Lecture 34: Basic probability theory

Nitin Saxena *

IIT Kanpur

1 The probabilistic method

1.1 Set families

Now we will see another clever usage of probability to prove something about extremal set families, that appear in many interesting applications.

Let $\mathcal{F} = \{(A_i, B_i) \mid i \in [h]\}$ be a family of pairs of subsets of an arbitrary set. We call $\mathcal{F}$ a (k, ℓ)-*system* if $|A_i| = k, |B_i| = \ell$, $A_i \cap B_i = \emptyset$ and $A_i \cap B_j \neq \emptyset$, for all $i \neq j \in [h]$.

For example, if we take the universe to be $U = [k + \ell]$, then $\mathcal{F} := \{(A, A^c) \mid A \in \binom{U}{k}\}$ is a (k, ℓ)-system. Note that it has size $h = \binom{k+\ell}{k}$. Could there be a system with a bigger h?

Theorem 1 (Bollobás, 1965). *If $\mathcal{F}$ is a (k, ℓ)-system then $h \leq \binom{k+\ell}{k}$.*

Proof. Let $U := \cup_{i\in[h]}(A_i \cup B_i)$ and consider a random order π of U. Let E_i be the event that the elements of A_i *precede* those in B_i wrt order π.

Clearly, $P(E_i) = \binom{k+\ell}{k}^{-1}$, for $i \in [h]$.

We claim that E_i, E_j, for $i \neq j$, are mutually *exclusive*. Suppose not, then the elements of A_i (resp. A_j) *precede* those in B_i (resp. B_j). Wlog, if the last element in A_i does not come after that of A_j, then all of A_i precedes all of B_j, implying $A_i \cap B_j = \emptyset$, which contradicts the definition of $\mathcal{F}$. Thus,

$$1 \;\geq\; P(\cup_{i\in[h]} E_i) \;=\; \sum_{i\in[h]} P(E_i) \;=\; h / \binom{k+\ell}{k} .$$

□

2 Super-concentrators

This is a graph theory term inspired from telecommunications. We will see a construction from [4].

A *super-concentrator* is a directed acyclic graph $G = (V, E)$, with n special input nodes $I \subset V$ and n output nodes $O \subset V$. It satisfies the property that for any $1 \leq k \leq n$, *any* k subset of I is connected with *any* other k subset of O with k disjoint paths.

Note 1. This captures the physical situation where k pairs of people want to talk to each other simultaneously on the telephone.

This is a very strong connectivity property and can be used to design robust networks. It is very easy to design a super-concentrator with $O(n^2)$ edges (Exercise). We will see how linear size (number of edges) super-concentrators can be constructed. We will do it recursively. But we first need "concentrators".

An (n_1, n_2, u) *concentrator* has n_1 input nodes and n_2 output nodes. For any $1 \leq k \leq u$ and an input k-subset S, *there exists* an output k-subset T which is connected to S using k disjoint paths.

Exercise 1. In concentrators, number of input and output nodes could be different and there is an upper limit on size k subset which we are allowed to choose. What is the other essential difference between concentrator and super-concentrator?

Quantification $\forall S \exists T$ has been weakened to $\exists S \forall T$.

* Edited from Rajat Mittal's notes.

References

1. N. Alon and J. H. Spencer. The Probabilistic Method. *Wiley*, 2008.
2. H. Tijms Understanding Probability. *Cambridge University Press*, 2012.
3. D. Stirzaker. Elementary Probability. *Cambridge University Press*, 2003.
4. U. Schöning. Gems of Theoretical Computer Science. *Springer-Verlag*, 1998.