Undergraduate Colloquium

Ref: The Probabilistic Method - Alon & Spencer

Example 1. Ramsey numbers.

Def. The Ramsey number $R(k,\ell)$ is the minimum n such that at any party with n or more people there must either be

① a group of k people all of whom know one another or

② a group of ℓ people all of whome are strangers to each other.

Thm. $R(k,\ell)$ exists

Focus on $R(k,k)$ and get a lower bound.

Idea. Consider a random party where any pair of people know each other with probability $1/2$.

Given a set of k people (out of n) the probability that either they all know each other or are all strangers is $\frac{1}{2^{\binom{k}{2}}} \cdot \frac{1}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$.

Hence the probability that some set of k people satisfy ① or ② is at most $p = \binom{n}{k} 2^{1-\binom{k}{2}}$.

So if n is such that $p<1$, then with positive probability neither ① or ② occurs, and thus $R(k,k) > n$.

For $k \geq 3$, $n = \lfloor 2^{k/2} \rfloor$ makes $p<1$. Hence $R(k,k) > \lfloor 2^{k/2} \rfloor$, so $R(k,k)$ has an exponential lower bound.

Expectation.

Let S be a finite set of outcomes where the probability of outcome $s \in S$ is $pr(s)$. Then for $A \subset S$ the probability of an outcome belonging to A is $\sum_{s \in A} p_s =: pr(A)$. $\quad (pr(S) = 1)$

Suppose $X : S \to \mathbb{R}$ is a function. Then X takes value r with probability $pr(\{s : X(s) = r\})$. The expected (or average) value of X is

$$E[X] = \sum_{r \in \mathbb{R}} r \cdot pr(\{s : X(s) = r\})$$

Key observations.

1. If $X, Y : S \to \mathbb{R}$ are functions then $E[X+Y] = E[X] + E[Y]$.
2. X must take a value less than or equal to $E[X]$ and also a value greater than or equal to $E[X]$.

Example 2. Sum-free sets

Def. A subset $A \subset \mathbb{Z}$ is <u>sum-free</u> if there do not exist $a, b, c \in A$ such that $a + b = c$.

Th. (Erdős) Any set of n nonzero integers contains a sum-free subset of size at least $n/3$.

Proof. Fix a prime number $p = 3k+2$ such that $p > 2\max\{|a| : a \in A\}$. Define a set

$$C = \{k+1, k+2, \dots, 2k+1\}$$

~~of residue classes~~ Note that C is sum-free modulo p even. Now choose $x \in \{1, 2, \dots, p-1\}$ uniformly at random. Let us write $A = \{a_1, \dots, a_n\}$. Then for each x let

$$F(x) = \#\{i : xa_i \in C \text{ modulo } p\}.$$

Notice that if we put $F_i(x) = \begin{cases} 1 & xa_i \in C \bmod p \\ 0 & \text{otherwise} \end{cases}$ then $F = \sum_{i=1}^{n} F_i$. The expected number of xa_i with residue classes in C is

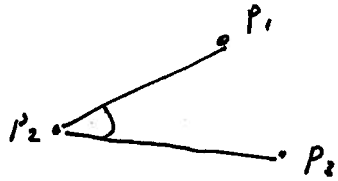
$$E[F] = \sum_{i=1}^{n} E[F_i] = \sum_{i=1}^{n} \frac{\#C}{p-1} = n\frac{k+1}{3k+1} > \frac{n}{3}.$$

Thus there must be some x such that more than $n/3$ of the numbers $\{xa_i\}$ belong to C modulo p. For this x let $S = \{a_i : xa_i \in C \bmod p\}$. We claim that S is sum free. If $a, b, c \in S$ satisfy $a+b=c$ then $xa + xb = xc \bmod p$. Since C is sum-free mod p this is impossible.

