In this lecture we treat perhaps the most celebrated application of the probabilistic method.

First, some context. Graph coloring is a very popular topic with a long history. It has connections with scheduling problems and raises very interesting algorithmic questions. Recall that the chromatic number $\chi(G)$ of a graph $G$ is the fewest number of colors required to color its vertices so that adjacent vertices receive different colors. We can estimate $\chi(G)$ in terms of some other basic graph parameters. A clique in $G$ is a subset of pairwise adjacent vertices, while an independent set in $G$ is a subset of pairwise non-adjacent vertices. The clique number $\omega(G)$ is the maximum size of clique in $G$, while the independence number $\alpha(G)$ is the maximum size of an independent set in $G$. Clearly, $\chi(G) \geq \omega(G)$, since the vertices of a clique must get different colors. Similarly, $\chi(G) \geq n/\alpha(G)$, where $n$ denotes the number of vertices in $G$, since a color class is an independent set, so it can contain at most $\alpha(G)$ vertices.

How good are these naive estimates? Notice that $r(n) \geq 2^{n/2}$ implies that there exists a (red) graph $G$ on $n$ vertices with $\omega(G) \leq n$ and $\alpha(G) \leq n$. Hence $\chi(G) \geq 2^{n/2}/n \gg \omega(G)$. On the other hand, taking a graph with large chromatic number and adding a huge independent set will give examples showing that $\chi(G) \gg n/\alpha(G)$; in fact, we can take $\chi(G) \to \infty$ while $n/\alpha(G)$ holds constant.

What about holding $\omega(G)$ fixed – can $\chi(G)$ grow without bound? Do there exist, for instance, graphs with large chromatic number and no $K_3$? Indeed, there do:

**Theorem 0.1** (Zykov (1949)). There exist triangle-free graphs of arbitrarily large chromatic number.

**Proof.** Suppose that we have a triangle-free graph $G$ with $\chi(G) = k$. We will show how to build a triangle-free graph $H$ with $\chi(H) = k + 1$. Induction completes the proof. Start by taking $k$ different copies of $G$, labeled $G_1, \ldots, G_k$. Pick one vertex from each $G_i$ and connect them all to a new vertex. Do this for each of the $|V(G)|^k$ choices of vertices. The resulting graph is our graph $H$. It has an independent set $A$ of $|V(G)|^k$ new vertices, in addition to $k|V(G)|$ vertices in the various $G_i$. This may seem like a lot, but perhaps not so much after van der Waerden’s theorem.

Let’s check that $H$ has the two stated properties. The set of vertices in $A$ is independent, and the neighbors of any vertex in $A$ are in different copies of $G_i$, which do not have any edges between them. Therefore, there is no triangle that uses any vertex in $A$. There is no triangle in any copy of $G_i$ either, by assumption, and there is no triangle using vertices from different $G_i$ since, as we remarked, there is no edge using vertices from different $G_i$. Consequently, $H$ is triangle-free.

We can properly $(k + 1)$-color $H$ by properly $k$-coloring each copy $G_i$ and giving all of the new vertices a new color. Suppose instead that we $k$-color $H$. We are looking for a monochromatic edge. If there isn’t one in any $G_i$, then each one has a proper $k$-coloring. Since $\chi(G_i) = k$, each of the $k$ colors appears on $G_i$. In particular, we can find a vertex $v_i$ in $G_i$ colored by the $i$-th color. But there is a vertex in $A$ connected to precisely these $v_i$, so it...
must have the same color as one of them. This gives the monochromatic edge and completes
the argument that $\chi(H) = k + 1$. □

A more efficient construction due to Mycielski constructs $H$ so that $|V(H)| = 2|V(G)| + 1$. A
related construction due to Tutte produces graphs of arbitrarily high chromatic number with
the further feature that there are no 3-, 4-, or 5-cycles. You can investigate these constructions
on the problem set.

The girth of a graph is the length of its shortest cycle. Note that if a graph has large girth,
say $2r + 1$ or larger, then we can get started coloring it nearby any vertex $v$ by coloring $v$ red,
coloring its neighbors blue, coloring the neighbors of its neighbors red, etc., out to the ball of
radius $r$ about $v$, which is a tree. This is the case even if the odd girth of $G$ is $2r + 1$ or larger,
since this ball is a bipartite graph.

If a graph can be locally colored with few colors, is its chromatic number low? Zykov’s,
Mycielski’s, and Tutte’s constructions indicated a negative answer, but saying anything beyond
girth 6 was a major challenge: could it be that graphs of girth 7 are 100-colorable?

Erdős gave a stunning proof to the contrary using the probabilistic method:

**Theorem 0.2** (Erdős (1965)). For all $l > 0$ and $k > 0$, there exists a graph of girth $>l$ and
chromatic number $>k$.

Constructions followed, for example due to Lovász (1968), but they took several more years
to be discovered, and they do not indicate the abundance of examples that Erdős’s proof
yields.

**Proof.** As before, we construct a graph $G$ on $n$ vertices by adding in edges with a carefully
chosen probability $p \in (0, 1)$.

