14. THE PROBABILISTIC METHOD

© Douglas B. West

14.1. EXISTENCE AND EXPECTATION

14.1.1. A random k-subset of [n] contains element 1 with probability k/n. Among the \( \binom{n}{k} \) possible k-sets, equally likely, \( \binom{n-1}{k-1} \) contain 1. Letting \( p \) be the desired probability, we have

\[
p = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{\binom{n-1}{k-1}}{\binom{n}{k}} = \frac{k}{n}.
\]

14.1.2. Club A randomly inviting club B; success has small probability. Invitations sent from members of club A to members of club B form a random function from A to B, and we want the probability that it is surjective. There are \( n^n \) functions, of which \( n! \) are surjective, so the probability is \( n!/n^n \). By Stirling’s Approximation, \( n!/n^n \sim e^{-n} \sqrt{2\pi n} \).

14.1.3. Discrete variable with infinite expectation. Let each positive integer \( k \) occur with probability \( 1/2^k \). Let the random variable \( Y \) have value \( 2^k \) when the outcome is \( k \).

Alternatively, let values that are not positive powers of 2 occur with probability 0, let \( 2^k \) occur with probability \( 2^{-k} \) when \( k \geq 1 \), and just take the expected value of the outcome of the experiment.

14.1.4. Binomial tail probabilities. The lot with nine space fails only when all ten permit holders arrive, with probability \( (1/10)^9 \). The larger lot fails if at least 19 holders arrive, with probability \( 20(1/10)^9(9) + (1/10)^{20} \). That probability is less than \( 1/10 \), which is much smaller than \( (1/10)^9 \). This is related to the notion that large deviations from the mean are less likely when there are many trials.

14.1.5. Probabilistic Quickies.

a) The expected number of fixed points in a random permutation of [n] is 1. Each element is a fixed point with probability \( 1/n \), since it is fixed in \( (n-1)! \) of the equally likely permutations. By the linearity of expectation, the expected number of fixed points is the sum of \( n \) such probabilities.

b) The expected number of male/female pairs in a random pairing of \( n \) men and \( n \) women is \( n^2/(n-1) \). There are \( \prod_{i=1}^{n} (2n-2i+1) \) possible pairings. The common notation for this value is \( (2n-1)!! \).

Proof 1 Let \( X \) be the number of male/female pairs. There are \( n^2 \) possible such pairs. A particular pair occurs in \( (2n-3)!! \) pairings. By linearity, \( \mathbb{E}(X) = n^2 (2n-3)!! / (2n-1)!! \).

Proof 2 Let \( X \) be the number of men paired with women. A particular man is equally likely to be paired with any other person. Hence the probability that the partner of a given man is a woman is \( n/(2n-1) \), and \( \mathbb{E}(X) = n^2/(2n-1) \).

c) The expected number of vertices of degree \( k \) in a random graph is \( n \binom{n-1}{k} p^k (1-p)^{n-k} \). The degree of \( v_i \), a vertex in a random graph is the number of incident edges that occur. Each such edge occurs with probability \( p \), and there are \( \binom{n-1}{k} \) sets of \( k \) incident edges. Thus we need the probability of having exactly \( k \) successes among \( n-1 \) trials with success probability \( p \). This is \( \binom{n-1}{k} p^k (1-p)^{n-1-k} \).

Letting \( X_i = 1 \) if \( v_i \) has degree \( k \) and \( X_i = 0 \) otherwise, we seek \( E(X) \), where \( X = \sum_{i=1}^{n} E(X_i) \). Thus we just multiply \( \mathbb{P}(X_i = 1) \) by \( n \).

14.1.6. The expected number of monochromatic copies of \( K_{r,r} \) in a random 2-coloring of \( E(K_{n,n}) \) is \( \binom{n}{r}^2 2^{n-2r} \). Each possible copy of \( K_{r,r} \) occurs with probability \( 2^{-2r} \), since it can be homogeneous in each color. Since there are \( \binom{n}{r} \) possible copies, linearity of expectation completes the proof.

14.1.7. A \( k \)-uniform hypergraph that is not \( t \)-colorable must have at least \( t^{k-1} \) edges. Color the vertices uniformly at random, independently. Each edge is monochromatic with probability \( t \cdot t^{-k} \). If there are \( m \) edges, and \( m t^{1-k} < 1 \), then some outcome of the experiment is a proper coloring. Hence \( m \geq t^{k-1} \) is necessary to obtain non-\( t \)-colorability.

14.1.8. Some \( n \)-vertex tournament (orientation of \( K_n \)) has at least \( n!/2^{n-1} \) spanning paths. The possible spanning paths correspond to permutations of \([n]\). Each such path occurs if \( n-1 \) edges are oriented as desired and thus has probability \( 2^{-(n-1)} \). By linearity of expectation, the expected number of spanning paths in a random tournament is \( n!/2^{n-1} \).

14.1.9. If \( G \) is a graph with \( p \) vertices, \( q \) edges, and automorphism group of size \( s \), and \( n = (sk^{q-1})^{1/p} \), then some \( k \)-coloring of \( E(K_n) \) has no monochromatic copy of \( G \). In a random k-coloring of \( E(K_n) \) (edges colored independently with probability \( 1/k \) of each color), each copy of \( G \) is monochromatic with probability \( (1/k)^{q-1} \). If the number of copies of \( Q \) is less than \( k^{q-1} \), then by the Union Bound there is a outcome of the coloring having no monochromatic copy of \( G \).
To form a copy of $G$ in $K_n$, map the vertices of $G$ injectively into $V(K_n)$. There are $n^p$ such injections, but in sets of size $s$ they give the same copy of $G$ in $K_n$. Thus it suffices to have $n^p/s < k^q - 1$. With $n^p = sk^q - 1$, as assumed, the condition holds.

14.1.10. **Sperner’s Theorem:** Every antichain of subsets of $[n]$ has size at most $(\binom{n}{\lfloor n/2 \rfloor})$. Let $F$ be such an antichain. For a random permutation $\sigma$ of $[n]$, let $X$ be the number of members of $F$ that appear as initial segments of $\sigma$. Since $F$ is an antichain, always $X \leq 1$. On the other hand, $X$ is the sum of indicator variables for the members of $F$: for $A \in F$, let $X_A = 1$ if $A$ appears as an initial segment of $\sigma$; otherwise $X_A = 0$. The number of permutations in which $A$ appears as an initial segment is $|A!|(|n - |A|)!$. Since there are $n!$ equally likely permutations of $[n]$, we have $E(X_A) = P(X_A = 1) = (\binom{n}{|A|})^{-1} \geq (\binom{n}{\lfloor n/2 \rfloor})^{-1}$. By the linearity of expectation, $|F| (\binom{n}{\lfloor n/2 \rfloor})^{-1} \leq E(X) = 1$, so $|F| \leq (\binom{n}{\lfloor n/2 \rfloor})$.

14.1.11. If $H$ is an $r$-uniform hypergraph with at most $(k - 1)^r/k^{r-1}$ edges, where $r \geq k$, then $H$ has a $k$-coloring in which every color appears in every edge. Color $V(H)$ uniformly at random using $k$ colors. The probability that a given edge fails to receive a particular color is $(k - 1)^r/k^r$. By the Union Bound, the probability that the edge is missing some color is less than $k(k - 1)^r/k^r$. When there are $m$ edges, the probability that some edge is missing some color is less than $mk(k - 1)^r/k^r$, again by the Union Bound. Hence if $m \leq (k - 1)^r/k^{r-1}$, then some outcome of the experiment produces the desired coloring.

14.1.12. The expected number of descents in a random permutation is $(n - 1)/2$. Let $X$ be the random variable that is the number of descents.

**Proof 1** (linearity of expectation). For $1 \leq i < j \leq n$, let $X_{i,j} = 1$ if $i$ immediately follows $j$ (forming a descent), and $X_{i,j} = 0$ otherwise. Now $X = \sum X_{i,j}$. With $ji$ kept together, this descent occurs in $(n - 1)!$ permutations of $[n]$. Thus $E(X_{i,j}) = P(X_{i,j} = 1) = (n - 1)!/n! = 1/n$. There are $\binom{n}{2}$ such variables, so $E(X) = (n - 1)/2$.

**Proof 2** (linearity of expectation). For $1 \leq i \leq n - 1$, let $X_i = 1$ if there is a descent at $i$, and $X_i = 0$ otherwise. Now $X = \sum X_i$. Interchanging positions $i$ and $i + 1$ pairs permutations having and not having a descent at $i$, so $E(X_i) = P(X_i = 1) = 1/2$. There are $n - 1$ such variables, so $E(X) = (n - 1)/2$.

14.1.13. If $S$ is a finite set of finite binary words such that none is a prefix of another, then $\sum_{a \in S} 2^{-|a|} \leq 1$. Generate a binary string by coin flips. For $a \in S$, the probability that $a$ occurs at the start of the string is $2^{-|a|}$. Since no word in $S$ is a prefix of another, these events are disjoint. Hence the probability that some word in $S$ occurs starting at the beginning of the string is $\sum_{a \in S} 2^{-|a|}$. Being a probability, this is at most 1.

14.1.14. For clauses $e_1, \ldots, e_m$ such that $\sum_{i=1}^m 2^{-|e_i|} < 1$, there is a satisfying truth assignment. Assign true/false values to the variables at random, with each having probability $1/2$ of being true, independently. Clause $e_i$ then fails to be satisfied with probability $2^{-|e_i|}$. Hence the probability that some clause fails to be satisfied is at most $\sum_{i=1}^m 2^{-|e_i|}$. If $\sum_{i=1}^m 2^{-|e_i|} < 1$, then in some outcome of the experiment every clause is satisfied.

**Sharpness:** Given $k$ variables, for each subset $S$ create a clause in which the variables in $S$ appear as TRUE and the others appear as FALSE. There are $2^k$ clauses. For each way of setting all the variables, every truth assignment leaves one clause clause unsatisfied, even though $\sum 2^{-|e_i|} = 1$.

14.1.15. If $n_k$ is the least number of vertices in a non-$k$-choosable bipartite graph, and $m_k$ is the least number of edges in a non-$2$-colorable $k$-uniform hypergraph, then $m_k \leq n_k \leq 2m_k$.

a) $n_k \leq 2m_k$. Let $H$ be a smallest non-$2$-colorable $k$-uniform hypergraph. From $H$ we construct a non-$k$-choosable complete bipartite graph $G$ with $2|E(H)|$ vertices. Use two copies of the edge set of $H$ as the parts $X$ and $Y$ of $G$. For each $v \in V(G)$, let $L(v)$ be the set of vertices of the edge in $H$ corresponding to $v$; note that $|L(v)| = k$.

Since $G$ is a complete bipartite graph, an $L$-coloring chooses each color in only one part. For each vertex of $H$ so chosen, put it into $S_X$ or $S_Y$ according to the part where it was chosen. Since each edge of $H$ generates vertices in both $X$ and $Y$, this produces a proper $2$-coloring of $H$. Since $H$ is not $2$-colorable, we conclude that $G$ is not $L$-colorable.

b) $m_k \leq n_k$. Let $G$ be a non-$k$-choosable $X, Y$-bigraph. From $G$ we construct a non-$2$-colorable $k$-uniform hypergraph $H$ with $|V(G)|$ edges. Since $G$ is not $k$-choosable, there is a list assignment $L$ with lists of size $k$ such that $G$ is not $L$-colorable. Let $H$ be the hypergraph with vertex set $\bigcup_{v \in V(G)} L(v)$ and edge set $\{L(v) : v \in V(G)\}$.

If $H$ is $2$-colorable, then a proper $2$-coloring of $H$ partitions the colors in the lists $L$ into two sets $A$ and $B$. If $v \in X$, then choose for $v$ an element of $L(v) \cap A$; if $v \in Y$, choose an element of $L(v) \cap B$. Since elements of $A$ are chosen only on $X$ and elements of $B$ only on $Y$, the result is an $L$-coloring of $G$. Since $G$ is not $L$-colorable, we conclude that $H$ is not $2$-colorable.

14.1.16. If $k > 1 + \lg n$, then every $n$-vertex bipartite graph $G$ is $k$-choosable. Let $X$ and $Y$ be the partite sets of $G$, with $L$ a $k$-uniform list assignment.

Let $S = \bigcup_{v \in V(G)} L(v)$. Let $H$ be the hypergraph with vertex set $S$ and edge set $\{L(v) : v \in V(G)\}$. In a proper $2$-coloring of $H$, view the two labels
sits in a random seat, and each subsequent passenger sits in his or her own seat if it is available and otherwise chooses a random seat from those that remain, then the probability that the last passenger sits in his or her own seat is $1/2$. Every passenger other than the first and the last already sat in his or her own seat unless it was already occupied, so none of those passengers’ seats are available when the last passenger arrives. Hence the only possible seats that can be the last seat are those assigned to the last passenger and the first passenger. As long as those two seats both remain available, they are equally likely to be chosen by each passenger looking for a seat. Hence they are equally likely to be the last seat available.

