Lecture 2

Ernie's class
Friday, 9/3/10

Let $(S, \Sigma, \mathbb{P})$ be a probability space.

Let $B \in \Sigma$ be an event.

We saw that the following is a σ-algebra:

$$\mathcal{F}_B = \{ A \cap B : A \in \Sigma \}.$$

When B is s.t. $\mathbb{P}(B) > 0$, turns out we can equip $\mathcal{F}_B$ with a natural (conditional) probability measure as follows: For $A \in \mathcal{F}_B$,

$$\text{"}\mathbb{P}(A \mid B)\text{"} := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \longrightarrow (1)$$

(*) Note that the LHS is technically a different probability measure (once we prove it is a probab. measure); but it is standard to abuse notation and use the same letter $\mathbb{P}$ and denote

the dependence on B, using $(\cdot|B)$.

(*) We leave it as an exercise to prove that (1) defines a probability measure over the σ-algebra $\mathcal{F}_B$.

Warning. Conditioning can cause confusion!

Example. Monty Hall.

Toy Example. (I failed to convince my family of this, e.g.): The chance of having two boys (conditioned on) given the fact that a couple had at least one boy is a $1/3$ (and not a $1/2$ or a $1/4$?)

$$\mathbb{P}(\{BB\}) = \frac{1}{4} \; ; \; \mathbb{P}(\{BB\} \mid \{BB, BG, GB\}) = \frac{1}{3}$$

while if you had no knowledge then at each birth (we assume) there is a $1/2$ chance of a boy (independently): Hence

* Interestingly enough, $\frac{1}{2} = \mathbb{P}(\{BB\} \mid \text{"the younger is a boy"})$.

Independence of events.

Defn. For a finite collection $A_1, A_2, \dots, A_k \in \Sigma$, we say $\{A_1, \dots, A_k\}$ are mutually independent if

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_\ell}) = \prod_{j=1}^{\ell} \mathbb{P}(A_{i_j}).$$

for each ~~subset~~ subcollection ...

- An infinite collection of events is independent, if every finite subcoll. is independent.

(*) Common error: to think A, B: indep. if $A \cap B = \emptyset$. (NOT TRUE)

**** WARNING **** Pairwise independence versus Mutual independence:

Example (1) Jane has 3 children each of which is equally likely to be a boy or a girl independently of the others.

$A = \{$ all children are of the same gender $\}$

$B = \{$ there is at most one boy $\}$

$C = \{$ the family includes a boy and a girl $\}$.

(i) show A, B: indep. and B, C: indep.,

Example ② Let  $\mathbb{P} \equiv 1/9$; $\Sigma' = \{aaa, bbb, ccc, abc, acb, bac, bca, cab, cab\}$.

$A_i = \{ i^{th} \text{ letter is } a \}$; Show $\{A_1, A_2, A_3\}$ pairwise indep. but not mutually ④

Exercise ① $\mathbb{P}(A \mid B \cap C) = \dfrac{\mathbb{P}(A \cap B \mid C)}{\mathbb{P}(B \mid C)}$.

More generally,

$$\mathbb{P}(A \mid B_1, B_2, \cdots, B_n) = \frac{\mathbb{P}(A\, B_1 B_2 \cdots B_k \mid B_{k+1} \cdots B_n)}{\mathbb{P}(B_1 B_2 \cdots B_k \mid B_{k+1} \cdots B_n)}.$$

(where "$\cap$" is suppressed for convenience).

§ The Lovász local lemma.

("BAD")

Let $A_1, A_2, \cdots, A_n$ be events in some probability space. Let $\bar{A}_i$ denote the complement of event A_i (and we say A_i did not occur). Since $A_i \cap \bar{A}_i = \emptyset$,

$$\mathbb{P}(\bar{A}_i) = 1 - \mathbb{P}(A_i).$$

OBSVN 1. If $\sum_i \mathbb{P}(A_i) < 1$, then $\mathbb{P}(\cap \bar{A}_i) > 0$.

For $\mathbb{P}(\cap \bar{A}_i) = 1 - \mathbb{P}(\cup_i A_i)$

$\geq 1 - \sum_i \mathbb{P}(A_i) > 0$, by hypo.

(2.) If $\mathbb{P}(A_i) < 1$ for each i, and

if $\{A_i\}$ mutually independent,

then $\mathbb{P}(\bigwedge_i \bar{A}_i) = \prod_i \mathbb{P}(\bar{A}_i) = \prod_{i=1}^{n} [1 - \Pr(A_i)]$

> 0, by the hypo.

