 .$$

Ex2. Universal Sets

Ex 2. Universal Sets

- Let a be a 0-1 string of length n.
  - For any subset $S = \{i_1, i_2, \dots, i_k\}$ *of* k coordinates,

    define the ***projection of a onto S*** to be

$$a\Big|_S := \left(a_{i_1}, a_{i_2}, \dots, a_{i_k}\right),$$

    i.e., the *substring* formed by the coordinates specified in S.

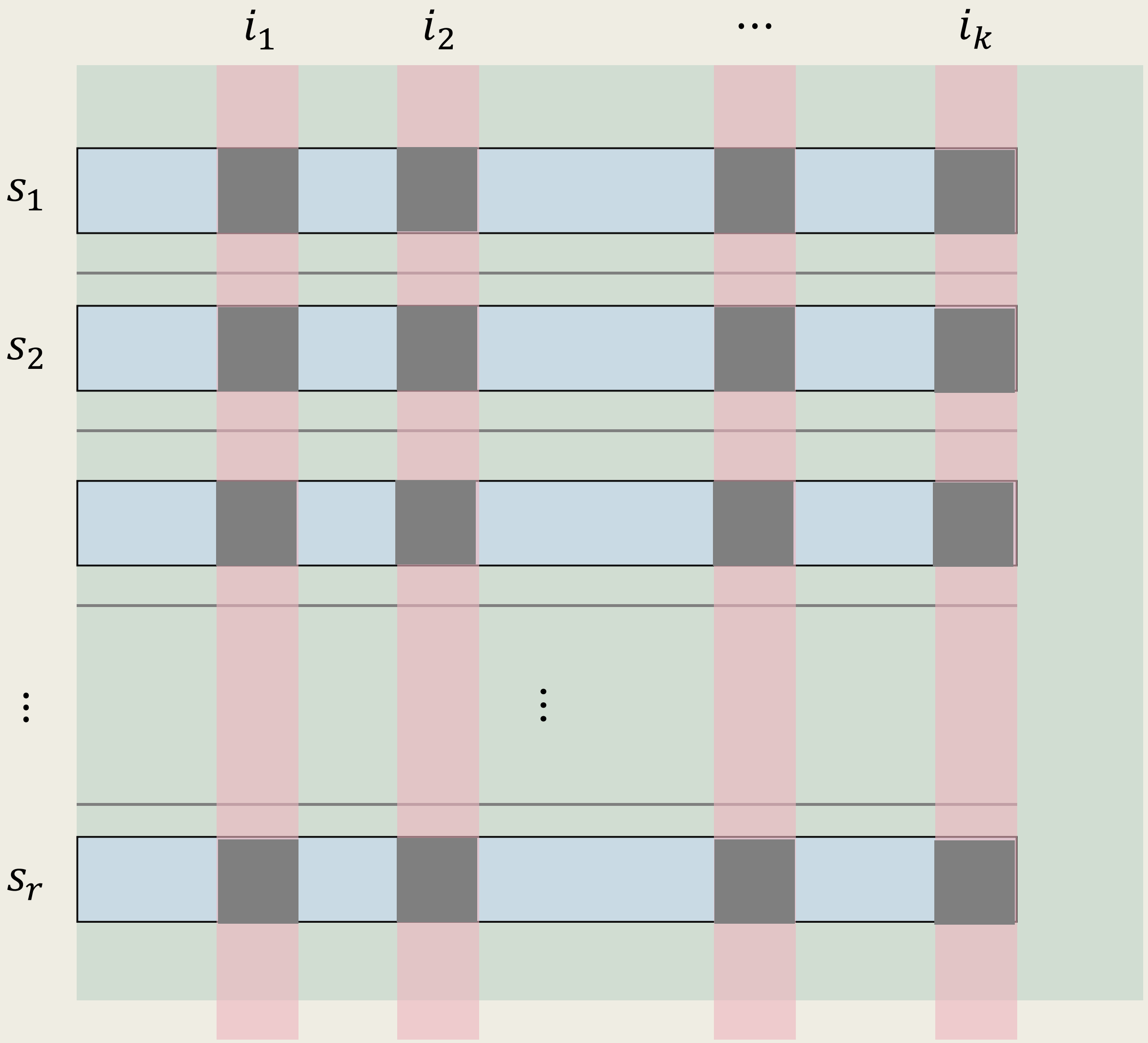
Ex 2. Universal Sets

- Let A be a set of 0-1 strings of length n.