14.1.20. A tournament cannot contain three 3-cycles on four vertices. Begin with a 3-cycle, say $[x, y, z]$. No matter how the edge joining $y$ and the fourth vertex $w$ is oriented, it cannot make a path both with $xy$ and with $yz$. Hence there can only be one additional 3-cycle.

The maximum number of edges in an $n$-vertex 3-uniform hypergraph having no set of four vertices containing three edges is at least $\frac{1}{4}(\binom{n}{3})$. Create a random tournament $T$ on $n$ vertices. Make a 3-uniform hypergraph $H$ by letting each triple that is a 3-cycle in $T$ be an edge in $H$. Each outcome has no four vertices containing three edges, since $T$ cannot contain three 3-cycles on four vertices.

Because the tournament is generated at random, each triple has probability $1/4$ of becoming a cycle. Hence the expected number of 3-cycles is $\frac{1}{4}(\binom{n}{3})$. All generated hypergraphs satisfy the 4-vertex restriction, and some outcome has at least $\frac{1}{4}(\binom{n}{3})$ edges.

14.1.21. Two letters drawn at random from an alphabet of size 26 have probability at least 1/26 of being the same. More generally, two samples from any distribution over $n$ outcomes have probability at least $1/n$ of being the same. If the distribution is $p_1, \ldots, p_n$, then the probability of drawing the same outcome twice is $\sum_{i=1}^{n} p_i^2$. By the convexity of the squaring function, $\frac{1}{n} \sum_{i=1}^{n} p_i^2 \geq \left(\frac{\sum_{i=1}^{n} p_i}{n}\right)^2 = 1/n^2$. Hence the stated probability is at least $1/n$.

14.1.22. In a complete $k$-ary tree with leaves at distance $l$ from the root, the vertices fail independently with probability $p$, then the expected number of nodes accessible from the root is $\frac{1 - p}{1 - kp}$. There are $k^j$ vertices at depth $j$ (distance $j$ from the root). A vertex at depth $j$ is accessible if and only if it and its ancestors are alive. Thus it is accessible with probability $(1 - p)^{j+1}$. Using linearity of the expectation, the expected number of accessible nodes is $\sum_{j=0}^{l} (1 - p)(k - kp)^j$. When $p = 1/k$, this is simply $1 - p$. Otherwise, the expectation is $\frac{(1 - p)(k - kp)^{j+1}}{1 - kp}$. 

as restricting usage of colors to $X$ or to $Y$. For each $v \in V(G)$ a color exists in $L(v)$ that has been restricted to usage on the partite set containing $v$, and hence $G$ is $L$-colorable. When $L$ is $k$-uniform, also $H$ is $k$-uniform, and hence $H$ is 2-colorable if it has fewer than $2^{k-1}$ edges. Thus $n < 2^{k-1}$ suffices, and this is equivalent to $k > 1 + \lg n$.

Every $n$-vertex bipartite graph $G$ is $k$-choosable when $k > \log_2 n$. Restrict each color in $S$ to usage on $X$ or on $Y$, with probability 1/2 each, independently. The probability that $L(v)$ has no color that has been restricted to the partite set containing $v$ is $2^{-k}$. If $n2^{-k} < 1$, then in some outcome of the experiment every vertex has an available color that is restricted to its partite set. Thus $G$ is $L$-colorable when $k > \lg n$.

The direct argument yields a stronger result because we don’t mind if all colors in $L(v)$ are restricted to the partite set containing $v$.

14.1.17. If a hypergraph $H$ has no edges $e_1, \ldots, e_k$ that pairwise intersect in precisely the same vertex $x$, then $H$ is $k$-choosable. Color the vertices in any fixed order. When $x$ is reached, if $x$ completes edges $e_i$ and $e_j$ in colors $i$ and $j$, respectively, then $e_i \cap e_j = \{x\}$, since no vertex has been given both colors $i$ and $j$. Hence there is a color in the list assigned to $x$ that if used on $x$ does not complete a monochromatic edge in that color, since if $x$ completes an edge in each of $k$ colors, then those $k$ edges pairwise share only $x$, which is forbidden.

Generalization of degeneracy in graphs. Define a hypergraph to be $k$-degenerate if every subhypergraph induced by a set of vertices has a vertex that does not belong to $k$ edges that are otherwise disjoint. Given such a hypergraph, order the vertices by iteratively placing such a vertex of the remaining induced subhypergraph last. Coloring the vertices in this order allows choosing a proper coloring from the lists by applying the argument in the previous paragraph.

14.1.18. Non-$k$-choosable tripartite 3-uniform hypergraph. Let $n = \binom{3k-1}{k}$, let $H$ be the hypergraph with three parts of size $n$ in which every triple obtained by using one vertex from each part is an edge, and let the lists on each part be the $k$-subsets of $[3k - 1]$. That is, each list consists of $k$ colors chosen from $[3k - 1]$.

Suppose that a proper coloring of $H$ can be chosen from the lists. No color can be used on vertices from all three parts. Hence there are at most $6k - 2$ incidences of colors with parts. This means that on at least one of the parts at most $2k - 1$ colors are used. Hence there are at least $k$ colors not used on that part. However, every set of $k$ colors is the list of some vertex in that part, so there is a vertex from which we did not choose a color.

The construction generalizes to $r$-partite $r$-uniform hypergraphs.

14.1.19. If the first passenger to board a plane loses her boarding pass and
14.1.23. In a random tree $T$ with $V(T) = [n]$, letting $X, Y, Z$ be the numbers of leaves, vertices of degree 2, and leaves with neighbors of degree 2, the asymptotic values of $E(X)/n, E(Y)/n,$ and $E(Z)/n$ are $e^{-1}, e^{-1}$, and $e^{-2}$.

The trees correspond bijectively to elements of $[n]^{n-2}$, with $i$ appearing $d_i(i-1)$ times in the $(n-2)$-tuple corresponding to $T$. Hence $X$ is the number of labels missing from a random member of $[n]^{n-2}$. Let $X_i$ be the indicator variable for the absence of $i$; note that $P(X_i = 1) = (1 - 1/n)^{n-2}$. Thus $E(X) = n(1 - 1/n)^{n-2} - 1/n - 1 - n^{-1}$, as desired.

Similarly, $Y$ is the number of labels appearing once. A given label $i$ appears once in the random $n$-ary string with probability $n!/(1 - 1/n)^{n-3}$. Hence $E(Y) = n!(1 - 1/n)^{n-3} - ne^{-1}$, as desired.

Let $A$ be the event that $n$ is a leaf. Let $B$ be the event that its neighbor has degree 2. By symmetry and linearity, $E(Z) = nP(A \cap B)$. By the definition of conditional probability, $P(A \cap B) = P(B \mid A)P(A)$. Note that $P(A) = E(X)/n = (1 - 1/n)^{n-2}$.

Given that $n$ is a leaf, $T - n$ is a tree with vertex set $[n-1]$. For each such tree, attaching $n$ yields $n - 1$ distinct trees on $[n]$, and each tree with $n$ as a leaf arises once in this way. Thus each element of $[n-1]$ is equally likely to be the neighbor of $n$ when $n$ is a leaf.

Let $C_i$ be the event that $i$ is adjacent to $n$ and $D_i$ be the event that $i$ is a leaf in $T - n$. We have $P(D_i \cap C_i \mid A) = P(D_i \mid C_i \cap A)P(C_i \mid A)$. We have computed $P(C_i \mid A) = 1/(n-1)$ and $P(D_i \mid C_i \cap A) = (1 - 1/(n-1))^{n-3}$. Since $C_1, \ldots, C_{n-1}$ are disjoint and equally likely, we compute

$$P(A \cap B) = (n-1)P(D_i \cap C_i \mid A) = (n-1)(1 - 1/n)^{n-2} - 1/(1 - 1/n)^{n-3} - n^{-1} - 1/(n - 1/n)^{n-2}.$$

Thus

$$\frac{1}{n}E(Z) = \left(1 - \frac{1}{n-1}\right)^{n-1} \left(1 - \frac{1}{n}\right)^{n-2} = e^{2}.$$

14.1.24. The pill bottle with $m$ large pills and $n$ small pills. Each day a random pill is chosen; a small pill disappears, and a large pill becomes a small pill. The expected number $X_{m,n}$ of small pills remaining when the last large pill is gone is $\frac{n}{m+1} + \sum_{k=1}^{m} \frac{1}{k}$. The expected number $Y_{m,n}$ of steps to reach that point is $2m + n - \frac{n}{m+1} - \sum_{k=1}^{m} \frac{1}{k}$.

Proof 1 (recurrence and induction). Note that $X_{0,n} = n$ and $Y_{0,n} = 0$. For $m > 0$, on the first step the probability is $\frac{m}{m+n}$ of picking a large pill and $\frac{n}{m+n}$ of picking a small pill. These probabilities weight the expectations over the remainder of the process. For $m > 0$, we compute

$$X_{m,n} = \frac{m}{m+n}X_{m-1,n+1} + \frac{n}{m+n}X_{m,n-1},$$

$$Y_{m,n} = 1 + \frac{m}{m+n}Y_{m-1,n+1} + \frac{n}{m+n}Y_{m,n-1}.$$

Computing with small values $(m = 0, 1, 2, \ldots)$ may suggest the answer, provable using induction on $m$ (and on $n$ for fixed $m$). The values for $m = 0$ satisfy the claimed formulas. The recurrence then yields

$$X_{m,n} = \frac{m}{m+n}x_{m-1,n+1} + \frac{n}{m+n}\sum_{k=1}^{m-1} \frac{1}{k} + \frac{n}{m+n}\sum_{k=1}^{n} \frac{1}{k}$$

$$= \frac{n+1}{m+n} + \sum_{k=1}^{m} \frac{1}{k} - \frac{1}{m+n} - \frac{n}{m+n}\left(\frac{1 - 1}{m + 1} + \frac{1}{m}\right)$$

$$= \frac{n+1}{m+n} + \sum_{k=1}^{m} \frac{1}{k} - \frac{1}{m+n} - \frac{n}{m+n}\left(\frac{1 - 1}{m + 1} + \frac{1}{m}\right)$$

$$= \frac{n}{m+n} + \sum_{k=1}^{m} \frac{1}{k} + \frac{n}{m+n} - 1$$

$$= \frac{n}{m+n} + \sum_{k=1}^{m} \frac{1}{k} + \frac{n}{m+n} - 1$$

$$= \frac{n}{m+n} + \sum_{k=1}^{m} \frac{1}{k} + \frac{n}{m+n} - 1$$

$$= 2m + n - X_{m,n} = 2m + n - \frac{n}{m+1} - \sum_{k=1}^{m} \frac{1}{k}.$$

Proof 2 (linearity of expectation). The expected number of small pills remaining is the sum of the probabilities that each pill remains as a small pill. The probability that an initial small pill remains when there are no large pills is the probability that it is chosen after all the large pills are first chosen; this probability is $1/(m + 1)$. The probability that the $k$th broken large pill remains as a small pill after the remaining large pills are broken is $1/(m + k + 1)$. Hence the answer is $\frac{n}{m+1} + \sum_{k=1}^{m} \frac{1}{k}$.

The number of steps to completely consume all the pills is $2m + n$. The number of steps to the point where only small pills remain is $2m + n$ minus the number of small pills remaining. Again by linearity, we thus have $Y_{m,n} = 2m + n - X_{m,n}$. 

Since the formulas for $X_{m,n}$ and $Y_{m,n}$ sum to $2m + n$, we compute

$$Y_{m,n} = 1 + \frac{m}{m+n}Y_{m-1,n+1} + \frac{n}{m+n}Y_{m,n-1}$$

$$= \frac{m}{m+n}(1 + Y_{m-1,n+1}) + \frac{n}{m+n}(1 + Y_{m,n-1})$$

$$= \frac{m}{m+n}(2m + n - X_{m-1,n+1}) + \frac{n}{m+n}(2m + n - X_{m,n-1})$$

$$= 2m + n - \frac{n}{m+n}X_{m-1,n+1} - \frac{n}{m+n}X_{m,n-1}$$

$$= 2m + n - X_{m,n} = 2m + n - \frac{n}{m+1} - \sum_{k=1}^{m} \frac{1}{k}.$$
14.1.25. When \( n \) distinct pairs of socks are put into the laundry, and \( k \) socks are drawn at random without replacement, the expected number of pairs is \( \frac{k(k-1)}{4n-2} \). What is drawn is a random \( k \)-subset of the \( 2n \) socks, with the socks in a pair being indistinguishable. By linearity of expectation, the answer is \( n \) times the probability that a particular pair is drawn.