(verify that $\{A_i\}$: mut. indep. $\iff$ $\{\bar{A}_i\}$: mut. indep.)

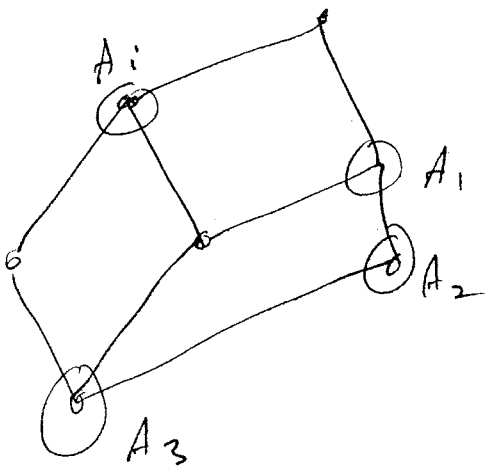
The following elegant, but powerful, lemma of Lovász (from a paper of Erdös-Lovász) asserts that under limited dependence, one can make a similar conclusion as in above.

Let $V = \{A_1, \ldots, A_n\}$

Defn. $G = (V, E)$ is a dependency graph for the events $A_1, \ldots, A_n$ if A_i is mutually independent of all A_j which are not adjacent to A_i in G.

(tricky to grasp, at first!)

G not unique, note!



$\mathbb{P}(A_i \mid$ any Boolean combin. of $A_1, A_2, A_3) = \mathbb{P}(A_i)$.

Lemma. $A_1, \cdots, A_n$: events with dependency graph G so that

$\mathbb{P}(A_i) \le p$, $\forall i$ and $\deg(A_i) \le d$, $\forall i$ in G.

and $4dp < 1 \Rightarrow \mathbb{P}(\bigwedge \overline{A}_i) > 0$.

Proof : See handout. Uses induction, Exercise ①, and $\mathbb{P}(\bigcup A_i) \le \sum_i \mathbb{P}(A_i)$.

□.

LECTURE 8

The Lovász Local Lemma

The Lemma. Let $A_1, \cdots, A_n$ be events in a probability space. In combinatorial applications the A_i are "bad" events. We wish to show $\Pr[\wedge \bar{A}_i] > 0$ so that there is a point (coloring, tournament, configuration) x which is good. The basic probabilistic method of Lecture 1 may be written:

Counting sieve. If $\sum \Pr[A_i] < 1$, then $\Pr[\wedge \bar{A}_i] > 0$.

There are other simple conditions that ensure $\Pr[\wedge \bar{A}_i] > 0$.

Independence sieve. If $A_1, \cdots, A_n$ are mutually independent and all $\Pr[A_i] < 1$, then $\Pr[\wedge \bar{A}_i] > 0$.

The Lovász Local Lemma is a sieve method which allows for *some* dependence among the A_i. A graph G on vertices $[n]$ (the indices for the A_i) is called a *dependency graph* for $A_1, \cdots, A_n$ if for all i A_i is mutually independent of all A_j with $\{i, j\} \notin G$. (That is, A_i is independent of any Boolean function of these A_j.)

LOVÁSZ LOCAL LEMMA (Symmetric case). *Let $A_1, \cdots, A_n$ be events with dependency graph G such that*

$$\Pr[A_i] \leq p \quad \textit{for all } i, \qquad \deg(i) \leq d \quad \textit{for all } i$$

and

$$4dp < 1.$$

Then

$$\Pr[\wedge \bar{A}_i] > 0.$$

Proof. We show by induction on s that if $|S| \leq s$, then for any i

$$\Pr\left[A_i \,\middle|\, \bigwedge_{j \in S} \bar{A}_j\right] \leq 2p.$$

For $S = \emptyset$ this is immediate. Renumber for convenience so that $i = n$, $S = \{1, \cdots, s\}$ and $\{i, x\} \notin G$ for $x > d$. Now

$$\Pr[A_n | \bar{A}_1 \cdots \bar{A}_s] = \frac{\Pr[A_n \bar{A}_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s]}{\Pr[\bar{A}_1 \cdots \bar{A}_d | \bar{A}_{d+1} \cdots \bar{A}_s]}.$$