- We say that $\boldsymbol{A}$ is $(\boldsymbol{n},\boldsymbol{k})$***-universal***, if for any subset $S=\{i_1,i_2,\ldots,i_k\}$ of k coordinates, the projection of A onto S,

$$A\Big|_S \coloneqq \left\{\ a\Big|_S\ :\ a\in A\ \right\}$$

always contains all possible 2^k combinations.



For an arbitrary choice of

k coordinates $i_1, i_2, \ldots, i_k$,

the projection of the strings onto the k coordinates contains all 2^k possible strings.

Ex 2. Universal Sets

$$\frac{0\ 0\ 0\ 0}{1\ 1\ 1\ 1}$$

$$\frac{0\ 1\ 0\ 1}{1\ 0\ 1\ 0}$$

are both $(4,1)$-universal.

Ex 2. Universal Sets

0 0 1
0 1 0
1 0 0
1 1 1

is $(3,2)$-universal, but

0 0 1
0 1 0
1 0 1
1 1 0

is not.

We are interested in knowing,
how many strings does it suffice to be universal.

Ex 2. Universal Sets

- When the entries of the strings are *determined randomly*,
  we can write down the probability that
  the generated strings are not (n, k)-universal.
  - By requiring the probability to be < 1, we get a simple bound.

Theorem 2 (Kleitman-Spencer 1973).

If $\binom{n}{k} \cdot 2^k \cdot \left(1 - 2^{-k}\right)^r < 1$,

then there is an (n, k)-universal set of size r.

- Let A be a set of r random 0-1 strings of length n, where each entry takes values 0 or 1 independently with probability 1/2.

- Fix a set S of k coordinates.

  For any vector $v \in \{0,1\}^k$,

$$\Pr\left[v \notin A\Big|_S\right] = \prod_{a \in A} \Pr\left[v \neq a\Big|_S\right] = \prod_{a \in A} \left(1 - 2^{-k}\right) = \left(1 - 2^{-k}\right)^r .$$

- Fix a set S of k coordinates. For any vector $v \in \{0,1\}^k$,

$$\Pr\left[v \notin A\Big|_S\right] = \prod_{a \in A} \Pr\left[v \neq a\Big|_S\right] = \prod_{a \in A} \left(1 - 2^{-k}\right) = \left(1 - 2^{-k}\right)^r .$$

- There are $\binom{n}{k} \cdot 2^k$ ways to choose such a pair (S, v).

  By union bound, the probability that A is not (n, k)-universal

  is at most

$$\sum_{S,v} \Pr\left[v \notin A\Big|_S\right] = \binom{n}{k} \cdot 2^k \cdot \left(1 - 2^{-k}\right)^r$$

  - When $\binom{n}{k} \cdot 2^k \cdot \left(1 - 2^{-k}\right)^r < 1$, $\Pr[\, A \text{ is } (n, k)\text{−universal} \,] > 0$.

2-Colorable Families

2-Colorable Families

- Let $\mathcal{F}$ be a family of subsets for some finite ground set N, and let

$$g : N \longrightarrow \{R, B\}$$

  be a coloring of the elements in N into *red* or *blue*.

  – A set $A \in \mathcal{F}$ is *monochromatic*, if $g(x) = g(y)$ for all $x, y \in A$, i.e., all the elements in A are colored the same.

  – g is said to be a *valid 2-coloring* for F, if ***none** of the sets* in $\mathcal{F}$ is *monochromatic*.

- A set family $\mathcal{F}$ is k-uniform if $|A| = k$ for all $A \in \mathcal{F}$.

Theorem 4 (Erdös 1963b).

Every k-uniform family with fewer than 2^{k-1} members (subsets) is 2-colorable.

- Suppose that we color the elements independent with a fair 0-1 coin.
  - For any $A \in F$, $\Pr[\, A \text{ is monochromatic} \,] = 2 \cdot 2^{-k} = 2^{1-k}$.
  - When $|F| < 2^{k-1}$,

    the expected number of monochromatic sets is $|F| \cdot 2^{1-k} < 1$.

Theorem 4 (Erdös 1963b).

Every k-uniform family with fewer than 2^{k-1} members (subsets) is 2-colorable.

