The Probabilistic Method

exercise solutions by

Marcel K. Goh

15 February 2022

Note. A star indicates a starred problem in the text; these are supposed to be harder. I have decided to put my work here even when I haven't solved the problem yet, so that when I come back later I can see what I've tried. I won't put the black square ▮ until I think the proof is actually done. Also I wrote log everywhere to mean $\log_e$ even though the text uses ln.

1. The Basic Method

Exercise 1.1. *Prove that if there is a real p, with $0 \le p \le 1$ such that*

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} < 1,$$

then the Ramsey number $r(k,t)$ satisfies $r(k,t) > n$. Using this, show that

$$r(4,t) \ge \Omega\bigl(t^{3/2}/(\log t)^{3/2}\bigr).$$

Proof. We follow the proof of Proposition 1.1.1 in the book. We consider a random graph on n vertices, where each edge is present with probability p. Let K be the event that there is a clique of size k in the graph, and let I be the event that there is an independent set of size t in the graph. By the union bound,

$$\mathbf{P}\{K \cup I\} \le \mathbf{P}\{K\} + \mathbf{P}\{I\} \le \sum_{|S|=k} p^{\binom{k}{2}} + \sum_{|S|=t} (1-p)^{\binom{t}{2}} = \binom{n}{k} p^{\binom{k}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} < 1.$$

This means that $\mathbf{P}\{\neg K \cap \neg I\} > 0$ and since the sample space is finite, there exists a graph on n vertices with no clique of size k and no independent set of size t and therefore $r(k,t) > n$.

Next we show that $r(4,t) > \bigl(t/(e\log t)\bigr)^{3/2}$ for large enough t. Note that

$$\binom{n}{4} p^{\binom{k}{2}} + \binom{n}{t}(1-p)^{\binom{t}{2}} \le n^4 p^6 + \frac{e^t n^t}{t^t}(1-p)^{t^2/4},$$

by the inequalities

$$\frac{n^k}{k^k} \le \binom{n}{k} \le \frac{e^k n^k}{k^k}.$$

Setting $n = t^{3/2}/(e\log t)^{3/2}$, we have

$$\begin{aligned}
n^4 p^6 + \frac{e^t n^t}{t^t}(1-p)^{t^2/4} &= \left(\frac{tp}{e\log t}\right)^6 + \frac{e^t t^{3t/2}}{t^t e^{3t/2}(\log t)^{3t/2}}(1-p)^{t^2/4} \\
&= \left(\frac{tp}{e\log t}\right)^6 + \frac{t^{t/2}}{e^{t/2}(\log t)^{3t/2}}(1-p)^{t^2/4} \\
&\le \left(\frac{tp}{e\log t}\right)^6 + \left(\frac{t(1-p)^{t/2}}{e(\log t)^3}\right)^{t/2} \\
&\le \left(\frac{tp}{e\log t}\right)^6 + \left(\frac{t}{e^{pt/2+1}(\log t)^3}\right)^{t/2},
\end{aligned}$$

where in the last line we used the inequality $1 - p \le e^{-p}$. Choosing $p = 2\log t/t$, we simply need t large enough such that

$$\left(\frac{t}{e^{\log t+1}(\log t)^3}\right)^{t/2} = \left(\frac{1}{e(\log t)^3}\right)^{t/2} < 1 - \left(\frac{2}{e}\right)^6,$$

which can be done since the left-hand side goes to 0. ∎

Exercise 1.2. *Suppose $n \ge 4$ and let H be an n-uniform hypergraph with at most $4^{n-1}/3^n$ edges. Prove that there is a colouring of the vertices of H by 4 colours so that in every edge all 4 colours are represented.*

Proof. Let each vertex of H be independently given one of the four colours uniformly at random. (If H is infinite, it does not matter what colour we give to vertices that do not appear in any edge, so it suffices to consider H finite, which makes the sample space finite.) Given some edge e of H with n vertices, there are 4^n total ways that e may be coloured, and for each of the four colours, 3^n total ways that e may be coloured using only the other three colours. Let $K(e)$ denote the event that e does not contain all four colours. By the inclusion-exclusion principle,

$$\mathbf{P}\big\{K(e)\big\} = 4 \cdot 3^n - 6 \cdot 2^n + 4$$

Since $6 \cdot 2^n \ge 96 > 4$, the probability that a given edge *does not* contain all four colours is (much) less than $3^n/4^{n-1}$. By the union bound,

$$\mathbf{P}\Big\{\bigcup_{e\in E(H)} K(e)\Big\} \le \sum_{e\in E(H)} \mathbf{P}\big\{K(e)\big\} < \frac{4^{n-1}}{3^n} \cdot \frac{3^n}{4^{n-1}} = 1.$$