The probability that a particular pair is drawn is the probability that both copies of 1 occur among the first \( k \) entries of a random permutation of the multiset \( \{1,2,\ldots,n\} \). In a random permutation, there are \( \binom{2n}{2} \) possible locations for the 1s, all equally likely, and \( \binom{k}{2} \) of those occur among the first \( k \) places. Hence the probability of getting both copies of the first pair is \( \frac{\binom{k}{2}}{\binom{2n}{k}} \). Multiplying by \( n \) yields the expected number of pairs.

There are other approaches to computing this expectation.

14.1.26. When a permutation of \([n]\) is formed by choosing the first \( k \) positions randomly and filling the remainder in ascending order, the expected number \( E_{n,k} \) of left-to-right maxima is

\[
\frac{n}{k+1} + \sum_{j=1}^{k} \frac{1}{j}.
\]

Note that with the given formula, \( E_{n+1,k} - E_{n,k} = \frac{1}{k+1} \).

Among the permutations that have a particular set \( S \) of values in a particular set of positions, the fraction that have a particular element of \( S \) in the last of those positions is \( 1/|S| \). By linearity of expectation, the expected number of left-to-right maxima is the sum of the probabilities of having left-to-right maxima in each of the positions. For \( j \leq k \), a maximum occurs at position \( j \) if and only if the value in position \( j \) is the largest among the first \( j \) values; the probability is \( 1/j \). For \( j > k \), a maximum occurs if and only if the value in position \( j \) is larger than the values in the first \( k \) positions; the probability is \( 1/(k+1) \). Summing these probabilities yields the formula claimed.

14.1.27. Bipartite subgraphs. Let \( G \) be an \( n \)-vertex graph with \( m \) edges.

a) \( G \) has a bipartite subgraph with at least \( m/2 \) edges. Select a random vertex subset \( A \) by choosing each vertex with probability \( 1/2 \), independently. For each edge \( e \), the probability is \( 1/2 \) that \( A \) contains exactly one endpoint of \( e \), placing \( e \) in the cut \([A,\overline{A}]\). By linearity of expectation, the expected number of edges in the cut is \( m/2 \). Furthermore, the edges in a cut form a bipartite subgraph.

b) \( G \) has a bipartite subgraph with at least \( m \frac{n/2}{n(n-1)} \) edges. Choose \( A \) at random from all \( \binom{n/2}{m} \) vertex subsets. The number of these subsets containing exactly one endpoint of a given edge \( e \) is \( 2 \binom{n/2-1}{m/2-1} \). Thus \( e \) belongs to the cut \([A,\overline{A}]\) with probability \( 2 \binom{n/2-1}{m/2-1}/\binom{n/2}{m} \). Since

\[
\binom{n}{n/2} = \frac{n}{\binom{n/2}{n/2}} \binom{n-1}{n/2} = \frac{n}{\binom{n/2}{n/2}} \binom{n-1}{n/2} \binom{n-2}{n/2} - 1,
\]

the probability is \( \frac{2n/2}{n(n-1)} \). Since \( \binom{n-1}{m/2} = 2 \binom{n/2}{m/2} - 1 \), linearity of expectation yields \( m \frac{n(n+1)}{2n/2} \) as the expected size of the cut, and some cut is at least this large. This improves part (a).

c) If \( G \) has a matching with \( k \) edges, then \( G \) has a bipartite subgraph with at least \((m+k)/2 \) edges. Let \( M \) be a matching of size \( k \) in \( G \). Choose a random vertex subset \( S \subseteq V(G) \) as follows: for \( e \in M \), put one endpoint of \( e \) in \( S \), choosing each with probability \( 1/2 \), and for each vertex not covered by \( M \), put it in \( S \) with probability \( 1/2 \). Let \( X = |[S,\overline{S}]| \); the cut is a bipartite subgraph. For edge \( e \), let \( X_e \) be the indicator variable that is 1 when \( e \in [S,\overline{S}] \); note that \( X = \sum X_e \). Also \( \mathbb{P}(X=1) = 1/2 \) when \( e \notin M \), while \( \mathbb{P}(X=1) = 1 \) when \( e \in M \). Hence \( \mathbb{E}(X) = (m-k)/2 + k = (m+k)/2 \). By the Pigeonhole Property of the expectation, some choice of \( S \) yields a bipartite subgraph with at least \((m+k)/2 \) edges.

14.1.28. Every \( n \)-vertex graph \( G \) with average degree \( d \) and minimum degree \( k \) has an induced subgraph with at least \( \frac{k^2}{d+1} \) vertices that does not contain \( K_{k+1} \). Generate a random vertex ordering. Let \( S \) be the set of vertices \( v \) such that \( v \) is one of the first \( k \) vertices of \( N[v] \) in the ordering. The subgraph induced by \( S \) cannot contain \( K_{k+1} \), because the vertex of a \( (k+1) \)-clique that appears last in the ordering has \( k \) earlier neighbors.

Let \( X = |S| \). The probability that a particular vertex \( v \) is chosen for \( S \) is \( \frac{k}{d+1} \). By the linearity of expectation, \( \mathbb{E}(X) = \sum_{v \in V(G)} \frac{k}{d+1} \). By the convexity of \( 1/(x+1) \) as a function of \( x \), we have \( \mathbb{E}(X) \geq kn/(d+1) \). Hence some outcome of the experiment yields the desired induced subgraph.

14.1.29. If \( \{A_i\}_{i=1}^m \) and \( \{B_i\}_{i=1}^m \) are subsets of \([n]\), with \( |A_i| = a_i \), \( |B_i| = b_i \), and \( A_i \cap B_j = \emptyset \) if and only if \( i = j \), then \( \sum_{i=1}^m \frac{\binom{n}{a_i+b_i}}{\binom{n}{a_i}} \leq 1 \). When \( n \) is ordered randomly, the probability that \( A_i \) appears completely before \( B_i \) is \( \binom{n}{a_i+b_i} \). In a given ordering, this event can occur for only one value of \( i \); if \( A_i \) occurs completely before \( B_i \), then some element of \( B_i \) in \( A_i \) occurs before some element of \( A_i \) in \( B_i \). The sum of the probabilities of disjoint events is bounded by 1, which yields the desired inequality.

The maximum size of an antichain in \( 2^n \) is \( \binom{n}{n/2} \). Let the elements of an antichain be \( A_1, \ldots, A_m \), and let \( B_i = \overline{A_i} \). If \( X \) and \( Y \) are incomparable, then \( X \) intersects \( Y \), and \( Y \) intersects \( X \). Hence the pairs \( (A_i, B_i) \) satisfy the condition for Bollobás' Inequality. Since \( a_i + b_i = n \), we obtain \( \sum_{i=1}^m \frac{\binom{n}{a_i}}{\binom{n}{n/2}} \leq 1 \). This is precisely the LYM Inequality for antichains, and multiplying by the middle binomial coefficient yields \( m \leq \binom{n}{n/2} \).


a) For independent trials with success probability \( p \), if \( Y \) is the number of trials to the first success, then \( \mathbb{E}(Y) = 1/p \).
Proof 1 (explicit summation). If \( Y = k \), then there were \( k - 1 \) failures followed by one success. Thus \( \mathbb{P}(Y = k) = p(1 - p)^{k-1}, \) and 
\[
\mathbb{E}(Y) = \sum_{k=1}^{\infty} kp(1 - p)^{k-1}.
\]
Since
\[
\sum_{k=1}^{\infty} kx^{k-1} = \frac{d}{dx} \sum_{k=0}^{\infty} x^k = \frac{d}{dx} \frac{1}{1-x} = \frac{1}{(1-x)^2}
\]
when \( 0 < x < 1 \), we obtain \( \mathbb{E}(Y) = p \frac{1}{p^2} = 1/p. \)

Proof 2 (conditional probability). Let \( q = \mathbb{E}(Y) \). If the first trial succeeds, then \( Y = 1 \). If it fails, then the expected number of additional trials is \( q \). Thus \( q = p \cdot 1 + (1-p)(1+q) \). Solving for \( q \) yields \( q = 1/p \).

b) The Coupon Collector: When trials have \( n \) equally likely outcomes, the expected number of trials to obtain all outcomes is \( n \sum_{i=1}^{n} \frac{1}{i} \). Let \( X \) be the number of trials used. Let \( X_i \) be the number of trials after \( i-1 \) of the outcomes have been obtained until \( i \) outcomes have been obtained. We have \( X = \sum_{i=1}^{n} X_i \). When we have obtained \( i - 1 \) of the outcomes, a given trial provides a new outcome with probability \( \frac{n-i+1}{n} \). Hence \( X_i \) is a geometric random variable with success probability \( \frac{n-i+1}{n} \). Using part (a), linearity of expectation, reversing the order of summation, and letting \( j = n + i - 1 \), we obtain \( \mathbb{E}(X) = n \sum_{i=1}^{n} \frac{1}{j} \).

14.1.31. The optimal pebbling number of the \( k \)-cube \( Q_k \) is at least \( (4/3)^k \). A pebbling move takes two pebbles from one vertex and puts one on a neighboring vertex. The optimal pebbling number is the minimum total size of a distribution \( D \) from which every vertex is reachable.

a) \( \sum_{r \geq 0} a_{r,t}+1 \geq 1 \) for any vertex \( r \) in \( Q_k \), where \( a_{r,t} \) is the number of pebbles in \( D \) on vertices at distance \( t \) from \( r \). At the end of a list of moves reaching \( r \), the inequality holds for the final distribution. For any move, \( a_{r,t} \) for some \( t \) is reduced by 2, and \( a_{r,t'} \) is increased by 1, where \( t' \geq t - 1 \). Hence the left size loses 2 \(-2t\) and gains at most \( 2t^{t-1} \), so the sum cannot increase. Therefore, if \( r \) is reachable, then the sum must be at least 1 at the beginning.

b) \( \mathbb{E}(a_{r,t}) = |D| \left( \frac{e}{p} \right)^{2t} \). We compute \( \mathbb{E}(a_{r,t}) \) directly, again viewing \( a_{r,t} \) as a sum of variables for each pebble. The probability that a fixed pebble in \( D \) has distance \( t \) from a random target \( r \) is \( \left( \frac{e}{p} \right)^{2t} \), since \( \left( \frac{e}{p} \right) \) of the \( 2^t \) vertices in \( Q_k \) have distance \( t \) from this pebble. Grouping the contributions to \( \sum_{r} a_{r,t} \) by the pebbles in \( D \) rather than by \( r \) and applying linearity yields \( \mathbb{E}(a_{r,t}) = |D| \left( \frac{e}{p} \right)^{2t} \).

\[ |D| \geq (4/3)^k \]. Substituting the equality above into the inequality from part (a) and simplifying yields \( |D| \sum_{t=0}^{\infty} \left( \frac{e}{p} \right)^{2t} \geq 2^k \). By the Binomial Theorem, \( |D| \left( 1 + \frac{1}{2} \right)^k \geq 2^k \), and hence \( |D| \geq (4/3)^k \).

14.1.32. An \( n \)-vertex graph \( G \) with \( \delta(G) > \frac{\ln n}{r+1} \) in \( n \) has an \( r \)-dynamic coloring with at most \( \Delta(G) + r + s \) colors. Let \( k = \Delta(G) + r + s \). If \( n \leq k \), then we can give the vertices distinct colors. Hence we may assume \( n > k \).

Color vertices of \( G \) in the order \( v_1, \ldots, v_n \) from a set of \( k \) colors, always giving the next vertex a uniformly random color among the colors not used on its earlier neighbors. Since \( k > \Delta(G) \), the procedure always produces a proper coloring.

Since \( \delta(G) > r \), the resulting coloring fails to be \( r \)-dynamic only if some neighborhood \( S \) is restricted to a set \( T \) of at most \( r - 1 \) colors. At every point where a vertex of \( S \) was colored, at least \( r + s \) colors were available. Hence the probability that the colors on \( S \) are restricted to \( T \) is at most \( \left( \frac{r-s}{r+1} \right)^{\delta(G)} \). We use the inequality \( 1 - x \leq e^{-x} \) with \( x = \frac{r-s}{r+1} \) to bound this by \( e^{-\frac{r-s}{r+1} \delta(G)} \), which is less than \( n^{-r} \).

There are \( \left( \frac{r}{r-1} \right)^n \) choices of the set \( T \) of \( r - 1 \) colors; this value is less than \( n^{-r} \) since \( k < n \). Hence by the Union Bound the probability that the coloring fails to be \( r \)-dynamic is less than \( n^{n-r} \), which equals 1. Hence some outcome of the experiment is \( r \)-dynamic.