- Suppose that we color the elements independent with a fair 0-1 coin.
  - When $|F| < 2^{k-1}$, the expected number of monochromatic sets is $|F| \cdot 2^{1-k} < 1$.
  - There must be a coloring whose value is at most the expectation. Since the number of monochromatic sets is integral, it must be 0.

Theorem 5 (Erdös 1964a).

If k is sufficiently large, then there exists a k-uniform family F with $|F| \leq k^2 2^k$ that is not 2-colorable.

- Let $r = \lfloor k^2/2 \rfloor$ and $N = \{1,2,\ldots,r\}$ be the ground set to consider.
- Consider a ***random family*** $F = \{A_1, A_2, \ldots, A_b\}$ generated as follows.
  - Let A_i be a set picked uniformly and independently from all size-k subsets of N,

    i.e., for any $A \subseteq N$, $\Pr[A_i = A] = \binom{r}{k}^{-1}$.

Imagine that we do this
before generating the set family.

- ***Fix a coloring***, say, χ, on N that uses $\boldsymbol{a}$ **reds** and $\boldsymbol{r-a}$ **blues**.
  - For any $1 \le i \le b$,

$$\Pr[A_i \text{ is monochromatic}] = \Pr[A_i \text{ is red}] + \Pr[A_i \text{ is blue}]$$

$$= \frac{\binom{a}{k} + \binom{r-a}{k}}{\binom{r}{k}} \ge 2 \cdot \frac{\binom{r/2}{k}}{\binom{r}{k}} := p\,.$$

$\binom{a}{k}$ ways to form a red set, each is chosen with probability $1/\binom{r}{k}$.

By Jensen's inequality

Refer to the jamboard for more details.

- Fix a coloring, say, χ, on N that uses a reds and $r-a$ blues.
  - For any $1 \le i \le b$,

    $\Pr[A_i \text{ is monochromatic}] = \Pr[A_i \text{ is red}] + \Pr[A_i \text{ is blue}]$

    $$= \frac{\binom{a}{k} + \binom{r-a}{k}}{\binom{r}{k}} \ge 2 \cdot \frac{\binom{r/2}{k}}{\binom{r}{k}} \coloneqq p\,.$$

  - By the asymptotic formula for binomial coefficient,

    $$p \approx e^{-1} \cdot 2^{1-k}\,.$$

  - Since A_i are independently chosen,

    $$\Pr[\,\chi \text{ is legal for } F\,] \le \prod_{1\le i\le b} (1-p) \le (1-p)^b\,.$$

Refer to the jamboard for more details.

- Since A_i are independently chosen,

$$\Pr[\,\chi \text{ is legal for } F\,] \;\le\; (1-p)^b\,.$$

- There are 2^r possible colorings on N.

  By the union bound,

$$\Pr[\text{ at least one coloring is legal for } F\,]$$

$$\le\; 2^r \cdot (1-p)^b \;<\; e^{r\cdot\log 2 - pb}\,,$$

  which is no more than 1 when

$$b \;=\; \frac{r\cdot\log 2}{p} \;=\; \big(1+o(1)\big)\cdot k^2 \cdot 2^{k-2}\cdot e\log 2 \;\le\; k^2\cdot 2^k\,.$$

- Since A_i are independently chosen,

$$\Pr[\, \chi \text{ is legal for } F\,] \le (1-p)^b .$$

- There are 2^r possible colorings on N.

By the union bound,

$$\Pr[\text{ at least one coloring is legal for } F\,] < e^{r \cdot \log 2 - pb} ,$$

which is no more than 1 when $b \le k^2 \cdot 2^k$.

- Hence, $\Pr[\text{ no coloring is legal for } F\,] > 0$ when $b \le k^2 \cdot 2^k$, and there must exist one set family that has no valid 2-coloring.

2-Colorability of Uniform Set Families

- Let $B(k)$ be the smallest size of k-uniform families that are ***not*** 2-colorable.
  - By Theorem 4 and Theorem 5, we know that

    $$2^{k-1} \leq B(k) \leq k^2 \cdot 2^k \,.$$

  - For the exact values,
    so far, only $B(2) = 3$ and $B(3) = 7$ are known.

Determine the exact value for $B(k)$ ---
A somewhat interesting question of unknown importance.

Theorem 6.

Let F be a set family, with $|A| \geq 2$ for all $A \in F$. If $A \cap B \neq \emptyset$ implies that $|A \cap B| \geq 2$ for any $A, B \in F$, then F is 2-colorable.

