Math 155 (Lecture 30)

November 15, 2011

In the last lecture, we described some elementary applications of the probabilistic method. In this lecture, we will introduce some of the language of probability theory, which will be important for discussing more sophisticated applications.

Definition 1. A *finite probability space* consists of a finite set Ω (the set of *outcomes*) together with, for each $\omega \in \Omega$, a nonnegative real number p_ω, called the *probability of* ω. We assume that $\sum_{\omega \in \Omega} p_\omega = 1$.

Given a finite probability space Ω, we define an *event* to be a subset $E \subseteq \Omega$. The *probability* of an event E is the real number $P(E) = \sum_{\omega \in E} p_\omega$.

In what follows, we will fix a finite probability space Ω. In the last lecture, we used the following very rudimentary fact: if E and E' are two events, then

$$P(E \cup E') \leq P(E) + P(E').$$

In fact, we can be more precise: we have $P(E \cup E') = P(E) + P(E') - P(E \cap E')$. If E and E' are disjoint, this gives $P(E \cup E') = P(E) + P(E')$. Another special case is worthy of mention:

Definition 2. Let E and E' be events. We say that E and E' are *independent* if $P(E \cap E') = P(E)P(E')$.

If E and E' are independent, we have $P(E \cup E') = P(E) + P(E') - P(E)P(E')$.

It is sometimes helpful to think about independent events in terms of *conditional probability.* If E and E' are events with $P(E') \neq 0$, we let $P(E|E')$ denote the quotient

$$\frac{P(E \cap E')}{P(E)}$$

We can think of $P(E|E')$ as the probability that the event E will occur, given that we know that E' will occur. Note that E and E' are independent if and only if $P(E|E') = P(E)$: that is, if the event E' tells us nothing about E.

We will need the following elaboration on Definition 2:

Definition 3. Suppose we are given events E and $E'_1, E'_2, \ldots, E'_k \subseteq \Omega$. We will say that E is *independent* of the set of events $\{E'_1, E'_2, \ldots, E'_k\}$ if E is independent of $\bigcap_{i \in S} E'_i$ for all subsets $S \subseteq \{1, \ldots, k\}$ (note that we do not assume that the events E'_i are independent of each other).

Example 4. Let Ω be the set of all graphs with vertex set $\{1, \ldots, m\}$. In the last lecture, we viewed Ω as a probability space (where each outcome has probability $2^{-\binom{m}{2}}$). For every subset $S \subseteq \{1, \ldots, m\}$, we can consider events $E_S, E'_S \subseteq \Omega$, where E_S is the set of graphs containing S as a clique, and E'_S is the set of graphs containing S as an anticlique. Recall that our basic goal was to obtain lower bounds for Ramsey numbers, by showing that

$$P(\bigcup_{|S|=n} E_S \cup \bigcup_{|S|=n} E'_S) \leq \sum_{|S|=n} P(E_S) + P(E'_S) < 1$$

unless m is large compared with n.

Notation 5. If E is an event, we let $E^c = \Omega - E$ denote the complement of E, so that $P(E^c) = 1 - P(E)$.

Suppose we are given a set $E_1, \ldots, E_k$ of mutually independent events (that is, each E_i is independent of the set of all the other E_j's). Then

$$P(E_1 \cup \cdots \cup E_k) = 1 - P(E_1^c \cap \cdots \cap E_k^c) = 1 - P(E_1^c) \cdots P(E_k^c) = 1 - (1 - P(E_1)) \cdots (1 - P(E_k)).$$

In particular, we see that $P(E_1 \cup \cdots \cup E_k)$ is strictly less than one provided that $P(E_i) < 1$ for each i. We therefore have two different ways to prove that $E_1 \cup \cdots \cup E_k \neq \Omega$:

(a) We can try to prove that $\sum P(E_i) < 1$.

(b) We can try to prove that the events E_i are independent and that $P(E_i) < 1$ for all i.

For many applications, it is useful to employ a sort of hybrid between these approaches. That is, we would like to say something about the probability $P(E_1 \cup \cdots \cup E_k)$ in the case where the E_i are "mostly" independent of one another.

Theorem 6 (Lovász Local Lemma)**.** *Suppose we are given a collection of events $E_1, \ldots, E_k$ and a graph G with vertex set $\{1, \ldots, k\}$. Assume that:*

($*$) *For each $1 \leq i \leq k$, if S is a set of vertices of G which are not adjacent to i (and does not include i), then E_i is independent of the set of events $\{E_j\}_{j \in S}$.*

Suppose further that we are given real numbers $0 \leq x_i < 1$ such that

$$P(E_i) \leq x_i \prod_{(i,j) \text{ adjacent}} (1 - x_j).$$

Then $P(E_1 \cup \cdots \cup E_k) \leq 1 - \prod_{1 \leq i \leq k}(1 - x_i)$. In particular, we have $E_1 \cup \cdots \cup E_k \neq \Omega$.