Since the sample space is finite this implies that there is some colouring of the vertices of H in which every edge has all four colours. ∎

***Exercise 1.3.** *Prove that for two independent, identically distributed real random variables X and Y,*

$$\mathbf{P}\big\{|X-Y| \le 2\big\} \le 3\,\mathbf{P}\big\{|X-Y| \le 1\big\}.$$

Proof. [Not done. Need to use independence somehow. I think by symmetry it is enough to show that $\mathbf{P}\{1 < X - Y \le 2\} \le \mathbf{P}\big\{|X-Y| \le 1\big\}$ or $\mathbf{P}\big\{X - Y > 1 \,\big|\, |X-Y| \le 2\big\} \le 1/3$. In the second expression, X and Y are no longer independent.]

***Exercise 1.4.** *Let $G = (V, E)$ be a graph with n vertices and minimum degree $\delta > 10$. Prove that there is a partition of V into two disjoint subsets A and B so that $|A| \le O\big((n\log n)/\delta\big)$, and each vertex of B has at least one neighbour in A and at least one neighbour in B.*

Proof. We follow the construction of the dominating set from the proof of Theorem 1.2.2, since the required A here is a dominating set with some extra conditions. Let p be chosen later and let X be a random set of vertices obtained by independently selecting each $v \in V$ with probability p. Then as in the proof from the textbook, we let Y be the set of all $v \in V \setminus X$ that have no neighbours in X. At this point $X \cup Y$ is a dominating set, but we are not yet done constructing A, since there may still be elements in $V \setminus (X \cup Y)$ all of whose neighbours belong to $X \cup Y$. Let $Z \subseteq V \setminus (X \cup Y)$ denote all of these elements. For any given $v \in V$, we have $\mathbf{P}\{v \in X\} = p$ and

$$\mathbf{P}\{v \in Y\} = (1-p)^{\deg(v)+1} \le (1-p)^{\delta+1} \le e^{-p(\delta+1)},$$

since for $v \in Y$ we need v itself as well as all $\deg(v)$ of its neighbours to not be in X. Lastly,

$$\mathbf{P}\{v \in Z\} = (1-p)\prod_{w\in N(v)}\big(p + (1-p)^{\deg(w)+1}\big) \le (1-p)\big(p + (1-p)^{\delta+1}\big)^{\delta} \le (1-p)(p + e^{-p(\delta+1)})^{\delta}.$$

Now we compute

$$\begin{aligned}\mathbf{E}\big\{|A|\big\} &= \sum_{v\in V} v \in \{X \cup Y \cup Z\} \\ &\le n\big(p + e^{-p(\delta+1)} + (1-p)(p + e^{-p(\delta+1)})^{\delta}\big).\end{aligned}$$

Since $p + e^{-p(\delta+1)} < 1$, we can remove the δ from its exponent for the looser but simpler bound

$$\begin{aligned}\mathbf{E}\{|A|\} &\le np + ne^{-p(\delta+1)} + np - np^2 + ne^{-p(\delta+1)} - npe^{-p(\delta+1)} \\ &= (2-p)(np + ne^{-p(\delta+1)}).\end{aligned}$$

Setting $p = \log(\delta+1)/(\delta+1)$ just as in the dominating set proof, we have

$$e^{-p(\delta+1)} = \frac{1}{\delta+1}$$

and

$$\mathbf{E}\{|A|\} \le \left(2 - \frac{\log(\delta+1)}{\delta+1}\right)\left(n\frac{\log(\delta+1)}{\delta+1} + \frac{n}{\delta+1}\right) = 2n\frac{\log(\delta+1)}{\delta+1} + o\left(\frac{n\log\delta}{\delta}\right),$$

which has the required asymptotics. It remains to choose an A with $|A|$ at least this average. ▮