We bound the numerator

$$\Pr[A_n\bar{A}_1\cdots\bar{A}_d|\bar{A}_{d+1}\cdots\bar{A}_s]\le\Pr[A_n|\bar{A}_{d+1}\cdots\bar{A}_s]$$
$$=\Pr[A_n]\le p$$

as A_n is mutually independent of $A_{d+1},\cdots,A_s$. We bound the denominator

$$\Pr[\bar{A}_1\cdots\bar{A}_d|\bar{A}_{d+1}\cdots\bar{A}_s]\ge 1-\sum_{i=1}^{d}\Pr[A_i|\bar{A}_{d+1}\cdots\bar{A}_s]$$
$$\ge 1-\sum_{i=1}^{d}2p \qquad \text{(Induction)}$$
$$=1-2pd\ge\tfrac{1}{2}.$$

Hence we have the quotient

$$\Pr[A_n|\bar{A}_1\cdots\bar{A}_s]\le p/\tfrac{1}{2}=2p,$$

completing the induction. Finally

$$\Pr[\bar{A}_1\cdots\bar{A}_n]=\prod_{i=1}^{n}\Pr[\bar{A}_i|\bar{A}_1\cdots\bar{A}_{i-1}]\ge\prod_{i=1}^{n}(1-2p)>0. \qquad \square$$

The proof is so elementary that it could, and I think should, be taught in a first course in probability. It has had and continues to have a profound effect on probabilistic methods.

The diagonal Ramsey function. A lower bound for $R(k,k)$, our first use of the probabilistic method in Lecture 1, provides a simple application of the Lovász Local Lemma. Consider a random two-coloring of K_n with A_S the event "S is monochromatic," S ranging over the k-sets of vertices. Define G by placing $\{S,T\}\in G$ if and only if $|S\cap T|\ge 2$. Then A_S is mutually independent of all A_T with T not adjacent to G, since the A_T give information only about edges outside of S. Hence G is a dependency graph. (When $|S\cap T|=2$ the events A_S, A_T are independent; note however that mutual independence from a family of A_T is far stronger than pairwise independence with each A_T. Recall the old chestnut: I have two children and at least one is a girl. What is the probability they are both girls. Conditional on the younger being a girl it is one half. Conditional on the older being a girl it is one half. Conditional on the disjunction it is one third. But I digress.) We apply the Lovász Local Lemma with $p=\Pr[A_S]=2^{1-\binom{k}{2}}$ and

$$d=|\{T:|S\cap T|\ge 2\}|\le\binom{k}{2}\binom{n}{k-2}.$$

COROLLARY. *If*

$$4\binom{k}{2}\binom{n}{k-2}2^{1-\binom{k}{2}}<1,$$

then $R(k,k)>n$.

The asymptotics are somewhat disappointing.

COROLLARY.

$$R(k,k)>\frac{\sqrt{2}}{e}k2^{k/2}(1+o(1)).$$

This improves the lower bound given in Lecture 1 by a factor of 2 and the improvement via the deletion method in Lecture 2 (which, oddly, was only published after the better bound) by a factor of $\sqrt{2}$. The gap between the upper and lower bounds has not really been effectively decreased. The lower bound of Erdös was found in April 1946 (published in 1947) and progress on this difficult problem has been slow.

The van der Waerden function. Here the improvement is more impressive. Color $[n]$ randomly. For each arithmetic progression S of size k let A_S be the event that S is monochromatic. Let S, T be adjacent in G if they intersect. (In all our applications the probability space will be a random coloring of some set Ω. For Ramsey's Theorem Ω was the edge set of K_n. Events will be not adjacent if they deal with the coloring on disjoint sets.) Now $p=2^{1-k}$ and $d\leq nk$ as one progression intersects (exercise) at most nk others. Hence we have the following theorem.

THEOREM. *If* $4nk2^{1-k}<1$, *then* $W(k)>n$. *That is,* $W(k)>2^k/8k$.

This greatly improves the bound $W(k)>2^{k/2}$ of Lecture 1. Still we must in all honesty mention that $W(p)\geq p2^p$, for p prime, has been shown by completely constructive means!

Algorithm? The Lovász Local Lemma proves the existence of an x satisfying $\bigwedge \bar{A}_i$ even when $\Pr[\bigwedge \bar{A}_i]$ may be exponentially small. We have seen in Lecture 4 that when $\sum \Pr[A_i]<1$ there often is an algorithm to find a specific "good" x.

Open Problem. Can the Lovász Local Lemma be implemented by a Good Algorithm?

Let us be more specific. Suppose $S_1,\cdots,S_n\subset[n]$ with all $|S_i|=10$ and all $\deg(j)=10$. Two color $[n]$ randomly and let A_i be the event that S_i is monochromatic. Let i, i' be adjacent in the dependency graph if and only if their corresponding sets intersect. Each S_i is intersected by at most 90 other $S_{i'}$. We apply the Lovász Local Lemma with $p=\Pr[A_i]=2^{-9}$ and $d=90$. As $4dp<1$, $\Pr[\bigwedge \bar{A}_i]>0$ and so there is a two-coloring χ for which no S_i is monochromatic. Is there a polynomial (in n) time algorithm for finding such a coloring?