Remark 7. We can summarize ($*$) more informally by saying that an event E_i is independent of those events which are not adjacent to E_i in the graph G.

Example 8. Suppose that the events E_i are independent. Then we can take G to be a graph with no edges, and $x_i = P(E_i)$. In this case, we have equality

$$P(E_1 \cup \cdots \cup E_k) = 1 - \prod (1 - x_i).$$

Proof. We first prove the following:

($*'$) If $S \subseteq \{1, \ldots, k\}$ is a set which does not contain some integer i, then

$$P(E_i | \bigcap_{j \in S} E_j^c) \leq x_i \prod_{i,j \text{ adjacent }, j \notin S} (1 - x_j).$$

The proof proceeds by induction on the number of elements of S. Suppose first that S contains no elements which are adjacent to i. Then E_i is independent of the set $\{E_j\}_{j \in S}$. Thus

$$\begin{aligned} P(E_i | \bigcap_{j \in S} E_j^c) &= P(E_i) \\ &\leq x_i \prod_{i,j \text{ adjacent}} (1 - x_j) \\ &\leq x_i \prod_{i,j \text{ adjacent}, j \notin S} (1 - x_j). \end{aligned}$$

To carry out the inductive step, assume that there exists an index $i' \in S$ which is adjacent to i. Let $S' = S - \{i'\}$. Then

$$P(E_i | \bigcap_{j \in S} E_j^c) = \frac{P(E_i \cap E_{i'} | \bigcap_{j \in S'} E_j^c)}{P(E_{i'}^c | \bigcap_{j \in S_+} E_j^c)}.$$

By the inductive hypothesis, the numerator is bounded above by

$$x_i(1 - x_{i'}) \prod_{i,j \text{ adjacent } ,j \notin S} (1 - x_j).$$

It will therefore suffice to show that the denominator is $\geq 1 - x_{i'}$. Equivalently, we must show that

$$P(E_{i'} | \bigcap_{j \in S'} E_j^c) \leq x_j.$$

This follows immediately from $(*')$, applied to the set S'.

We now compute

$$\begin{aligned} P(\bigcap_{1 \leq i \leq k} E_i^c) &= P(E_1^c)P(E_2^c|E_1^c)P(E_3^c|E_1^c \cap E_2^c) \cdots \\ &= (1 - P(E_1))(1 - P(E_2|E_1^c)) \cdots \\ &\geq (1 - x_1)(1 - x_2) \cdots (1 - x_k). \end{aligned}$$

It follows that that

$$P(E_1 \cup \cdots \cup E_k) \leq 1 - \prod_{1 \leq i \leq k} (1 - x_i).$$

□

Corollary 9 (Lovász Local Lemma, Symmetric Version)**.** *In the situation of Theorem 6, suppose that the graph G has valence $\leq d$ at each vertex (that is, each vertex is adjacent to at most d other vertices). If each of the events E_i has probability $\leq \frac{1}{e(d+1)}$, then $P(E_1 \cup \cdots \cup E_k) < 1$. Here e denotes Euler's constant.*

Proof. We take $x_i = \frac{1}{d+1}$ for each i. According to Theorem 6, it suffices to show that

$$P(E_i) \leq x_i \prod_{i,j \text{ adjacent}} (1 - x_j).$$

Since $P(E_i) \leq \frac{1}{e(d+1)}$ and

$$\frac{1}{d+1}(\frac{d}{d+1})^d \leq x_i \prod_{i,j \text{ adjacent}} (1 - x_j),$$

we are reduced to proving that

$$\frac{1}{e(d+1)} \leq \frac{1}{d+1}(\frac{d}{d+1})^d.$$

Multiplying by $d + 1$ and taking logarithms, we want

$$-1 \leq d \log \frac{d}{d+1}$$

or equivalently

$$\frac{1}{d} \geq \log(1 + \frac{1}{d}).$$

This can be read off immediately from the power series expansion

$$\log(1 + \frac{1}{d}) = \frac{1}{d} - \frac{1}{2d^2} + \frac{1}{3d^3} - \cdots$$

□

Remark 10. Note that the assumption $P(E_i) \leq \frac{1}{e(d+1)}$ in Corollary 9 does not depend on the number of events k: it depends only on the number d, which measures the failure of these events to be independent.