14.1.33. When two players alternately flip a coin that has probability \( p \) of landing heads, the probability that Player 1 is the first to obtain heads is \( 1/(2-p) \). Let \( q \) be the desired probability. Player 1 can win on flip 1 or by winning the game after two tails. Hence \( q = p + q(1-p)^2 \). Solving for \( q \) yields \( q = \frac{p}{1-(1-p)^2} = \frac{p}{2p-p^2} = \frac{1}{2-p} \).

Alternatively, summing the probabilities that Player 1 wins after various numbers of flips yields the geometric sum \( q = p \sum_{t=0}^{\infty} (1-p)^2 \), which leads to the same computation.

14.1.34. Expected length of random walk. The walk is on a path. If it reaches an endpoint of the path, it terminates. Otherwise, it moves left or right by one unit with probability 1/2 each.

a) The expected number of steps to move one step left or n steps right is \( n \). More generally, suppose that the path has \( n + 2 \) positions, where there are \( a \) positions to the left of the start and \( b \) positions to the right, with \( a + b = n + 1 \). Let \( E_a \) denote the expected number of steps to the end. Note that \( E_0 = E_{n+1} = 0 \) and \( E_a = E_{n+1-a} \) for all \( a \).
When 1 \leq a \leq n, the walk must take a step, and then it follows another walk from its new position. Thus \( E_a = 1 + \frac{1}{2} E_{a-1} + \frac{1}{2} E_{a+1} \) for \( 1 \leq a \leq n \). Let \( T = \sum_{a=1}^{n} E_a \). Summing the equations yields
\[
T = n + \frac{1}{2} \sum_{a=0}^{n-1} E_a + \frac{1}{2} \sum_{a=2}^{n+1} E_a = n + \frac{1}{2} (T - E_n) + \frac{1}{2} (T - E_1) = n + T - E_1,
\]
where we have used \( E_1 = E_n \). Thus \( E_1 = n \), as desired.

Comment: This argument derives the one desired value, but in fact the equations are all satisfied by letting \( E_a = (n+1-a) \) for all \( a \).

b) \((n+1)\) is the expected number of steps until the first time when the maximum and minimum values reached by the walk differ by \( n \). Before the difference reaches \( n \), it must reach \( n-1 \). Let \( Y_n \) be the number of steps after the difference first reaches \( n-1 \) until it first reaches \( n \). Thus the number \( X_n \) of steps until the difference first reaches \( n \) is \( \sum_{k=1}^{n} Y_k \), and by the linearity of expectation \( \mathbb{E}(X_n) = \sum_{k=1}^{n} \mathbb{E}(Y_k) \).

When the difference first equals \( k \), the walk is at the current minimum or maximum value. In order to increase the difference, the walk must move one step farther away from 0 or move \( k \) steps in the other direction. Thus \( \mathbb{E}(Y_k) = k \) by part (a). Hence \( \mathbb{E}(X_n) = \sum_{k=1}^{n} k = \binom{n+1}{2} \).

14.1.35. For a random walk on a cycle, the last vertex reached is equally likely to be any of the non-initial vertices. For the walk from \( v \) to reach \( u \) last, it must first reach one of the neighbors of \( u \) and then reach the other neighbor of \( u \) before reaching \( u \). Indexing the vertices as \( v_1, \ldots, v_n \), with indices modulo \( n \), by symmetry the probability of reaching \( v_{i+1} \) before \( v_i \) when starting from \( v_{i-1} \) is the same as the probability of reaching \( v_{i-1} \) before \( v_i \) when starting from \( v_{i+1} \); call this probability \( p \). Hence no matter which neighbor of \( u \) is reached first, the probability that \( v_j \) will be the last vertex reached is \( p \). Furthermore, this value \( p \) in the computation is independent of the initial vertex \( v_1 \), as long as \( i \neq j \).

14.1.36. If an unbiased coin is flipped until \( k \) heads occur, yielding \( X \) isolated heads and \( Y \) runs, then \( \mathbb{E}(X) = (k+2)/4 \) and \( \mathbb{E}(Y) = (2k+1)/2 \); except \( \mathbb{E}(X) = 1 \) when \( k = 1 \).

For \( 1 \leq i \leq k \), let \( t_i \) be the random variable having value 1 if the \( i \)th head is immediately preceded by a tail and value 0 otherwise. Note that \( \mathbb{P}(t_1 = 1) = 1/2 \) and that \( t_1, \ldots, t_r \), are independent.

The first head is isolated if and only if \( t_2 = 1 \), and the last head is isolated if and only if \( t_k = 1 \). For \( 2 \leq i \leq k-1 \), the \( i \)th head is isolated if and only if \( t_i = t_{i+1} = 1 \). Thus \( X = t_2 + t_k + \sum_{i=2}^{k-1} t_i t_{i+1}, \) so
\[
\mathbb{E}(X) = \left( \frac{1}{2} \right) + \left( \frac{1}{2} \right) (k-2) \frac{1}{4} = k \frac{1}{2}.
\]

14.1.37. A random pairing of 2n people who are grouped in \( n \) sets, such as both Senators from \( n \) states.

a) With probability \( \sum_{k=0}^{n} (-1)^k \frac{k!}{(n-k)! \binom{2n}{k}} \), no two people from the same set (Senators from the same state) are paired. The number of pairings of 2n people is \( \frac{(2n)!}{2^n n!} \); denote this by \( M(n) \). Let \( A_n \) be the set of pairings containing the \( k \)th original pair. By the Inclusion-Exclusion Principle, the number of pairings containing none of the original pairs is \( \sum_{k=0}^{n} (-1)^k \binom{n}{k} M(n-k) \). To compute the probability, we divide this by \( M(n) \), since the pairings are equally likely. Since
\[
\binom{n}{k} M(n-k) = \frac{n!}{k!(n-k)!} \frac{(2n-2k)!}{(n-k)!2^{n-k}} \frac{n! 2^n}{k! 2^n} = \frac{2}{k!} \binom{2n}{k},
\]
the formula simplifies as claimed.

b) The limit of the probability as \( n \to \infty \) is \( e^{-1/2} \). By Stirling’s Formula, \( \binom{2n}{n} \approx \frac{2^{2n}}{\sqrt{\pi n}} \). Hence the sum in part (a) is approximately \( \sum_{k=0}^{n} (-1/2)^k (1/k!) \sqrt{n/(n-k)} \). Breaking this sum after \( n^{1/3} \) terms yields a sum approaching \( e^{-1/2} \), because it is bounded by the expression
\[
\left( 1 + \frac{1}{2} \right) \frac{n^{1/3}}{2} + \mathcal{O}(n^{-4/3}) \left( \sum_{k=0}^{n^{1/3}} (-1/2)^k \right),
\]
while the sum of the rest approaches 0.

14.1.38. Drawing from an urn with \( n \) balls, of which \( k \) are white. 

a) If drawn balls are replaced, then the expected number of draws to the first white ball is \( n/k \). Let \( X \) be the number of the draw when the first white ball is drawn. At each draw the probability of a white ball is \( k/n \). Let \( p = k/n \). It suffices to show that \( \mathbb{E}(X) = \frac{1}{p} \). (Note: This is the same as part (a) of Exercise 14.1.30.)

Proof 1 (direct computation). If \( j \geq 1 \), then \( \mathbb{P}(X = j) = (1-p)^{j-1}p \).

Computing the expectation,
\[
\mathbb{E}(X) = \sum_{j=1}^{\infty} j(1-p)^{j-1}p = p \frac{d}{dx} \left( \frac{1}{1-x} \right) \bigg|_{x=1-p} = \frac{p}{1-x} \bigg|_{x=1-p} = \frac{1}{p}.
\]

Proof 2 (rephrasing of expectation). For a nonnegative integer-valued variable \( X \), always \( \mathbb{E}(X) = \sum_{j \geq 1} \mathbb{P}(X \geq j) \). Since \( \mathbb{P}(X \geq j) = (1-p)^{j-1} \), we have \( \mathbb{E}(X) = \sum_{j \geq 1} (1-p)^{j-1} = \frac{1}{1-(1-p)} = \frac{1}{p} \).
Given unit vectors \( v_1, \ldots, v_n \) in \( \mathbb{R}^n \) and \( w = \sum p_i v_i \), where \( p_1, \ldots, p_n \in [0, 1] \), there is some subset of \( v_{i_1}, \ldots, v_{i_m} \), whose sum differs from \( w \) by a vector with magnitude at most \( \sqrt{m}/2 \). Let \( u = \sum_{i=1}^n \varepsilon_i v_i \) with \( \varepsilon_1, \ldots, \varepsilon_n \) chosen independently from \( \{0, 1\} \). To approximate \( w \), it is natural to design the experiment so that \( \mathbb{E}(u) = w \). To achieve this, set \( \mathbb{P}(\varepsilon_i = 1) = p_i \) and \( \mathbb{P}(\varepsilon_i = 0) = 1 - p_i \).

Now let \( X = |w - u|^2 \); it suffices to prove \( \mathbb{E}(X) \leq n/4 \). Since \( w - u = \sum (p_i - \varepsilon_i) v_i \), we have

\[
\mathbb{E}(X) = \sum \mathbb{E}((p_i - \varepsilon_i) v_i \cdot (p_j - \varepsilon_j) v_j) = \sum (v_i \cdot v_j) \mathbb{E}((p_i - \varepsilon_i)(p_j - \varepsilon_j)).
\]

Since \( p_i - \varepsilon_i \) is \( p_i \) or \( 1 - p_i \) with probabilities \( 1 - p_i \) and \( p_i \), it follows that

\[
\mathbb{E}((p_i - \varepsilon_i)(p_j - \varepsilon_j)) = p_i(1 - p_j) + (1 - p_i)p_j - 2p_i(1 - p_j) = (1 - 2p_i)p_j - p_i(1 - p_j) = (1 - 2p_i)p_j - p_i(1 - p_i).
\]

Since \( p_i \) is \( p_i \) or \( 1 - p_i \) with probabilities \( 1 - p_i \) and \( p_i \), it follows that \( \mathbb{E}((p_i - \varepsilon_i)(p_j - \varepsilon_j)) \leq p_i(1 - p_i) + (1 - p_i)p_i = 1 - 2p_i \). Hence for any \( p_1, \ldots, p_n \), there is a choice for \( \varepsilon_1, \ldots, \varepsilon_n \) such that \( X \leq n/2 \).
number of elements \( a \in A \) such that \( ba \in C \) is greater than \( |A|/3 \). Hence for some choice of \( b \), we obtain a subset \( S \) of \( A \), with \( |S| > |A|/3 \), such that \( ba \in C \) for all \( a \in S \). We cannot have \( a_1 + a_2 = a_3 \) for \( a_1, a_2, a_3 \in S \), because then \( ba_1 + ba_2 = ba_3 \), which implies \( ba_1 + ba_2 \equiv ba_3 \) (mod \( p \)), which is impossible because \( C \) is sum-free.

14.1.44. Let \( A_1, \ldots, A_t \) be events in a probability space of \( n \) equally likely outcomes, with \( p_i = \mathbb{P}(A_i) = |A_i|/n \) and \( s = \sum_{i=1}^t p_i \). Given \( r \in \mathbb{N} \) with \( 1 \leq r \leq s \), there is a set of \( r \) events in \( \{A_1, \ldots, A_t\} \) such that the probability that they all occur is at least \( \binom{t}{r} / s^{r/r!} \), where \( \binom{t}{r} = s^r / r! \).

It suffices to show that \( \binom{s-r}{r} / \binom{s}{r} \) is the average probability of the intersection over all \( \binom{t}{r} \) choices of \( r \) of the events. That is, we show that the average size of the intersection is at least \( n \binom{s}{r} / \binom{s}{r} \), or rather that the sum \( S \) of the intersection sizes is at least \( n \binom{s}{r} \).

Let \( b_i \) be the number of sets in \( A_1, \ldots, A_t \) containing the \( j \)th element of the space; this element contributes once to each of \( \binom{b_i}{r} \) intersections. Note also that \( \sum_{j=1}^n b_j = \sum_{i=1}^t |A_i| = ns \). Over \( b_1, \ldots, b_n \) with fixed sum \( ns \), the convexity of \( \binom{b_i}{r} \) yields the desired result:

\[
S = \sum_{i=1}^n \binom{b_i}{r} \geq n \left( \binom{n}{r} / \binom{s}{r} \right) = n \binom{s}{r}.
\]

14.1.45. The probability that \( n \) points independently chosen uniformly at random on a circle all lie in a semicircle is \( n/2^{n-1} \). Let \( x_1, \ldots, x_n \) be the chosen points. Let \( A \) be the event that the half-open semicircle starting immediate after \( x_i \) contains none of the other points. Since a semicircle occupies half of the perimeter and the placement of the points is independent, \( \mathbb{P}(A_i) = (1/2)^{n-1} \).