- The given condition is strong enough for a greedy algorithm to work.
  - Let $N = \{x_1, x_2, \ldots, x_n\}$ be the ground set.
  - The algorithm proceeds as follows.
    - For $i = 1,2, \ldots, n$, do
      - If coloring x_i red does not make any set monochromatic, then color x_i red.

        Otherwise, color x_i blue.

Theorem 6.

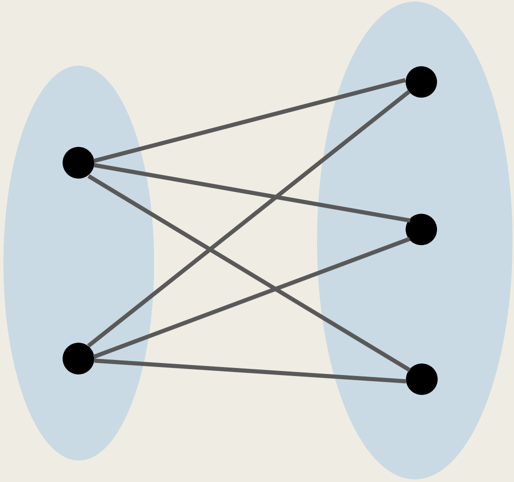
Let F be a set family, with $|A| \geq 2$ for all $A \in F$. If $A \cap B \neq \emptyset$ implies that $|A \cap B| \geq 2$ for any $A, B \in F$, then F is 2-colorable.

- For the correctness of the algorithm, observe the following.
  - If x_i cannot be colored red, then there exists some set $A \subseteq \{x_1, x_2, \dots, x_i\}$ with $x_i \in A$ and $\boldsymbol{A \setminus \{x_i\}}$ **are all red**.
  - If x_i cannot be colored blue, then there exists some $B \subseteq \{x_1, x_2, \dots, x_i\}$ with $x_i \in B$ and $\boldsymbol{B \setminus \{x_i\}}$ **are all blue**.
  - If both red and blue are not possible,
    then $x_i \in A \cap B \neq \emptyset$,
    which implies that $|A \cap B| \geq 2$, a contradiction.

Covering by Bipartite Cliques

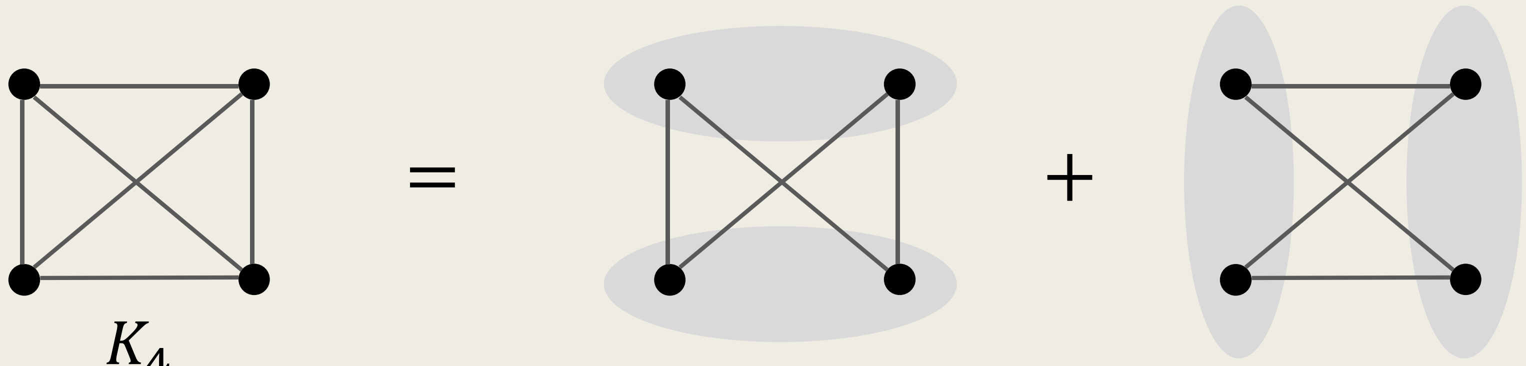
Bipartite Cliques

- A *bipartite clique*, or, *biclique*, is a ***complete bipartite graph***.
  - It is a bipartite graph.
  - There is an edge between every pair of vertices from the two partite sets.