Here is a first approximation to the strategy. If we choose $p$ small enough, then $G$ will
avoid short cycles, i.e. cycles of length $\leq l$, with high probability. On the other hand, if we
choose $p$ large enough, then $G$ will have small independence number, say $\alpha(G) < n/k$, with
high probability. If there is a common parameter $p$ that is both small and large enough to
serve the two purposes, then with positive probability this method selects a graph with girth
$>l$ and $\alpha(G) < n/k \implies \chi(G) > k$, as desired.

This turns out to be too tall of an order: there is no way to select $p$ to simultaneously obtain
both of the desired outcomes, at least as the proof below proceeds. You should think about
why this is after going through the proof (this is one of the homework problems). Instead, we
will use the alteration method. We will first choose $p$ so that

(♠) with $n \gg 0$ and probability $>1/2$, $G$ contains no more than $n/2$ short cycles.

This lets us take $p$ significantly larger than in the first strategy. With this $p$ we can obtain
the second part of our order, and even with a little more room:

(♣) with $n \gg 0$ and probability $>1/2$, $\alpha(G) \leq n/(2k)$.
Then with positive probability this method selects a graph with \( \leq \frac{n}{2} \) short cycles and \( \alpha(G) < \frac{n}{(2k)} \). We take such a graph and remove one vertex from each short cycle, creating a graph \( G' \) with \( > \frac{n}{2} \) vertices, no short cycles, and \( \alpha(G') < \frac{n}{(2k)} \). This \( G' \) therefore has girth \( > l \) and \( \chi(G') \geq (\frac{n}{2})/\alpha(G') > k \), as desired.

Our proof now breaks into two steps to show that the appropriate choice of \( p \) can be made to obtain both (♠) and (♣) in tandem.

**Step 1:** ensuring few short cycles. What is the expected number of short cycles in \( G \)?

There are \( \frac{n(n-1) \cdots (n-k+1)}{2k} \) ways to specify the vertices of a \( k \)-cycle in the order they are connected around the cycle, and these vertices have probability \( p^k \) of spanning a \( k \)-cycle in this order in \( G \). Thus, the expected number of short cycles is

\[
\sum_{k=3}^{l} \frac{n(n-1) \cdots (n-k+1)}{2k} \cdot p^k \leq \sum_{k=3}^{l} (np)^k.
\]

We can make this quite small by taking

\[
p = n^{\theta-1} \text{ with } 0 < \theta < 1/l; \text{ say } \theta = 1/(2l).
\]

Then each term in the sum is at most \( n^{\theta l} \ll n \), so the sum itself is \( \ll n \). Thus, for \( n \gg 0 \), the expected number of cycles is \( \leq \frac{n}{4} \), say. Now we use a deviation bound, a version of Markov’s inequality, which is much simpler than the Chernoff-style bound we used in the last lecture notes. It states that if a non-negative random variable has expected value \( \mu \), then the probability it assumes a value larger than \( 2\mu \) is less than \( 1/2 \). Compare: if a bag contains 100 potatoes, and a potato weighs on average 1 pound, then fewer than half of the potatoes weigh more than 2 pounds. Applied in this context, since the expected number of short cycles in \( G \) is \( \leq \frac{n}{4} \) for \( n \gg 0 \), we obtain (♠).

**Step 2:** ensuring low independence number. Let us determine the probability that \( G \) contains an independent set of size \( x \), for a given value \( x \). For a fixed set of \( x \) vertices, the probability that we do not put in any edges between them is \( (1-p)^{\binom{x}{2}} \). Using the union bound, it follows that the probability that \( G \) contains an independent set of size \( x \) is no more than \( \binom{n}{x} (1-p)^{\binom{x}{2}} \). That is,

\[
\Pr(\alpha(G) \geq x) \leq \binom{n}{x} (1-p)^{\binom{x}{2}}.
\]

We wish to argue that this is small (less than \( 1/2 \)) while keeping \( x \) still fairly small (remember our target is \( x = \frac{n}{(2k)} \)). To get a handle on it, we utilize a handy inequality: \( 1 - p < \exp(-p) \) for all positive \( p \). This is an easy fact from calculus. It works especially well in this situation, because \( p \) is quite small, so the Taylor series approximation \( \exp(-p) \approx 1 - p \) is quite good. The other advantage is that it allows us to manipulate terms more easily. We therefore have

\[
\Pr(\alpha(G) \geq x) < \binom{n}{x} \exp(-p^{\binom{x}{2}}) < (n \exp(-p(x-1)/2))^{x}.
\]
If we take $x = 1 + 3 \ln(n)/p$, then the right side becomes $n^{-x/2}$. With $p = n^{\theta-1}$, we have $x \sim n^{1-\theta} \cdot \ln(n)$, from which we see that $n^{-x/2} \to 0$ as $n \to \infty$. Thus, we obtain

$$\Pr(\alpha(G) > 3n^{1-\theta} \cdot \ln(n)) < 1/2$$

for $n \gg 0$.

That is, for $n \gg 0$, with probability $> 1/2$, $\alpha(G) \leq 3n^{1-\theta} \cdot \ln(n)$. Choosing $n \gg\gg 0$, this function is $< n/(2k)$. Thus, we obtain (♣).

The proof is now complete! □

This proof comes directly from Alon and Spencer’s The Probabilistic Method.