**Ramsey number** The Ramsey number $R(k, l)$ is the smallest number $n$ such that any red-blue two-coloring of $K_n$'s edges will always create a red $K_k$ or a blue $K_l$.

**Finite sample space** A finite sample space is just some finite set $\Omega$.

**Probability function** Given a finite sample space $\Omega$, a probability function $Pr$ on $\Omega$ is just a map $Pr : \Omega \to [0, 1]$ with the property that

$$\sum_{\omega in \Omega} Pr(\omega) = 1.$$

**Finite probability space** A pair $(\Omega, Pr)$, where $\Omega$ is a finite sample space and $Pr$ is a probability function.

**Uniform distribution** A pair $(\Omega, Pr)$, where $\Omega$ is a finite sample space and $Pr$ is the probability function given by $\Pr(\omega) = 1/|\Omega|$, for every $\omega \in \Omega$.

**Event** An event $A$ is just some subset of a finite sample space.

**Random variable** A random variable $X$ on some finite sample space $\Omega$ is just a map from $\Omega$ to $\mathbb{R}$.

**Expectation** The expectation of a random variable $X$ is the integral of $X$ over $\Omega$. For finite spaces, this is just the sum

$$\sum_{\omega in \Omega} Pr(\omega) \cdot X(\omega).$$

The probabilistic method in combinatorics first arose in 1947, when Erdös used it to prove the following claim:

**Theorem 1** $R(k, k) > \lfloor 2^{k/2} \rfloor$.

**Proof.** Fix some value of $n$, and consider a random uniformly-chosen 2-coloring of $K_n$'s edges: in other words, let us work in the probability space $(\Omega, Pr) =$ (all 2-colorings of $K_n$'s edges, $Pr(\omega) = 1/2^{\binom{n}{2}}$.)

For some fixed set $R$ of $k$ vertices in $V(K_n)$, let $A_R$ be the event that the induced subgraph on $R$ is monochrome. Then, we have that

$$Pr(A_R) = 2 \cdot \left(2^{\binom{n}{2} - \binom{k}{2}}\right) / 2^{\binom{n}{2}} = 2^{1-\binom{k}{2}}.$$

Thus, we have that the probability of at least one of the $A_R$'s occuring is bounded by

$$Pr(\bigcup_{|R|=k} A_R) \leq \sum_{R \subset \Omega, |R|=k} Pr(A_R) = \binom{n}{k} 2^{1-\binom{k}{2}}.$$

If we can show that $\binom{n}{k} 2^{1-\binom{k}{2}}$ is less that 1, then we know that with nonzero probability there will be some 2-coloring $\omega \in \Omega$ in which none of the $A_R$'s occur! In other words, we know that there is a 2-coloring of $K_n$ that avoids both a red and a blue $K_k$.

Solving, we see that

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

whenever $n = \lfloor 2^{k/2} \rfloor, k \geq 3$. So we're done!

# 3 Example 2: Splitting Graphs

We close here with one last example of the probabilistic method:

**Theorem 2** *If $G$ is a graph, then $G$ contains a bipartite subgraph with at least $E/2$ edges.*

**Proof.** Pick a subset of $G$'s vertices, $T$, uniformly at random (i.e. select $T$ by flipping a coin for each of $G$'s vertices, and placing vertices in $T$ iff our coin comes up heads.) Let $B = V(G) \setminus T$.

Call an edge $\{x, y\}$ of $E(G)$ **crossing** iff exactly one of $x, y$ lie in $T$, and let $X$ be the random variable defined by

$$X(T) = \text{number of crossing edges for } T.$$

Then, we have that

$$X(T) = \sum X_{x,y}(T),$$

where $X_{x,y}(T)$ is the 0-1 random variable defined by $X_{x,y}(T) = 1$ if $\{x, y\}$ is an edge of $G$ that's crossing, and 0 otherwise.

The expectation $\mathbb{E}(X_{x,y})$ is clearly $1/2$, because we chose $x$ and $y$ to be in $T$ at random. Thus, by the linearity of expectation, we have that

$$\mathbb{E}(X) = \sum \mathbb{E}(X_{x,y}) = E/2.$$

so the expected number of crossing edges for a random subset of $G$ is $E/2$. Thus, there must be some $T \subset V(G)$ such that $X(T) \geq E/2$; taking the collection of crossing edges this set creates then gives us a bipartite graph $(B, T)$ with $\geq E/2$ edges in it.

example 3 In an university, there are 1600 delegates, who have formed 16000 committees of 80 person each. Prove that one can find two committed having at least four common members.

**Solution.** Sample a pair of committees uniformly at random (i.e., randomly pick one of the $\binom{16000}{2}$ possible pairs). Let $X$ be the number of people who are in both chosen committees. Note that $X = X_1 + \cdots + X_{1600}$, where each $X_i$ is the $\{0,1\}$-random variable telling whether the $i$-th person was in both chosen committees. By linearity of expectation,

$$\mathbb{E}[X] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_{1600}].$$

The magic is that each $\mathbb{E}[X_i]$ is easy to calculate! Let $n_i$ be the number of committees that the $i$-th person belongs to. Then, each $\mathbb{E}[X_i] = \mathbb{P}[i\text{-th person is in both picked committees}] = \binom{n_i}{2} / \binom{16000}{2}$. The only piece of information we know about the $\{n_i\}$ is that their sum $\sum_i n_i = 16000 \cdot 80$, so this suggests that we use convexity to bound $\mathbb{E}[X]$ in terms of the average of $\{n_i\}$, which we denote by $\overline{n} = (16000 \cdot 80)/1600 = 800$:

$$\mathbb{E}[X] \geq 1600 \cdot \binom{\overline{n}}{2} \Big/ \binom{16000}{2} = 1600 \cdot \binom{800}{2} \Big/ \binom{16000}{2} = 1600 \cdot \frac{800 \cdot 799}{16000 \cdot 15999} = 3.995.$$

(One could see that since $799 \approx 800$ and $15999 \approx 16000$, the last fraction is roughly $1/400$.) But by the Lemma, we know that **some outcome of the probabilistic sampling produces an $\boldsymbol{X \geq 3.995}$.** Since $X$ is always an integer, that outcome must in fact have $X \geq 4$. In particular, we conclude that some pair of committees has $\geq 4$ common members. $\square$