Covering by Bipartite Cliques

- Let $G = (V, E)$ be a graph.
- A ***biclique covering*** of G is a set of subgraphs $H_1, H_2, \ldots, H_t$ of G such that
  - H_i is a bipartite clique, for all $1 \leq i \leq t$.
  - Each edge in E belongs to H_i for some $1 \leq i \leq t$.



K_4

Covering by Bipartite Cliques

- The ***weight*** of a biclique covering $H_1, H_2, \ldots, H_t$ is defined to be

$$\sum_{1 \le i \le t} |V(H_i)|,$$

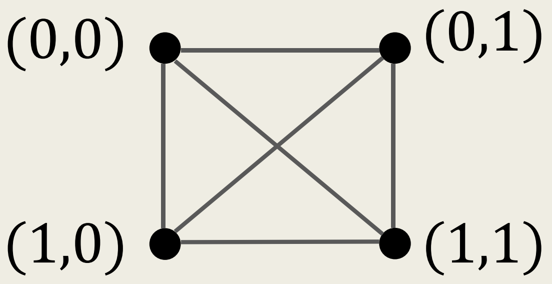
i.e., the *total number of vertices* used in the cover.

  – Let $\mathrm{bc}(G)$ denote the minimum weight of biclique coverings of G.

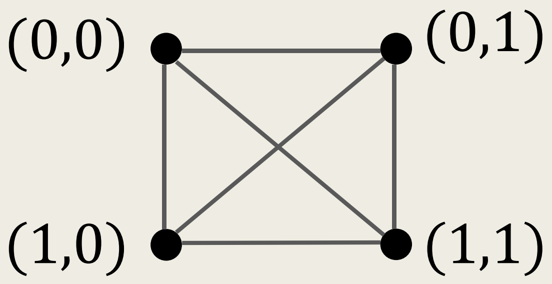
Theorem 3.

If $n = 2^m$, then $\mathrm{bc}(K_n) = n \cdot \log_2 n$.

- Let's prove the two directions "$\leq$" and "$\geq$" separately.
- For "$\leq$", we will construct a covering of weight $nm = n \cdot \log_2 n$.
  - This shows that, the minimum weight of K_n, $\mathrm{bc}(K_n)$, is ***at most*** $n \cdot \log_2 n$.
- Label the vertices K_n with a coordinate $\{0,1\}^m$.



- Label the vertices of K_n with a coordinate $\{0,1\}^m$.
- For any $1 \leq i \leq m$, define H_i as follows.
  - $V(H_i) = V(K_n)$.
  - For any $u, v \in V(K_n)$,
    $(u, v) \in E(H_i)$ if the i^{th}-coordinates of u and v differ.



- Then, each edge belongs to some H_i (why?),

  and the total weight is $nm = n \log_2 n$.

You will prove in HW#2 that H_i is a biclique.

Theorem 3.

If $n = 2^m$, then $\mathrm{bc}(K_n) = n \cdot \log_2 n$.

- To prove the other direction, i.e., $\mathrm{bc}(K_n) \geq n \cdot \log_2 n$, we use a probabilistic argument.
  - No matter how we organize the bicliques, the total weight is always at least $n \log_2 n$.

This is the harder part.

How can we prove a statement like this?

Is it because we're not smart enough to do this,
or there is no such way at all??

Derive ***properties***
for any biclique covering.

- To prove the other direction, i.e., $\mathrm{bc}(K_n) \geq n$ we use a probabilistic argument.

- Let $(A_i \times B_i)_{1 \leq i \leq t}$ be an arbitrary biclique covering for K_n, and let m_v be the number of bicliques that contains v.

By the double-counting principle on the total weight, we have

$$\sum_{1 \leq i \leq t} (|A_i| + |B_i|) = \sum_{1 \leq v \leq n} m_v \ .$$

It suffices to show that $\sum_{1 \leq v \leq n} m_v \geq n \cdot \log_2 n$.

