Winter School 2011

Thursday, January 13, 2011
11:51 AM

Introduction to Probabilistic Methods

Outline

① Introduction , Ramsey number

② Extremal combinatorics : Turan's theorem
Erdös-ko-Rado theorem

③ Alternation : dominating set , graph coloring

④ Shannon's coding theorem

Ramsey number

Given a BIG complete graph where each edge is either red or blue , is it true that there must be a big red clique or a big blue clique ?

Let $R(k,k)$ be the minimum number so that if a complete graph has at least $R(k,k)$ vertices then there must be a red clique of size at least k or a blue clique of size k for any two coloring of its edges .

e.g. $R(3,3) = 6$

<u>Question</u> : Is $R(k,k)$ finite ?

[Ramsey 1929] $R(k,k)$ is finite for any k.

<u>Question</u> : How big is $R(k,k)$?

Exercise : $R(k,k) \leq 2^{2k}$

Is it tight ?

Can we find a 2-coloring of the edges of a large

Can we find a 2-coloring of the edges of a large graph so that there is no red clique of size k and no blue clique of size k?

It turns out that this is an extremely difficult question

But a simple probabilistic method shows $R(k,k) \geq \lfloor 2^{k/2} \rfloor$

Theorem: $R(k,k) \geq \lfloor 2^{k/2} \rfloor$

Proof: Consider a "random" 2-coloring of the edges of a complete graph with n vertices.

What is the probability that a specific subset R of k vertices is "monochromatic" (i.e. all red or all blue)?

This probability is $2 \cdot 2^{-\binom{k}{2}}$.

Since there are $\binom{n}{k}$ subsets of size k, the probability that **some** subset of size k is monochromatic is at most $\binom{n}{k} \cdot 2^{1-\binom{k}{2}}$.

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

Then it means that there is a coloring with no monochromatic clique of size k.

Since $\binom{n}{k} 2^{1-\binom{k}{2}} < \frac{n^k}{k!} \cdot \frac{2^{1+\frac{k}{2}}}{2^{k^2/2}} = \frac{2^{1+\frac{k}{2}}}{k!} \cdot \left(\frac{n}{2^{k/2}}\right)^k$,

any $n \leq 2^{k/2}$ will imply that $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$.

Therefore, $R(k,k) \geq \lfloor 2^{k/2} \rfloor$ ■

The proof shows that the number of 2-colorings is more than the number of 2-colorings with a monochromatic clique of size k.

The proof is non-constructive.

Extremal Combinatorics:

Characterize the "extreme" objects satisfying some constraints.

e.g. How many edges can a graph have if there is no triangle?

How many edges can a graph have if there is no 4-cycle?

(The answer to the first question is $n^2/4$,

while the answer to the second question is $\Theta(n^{1.5})$.)

Question: What is the minimum number of edges in an n-vertex graph such that there is no independent set of size $k+1$?

Example: Divide the vertices into k cliques of same size.
(assuming n/k is an integer)

Then this graph has no independent set of size $k+1$, and the number of edges is $k\binom{n/k}{2} = \frac{n(n-k)}{2k}$

[Turán 1941] This is the extreme example.

To prove Turán's theorem, we first prove the following using a probabilistic argument.

Theorem: Let $\alpha(G)$ be the size of maximum independent set in G, and d_v be the degree of v.

Then $\alpha(G) \geq \sum_v \frac{1}{d_v+1}$

Proof: Consider a random ordering of the vertices.

Take the subset S of vertices with no neighbor "in front" in the ordering.

"in front" in the ordering.

Then S is an independent set.

What is the size of S?

It is a random variable (depending on the ordering) and let's compute its expected value.

Let $S_v = \begin{cases} 1 & \text{if } v \text{ is in } S \\ 0 & \text{otherwise} \end{cases}$

Then $E[|S|] = E[\sum_v S_v] = \sum_v E[S_v]$.

where the last equality is called "linearity of expectation."

Since we pick a random ordering, the probability that no neighbor of v is in front of v is $\frac{1}{d_v+1}$, and hence $E[S_v] = \frac{1}{d_v+1}$, and the theorem follows. ■

<u>Proof of Turán</u>: Note that $\sum_v \frac{1}{d_v+1}$ is minimized when d_v are as close as possible.

So, for a graph with m edges,

$$\alpha(G) \geq \sum_v \frac{1}{d_v+1} \geq n\left(\frac{1}{\frac{2m}{n}+1}\right) = \frac{n^2}{2m+n}$$

Since $\alpha(G) \leq k$, this implies that

$$k \geq \frac{n^2}{2m+n} \Rightarrow m \geq \frac{n(n-k)}{2k}$$ ■

There are many other beautiful proofs of this theorem.

Alternation: probabilistic method + deterministic changes

Dominating set: a subset of vertices S such that

Dominating set: a subset of vertices S such that every vertex has a neighbor in S.

Theorem: Let G be a graph with n vertices and minimum degree d. Then G has a dominating set of at most $n\left(\frac{1+\ln(d+1)}{d+1}\right)$ vertices.

Proof: Construct S by picking each vertex with probability p.

Then $E[|S|] = np$.

Let T be the set of vertices with no neighbor in S.

