

We have already seen a proof of the existence of splitter graphs that uses a probabilistic construction. This basic idea, often called the *Probabilistic Method*, is very powerful, and can be used to derive several results in a surprisingly simple manner.

We see a few examples of this.

1 Lower bound on some Ramsey Numbers

The Ramsey number $R(k, l)$ is defined as the minimum n such that every edge colouring of K_n , the complete graph on n vertices using 2 colours blue and red contains either a blue K_k or a red K_l as a subgraph.

As an example, consider $R(3, 3)$. You should be able to prove by trial and error that $R(3, 3) = 6$. Incidentally this is related to the following puzzle: in any party of 6 persons there are either 3 persons who know each other, or 3 none of whom knows the other two.

The exercises ask you to prove that $R(k, l)$ is well defined and finite. Here we only show that the symmetric Ramsey numbers, i.e. with $k = l$ are rather large using a probabilistic argument.

Theorem 1 $R(k, k) > 2^{k/2}$ for $k \geq 3$.

Proof: Assume $n \leq 2^{k/2}$. For any such n we will show that there exists a colouring in which there is neither blue nor red K_k . This proves the theorem.

For this it suffices to argue that if we choose a colouring at random, with non-zero probability, it will not contain a monochromatic K_k .

We instead consider the complementary event, i.e. in a random colouring there is a subgraph K_k that is monochromatic. It suffices if we prove that this event happens with probability strictly less than 1. The probability space we consider is the one in which each edge is coloured blue or red independently with equal probability.

Consider the probability that any fixed size k subgraph is blue. For this to happen, all the $\binom{k}{2}$ edges in it must be coloured blue. This happens with probability $2^{-\binom{k}{2}}$. With the same probability the subgraph could be red.

Thus the probability that the subgraph is monochromatic is simply the sum, i.e. $2^{1-\binom{k}{2}}$.

Since there are $\binom{n}{k}$ different subgraphs of size k , the probability that any one of them is monochromatic is at most

$$\binom{n}{k} 2^{1-\binom{k}{2}} \leq \frac{n^k}{k!} \frac{2^{1+k/2}}{2^{k^2/2}} \leq \frac{2^{1+k/2}}{k!}$$

where the last inequality follows using $n \leq 2^{k/2}$. Clearly, this probability is less than one for $k \geq 3$. ■

In fact note that the probability that the randomly constructed graph has a monochromatic K_k as evaluated above tends to zero as k tends to infinity.

2 Chromatic number vs. Clique number

The chromatic number $\chi(G)$ of a graph G is defined as the minimum number of colours needed to colour its vertices such that adjacent vertices receive distinct colours. The clique number $\omega(G)$ of a graph G is defined as the size of the largest clique, i.e. complete subgraph, in G . Since every vertex in a clique must have a different colour in a valid colouring of a graph, clearly $\chi(G) \geq \omega(G)$.

An interesting question is, how much larger can $\chi(G)$ be than $\omega(G)$? It is not easy to construct graphs in which these numbers differ a lot. A simple example, though, is any odd cycle on 5 or more vertices. Clearly for this the clique number is 2, while the chromatic number is 3. It is tempting to regard the presence of a clique as a “fundamental” reason for having a large chromatic number. But this is very far from the truth!

For proving this it is convenient to define another graph parameter $\alpha(G)$ which gives the size of the largest independent set of the graph, i.e. the largest subset of vertices which do not contain an edge among them. Clearly, in any legal vertex colouring, the vertices having the same colour must be independent. Thus every graph G must have an independent set of size at least $n/\chi(G)$. In other words $\alpha(G) \geq n/\chi(G)$.

We will say that a graph property holds for almost all graphs if the probability of a random graph of a given size n tends to 1 as n tends to ∞ .