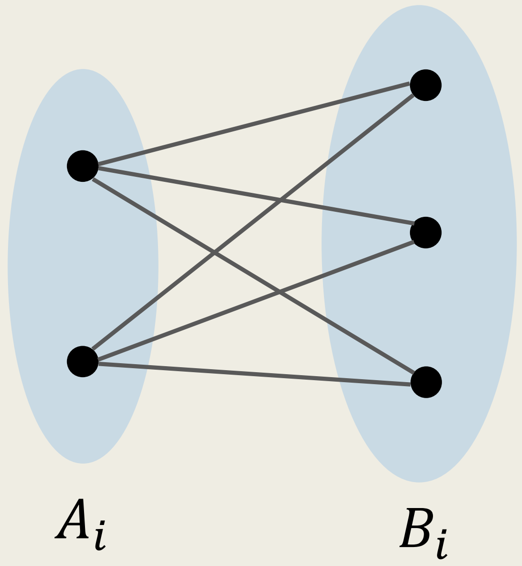
It suffices to show that $\sum_{1 \leq v \leq n} m_v \geq n \cdot \log_2 n$.

- Note that, this inequality to prove says that,

  the average number of bicliques that contain each vertex is at least $\log_2 n$.

It suffices to show that $\sum_{1 \le v \le n} m_v \ge n \cdot \log_2 n$.

- Toss a fair 0-1 coin for each biclique $A_i \times B_i$ in any order.
  - If 0 pops up, remove the *vertex set* A_i from K_n.
  - If 1 pops up, remove B_i from K_n.



Remove one of A_i, B_i from K_n.

Let a fair coin make the decision.

- Toss a fair 0-1 coin for each biclique $A_i \times B_i$ in any order.
  - If 0 pops up, remove the *vertex set* A_i from K_n.
  - If 1 pops up, remove B_i from K_n.

- **Claim:** When the process ends, *at most one vertex* will remain in K_n.
  - If there are more than two vertices, say, u, v,
    they are connected by edge (u, v) in K_n and will have gone through the process, since at least one of $(A_i \times B_i)$ covers (u, v).

    This means that,
    at most one of them can survive when the coin is tossed.

    A contradiction.

- Toss a fair 0-1 coin for each $A_i \times B_i$ in any order.
  If 0 pops up, remove A_i from K_n. Otherwise, remove B_i from K_n.
- **Claim:** At most one vertex will remain when the above process ends.

- For any $1 \leq v \leq n$,
  let X_v be the indicator variable for the event that vertex v survives after the process, and let $X = \sum_{1 \leq v \leq n} X_v$.
  - By the above claim, $\mathrm{E}[X] \leq 1$.
  - Moreover, for each vertex v,

$$\Pr[\, v \text{ survives} \,] = 2^{-m_v}\,.$$

$X \leq 1$ always holds,
no matter what the toss outcomes are.

v survives with probability $1/2$
for each biclique that contains it.

- We have

$$\sum_{1 \le v \le n} 2^{-m_v} = \sum_{1 \le v \le n} \Pr[\, v \text{ survivies} \,] = \sum_{1 \le v \le n} \mathrm{E}[X_v] = \mathrm{E}[X] \le 1 \,.$$

- By the arithmetic-geometric mean inequality,

$$\frac{1}{n} \ge \frac{1}{n} \sum_{1 \le v \le n} 2^{-m_v} \ge \left(\prod_{1 \le v \le n} 2^{-m_v} \right)^{1/n} = 2^{-\frac{1}{n} \cdot \sum_{1 \le v \le n} m_v} \,.$$

This implies that $2^{\frac{1}{n} \cdot \sum_{1 \le v \le n} m_v} \ge n$, and $\sum_{1 \le v \le n} m_v \ge n \cdot \log_2 n$.

Some Useful Tools & Bounds

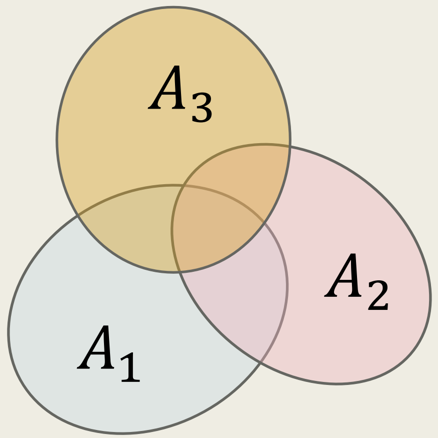
Common tools for upper- / lower- bounding the probabilities.