The events are pairwise disjoint, because if \( x_j \) is not in the semicircle starting after \( x_i \), then \( x_j \) is in the semicircle starting after \( x_j \). Hence the desired probability is \( \sum_{i=1}^n \mathbb{P}(A_i) \).

14.1.46. When each of \( n \) balls is placed in some positive integer cell, independently, by the distribution choosing cell \( j \) with probability \( 2^{-j} \) for \( j \in \mathbb{N} \), the expected number of empty cells below the highest-indexed occupied cell is 1. Let \( X \) be the number of empty cells below the highest-indexed occupied cell. We have \( X = \sum_{k=0}^\infty X_k \), where \( X_k = 1 \) if cell \( k \) is empty and some higher cell is occupied, and \( X_k = 0 \) otherwise. By linearity of expectation, the desired answer is \( \sum_{k=1}^\infty \mathbb{P}(X_k = 1) \).

Note that a ball is place in cell \( k \) with probability \( 2^{-k} \) after cell \( k \) with probability \( 2^{-k} \), and before cell \( k \) with probability \( 1 - 2^{-k} \). In order to have \( X_k = 1 \), cell \( k \) must be empty, which occurs with probability \( (1 - 2^{-k})^n \). Therefore, the expected number of empty cells below the highest-indexed occupied cell is 1.

Hence, the probability of not reaching \( v \) in the first \( n-1 \) steps when starting from \( u \) is at most \( 1 - 2^{-n} \). If \( u \) is not reached, then after \( n-1 \) steps the walk is at some vertex \( w \), and again there is a path of length at most \( n-1 \) to reach \( v \). In each segment of \( n-1 \) steps, the probability of failing to reach \( v \) is at most \( 1 - 2^{-n} \). Hence the probability of never reaching \( v \) is at most \( (1 - 2^{-n})^{m/(n-1)} \).

This expression bounds the probability of failure for any triple \( (G, u, v) \). Since there are at most \( (n^2)^n \) mazes on \( n \) vertices and at most \( n^2 \) pairs of vertices in each such maze, there are at most \( n^2 + 2 \) possible bad triples. Using linearity of expectation, it suffices to pick \( m \) so that

\[
n^{2n+2} (1 - 2^{-n})^{m/(n-1)} < 1.
\]

Taking logarithms and rearranging reduces the desired inequality to \( (2n+2)(n-1)\ln n < -m \ln(1-x) \), where \( x = 2 \cdot 2^{-n} \). With \( -\ln(1-x) = x + O(x^2) \), it suffices to make \( m \) a bit larger than \( n^2 \ln n \). Indeed, since \( -m \ln(1-x) \) only needs to exceed \( 2(n^2 - 1) \ln n \) rather than \( 2n^2 \ln n \), and \( 1 + n^{-2} > 1 + 2^{-n} \), already \( m = n^2 \ln n \) suffices.
Upper bound. We show that an $n$-universal search list must contain each binary string of length $n-1$ as a consecutive sublist. Give such a string $a_1, \ldots, a_{n-1}$, construct a maze $G$ consisting of the path $(1, \ldots, n)$ plus one edge from $i$ to $1$ for $1 \leq i \leq n-1$ and two edges from $n$ to $1$. For $1 \leq i \leq n-1$, the edge from $i$ to $i+1$ has label $a_i$, and the edge from $i$ to $1$ has label $1-a_i$. To reach $n$ from $1$, the search list must at some point have the bits $a_1, \ldots, a_{n-1}$ in order.

There are $2^{n-1}$ binary lists of length $n-1$, and each successive bit in a search list (after the first $n-2$ bits) ends only one of them. Hence at least $2^{n-1} + n - 2$ bits are needed.

14.2. REFINEMENT OF BASIC METHODS

14.2.1. There is a 2-coloring of the edges of $K_{m,m}$ having no monochromatic copy of $K_{t,t}$, where $m = n - (\binom{n}{t})^2 2^{1-t}$ for some natural number $n$. Begin with a random 2-coloring of $K_{n,n}$. There are $\binom{n}{t}^2$ copies of $K_{t,t}$, and for each the probability of being monochromatic is $2^{1-t}$. Hence the expected number of monochromatic copies of $K_{t,t}$ is $\binom{n}{t}^2 2^{1-t}$. Deleting one vertex (from each side) from each monochromatic copy of $K_{t,t}$ in a 2-coloring having at most the expected number of monochromatic copies leaves a 2-coloring of $K_{m,m}$ (or a graph containing it) with no monochromatic copy.

14.2.2. If $H$ is a hypergraph in which every edge has size at least $k$ and intersects fewer than $\left\lfloor \frac{k^{k-1}}{e} \right\rfloor$ other edges, then $H$ is 2-colorable. We show that in a random coloring, the probability is nonzero that no edge is monochromatic. Let $A_i$ be the event that edge $e_i$ is monochromatic. With edges of size at least $k$, each event $A_i$ occurs with probability at most $2^{-k(k-1)}$. By the Mutual Independence Principle, each event $A_i$ is mutually independent of the set of all edges not intersecting $e_i$, that is, all but at most $2^{k-1}$ edges. That is, the condition $e p d \leq 1$ holds so that the Symmetric Local Lemma applies to guarantee that $H$ is 2-colorable.

14.2.3. A $k$-uniform hypergraph in which every vertex appears in exactly $k$ edges is 2-colorable if $k \geq 9$. Color the vertices randomly. Each edge has probability $2^{1-k}$ of being monochromatic. Let $A_e$ be the event that edge $e$ is monochromatic. By the Mutual Independence Principle, $A_e$ is mutually independent of the set of events for all edges that do not intersect $e$. This excludes $A_e$ and the at most $k^2$ events that it intersects. Hence the Local Lemma applies if $e 2^{1-k}(1 + k^2) \leq 1$, which is true when $k \geq 9$.

14.2.4. $(1 - 1/d)^{d-1} > e^{-1}$ for $d \geq 2$. Note that $1/2 > e^{-1}$, so the inequality holds when $d = 2$. The left side is decreasing for $d \geq 2$, because its derivative is $\frac{1}{d} + \ln(1 - \frac{1}{d})$, which equals $-\sum_{k=2}^{\infty}(1/d)^k$, all terms of which are negative. Thus it suffices to show that the limit is $e^{-1}$. This follows because $(1 - 1/d)^{-1} \to 1$ and in general $(1 + x/n)^n \to e^x$ (by applying l’Hôpital’s Rule to $n \ln(1 + x/n)$).

14.2.5. Given $R(3, k) \geq ck^2/(\ln k)^2$ for some constant $c$, there is a triangle-free $r$-chromatic graph with at most $O((r \ln r)^2)$ vertices. We ignore lower-order terms in the computations. When $n = ck^2/(\ln k)^2$, we have $k \sim \sqrt{c} \ln c \ln n$, shown by substituting into this the formula for $n$ in terms of $k$. By the definition of $R(3, k)$, there is a graph $G$ with $n$ vertices having no triangle and no independent set of size $k$. Its chromatic number is at least $n/k$. Let $r = n/k$, so $r \sim 2\sqrt{\ln n}/\ln n$. In terms of $r$, and ignoring lower-order terms, we have $n \sim \frac{1}{e} r^2(\ln r)^2$, shown by using the formula for $r$ in terms of $n$. Thus $G$ is a triangle-free $r$-chromatic graph with at most $c'(r \ln r)^2$ vertices.

When the lower bound for $R(3, k)$ is improved to $ck^2/\ln k$, the resulting formula for $n$ in terms of $r = \frac{1}{e} r^2 \ln r$.

14.2.6. A graph with $n$ vertices and average degree $d$ has an independent set with at least $n/(2d)$ vertices. This is the lower bound for $R(3, k)$.

14.2.7. If an $n$-vertex graph $G$ has a proper coloring such that every vertex has neighbors in at most $k$ color classes, then $\alpha(G) \geq n/(k+1)$. If $G[S]$ has $X$ vertices and $Y$ edges, then $S$ contains an independent set of size at least $X - Y$, by deleting a vertex of each edge in $G[S]$. We guarantee a large independent set by choosing $p$ to maximize $E[X - Y]$. By linearity of expectation, $E[X - Y] = E[X] - E[Y]$. Note that $E[Y] = np$. Similarly, the probability that a specified edge of $G$ is induced by $S$ is $p^2$, since both its endpoints must be included, so $E[Y] = p^2 nd/2$. Hence $E[X - Y] = np(1 - pd/2)$. We choose $p = 1/d$ to maximize this, which is valid since $d \geq 1$, obtaining $E[X - Y] = n/(2d)$.

14.2.8. Dominating set algorithm. Assume that $G$ has order $n$ and minimum degree $k$.

a) Given $S \subseteq V(G)$ and $U = V(G) - \bigcup_{v \in S} N[v]$, there exists $v \in V(G) - S$
such that $|N[v] \cap U| \geq |U|(k+1)/n$. The set $U$ is the set of vertices not dominated by $S$. Each vertex in $U$ has at least $k$ neighbors, so the sizes of these closed neighborhoods sum to at least $|U|(k+1)$. Since $G$ has $n$ vertices, some vertex $w$ appears in at least $|U|(k+1)/n$ of these sets. Thus $N[w]$ has at least $|U|(k+1)/n$ vertices in $U$. Note that $w$ may be any vertex of $V(G) - S$.

b) After $n \ln(k+1)/(k+1)$ steps of iteratively selecting a vertex with the maximum number of neighbors among undominated vertices, at most $n(k+1)$ vertices remain undominated. By part (a), if there are $r$ undominated vertices at a certain stage, then there are at most $r(1 - (k+1)/n)$ undominated vertices after the next selection. Hence at most $n(1 - (k+1)/n)^{n \ln(k+1)/(k+1)} = n/(k+1)$ undominated vertices remain after $n \ln(k+1)/k + 1$ steps. These vertices combine with the selected vertices form a dominating set of size at most $n1+\ln(k+1)/k+1$.

14.2.9. An $n$-vertex graph with minimum degree $k$ has $k+1/2\ln n$ pairwise disjoint dominating sets. Color the vertices randomly from a set of $r$ colors, where $r = \lceil k+1/2\ln n \rceil$. The probability that a given closed neighborhood omits a given color is at most $(1 - 1/r)^{k+1}$. By the Union Bound, the probability that some closed neighborhood omits some color is at most $nr(1 - 1/k)^{k+1}$, which is bounded by $nre^{-(k+1)/r}$ since $1 + x \leq e^x$. We compute $nre^{-(k+1)/r} = nre^{-2n/k} = nre^{-2} < r/n < 1$, where $r < n$ follows from $k < n$. Hence with positive probability the $r$ sets produced are disjoint dominating sets.

14.2.10. A graph is locally linear if every edge lies in exactly one triangle (every vertex neighborhood induces a matching).

a) A union of disjoint triangles is locally linear if and only if it has no two triangles on four vertices and no three triangles on six vertices ((6,3)-configuration). Let $G$ be locally linear. Two triangles on four vertices share have a common edge, so that cannot occur. Suppose that there are no two such triangles, but there are three triangles on six points. If two of the triangles are disjoint, then the third triangle shares two points with one of them. Otherwise, we have two triangles on five points sharing one vertex $v$. To avoid having two triangles with a common edge, the third triangle must use the sixth point and one from each of the other two triangles, but now the neighborhood of $v$ does not induce a matching.

Conversely, assume that these configurations are forbidden in a union $G$ of disjoint triangles. Suppose that $N_G(v)$ has incident edges $xy$ and $yz$. The edge $vy$ must arise from a triangle involving $v$; we may assume that this triangle is $xvy$, since otherwise we can choose another pair of incident edges to obtain that configuration. The triangle $yxw$ is now forbidden, so $yz$ arises from another triangle $yzw$. However, we also have the edge $vz$, since $z \in N_G(v)$. Since there are no two triangles on four vertices, the triangle containing $vz$ is not completed by any of $\{x, y, w\}$, so it has the form $vzu$. Now there are three triangles on six vertices.

b) There is a locally linear graph on $n$ vertices with at least $(1 + o(1))1/17n^{3/2}$ edges. Generate $m$ triangles $T_1, \ldots, T_m$ at random by choosing triples of vertices from $[n]$, uniformly and independently.