Notice that the Lovász Local Lemma here guarantees the existence of a "needle in a haystack." If, say, $S_1,\cdots,S_{n/10}$ are disjoint a random χ is good with probability at most $(1-2^{-9})^{n/10}$. Can we actually find this exponentially small needle in polynomial time?

Addendum: Joke. Here we give an ironic demonstration of the power of the Lovász Local Lemma.

THEOREM. *Let* S, T *be finite sets with* $|T|\geq 8|S|$. *Then there exists a function* $f: S\to T$ *which is injective.*

Proof. Let f be a random function. For each $\{x, y\} \subset S$ let A_{xy} be the event $f(x) = f(y)$. Then $\Pr[A_{xy}] = 1/|T| = p$. Let $\{x, y\}$ and $\{u, v\}$ be adjacent in the dependency graph if and only if they intersect. Then the maximal degree in the dependency graph is $d = 2(|S|-1)$. As $4dp < 1$, $\Pr[\bigwedge \bar{A}_{xy}] > 0$ and so there exists an f for which $\bar{A}_{xy}$ for all x, y. That is, f is injective.

When $|T| = 365$ and $|W| = 23$ the "birthday problem" says that f has probability less than $\frac{1}{2}$ of being injective. When $|T| = 8|S|$ the probability of a random f being injective is exponentially small. The Counting Sieve proves the existence of an injective f only when $\binom{|S|}{2} < |T|$. □

Anti van der Waerden. Here is the original use of the Lovász Local Lemma.

THEOREM. *Let k, m satisfy $4m(m-1)(1-1/k)^m < 1$. Let $S \subset R$ with $|S| = m$. Then there exists a k-coloring $\chi: R \to [k]$ so that every translate $S + t$ is k-colored. That is, for all $t \in T$ and $1 \le i \le k$ there exists $s \in S$ with $\chi(s + t) = i$.*

Without use of the Lovász Local Lemma no proof is known that gives the existence of an $m = m(k)$ with this property. Notice a fundamental difference between the translation and homothety groups. Gallai's Theorem, a consequence of van der Waerden's Theorem, states that for all finite S and all finite colorings of R there is a monochromatic $S' = aS + t$.

Proof. First we let $B \subset R$ be an arbitrary finite set and k-color B so that all $S + t \subset B$ have all k colors. Color B randomly. For each t such that $S + t \subset B$ let A_t be the event that $S + t$ does not have all k colors. Then

$$\Pr[A_t] \le k(1-1/k)^m = p.$$

Let t, t' be adjacent in the dependency graph if and only if $S + t$ and $S + t'$ intersect. With given t this occurs only if $t' = +s - s'$ for distinct $s, s' \in S$, so that the dependency graph has degree at most $d = m(m-1)$. The conditions on k, m are precisely that $4dp < 1$. The Lovász Local Lemma applies and $\bigwedge \bar{A}_t \neq \emptyset$; there is a k-coloring of B for which all translates of S lying in B have all k colors.

Compactness. To color all of R we need the Compactness Principle. We state this in a form convenient for us.

COMPACTNESS PRINCIPLE. *Let Ω be an infinite set, k a positive integer, and let U be a family of pairs (B, χ) where $B \subset \Omega$ is finite, $\chi: B \to [k]$ such that*

(i) *U is closed under restriction. That is, if $(B, \chi) \in U$ and $B' \subset B$ then $(B', \chi|_{B'}) \in U$;*

(ii) *For all B some $(B, \chi) \in U$.*

Then there exists $\chi: \Omega \to [k]$ such that

$$(B, \chi|_B) \in U \quad \textit{for all finite } B \subset U.$$

Proof. Let X be the topological space of all $\chi: \Omega \to [k]$. Here we consider $[k]$ discrete and X has the usual product topology. That is, a basis for the open sets is given by the sets $\{\chi: \chi(b_i) = a_i, 1 \le i \le s\}$ over all $s, b_1, \cdots, b_s, a_1, \cdots, a_s$. For every finite B let X_B be the set of $\chi \in X$ with $(B, \chi|_B) \in U$. By (ii), $X_B \neq \emptyset$. Splitting $\chi \in X$ according to $\chi|_B$ gives a finite $(|B|^k)$ partition of X into sets both open