Example 3. Small angles

Th. (Erdős & Füredi)

For every $d \geq 1$ there is a set of $m = \lfloor \frac{1}{2}(\frac{2}{\sqrt{3}})^d \rfloor$ points in $\mathbb{R}^d$ such that every triple of points determines an acute angle:

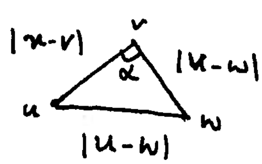


Proof. Let's look for these points in the set

$$V = \{(\varepsilon_1, \dots, \varepsilon_d) : \varepsilon_i = 0, 1\}$$

of vertices of the d-dimensional unit cube. Given three such points $u, v, w \in V$ observe that

$$|u-w|^2 \leq |u-v|^2 + |v-w|^2$$



$\Rightarrow$ α is at most $\pi/2$ by the law of cosines

After all $|u-w|^2 = \#\{i : u_i \neq w_i\}$ and similarly for the other two. If $u_i \neq w_i$ then either $u_i \neq v_i$ or $v_i \neq w_i$, and if $u_i = w_i$ then it can be the case that $u_i \neq v_i$ and $v_i \neq w_i$.

Hence u, v, w make a right angle unless there is some i such that $u_i = w_i$ but $u_i \neq v_i$ and $v_i \neq w_i$.

Choose $2m$ random vectors $v_1, \dots, v_{2m}$. This means for each i and j flip a coin to determine if $(v_i)_j$ is 0 or 1. Given three vectors v_i, v_j, v_k, they form a right angle ~~unless~~ if for each ℓ either $(v_i)_\ell \neq (v_k)_\ell$ or $(v_i)_\ell = (v_k)_\ell = (v_j)_\ell$. Out of the eight possibilities for $(v_i)_\ell, (v_j)_\ell, (v_k)_\ell$, six form a right angle.

Hence the probability that v_i, v_j, v_k form a right angle is $(3/4)^d$. This means that the expected number of right angles is $3\binom{2m}{3} \cdot (3/4)^d$. The value of m was chosen so that $3\binom{2m}{3}(3/4)^d \leq m$.

This implies that there is some collection of vectors $v_1, \dots, v_{2m}$ with at most m right angles. Removing at most m vectors, one from each right angle gives a set of m vectors where all of the angles are acute.

Note: If there is a duplication v_i, v_j, v_k where $v_i = v_j$ then by our criterion the angle is considered to be a right angle. So all of our vectors above are distinct.

Example 4. ~~Game theory~~ The Liar Game

An imp challenges you to guess a number it has chosen between 1 and n. You may ask q questions. Each question is of the form "is $x \in S$?" where $S \subset \{1, \dots, n\}$ is a subset you specified. The imp has a number $k \le q$ on its forehead which means it can lie in its answers up to k times. Also being an imp, it can answer consistently with several numbers, then at the end if there are more than one ~~choose~~ claim to choose one that you did not guess.

Is it worth accepting the imp's challenge? (If $k=0$ then you have a winning strategy when $q > \log_2 n$)

Theorem. If $n > \dfrac{2^q}{\sum_{i=0}^{k} \binom{q}{i}}$ then the imp has a winning strategy for the Liar's game.

Proof. Suppose the imp plays randomly, flipping a coin to determine whether or not to lie. We want to see if there is a chance the imp can beat you without breaking the rules.

For each $x \in \{1, \dots, n\}$ let $I_x = \begin{cases} 1 & \text{the imp lies about whether } x \in S \text{ at most } k \text{ times} \\ 0 & \text{otherwise} \end{cases}$

For a given pattern of questions and ~~answers~~ $I_x = 1$ iff x is consistent ~~and the~~ with the imps answers given the rules of the game.

The imp can win if the expected number of such x is greater than 1. Notice that

$$E[I_x] = \underbrace{2^{-q}}_{\text{probability of each outcome}} \underbrace{\sum_{i=0}^{k} \binom{q}{i}}_{\text{which ones are lies}}$$

Hence the expected number of consistent x after a random game is

$$E\Big[\sum_{x=1}^{n} I_x\Big] = \sum_{x=1}^{n} E[I_x] = n\, 2^{-q} \sum_{i=0}^{k} \binom{q}{i},$$

So if $n > \dfrac{2^q}{\sum_{i=0}^{k} \binom{q}{i}}$ there is always a pattern of truth and lies which results in more than one consistent x. So the imp can always win.