The probability that $T_i$ and $T_j$ are chosen from a given set of four vertices is $3^2/3^2 = 1/3$. By the Union Bound, the probability that $T_i$ and $T_j$ are chosen from some set of four vertices is bounded by $16/3^2$, which is asymptotic to $24/n^2$. The probability that each of $\{T_i, T_j, T_k\}$ is chosen from a given set of six vertices is $3^3/3^3 = 1/3$, so the probability that they are all chosen from some set of six vertices is bounded by $16/3^3 = 2000$, which is asymptotic to $2400/n^3$.

Over the $(m^2)$ sets of size 2 and $(m^3)$ sets of size 3 in $T_1, \ldots, T_m$, by linearity the expected number of bad 2-sets or 3-sets or triples is bounded (asymptotically) by $12m^2/n^2 + 400m^3/n^3$. By deleting one triple from each bad set in an outcome where the expected number of bad sets is at most the expectation, we obtain a set of at least $m - 12m^2/n^2 - 400m^3/n^3$ edge-disjoint triangles whose union is a locally linear graph. Since $(m/n)^3 > (m/n)^2$ when $m > n$, we want to maximize $m - 400m^3/n^3$. The maximum occurs when $1 = 1200m^2/n^3$, so we set $m = n^{3/2}/(20\sqrt{3})$. The subtraction from the middle term is then only linear and does not affect the asymptotics. We obtain $n^{3/2}/(20\sqrt{3}) - 400/(20\sqrt{3})$ [triples, which simplifies to $n^{3/2}/(30\sqrt{3})$] triples, which is asymptotic to $2400/n^3$.

14.2.11. Turán problem for cycles: $\text{ex}(n, C_k) \in \Omega(n^{1+1/(k-1)})$. Generate a random graph with edge probability $p$ (to be specified later). Let $X$ be the number of edges and $Y$ the number of $k$-cycles. If $E[X] < E[Y]$, then deleting an edge from each $k$-cycle in some graph with $X - Y \geq E[X] - E[Y]$ leaves a graph with at least $E[X] - E[Y]$ edges and no $k$-cycle.

We have $E[X] = \binom{n}{2}p$ and $E[Y] = \binom{n}{k}p^{k-1} - \binom{n}{k-1}p^{k-2}$. If $E[Y] < \frac{1}{2}E[X]$, then $\text{ex}(n, C_k) > \frac{1}{2}E[X]$. Since $\binom{n}{k} < n^{k-1}e^{-1/k}$, we have $E[Y] < \frac{n}{e^{k-1}}p^{k-1}$. It suffices to have $\frac{n}{e^{k-1}}p(np)^{-k+1} \leq \frac{n(n-1)}{4}p$, which is implied by $n^{k-2}p^{k-1} \leq 1$. Hence we choose $p = n^{-(k-2)/(k-1)}$. Now there is a $C_k$-free graph with more than $\frac{1}{2}\binom{n}{2}$ edges. This lower bound is asymptotic to $\frac{1}{4}n^{1+1/(k-1)}$.

Comment: For $k = 4$, this lower bound of $\Omega(n^{4/3})$ compares with an easy upper bound of $O(n^{5/2})$. A graph with $m$ edges avoids $C_4$ if and only if any two vertices have at most one common neighbor. Counting and convexity of quadratic functions yield $\binom{n}{2} \geq \sum_{v \in V(G)} \left( \frac{d(v)}{2} \right) \geq n(2m/n)$. Then the resulting quadratic inequality yields $m \leq \frac{n}{2}(1 + \sqrt{4n-3})$. 

Section 14.2: Refinement of Basic Methods
14.2.12. Every $n$-vertex graph with $cn^2$ edges contains a 1-subdivision of $K_a$ with $a = \lceil e^{3/2} \sqrt{n} \rceil$. With $m = |E(G)| = cn^2$, we have $2m/n = 2cn$. Let $r = 2$ and let $t = \lceil \frac{1}{2} \log_{1/2} n \rceil$ (we ignore fractional parts, similarly taking $a = e^{3/2} \sqrt{n}$). The number of vertices in the 1-subdivision of $K_a$ is \( \binom{n}{a} + a \), which is less than $a^2$. Let $b = a^2$. By the choice of $t$, we have $e^t \approx n^{-1/2}$. To apply the Dependent Random Choice Lemma, we compute

$$\frac{(2m/n)^t}{n^{t-1}} - \binom{n}{b} \left( \frac{b}{n} \right)^t = n \left( \frac{2en}{n} \right)^t - \binom{n}{a^2} \left( \frac{a^2}{n} \right)^t$$

$$= n(2e)^t - \binom{n}{2} e^t \approx 2^t \sqrt{n} - \sqrt{n}/2 > \sqrt{n} \geq a.$$

The last inequality uses $\epsilon < 1/2$; the previous inequality uses $t \geq 1$.

Thus $G$ contains a set $U$ of $a$ vertices such that any two vertices in $U$ have at least $a^2$ common neighbors in $G$. Letting $U$ serve as the branch vertices in the 1-subdivision $H$, each pair has enough common neighbors so we can choose distinct common neighbors for the pairs in $U$.

14.2.13. The Ramsey number of the $k$-dimensional cube $Q_k$, satisfies $R(Q_k, Q_k) \leq 2^{3k}$. Consider a red/blue coloring of $K_N$, were $N = 2^{3k}$. Let $m$ be the number of edges in the more plentiful color, which we may assume is red. We have $m \geq \frac{1}{2} \binom{N}{2} \geq \epsilon N^2$, where $\epsilon$ can be taken to be something a bit less than $1/2$, say $\epsilon = 2^{-7/8}$. Note that $2m/N \geq 2^{-4/3} N$.

Let $r = k$, $a = 2^{k-1}$, $b = 2^{k}$, and $t = 3k/2$. To apply the Dependent Random Choice Lemma, we compute

$$\frac{(2m/n)^t}{N^{t-1}} - \binom{N}{b} \left( \frac{b}{N} \right)^t = N \left( \frac{2^{4/3}N}{N} \right)^t - \binom{N}{k} \left( \frac{2^k}{N} \right)^t$$

$$\geq 2^{3k} 2^{-4/3} - \frac{N^k}{k!} 2^{-2kt} = 2^k - \frac{2^{3k^2}}{k!} 2^{-3k} > 2^k - 1 > 2^{k-1} = a.$$

Thus the red graph $G$ contains a set $U$ of $2^{k-1}$ vertices such that every set of $k$ vertices in $U$ has $2^{3k}$ common neighbors in $G$. Using $U$ as one part of the bipartition in $Q_k$, each vertex $v$ in the other part has a desired set $S$ of $k$ neighbors in $U$. Since $S$ has $2^k$ common neighbors in $G$ we can find an unused common neighbor to serve as $v$ until all the vertices of $Q_k$ are placed.

14.2.14. Triangle-free with large chromatic number. Construct a random graph with $n$ vertices by letting each pair be an edge with probability $p$, where $p = n^{-2/3}$. The probability that any given three vertices induce a triangle is $n^{-2}$. Hence the expected number of triangles is $\binom{n}{3} n^{-2}$, which is less than $n/6$.

Section 14.2: Refinement of Basic Methods

In a graph generated with edge probability $p$, the probability that $t$ vertices form an independent set is $(1 - p)^t$, which is bounded by $e^{-pt^2/2}$. By the Union Bound, the probability of having some independent set of size $t$ is bounded by $n^t e^{-pt^2/2}$, which we write as $e^{t(\ln n - pt^2/2)}$. When $t = (2 + \epsilon)^{1/2} \ln n$, then whp there is no such independent set.

We need an outcome with both no independent $t$-set and not too many triangles. When we exclude from the computation of the expected number of triangles the fraction of outcomes with large independent sets, the leading behavior of the expected number of triangles does not change, because the excluded outcomes tend to probability 0 exponentially fast. Being generous, we can guarantee an outcome with no independent $t$-set and at most $n/3$ triangles.

Deleting vertices cannot enlarge independent sets. If we delete $n/3$ vertices from such an outcome (including a vertex from each triangle), we are left with a triangle-free graph having $2n/3$ vertices and chromatic number at least $\frac{n}{2n/3} = \frac{n^2}{6n/3} - \frac{n/3}{2n/3} = \frac{n^2}{6n}$, which equals $\frac{n^{1/3}}{(2+c)\ln n}$.

14.2.15. Asymptotic sharpness of Reed’s Conjecture.

a) Whp, a random $n$-vertex graph generated with independent edge-probability $p$ has no independent set of size $\frac{2^{4/3}}{\ln n}$. The probability that $t$ vertices form an independent set is $(1 - p)^t$, which is bounded by $e^{-pt^2/2}$. By the Union Bound, the probability of having some independent set of size $t$ is bounded by $n^t e^{-pt^2/2}$, which we write as $e^{t(\ln n - pt^2/2)}$. When $t = (2 + \epsilon)^{1/2} \ln n$, then whp there is no such independent set.

b) There is a graph $H'$ with at least $n - \frac{1}{3} n^{3/4}$ vertices having independence number 2 and clique number less than $(2 + c) n^{3/4} \ln n$. For the graph generated in part (a), set $p = n^{-3/4}$. Whp there is no independent set of size $(2 + c) n^{3/4} \ln n$, and the expected number of triangles is less than $(1/6) n^{3/4} p^3$, by the linearity of expectation. When we restrict to the outcomes that have no larger independent set than desired, we lose an exponentially small portion of the probability. Hence this restriction does not asymptotically increase the expectation. To be generous, we may assume an outcome of the experiment with no independent set of size $(2 + c) n^{3/4} \ln n$ and at most $(1/3) n^{3/4}$ triangles. Deleting $(1/3) n^{3/4}$ vertices (including at least one vertex of each triangle) and taking the complement yields the desired graph $H'$, since deleting vertices cannot create larger independent sets.

c) Any $\epsilon$ such that Reed’s Conjecture holds for all graphs satisfies $\epsilon \leq 1/2$. Let $H'$ be the graph produced in part (b). Since $\alpha(H') = 2$, we have $\chi(H') \geq |V(H')| \geq \frac{1}{2} n - \frac{1}{3} n^{3/4}$. Let $G = H' + K_1$; this makes the maximum degree large while increasing the chromatic number and clique number
only by 1. If Reed’s Conjecture holds, then
\[ \frac{1}{2} n - \frac{1}{6} n^{3/4} + 1 \leq \chi(G) \leq e(\omega(H') + 1) + (1 - e)(n - \frac{1}{3} n^{3/4} + 1) \]
This simplifies to \( \frac{1}{2} n \leq (1 - e)n + O(n^{3/4} \ln n) \), which requires \( e \leq 1/2 \) when \( n \) is large.

**14.2.16. Off-diagonal Ramsey numbers.**

a) If \( \binom{n}{p}^2 + \binom{n}{1-p}^2 (1-p)(\frac{1}{p}) < 1 \) for some \( p \in (0, 1) \), then \( R(k, l) > n \). Also \( R(k, l) > n - \binom{n}{p}^2 - \binom{n}{1-p}^2 (1-p)(\frac{1}{p}) \) for all \( n \in \mathbb{N} \) and \( p \in (0, 1) \). Generate a random edge-coloring of \( K_n \) with probability \( p \) of red, \( 1 - p \) of blue, independently. The expected number of red \( k \)-cliques and blue \( l \)-cliques is \( \binom{n}{p}^2 + \binom{n}{1-p}^2 (1-p)(\frac{1}{p}) \). If \( p \) can be chosen to make this less than 1, then there exists an edge-coloring of \( K_n \) with no such monochromatic clique. In any case, for any \( p \), we can choose a coloring in which the number of these monochromatic cliques is at most the expectation and delete one vertex from each to obtain the lower bound in the second statement.

b) \( R(3, k) > k^{3/2 - o(1)} \). In order to obtain a positive lower bound from \( R(3, k) > n - \binom{n}{p}^2 - \binom{n}{1-p}^2 (1-p)(\frac{1}{p}) \), we choose \( p \) and \( n \) in terms of \( k \) so that the subtracted terms are less than \( n/2 \) (constant factors won’t matter). Using upper bound on these terms, we have \( R(3, k) > n(1 - \frac{1}{2} n^{2/3} - \frac{1}{k})(ne^{-pk^2})^{-1} \). The first term suggests setting \( p = cn^{-2/3} \) (it suffices to make \( c \) a constant as small as 1). Since \( k \) may be large, we also want \( e^{pk^2} > n \) (again we can make \( e^{pk^2} = c n \), with \( c \) a constant as large as 1).

Letting \( c = c' = 1 \) and taking the natural logarithm, we want to choose \( n \) so that \( k = 2n^{2/3} \ln n \). We want \( n^{3/2} \) to cancel 2 \( \ln n \) and leave \( k \), so we set \( n = \left( \frac{2}{3 \ln k} \right)^{3/2} \) to obtain \( 2n^{2/3} \ln n = k(1 - c'' \ln \ln k) \).