Some Useful Tools & Bounds

- **Union Bound.**

  Let $A_1, A_2, \ldots, A_n$ be events. Then

$$\Pr\left[\bigcup_{1 \le i \le n} A_i\right] \le \sum_{1 \le i \le n} \Pr[A_i].$$



Some Useful Tools & Bounds

- **Two useful inequalities.**
  - For any $t \neq 0$,
  $$1 + t < e^t.$$
  - For any $0 < t < 0.6838\ldots$,
  $$1 - t > e^{-t-t^2}.$$

By Taylor's expansion on e^t.

By Taylor's expansion on $\ln(1-t)$.
See the jamboard for further details.

- **Stirling's Approximation for $n!$.**

$$n! = \left(\frac{n}{e}\right)^n \cdot \sqrt{2\pi n} \cdot e^{\alpha_n}, \qquad \text{where } \frac{1}{12n+1} < \alpha_n < \frac{1}{12n}.$$

- The Stirling formula is a very tight approximation for $n!$.
  - It leads to the following formula for k^{th} factorial.

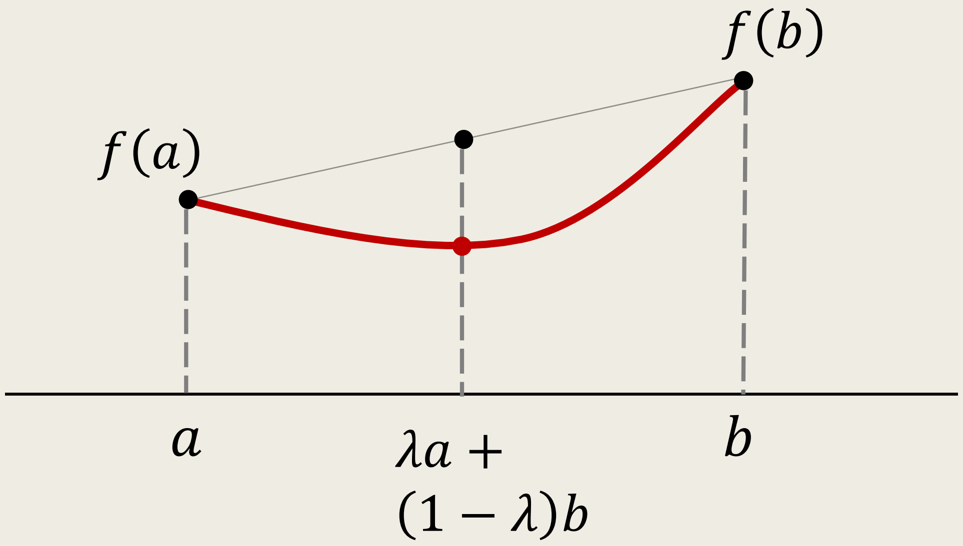
$$(n)_k \coloneqq n \cdot (n-1) \cdot \cdots \cdot (n-k+1)$$

$$= n^k \cdot e^{-\frac{k^2}{2n} - \frac{k^3}{6n^2} + o(1)}, \qquad \forall\, k = o\left(n^{\frac{3}{4}}\right).$$

- **Convex Function.**

  A real-valued function $f(x)$ is convex between $[a, b]$, if

  $$f(\lambda a + (1-\lambda)b) \;\le\; \lambda \cdot f(a) + (1-\lambda) \cdot f(b), \qquad \forall\, 0 \le \lambda \le 1\,.$$



The curve always ***falls under*** the linear function between $(a, f(a))$ and $(b, f(b))$.

- **Jensen's Inequality for Convex Functions.**

  If $\lambda_i \geq 0$, $\sum_{1\leq i\leq n} \lambda_i = 1$, and f is a real-valued *convex function*,

  then

$$f\left(\sum_{1\leq i\leq n} \lambda_i \cdot x_i\right) \leq \sum_{1\leq i\leq n} \lambda_i \cdot f(x_i) \ .$$

  – Refer to the jamboard for the proof.

- The Jensen's inequality is a very useful tool for obtaining *bounds that "**behaves linearly**" for convex functions*.

- **Arithmetic-Geometric Mean Inequality.**

  For any $a_i \geq 0$, we have

$$\frac{1}{n} \cdot \sum_{1 \leq i \leq n} a_i \geq \left(\prod_{1 \leq i \leq n} a_i \right)^{\frac{1}{n}} .$$

  - Refer to the jamboard for the proof.

- This is yet another fundamental & useful inequality (for obtaining nontrivial lower-bounds).