Some Applications of Probability

COS 341 Fall 2005

1 Ramsey Numbers

Our first application we discuss is to proving bounds on Ramsey numbers, which are of great interest in combinatorics. As an introduction to what they are, consider the following theorem.

Theorem 1 *In any group of 6 people, there are either 3 mutual acquaintances or 3 mutual strangers.*

Proof: Consider a complete graph on 6 vertices, one vertex corresponding to each of the 6 people. We color the edges of the graph with 2 colors, red and blue. An edge between two vertices is colored red if the corresponding people are strangers. If they are mutual acquaintances, the edge is colored blue. Note: We assume that one of the two cases always occurs. We need to prove that there is either a red triangle or a blue triangle in the graph.

Pick any vertex, say u. Of the five edges incident on u, at least three must be of the same color. Without loss of generality, suppse there are three edges colored blue incident on u. Let the other end points on these edges be v_1, v_2, v_2. If any of the edges (v_i, v_j) is colored blue, then the vertices u, v_i, v_j form a blue triangle. If this is not the case, then all edges (v_i, v_j) must be red. In this case, v_1, v_2, v_3 form a red triangle. Thus there is always a monochromatic triangle. ■

The above theorem is false if 6 is replaced by 5. (Verify this !) It is possible to color the edges of K_5 such that there is neither a red triangle nor a blue triangle. Thus, 6 is the smallest number for which the property holds.

Definition 1 *The Ramsey number $R(k, l)$ is the smallest value of n such that any edge coloring of K_n with two colors, say red and blue, has either a set of k vertices such that all edges between them are colored red, or a set of l vertices such that all edges between them are colored blue.*

In order to show that $R(k, l) > n$, we need to show that there is a way to color the edges of K_n with colors red and blue such that there the graph does not contain a red K_k or a blue K_l. The discussion above shows that $R(3, 3) = 6$. Exact values for the Ramsey numbers are hard to come by. It is known that $R(4, 4) = 18$, but the current best bounds on for $R(5, 5)$ are $42 \leq R(5, 5) \leq 56$.

We describe a lower bound on $R(k,k)$ due to Erdös. The proof uses probability in an interesting way and is an example of the probabilistic method. This is a remarkable technique to prove the existence of certain objects without explictly constructing them.

Theorem 2 *If n, k are integers such that*

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

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

Proof: Suppose we color the edges of K_n with colors red and blue at random, i.e. for each edge, we assign color red with probability 1/2 and assign color blue with probability 1/2. We make these choices independetly and at random for every edge. We will bound the probability that there is a monochromatic copy of K_k in the graph.

First, consider a set S of k vertices and consider the event M_S that the subgraph induced by S is monochromatic. Note that there are $kchoose2$ edges in the subgraph. The probability that they are all red is $2^{-\binom{k}{2}}$ and the probability that they are all blue is $2^{-\binom{k}{2}}$. Thus $\mathbf{Pr}[M_S] = 2 \cdot 2^{-\binom{k}{2}} = 2^{1-\binom{k}{2}}$.

The event that there is some monochromatic complete graph on k vertices is the union of the events M_S for all sets S of size k. By the union bound, the probability of this event is at most $\binom{n}{k} 2^{1-\binom{k}{2}}$. By the assumption in the statement of the theorem, this probability is strictly less than 1. Hence the probability of the complementary event, i.e. that there is no monochromatic K_k, is strictly positive. This implies that there must be some point in the sample space, i.e. some coloring of K_n such that there is no monochromaic K_k. Thus $R(k,k) > n$, establishing the theorem. ∎

Notice that the proof shows that the required coloring exists without giving an explicit construction of such a coloring.

Corollary 1 *$R(k,k) \geq 2^{k/2}$ for $k \geq 2$*

Proof: We prove that the condition in Theorem 2 holds for any $n < 2^{k/2}$. This is easy to verify for $k \geq 4$ using the inequality $\binom{n}{k} \leq n^k/2^{k-1}$. (See Exercise 17 in Section 4.4 of the Rosen text). We omit the details.

Also the statement of the theorem holds for $k = 2$ and $k = 3$ since $R(2,2) = 2$ and $R(3,3) = 6$. ∎

Next we consider a slight generalization of the idea used in the proof. In particular, we use the fact that if there is a random variable X with $E[X] = \mu$, then there must be a point s in the sample space such that $X(s) \geq \mu$.

Given a graph $G(V,E)$, a cut is a partition of the vertices into two disjoint sets V_1 and V_2. Let $E(V_1,V_2) = \{(u,v) \in E | u \in V_1 \text{ and } v \in V_2\}$. The size of the cut is defined to be $|E(V_1,V_2)|$. We say that an edge e belongs to the cut if $e \in E(V_1,V_2)$.

Theorem 3 *For any graph $G(V,E)$ there is a cut of size at least —E—/2*

Proof: Suppose we produce a cut at random, i.e. for evey vertex u, we place u in V_1 with probability $1/2$ and place u in V_2 with probability $1/2$. We make these random choices for every vertex $u \in V$ independently of the other vertices. Then, the probability that a particular edge (u, v) belongs to the cut (V_1, V_2) is exactly $1/2$.

The size of the cut can be expressed as the sum of indicator random variables X_e, one for each edge $e \in E$. The indicator random variable X_e represents the event that e belongs to the cut, i.e. it is 1 if e is in the cut and 0 otherwise. Then $\mathbf{E}[X_e] = 1/2$.