Theorem 2 *Almost all graphs have $\alpha \leq 2 \log n$, $\omega \leq 2 \log n$, and $\chi \geq n/2 \log n$.*

Proof: . Actually, we have essentially proved this already. Choosing a random colouring in the last theorem is equivalent to choosing a random graph in the following way. If an edge of K_n is coloured blue, we will consider it to be present in the random graph being constructed, and if it was coloured red, we will consider it to be absent. The probability of finding a blue K_k is simply the probability of having K_k in the graph constructed; while the probability of finding a red K_k is simply the probability of finding an independent set! Thus the probability of having a monochromatic K_k is simply the probability of having either a clique of size k or an independent set of size k . But this probability we have shown is $2^{1+k/2}/k!$, where we have $n \leq 2^{k/2}$. Thus choosing $k = 2 \log n$ we indeed have the probability tending to 0 and n tends to infinity. ■

3 Dominating set sizes

A subset S of vertices in a graph $G = (V, E)$ is said to be a dominating set if every vertex in $V - S$ has a neighbour in S .

Theorem 3 *Every graph G on n vertices and minimum degree δ has a dominating set of size at most $n[1 + \ln(\delta + 1)]/(\delta + 1)$.*

Proof: Here is a random construction for a dominating set S . First we select each vertex into S independently with probability p to be determined later. Let X denote the set of vertices thus selected. Let Y denote the vertices dominated by X , i.e. the vertices in X as well as their neighbours. Let $Z = V - Y$. Set $S = X \cup Z$. Clearly S is a dominating set.

We next estimate $|S|$. Since each vertex has a probability p of being in X , the expected size of X is np . Next, let us consider the probability that a vertex belongs to Z . For this, neither it nor its neighbours should get into X . There are at least δ neighbours, and thus the probability that any vertex nor its neighbours gets selected into X is at most $(1 - p)^{\delta+1}$. Thus the expected size of Z is $n(1 - p)^{\delta+1}$. So the expected size of S is at most:

$$np + n(1 - p)^{\delta+1}$$

Now we simply select p so as to minimize this number. It is convenient to approximate $1 - p \leq e^{-p}$, which is a good approximation for small p . Thus we wish to minimize $np + ne^{-p(\delta+1)}$. Setting the derivative to zero we get $n - n(\delta + 1)e^{-p(\delta+1)} = 0$. In other words $p = [\ln(\delta + 1)]/(\delta + 1)$. Thus we get

$$E[|S|] \leq np + ne^{-p(\delta+1)} = n[1 + \ln(\delta + 1)]/(\delta + 1)$$

But then there must be at least one choice in which S has no more than its expected size. ■

4 Summary

It is worth noting that We have used two different probabilistic arguments.

In the first, we used a counting argument to estimate the probability of a certain property, and from that we asserted that objects with that property must exist (if the probability was shown to be strictly non-zero), or objects without that property must exist (if the probability was shown to be strictly smaller than 1).

In the second, we considered the existence of objects having a certain value for certain property, e.g. the size of a certain set. We then computed the expected value over randomly chosen objects. We could then assert that objects with value at least (or at most) equal to the expected value had to exist.

So please brush up your probability theory, and your knowledge of random variables. Not much is needed, mostly only linearity of expectations. But you should fully follow why the expected number of nodes in Z was $n(1 - p)^{\delta+1}$, for example.

Finally dont be afraid if there arise big quantities in your calculations – they may get annihilated by very small quantities and in the end you may have a sensible result.

Exercises

1. Let $G = (V, E)$ be a graph with n vertices and e edges. Show that G contains a bipartite subgraph having at least $e/2$ edge. Hint: Randomly divide V into two parts, and estimate the number of edges that must go from one part to the other.

2. Suppose we have a graph G which we want to assign one of 3 colours to each vertex. It may not be possible to assign the colours such that adjacent vertices get different colours. But our goal is more modest. We are happy if two thirds of the edges have distinct colours on their ends. Show that such a colouring exists. Hint: Try the simplest idea first.
3. Prove that almost all graphs on n vertices have a connected component of size at least $n - 2\log n$. More credit if you show an even bigger component. Hint: Several ways of doing this. What is the probability that vertex 1 does not have a short path to vertex k ? What is the probability that there is no edge from vertices $1, \dots, k$ to $k + 1, \dots, n$, for example?