This choice of \( n \) (and \( p \)) yields \( R(3, k) > n(1 - \frac{1}{6} - \frac{1}{k}) \). Since \( (\ln k)^{-1} = -\ln k \ln k \), we can write this as
\[ R(3, k) > \frac{2n}{3} = \frac{2}{9} k^{3/2 - o(1)} \]
We could optimize the argument by letting \( c \) depend on \( n \) and \( c' \) depend on \( k \), but this won’t improve the exponent.

The first part of (a) yields no useful lower bound on \( R(3, k) \). To obtain \( p \) such that \( \binom{n}{p}^2 < 1 \), we need \( p < cn \). We also need \( e^{p(k-1/2)} > n \), which leads to \( k > 1 + (2/c)n \ln n \). Unfortunately, this works only when \( n \) is smaller than \( k \), and we already know trivially that \( R(3, k) > k \).

**14.2.17. Lower bounds on \( R(4, k) \).** To study \( R(4, k) \), color \( E(K_n) \) with red and blue, each edge being blue with probability \( p \), independently. We want to avoid \( \binom{n}{2} \) bad events of blue \( K_4 \) and \( \binom{n}{2} \) bad events of red \( K_4 \).

Section 14.2: Refinement of Basic Methods

a) \( R(4, k) \geq \Omega((k/ \ln k)^{3/2}) \). Using the Union Bound on the probability of having some bad event, \( \binom{n}{p}^2 + \binom{n}{1-p}^2 (1-p)(\frac{1}{p}) < 1 \) implies \( R(4, k) > n \). It suffices to choose \( n \) and \( p \) so that \( (\binom{n}{2})^2 p^4 + (\binom{n}{2})^2 e^{-r(\frac{1}{p})^2} \) < 1. We set \( p = \frac{\ln k}{c k} \) and \( n = \frac{k}{c} \left( k^{3/2} \right), \) with \( c \) to be chosen later. The bound on the probability becomes \( (4c)^{-4} + \frac{k^{1/2}}{c^{(3/2)}} e^{-r(\frac{1}{p})^2} \). Since \( k^{1/2} = \frac{1}{4} \ln k \ln k \), the second term tends to 0 as \( k \) tends to \( \infty \). Hence when \( k \) is sufficiently large, choosing \( c > 1/4 \) yields \( R(4, k) \geq \frac{k}{c} \cdot \left( k^{3/2} \right)^{3/2} \).

b) \( R(4, k) \geq \Omega((k/ \ln k)^{5/2}) \). We obtain a lower bound on \( R(4, k) \) by choosing \( p \) and \( n \) so that the expected number of blue copies of \( K_4 \) and red copies of \( K_4 \) are each at most \( n/4 \). From a coloring that achieves this, we remove one vertex from each blue \( K_4 \) and each red \( K_4 \) to obtain a coloring that proves \( R(4, k) > n/2 \). Let \( n = \frac{k^2}{10^4 (\ln k)^2} \) and \( p = \frac{1}{30} n^{-1/2} \). We compute
\[ \left( \frac{n}{4} p^8 < \frac{k^8}{24 \cdot 10^{12} (\ln k)^6} 50^8 k^6 \right) \]
and
\[ \left( \frac{n}{k} (1-p)^{\frac{1}{2}} < \left( \frac{ke}{10^4 (\ln k)^2} \right) k^{-10(k-1)/100} \right) \]
\[ \leq \left( \frac{e}{10^4 (\ln k)^2} k < \frac{n}{4} \right) \]
c) \( R(4, k) \geq \Omega((k/ \ln k)^{5/2}) \). To apply the Local Lemma, we have events \( A_5 \) for blue copies of \( K_4 \) and \( B_T \) for red copies of \( K_4 \). We follow the computations of Theorem 14.2.22. We want weight \( y \) for each \( A_5 \) and \( z \) for each \( B_T \) such that \( \mathbb{P}(A_5) \leq y w(D_S) \) and \( \mathbb{P}(B_T) \leq z w(D_T) \), where \( A_5 \) is mutually independent of the events outside \( D_S \) and \( w(D_T) \) is the product of the quantities “1 − weight” for the events in \( D_S \), and similarly for \( D_T \).

In \( D_S \), we include the events for all \( k \)-sets and the events for all 4-sets sharing at least two vertices with \( S \). There are fewer than \( 6n^2 \) of the latter. We do the same for \( D_T \); there are fewer than \( \binom{\frac{k}{2}}{2} \) 4-sets in \( D_T \).

Hence it suffices to choose \( p \), \( y \), and \( z \) so that \( p^6 < y(1 - y)^{n/2} (1 - z)^{\frac{1}{2}} \) and \( (1-p)^{\frac{1}{2}} < z(1-z)^{n/2} (1-z)^{\frac{1}{2}} \).

Use \( p = c_1 n^{-2/5} \) to produce the coloring. Let \( y = c_2 n^{-12/5} \) and \( z = c_3 e^{-n^{-2/5} (\ln n)^2} \). Note that \( n = c_4 (k/ \ln k)^{5/2} \), then \( k \sim c_5 n^{3/2} \ln n \). With these choices, \( y \) is a constant multiple of \( p^6 \). Since \( (1-p)^{\frac{1}{2}} < e^{-pk^{1/2}} \), it suffices to show \( c < (1 - y)^{n/2} (1 - z)^{\frac{1}{2}} \) and \( e^{-pk^{1/2}} < z(1-z)^{n/2} (1-z)^{\frac{1}{2}} \).

In the upper bounds, we can use \( \binom{n}{2}^2 < n^k \) and \( e^{\ln n} \). We omit the remaining details of the computation.
14.2.18. Existence of injective functions from \([m]\) to \([n]\). Make all \(n^m\) functions \(f : [m] \rightarrow [n]\) equally likely.

The Union Bound guarantees positive probability of an injection if \(n > \binom{m}{2}\). The bad events \(A_{i,j}\) have the form \(f(i) = f(j)\), each of which has probability \(1/n\). The number of such events is \(\binom{m}{2}\). By the Union Bound, the probability of having some bad event is at most \(\frac{\binom{m}{2}}{n^2}\). If this quantity is less than 1, then some function has no bad events. Thus there exists an injective function when \(n > \binom{m}{2}\).

The Local Lemma guarantees positive probability of an injection if \(n \geq 2en\). Consider the same bad events, with the same probability \(1/n\). By the Mutual Independence Principle, the event \(A_{i,j}\) is mutually independent of the set \(D_{i,j}\) of all events \(A_{k,l}\) such that \(\{i, j\} \cap \{k, l\} = \emptyset\). (Note that if \(A_{i,k}\) and \(A_{j,k}\) occur, then \(A_{i,j}\) must also occur, even though the events are pairwise independent.) Letting \(d\) be the number of events having \(i\) or \(j\) in their index, the Local Lemma yields positive probability of an injection if \(epd \leq 1\), where \(p = 1/n\). In fact, \(d = 2m - 1\), since the indices are chosen from \([m]\). Thus \(n \geq 2em\) suffices. This is a much smaller threshold than \(\binom{m}{2}\).

14.2.19. Two dependent events \(A_1\) and \(A_2\) with probability \(p\) can satisfy the hypotheses of the Local Lemma if and only if \(p < 25\). We seek weights \(x_1\) and \(x_2\) such that \(p \leq x_1(1-x_2)\) and \(p \leq x_2(1-x_1)\). If these inequalities both hold, but \(x_1 < x_2\), then \(x_1 - x_1x_2 < x_2 - x_1x_2\), and reducing \(x_2\) will increase the smaller bound while decreasing the large bound, so both bounds will still hold. Hence we may assume \(x_1 = x_2\). Now we need \(p \leq x(1-x)\), which is satisfiable if and only if \(p < 1/4\).

Comment: The value 1/4 here is closely related to the Neighborhood Local Lemma.

14.2.20. If \(H\) is a graph with maximum degree \(k\), and \(k \leq s^{3/2}/\sqrt{4e}\), then \(H\) has an \(s\)-edge-coloring in which no 4-cycle is monochromatic. Color \(E(H)\) uniformly and independently at random from a set of \(s\) colors. The event that a particular copy of \(C_4\) is monochromatic has probability \(s^{-3}\).

The number of 4-cycles through a given edge is at most \((k-1)^2\), since this is the maximum number of copies of \(P_4\) containing it. By the Mutual Independence Principle, the event that a particular copy of \(C_4\) occurs is mutually independent of the set of all events for copies of \(C_4\) sharing no edges with it. There are fewer than \(4k^2\) such events deleted.

With \(d = 4k^2\) and \(p = s^{-3}\), it suffices to have \(epd \leq 1\) (by the Local Lemma). This holds when \(k \leq s^{3/2}/\sqrt{4e}\).

14.2.21. Let \(S\) be a set of pairwise-disjoint cliques in a graph \(G\) with maximum degree \(k - 1\). If each member of \(S\) has at least \(ck\) vertices, and \(c > 2e/(1+2e)\), then \(G\) has an independent set consisting of one vertex from each member in \(S\). For example, \(c = 11/13\) is large enough.

From each member of \(S\), pick one vertex at random (equally likely). Define a (bad) event \(A_e\) for each edge \(e\) whose endpoints \(x\) and \(y\) lie in distinct members of \(S\); the event occurs if \(x\) and \(y\) are both selected. The desired independent set exists if there is positive probability of avoiding all such events. Event \(A_e\) occurs with probability at most \((ck)^{-2}\).

The occurrence of \(A_e\) is determined only by the random choices that select vertices from the cliques containing its endpoints. If such a clique \(Q\) has size \(tk\), then a vertex of \(Q\) has \(tk-1\) neighbors in \(Q\) and at most \((1-t)k\) neighbors outside \(Q\). Hence at most \(tk(1-t)k\) events depend on sets of random choices that include the choice made in \(Q\). Since \(t \geq c \geq 1/2\) (by the choice of \(c\)), we have \(t(1-t) \leq c(1-c)\). Therefore, accounting for both endpoints of the edge \(e\) and eliminating \(e\) itself, the Mutual Independence Principle implies that each event is mutually independent of a set of all but at most \(2e(1-c)k^2\) other events. With \(d = 2e(1-c)k^2\), the Local Lemma guarantees the desired independent set when \(epd < 1\), where \(p = (ck)^{-2}\). Hence \(2e(1-c)/c < 1\) suffices, and this occurs when \(c > 2e/(1+2e)\).

14.2.22. Van der Waerden numbers. We show that when \(n\) is too small, there exists a \(k\)-coloring of \([n]\) with no monochromatic \(l\)-term arithmetic progression. Hence we generate a random \(k\)-coloring of \([n]\).

a) \(w(l,k) \geq (lk^{l-1})^{1/2}\). We use the existence method; suppose \(n < (lk^{l-1})^{1/2}\). For each \(l\)-term arithmetic progression, the probability of being monochromatic in a random \(k\)-coloring is \(k \cdot k^{-l}\).

There are fewer than \(n^2/(l-1)\) such progressions, since there are fewer than \(n\) possible starting values and fewer than \(n/(l-1)\) ways to choose the constant difference. In fact, with constant difference \(d\) there are \(n-(l-1)d\) possible starting values, so we could replace \(n^2/(l-1)\) with \(\sum_{d=1}^{n} \lfloor n/(l-1)d \rfloor\), which is about half as big. We are content with the loose upper bound of \(n^2/l\) on the sum.

If \(k^{l-1}n^2/l < 1\), then there exists a \(k\)-coloring of \([n]\) with no monochromatic \(l\)-term arithmetic progression, and thus \(w(l,k) > n\). This holds when \(n < (lk^{l-1})^{1/2}\).

b) \(w(l,k) \geq (1-o(1))k^l/(ekl)\). We use the Local Lemma. For each \(l\)-term arithmetic progression \(S\), let \(A_S\) be the event that \(S\) is monochromatic. Events \(A_S\) and \(A_T\) are independent when \(|S \cap T| \leq 1\), but \(A_S\) is not mutually independent of the set of all events for progressions intersecting \(S\) at most once. If they are all monochromatic with the same color, then \(S\) is likely also to be monochromatic with that color.

By the Mutual Independence Principle (Proposition 14.2.12), \(A_S\) is mutually independent of the set of events for progressions that \(S\) does not
14.2.23. If \( d^+(v) = d^-(v) = k \) for every vertex \( v \) in a digraph \( G \), then \( G \) has \( \left\lfloor \frac{k/(2.5 + 2 \ln k)}{k} \right\rfloor \) pairwise disjoint cycles. Let \( r = \left\lfloor \frac{k/(2.5 + 2 \ln k)}{k} \right\rfloor \). If \( V(G) \) partitions into \( r \) nonempty sets such that each vertex has a successor in its own set, then there is a cycle in each subgraph induced by a set in the partition. If we put all vertices in \( V_1, \ldots, V_r \) at random, then some \( V_i \) may be empty. We avoid this by first placing one vertex into each of the \( r \) sets, choosing these vertices arbitrarily. Each remaining vertex is put into one of \( V_1, \ldots, V_r \) uniformly at random, independently.