Exercise 1.8. *Let F be a finite collection of binary strings of finite lengths and assume no member of F is a prefix of another one. Let N_i denote the number of strings of length i in F. Prove that*

$$\sum_{i\ge1} \frac{N_i}{2_i} \le 1.$$

Proof. This inequality looks an awful bit like the LYM inequality, so our proof is informed by the proof of that statement. Let m be the maximum length of any string in F (such a maximum exists because F is finite). Let $b_1, b_2, \ldots, b_m$ be a sequence of independent random variables with $\mathbf{P}\{b_i = 0\} = \mathbf{P}\{b_i = 1\} = 1/2$ for all $1 \le i \le m$. Consider the set of strings

$$C = \{b_1, b_1b_2, \ldots, b_1b_2\cdots b_m\}.$$

This is a random set, so we may study the random quantity $|C \cap F|$. Note that C contains exactly one bitstring of each length, and every string of length i has an equal chance of being in C, namely 2^{-i}. So

$$\mathbf{E}\{|C\cap F|\} = \sum_{s\in F} \mathbf{P}\{s \in C\} = \sum_{s\in F} 2^{-|s|} = \sum_{i\ge1}\frac{N_i}{2^i},$$

where $|s|$ is the length of the string s. In the last equality we simply grouped strings of the same length together. On the other hand, $|C \cap F|$ cannot possibly be greater than 1, since for any two members of C, one must be a prefix of the other. So $\mathbf{E}\{|C \cap F|\} \le 1$ as well, and this observation completes the proof. ▮

2. Linearity of Expectation

Exercise 2.1. *Suppose $n \ge 2$ and let $H = (V, E)$ be an n-uniform hypergraph with $|E| = 4^{n-1}$ edges. Show that there is a colouring of V by 4 colours so that no edge is monochromatic.*

Proof. Let ϕ be a random function from V to $[4]$, so that all 4^n possible colourings are equally likely. Let X be the number of monochromatic edges in H under this colouring; that is,

$$X = \sum_{e\in E} \mathbf{1}_{[e \text{ is monochromatic}]}\,.$$

The probability that a given edge $e = \{v_1, \ldots, v_n\}$ is monochromatic is $1/4^{n-1}$, since $\phi(v_1)$ can be anything, but then $\phi(v_2)$ through to $\phi(v_n)$ each have a $1/4$ chance of matching $\phi(v_1)$. So

$$\mathbf{E}\{X\} = \sum_{e\in E} \mathbf{P}\{e \text{ is monochromatic}\} = \frac{4^{n-1}}{4^{n-1}} = 1.$$

We can produce a colouring with 4^{n-1} monochromatic edges by setting $\phi(v) = 1$ for all $v \in V$, so $\mathbf{P}\{X > 1\} > 0$. Thus for $\mathbf{E}\{X\}$ to equal 1, we must also have $\mathbf{P}\{X < 1\} > 0$, and since X is nonnegative, this means that $\mathbf{P}\{X = 0\} > 0$ and there exists some colouring with no monochromatic edges. ▮

3. Alterations

Exercise 3.1. *As shown in Section 3.1, the Ramsey number $R(k,k)$ satisfies*

$$R(k,k) > n - \binom{n}{k} 2^{1-\binom{k}{2}}$$

for every integer n. Conclude that

$$R(k,k) \geq \bigl(1 - o(1)\bigr)\frac{k}{e}2^{k/2}.$$

Proof. We set $n = (k-1)2^{k/2}/e$ in the above, to get the bound

$$R(k,k) > \frac{k}{e}2^{k/2} - \frac{1}{e}2^{k/2} - \binom{(k-1)2^{k/2}/e}{k} 2^{1-\binom{k}{2}}.$$

Now we apply the inequality $\binom{n}{k} \leq e^k n^k / k^k$, which yields

$$\begin{aligned} R(k,k) &> \frac{k}{e}2^{k/2} - \frac{1}{e}2^{k/2} - \frac{e^k (k-1)^k 2^{k^2/2}/e^k}{k^k} 2^{1-k^2/2+k/2} \\ &= \frac{k}{e}2^{k/2} - \frac{1}{e}2^{k/2} - \left(1 - \frac{1}{k}\right)^k 2^{1+k/2}. \end{aligned}$$

Since $(1-1/k)^k \to 1/e$ as $k \to \infty$,

$$R(k,k) > \frac{k}{e}2^{k/2}\left(1 - \frac{1}{k} - \frac{2e(1-1/k)^k}{k}\right) = \frac{k}{e}2^{k/2}\bigl(1 - o(1)\bigr). \quad \blacksquare$$