By linearity of expectation, the expected size of the cut (V_1, V_2) is $|E|/2$. Hence there must exist some cut of value at least $|E|/2$. ∎

Next we consider an example involving boolean expressions. A 3-SAT formula is an AND of clauses, where each clause is an OR of three literals. Each literal is one of the boolean variables $x_1, \ldots, x_n$ or their negations $\overline{x_1}, \ldots, \overline{x_n}$. Testing whether there is an assignment to the variables so as to satisfy all clauses is an NP-hard problem. However, the next theorem shows that it is always possible to satisfy a large fraction of clauses in such a formula.

Theorem 4 *For any 3-SAT formula with m clauses, there exists an assignment to the variables which satisfies $(7/8)m$ clauses.*

Proof: Consider assigning truth values to variables at random, i.e. for every variable, we pick value T (true) with probability $1/2$ and value F (false) with probability $1/2$.

It is easy to verify that for any clause C, the probability that C is satisfied is $7/8$. Arguing along similar lines as in the previous proof, the expected number of clauses satisfied by a random assignment is $(7/8)m$. Thus there must exist some truth assignment that satisfies at least $(7/8)m$ clauses. ∎

The proof gives a randomized algorithm to construct a truth assignment with expected number of clauses satisfied being $(7/8)m$. We can actually give a deterministic algorithm to find a truth assignment satisfying at least $(7/8)m$ clauses. Note that we can always do this by checking all possible 2^n truth assignments, but such an algorithm would be terribly inefficient. The point is that such a truth assignment can be constructed deterministically in polynomial time.

Consider producing a truth assignment at random by setting the variables in the order $x_1, \ldots, x_n$. We will gradually replace the random choices by deterministic choices. Let X denote the number of clauses satisfied by an assignment. Then $E[X] = (7/8)m$, where the expectation is over the random choice of an assignment to $x_1, \ldots, x_n$. We will be interested in the the value of X given an assignment $v_1, \ldots v_i$ to $x_1, \ldots, x_i$, i.e. the value of X when the $x_1 = v_1, \ldots, x_i = v_i$ and the remaining variables have random values. We will show that there exists a partial assignment $v_1, \ldots, v_i$ such that

$$E[X | x_1 = v_1, \ldots, x_i = v_i] \geq (7/8)m \tag{1}$$

Given such a partial assignment for $x_1, \ldots, x_i$, we show how to extend this to such a partial assignment for $x_1, \ldots, x_i, x_{i+1}$ by picking an appropriate value for x_{i+1}. If we show this we are done. Initially, we start with an empty assignment. Clearly $E[X] \geq (7/8)m$. By repeating this procedure for n steps, we construct an assignment $v_1, \ldots v_n$ for $x_1, \ldots, x_n$. Since all variables are assiged values $E[X|x_1 = v_1, \ldots, x_n = v_n]$ is simply the number of clauses satisfied by the assignment $v_1, \ldots, v_n$. By the property (1) of this assignment, the number of clauses satisfied is at least $(7/8)m$.

It remains to show how we can extend the assignment at every step. Suppose we have an assignment $v_1, \ldots, v_i$ satisfying (1). Note that

$$\begin{aligned} E[X|x_1 = v_1, \ldots, x_i = v_i] &= \frac{1}{2}E[X|x_1 = v_1, \ldots, x_i = v_i, x_{i+1} = T] \\ &\quad + \frac{1}{2}E[X|x_1 = v_1, \ldots, x_i = v_i, x_{i+1} = F] \qquad (2) \end{aligned}$$

We pick $v_{i+1} = T$ or $v_{i+1} = F$ depending on which term on the LHS of (2) is higher. In other words, we pick $v_{i+1} = T$ if

$$E[X|x_1 = v_1, \ldots, x_i = v_i, x_{i+1} = T] \geq E[X|x_1 = v_1, \ldots, x_i = v_i, x_{i+1} = F]$$

and $v_{i+1} = F$ otherwise. The choice of v_{i+1} ensures that

$$E[X|x_1 = v_1, \ldots, x_i = v_{i+1}] \geq E[X|x_1 = v_1, \ldots, x_i = v_i] \geq (7/8)m.$$

In order to make this choice efficiently at every step, we need to be able to compute the value of terms of the form $E[X|x_1 = v_1, \ldots, x_i = v_i]$ in polynomial time. This is easy to do. We first apply the assignment to the variables $x_1, \ldots, x_i$ and simplify the clauses in which they appear.

The expected number of clauses satisfied by completing the assignment at random can be computed (using linearity of expectation) by adding up a contribution from every clause C. The contribution from clause C is the probability that clause C is satisfied by a random completion of the partial assignment. The probability that clause C is satisifed is 0 if all literals have already been set to F. It is 1 if any literal has already been set to T. If neither of these happen, and the clause has r literals unassigned, then the probability that the clause is satisfied is $1 - 1/2^r$.

Clearly this expectation can be computed in polynomial time. Thus, we obtain an efficient deterministic algorithm for finding a truth assignment to variables satisfying $(7/8)m$ clauses. The deterministic algorithm was obtained by derandomizing the random choices in the original randomized algorithm. This derandomization method is called *the method of conditional probabilities.*