Honors Combinatorics and Probability CS-284/Math-274
Instructor: László Babai

The Erdős - Ko - Rado Theorem

The proof given in class had a conceptual error in the use of indicator variables. The actual proof is simpler and does not use indicatior variables, while the basic idea (averaging over cyclic permutations) remains the same.

Theorem (Erdős - Ko - Rado) Let $1 \le k \le n/2$ and let $\mathcal{F} = \{A_1, \ldots, A_k\}$ be a k-uniform intersecting family of subsets of $[n]$ (i. e., $(\forall i \neq j)(A_i \neq A_j$ and $A_i \cap A_j \neq \emptyset)$. Then

$$m \le \binom{n-1}{k-1}.$$

We proved the following Lemma in class.

Lemma Let $1 \le k \le n/2$. Given a cyclic permutation σ of $[n]$, at most k of the k-arcs of σ belong to $\mathcal{F}$. In particular, if we pick a k-arc at random then the probability that it belongs to $\mathcal{F}$ is at most k/n.

(Prove the Lemma.)

Proof of the Theorem (based on the Lemma). We note that

$$\frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{k}{n} \tag{1}$$

(prove!) So what we need to prove is that at most a k/n fraction of the k-subsets of $[n]$ belong to $\mathcal{F}$. In other words, what we need to prove is that if we pick a k-subset $X \subset [n]$ at random then

$$P(X \in \mathcal{F}) \le \frac{k}{n}. \tag{2}$$

We prove this by an "averaging argument:" the Lemma tells us that the k/n bound on the proportion of members of $\mathcal{F}$ holds for certain n-tuples of k-subsests; we average this inequality over all choices of those n-tuples.

Let us generate $X \subset [n]$ in the following way: first pick a random cyclic permutation σ, and then pick a random k-arc of σ. This process generates a k-subset X from the uniform distribution over all k-subsets of $[n]$ (by symmetry).

Let S denote the set of all cyclic permutations of $[n]$. (Note that $|S| = (n-1)!$.) Now by the Theorem of Complete Probability (review from online text!), we have

$$P(X \in \mathcal{F}) = \sum_{s \in S} P(X \in \mathcal{F} \,|\, \sigma = s) P(\sigma = s).$$

But the conditional probability $P(X \in \mathcal{F} \,|\, \sigma = s)$ is at most k/n for every $s \in S$ (by the Lemma); so the right hand side is at most $(k/n) \sum_{s \in S} P(\sigma = s) = k/n$, completing the proof of inequality (2). Q.E.D.