Let $T_v = \begin{cases} 1 & \text{if } v \in T \\ 0 & \text{otherwise.} \end{cases}$

Then $E[|T|] = E[\sum_v T_v] = \sum_v E[T_v]$

The probability that $v \in T$ is at most $(1-p)^{d+1}$

So, $E[|T|] \leq n(1-p)^{d+1}$

Let $D = S \cup T$. Note that D is a dominating set.

$$|D| = |S| + |T| \leq np + n(1-p)^{d+1}$$

$$\leq np + ne^{-p(d+1)} \quad \text{since } 1-x \leq e^{-x}$$

Set $p = \frac{\ln(d+1)}{d+1}$

Then $|D| \leq n\left(\frac{1+\ln(d+1)}{d+1}\right)$ ■

This bound can be shown to be near optimal.

Graph coloring: One of the best examples in probabilistic method

Task: use minimum number of colors to color all the vertices so that adjacent vertices have different colors.

Question: When is the chromatic number high?

One possible reason: the graph has a large clique.

It turns out that there are examples with no triangle and have arbitrarily high chromatic number (see exercise).

The following theorem is even more surprising!

Theorem: For all k, ℓ, there exist graphs with no cycles of length at most ℓ and the chromatic number $> k$.

Proof: Consider a random graph where each edge is chosen with probability p.

The idea to prove that the graph has high chromatic number is to establish that every independent set is small.

Note that $\chi(G) \geq \frac{n}{\alpha(G)}$.

First, if we set $t = \lceil \frac{3}{p} \ln n \rceil$, we show that

$$\begin{aligned} \Pr(\alpha(G) \geq t) &\leq \binom{n}{t} (1-p)^{\binom{t}{2}} \\ &< n^t e^{-p\binom{t}{2}} \\ &< \left(n e^{-p(t-1)/2}\right)^t \\ &< \frac{1}{2} \quad \text{for sufficiently large } n. \end{aligned}$$

Now, let's bound the number of cycles of length $\leq \ell$

$$E[X] = \sum_{i=3}^{\ell} \binom{n}{i} \frac{i!}{2i} p^i \leq \sum_{i=3}^{\ell} \frac{n^i}{2i} p^i$$

Set $p = n^{\varepsilon - 1}$ where $\varepsilon < \frac{1}{\ell}$

Then $E[X] \le \sum_{i=3}^{\ell} \frac{n^{\varepsilon i}}{2i} = o(n)$.

In particular, $\Pr(X \ge \frac{n}{2}) < \frac{1}{2}$,

because otherwise $E[X] = \Omega(n)$.

So, with positive probability, the graph has less than $n/2$ "short" (length $\le \ell$) cycles and $\alpha(G) < 3n^{1-\varepsilon} \ln n$.

Now, delete one vertex from each short cycle to obtain G^*.

Then G^* has at least $n/2$ vertices, $\alpha(G^*) < \alpha(G)$, and G^* has no short cycles.

$$\chi(G^*) \ge \frac{|G^*|}{\alpha(G^*)} \ge \frac{n/2}{3n^{1-\varepsilon} \ln n} = \frac{n^{\varepsilon}}{6 \ln n}$$

It is larger than any k for n sufficiently large. ■

A warmup for Shannon's coding theorem

Setting: Alice wants to send a message to Bob through a noisy channel, where each bit is flipped independently with probability p.

<u>Question:</u> Can they come up with a scheme so that Bob can figure out what Alice wanted to send?

Naive solution: Alice sends each bit multiple times, and Bob takes the majority of each bit as the original message

This works to decrease the decoding error.

But, as the decoding error tends to zero, the rate tends to zero too! (rate = # of bits received / # of bits sent) too ! (rate = # of bits received / # of bits sent)

Question: Is there a scheme with decoding error tends to zero where the rate is a positive constant ?

[Shannon] : Yes, and prove matching upper and lower bounds.

We will give a positive answer to a modified setting.

Modified setting: ① at most pn errors when transmitting n bits
(stronger assumption)
② no decoding error
(stronger requirement)

Suppose Alice wants to send a message of m bits.
Instead of sending it directly, Alice will "encode" this to an n-bit message and send it through the channel, where $n > m$.
The encoding function $f : \{0,1\}^m \to \{0,1\}^n$ is a "random" function.
That is, each $f(x)$ uniformly and independently chosen.

Assumption: Alice and Bob both know the function f, say as an explicit table.

For two strings, the Hamming distance is the # of bits different.
If we can prove that the Hamming distance between any two codewords (i.e $f(x)$ & $f(y)$) are greater than $2pn$.
Then Bob can simply decode by finding the "nearest" codeword.

So, what is the probability that for a random function f, every two codewords have distance greater than $2pn$?

There are $\binom{2^m}{2}$ pairs of codewords.

For each pair, the probability that the distance $\leq 2pn$ is at most $\left(\sum_{i \leq 2pn} \binom{n}{i}\right) 2^{-n}$

So if $\binom{2^m}{2} \cdot \left(\sum_{i \leq 2pn} \binom{n}{i}\right) \cdot 2^{-n} < 1$

then there is a function f satisfying this property.

Note that $\sum_{i \leq 2pn} \binom{n}{i} \approx 2^{n H(2p)}$

where $H(q) = -q \log_2 q - (1-q) \log_2 (1-q)$ is the entropy function.

So, $\binom{2^m}{2} \cdot \left(\sum_{i \leq 2pn} \binom{n}{i}\right) \cdot 2^{-n}$

$$\approx 2^{2m} \cdot 2^{n H(2p)} \cdot 2^{-n}$$

$$= 2^{2m + n H(2p) - n} < 1$$

when $m < \frac{n}{2}(1 - H(2p))$

When $p < \frac{1}{4}$, the rate is a constant.

References: Alon, Spencer, "The Probabilistic Method"

The material here is just the most basic in the book.

There are many other techniques in the book.