4. The Second Moment

Exercise 4.1. *Let X be a random variable taking integral nonnegative values, let $\mathbf{E}\{X^2\}$ denote the expectation of its square, and let $\mathbf{V}\{X\}$ denote its variance. Prove that*

$$\mathbf{P}\{X = 0\} \leq \frac{\mathbf{V}\{X\}}{\mathbf{E}\{X^2\}}.$$

Proof. Since X is integer and nonnegative, we have $\mathbf{P}\{X = 0\} = 1 - \mathbf{P}\{X \geq 1\}$ and since $\mathbf{V}\{X\} = \mathbf{E}\{X^2\} - \mathbf{E}\{X\}^2$, to get our result it suffices to show that

$$\mathbf{P}\{X \geq 1\} \geq \frac{\mathbf{E}\{X\}^2}{\mathbf{E}\{X^2\}}.$$

We start by noting that

$$\mathbf{E}\{X\} = \sum_{k=0}^{\infty} k\,\mathbf{P}\{X = k\} = \mathbf{P}\{X \geq 1\} \sum_{k=1}^{\infty} \frac{k\,\mathbf{P}\{X = k\}}{\mathbf{P}\{X \geq 1\}} = \mathbf{P}\{X \geq 1\}\,\mathbf{E}\{X \mid X \geq 1\}.$$

Since the function $x \mapsto x^2$ is convex, we have, by Jensen's inequality,

$$\mathbf{E}\{X\}^2 = \mathbf{P}\{X \geq 1\}^2\,\mathbf{E}\{X \mid X \geq 1\}^2 \leq \mathbf{P}\{X \geq 1\}^2\,\mathbf{E}\{X^2 \mid X \geq 1\} = \mathbf{P}\{X \geq 1\}\,\mathbf{E}\{X^2\}.$$

Dividing both sides by $\mathbf{E}\{X^2\}$ gives us what we want. $\blacksquare$

Exercise 4.4. *Let X be a random variable with expectation $\mathbf{E}\{X\} = 0$ and variance $\mathbf{V}\{X\} = \sigma^2$. Prove that for all $\lambda > 0$,*

$$\mathbf{P}\{X \geq \lambda\} \leq \frac{\sigma^2}{\sigma^2 + \lambda^2}.$$

Proof. Let $\lambda > 0$. Note that if $X \geq \lambda$, then X is also positive so in particular, for any $a > 0$, $X + a$ and $\lambda + a$ are both positive. This implies that $(X + a)^2 \geq (\lambda + a)^2$. Note that since $\mathbf{V}\{X\} = \mathbf{E}\{X^2\} - \mathbf{E}\{X\}^2 = \sigma^2$ and $\mathbf{E}\{X\} = 0$, we have $\mathbf{E}\{X^2\} = \sigma^2$, which in turn implies that

$$\mathbf{E}\big\{(X + a)^2\big\} = \mathbf{E}\{X^2\} + 2a\,\mathbf{E}\{X\} + \mathbf{E}\{a^2\} = \sigma^2 + a^2.$$

Now we take the infimum over all $a > 0$ and apply Markov's inequality to get

$$\mathbf{P}\{X \geq \lambda\} \leq \inf_{a>0} \mathbf{P}\big\{(X + a)^2 \geq (\lambda + a)^2\big\} \leq \inf_{a>0} \frac{\sigma^2 + a^2}{(\lambda + a)^2}.$$

We now optimise over $a > 0$. We have

$$\frac{d}{da}\frac{\sigma^2 + a^2}{(\lambda + a)^2} = \frac{(\lambda + a)^2 2a - (\sigma^2 + a^2)2(\lambda + a)}{(\lambda + a)^4},$$

which is zero when $a = \sigma^2/\lambda > 0$. Plugging in this value of a gives

$$\mathbf{P}\{X \geq \lambda\} \leq \frac{\sigma^2 + \sigma^4/\lambda^2}{(\lambda + \sigma^2/\lambda)^2} = \frac{\sigma^2(\lambda + \sigma^2/\lambda)}{\lambda(\lambda + \sigma^2/\lambda)^2} = \frac{\sigma^2}{\sigma^2 + \lambda^2},$$

exactly what we need. ∎