Let \( A_j \) be the event that the \( j \)th vertex \( v_j \) has no successor in its own set. We may have forced as many as \( r - 1 \) successors of \( v_j \) already into sets other than the one containing \( v_j \). The probability that no other successors distributed at random are in the same set as \( v_j \) is at most \((1 - 1/r)^{k-r+1}\). Let \( p = (1 - 1/r)^{k-r} \), so \( \mathbb{P}(A_j) < p \). We use the Local Lemma to prove \( \mathbb{P}(\bigcap_j A_j) > 0 \), implying that the desired partition exists.

Let \( S_j = N^+(v_j) \cup \{v_j\} \). Event \( A_j \) is determined by the placement of \( S_j \), so \( A_j \) is mutually independent of the set of all \( A_i \) such that \( S_i \cap S_j = \emptyset \) (by the Mutual Independence Principle). Since all vertices in \( S_j \) have \( k \) predecessors, at most \( k + k(k-1) \) vertices other than \( v_j \) have successors in \( S_j \) (we count predecessors of \( v_j \) plus an upper bound on the number of vertices other than \( v_j \) with successors in \( N^+(v_j) \)). Also we have the vertices of \( S_j \setminus \{v_j\} \) itself. Thus \( A_j \) is mutually independent of a set of all but at most \( k^2 + k \) other events.

Let \( d = k^2 + k + 1 \). If \( k = 2 \), then \( r = 1 \), and the lack of sinks yields a cycle. If \( k \geq 3 \), then \( k^2 + k + 1 \leq \frac{3}{2} k^2 \). If \( epd < 1 \), then the Local Lemma completes the proof. We compute \( epd < e(1 - 1/r)^{k-r+1} < e^{2-k/r} \frac{3}{2} k^2 \). Taking logarithms, \( epd < 1 \) follows from \( k/r \geq 2 + \ln \frac{3}{2} + 2 \ln k \). Since \( \ln \frac{3}{2} < .5 \), this holds when \( r = \left\lfloor \frac{k/(2.5 + 2 \ln k)}{k} \right\rfloor \). (The additive constant in the denominator can be pushed closer to 2, since \( k/(2 + 2 \ln k) \) is 1 for \( k \leq 14 \). Thus it suffices to make the constant 2 + \ln \frac{341}{225} \), which is less than 2.07.)

### 14.2.24. When \( L \) is a \( k \)-uniform list assignment for a graph \( G \), and vertices are colored uniformly at random from their lists, independently, edge-violation events are not helpful for guaranteeing an \( L \)-coloring. Let \( A_{xy} \) be the event that \( x \) and \( y \) receive the same color. The lists on \( x \) and \( y \) may or may not be identical; we can only guarantee \( \mathbb{P}(A_{xy}) \leq 1/k \).

By the Mutual Independent Principle, \( A_{xy} \) is mutually independent of the events for edges not incident to \( xy \). If we exclude the remaining edges, we may be excluding as many as \( 2D - 1 \), where \( D = \Delta(G) \). In order to ensure \( epd \leq 1 \), where \( p = 1/k \) and \( d = 2D - 1 \), we may need \( k \geq 2eD \).

This is useless, because we know trivially that \( G \) is \((\Delta(G) + 1)\)-choosable.

### 14.2.25. Guaranteeing many colors on small subgraphs.

a. \( c_{r,s} n^{r-2}/(r-s-1) \) colors suffice for a coloring of \( E(K_n) \) such that every copy of \( K_r \) has at least \( s \) colors on its edges. Form a random edge-coloring of \( K_r \) with \( k \) colors. For each \( r \)-set \( S \), let \( A_S \) be the event that fewer than \( s \) colors appear on edges within \( S \). We have \( \mathbb{P}(A_S) < (s-1)/(r-1) \).

Also, the occurrence of \( A_S \) is unaffected by the colors on any edges not contained in \( S \). By the Mutual Independence Principle, \( A_S \) is mutually independent of the set of all events for \( r \)-sets sharing at most one vertex with \( S \). For the application of the Local Lemma, this yields \( d \leq \left(\frac{r}{2}\right)^{r-2} \) (there is a lot of overcounting).

If \( k \) is chosen so that the product of these two upper bounds is less than \( 1/e \), then the Symmetric Local Lemma implies that \( k \) colors suffice to guarantee a coloring with the desired property. We use \( n^{r-2}/(r-s+1) \) and \( (s-1)/(r-1) \) to conclude that \( n^{-2} k^{r-1} - k^{-s} \leq k^{-s} \), where \( k \) depends only on \( r \) and \( s \). Solving for \( k \), it suffices to let \( k \) be as large as \( c_{r,s} n^{r-2}/(r-s+1) \), where \( c_{r,s} \) is small enough so that \( e^{-s} < c_{r,s} < e^{s} \).

b. \( c_{r,s} n^{2(r-1)}/[r^{s-1}] \) colors suffice for a coloring of \( E(K_{n,n}) \) such that every copy of \( K_r \) has at least \( s \) colors on its edges. The argument is analogous to part (a). Form a random edge-coloring of \( K_{n,n} \) using \( k \) colors. For each choice of \( r \)-sets \( S \) and \( T \) chosen from the two parts, let \( A_{S,T} \) be the event that fewer than \( s \) colors appear on edges joining \( S \) and \( T \). We have \( \mathbb{P}(A_{S,T}) < \left(\frac{k}{n}\right)^{r-2} \).

Again, the occurrence of \( A_{S,T} \) is unaffected by the colors on any edges not joining \( S \) and \( T \). By the Mutual Independence Principle, \( A_S \) is mutually independent of the set of all events for \( S' \) such that \( S' \cap T' \cap S \cup T = \emptyset \) (we could allow more pairs, but it doesn’t gain much). For the application of the Local Lemma, this yields \( d < \left[r^{-2} \right] \). With the analogous simplified upper bounds, setting \( k = c'_{r,s} n^{2(r-1)}/[r^{s-1}] \) suffices for a sufficiently small \( c'_{r,s} \) in terms of \( r \) and \( s \).
14.2.26. Let $H$ be a hypergraph in which every edge has size at least 3, each edge of $H$ intersects at most $a_r$ (other) edges of size $r$. If $\sum a_r 2^{-r} \leq \frac{1}{8}$, then $H$ is 2-colorable. Color the vertices uniformly at random, independently. Let $A_i$ be the event that edge $i$ is monochromatic: $P(A_i) \leq \frac{1}{2}$. By the Mutual Independence Principle, $A_i$ is mutually independent of the set of all events for edges disjoint from $e_i$. Edge of size $r$ has probability $2^{1-r}$ of being monochromatic. Hence $\sum_{j \in E} P(A_j) \leq \sum_{r \geq 3} a_r 2^{-1-r} \leq \frac{1}{4}$, by the hypothesis. The Neighborhood Local Lemma now implies that $H$ is 2-colorable.

The symmetric Local Lemma requires the condition $edp \leq 1$, where $d$ is a uniform upper bound on the number of events not in the set of events mutually independent of $A_i$. For a hypergraph as described above, since $p = \frac{1}{2}$, applying the symmetric Local Lemma would require $d \leq 1$; if there is any edge of size 3, then the edges would have to be pairwise disjoint, that is, $\sum a_r = 0$.

14.2.27. Weighted Neighborhood Local Lemma. Let $A_1, \ldots, A_n$ be events in a probability space. For each $i$, let $D_i \subseteq [n] - \{i\}$ be such that $A_i$ is mutually independent of $\{A_j : j \not\in D_i \cup \{i\}\}$. Let $t_1, \ldots, t_n$ and $z$ be real numbers such that $t_i \geq 1$ and $0 \leq z \leq \frac{1}{2}$.

a) Proof of the lemma. With $P(A_i) \leq z^{t_i} \leq \frac{1}{2}$, we have $0 \leq x_i \leq \frac{1}{2}$, where $x_i = (2z)^{t_i}$. By the General Local Lemma, it suffices to prove $P(A_i) \leq x_i \prod_{j \in D_i} (1 - x_j)$ for these weights.

Let $\alpha = 2\ln 2$. When $0 \leq x \leq \frac{1}{2}$, we again have $e^{-\alpha x} \leq 1 - x$. Using $(2z)^{t_i} = x_i$ and $t_i/2 \geq \sum_{j \in D_i} (2z)^{t_j}$,

$$P(A_i) \leq z^{t_i} = (2z)^{t_i} e^{-\alpha t_i/2} \leq x_i e^{-\alpha \sum_{j \in D_i} (2z)^{t_j}} = x_i \prod_{j \in D_i} e^{-\alpha (2z)^{t_j}} \leq x_i \prod_{j \in D_i} (1 - x_j).$$

b) If $H$ is a hypergraph in which every edge has size at least 3 and each vertex of $H$ lies in at most $b_i$ edges of size $i$, where $\sum b_i 2^{-i/2} \leq 1/(6\sqrt{2})$, then $H$ is 2-colorable. Generate a 2-coloring of the vertices at random, uniformly and independently. Let $A_i$ be the event that edge $e_i$ is monochromatic, so $P(A_i) = 2^{k-1}$. Let $D_i$ be the set of indices for edges other than $A_i$ that intersect $e_i$. By the Mutual Independence Principle, $A_i$ is mutually independent of the events outside $\{a_j : j \not\in D_i \cup \{i\}\}$.

Set $z = 1/4$ and $t_i = (|e_i| - 1)/2$. Now $P(A_i) = z^{t_i}$. Summing over all edges that intersect $e_i$,

$$\sum_{j \in D_i} (z^2)^j \leq \sum_{j \in D_i} \sum_{j \geq 3} b_j (z^2)^j \leq \sum_{j \in D_i} \sum_{j \geq 3} b_j \left(\frac{1}{3}\right)^{j-1} \leq \sqrt{2} |e_i| \sum_{j \geq 3} b_j 2^{-j/2} \leq |e_i| \left(\frac{2t_i + 1}{6}\right) \leq \frac{t_i}{2},$$

where the last inequality uses $|e_i| \geq 3$. Hence the Weighted Neighborhood Local Lemma applies, and some coloring has no monochromatic edge.

14.3. Moments and Thresholds

14.3.1. For any fixed graph $H$, when $p$ is constant, whp $G^p$ contains $H$ as an induced subgraph. Let $k$ and $l$ be the numbers of vertices and edges of $H$. As in Theorem 14.3.18, there are $n(k_l)/A$ potential copies of $H$ in $G^p$ on $n$ vertices. Each arises as an induced subgraph of $G^p$ with probability $p^l (1 - p)^{(k_l - 1)}$, which is a constant. Hence $E(X) \sim n k_l / A \to \infty$.

We seek $P(X = 0) \to 0$. For this, it suffices by the Second Moment Method to show $E(X^2) \to E(X)^2$. As in Theorem 14.3.18, the contribution for each possible subgraph $H'$ of $H$ as an overlap of two copies of $H$ with $r$ common vertices grows as a multiple of $n^{2k-4}$. The ratio of this to $E(X)^2$ tends to 0. Since there is only a constant number of such subgraphs $H'$, their total contribution is still $O(E(X)^2)$. On the other hand, the contribution for disjoint copies of $H$ is asymptotic to $E(X)^2$, since the number of ways to choose the vertices for two disjoint copies is asymptotic to $n^{2k}$.

14.3.2. Generalization of Theorem 14.3.10.

a) For fixed $k, s, t, p$, whp $G^p$ has the following property: for every choice of disjoint vertex sets $S$ and $T$ of sizes $s$ and $t$, there are at least $k$ vertices that are adjacent to every vertex of $S$ and to no vertex of $T$. For vertices outside given sets $S$ and $T$, the probability of failure is $1 - p^s (1 - p)^t$, as in Theorem 14.3.10. A pair $(S, T)$ is bad if at least $n - s - t - k + 1$ vertices fail; the probability of this event consists of the top tail of $k$ terms in the binomial distribution. The probability is bounded by $k \binom{n-s-t-k+1}{n-s-t-1} (1 - p)^{k-1}$, where $q = 1 - p^s (1 - p)^t$.

With $k, s, t, p$ fixed, this probability is a polynomial in $n$ times $q^n$, where $q < 1$. Multiplying by $(s,t,n-s,t-k)$ (again a polynomial in $n$) yields the expected number of bad pairs. Hence the expectation tends to 0 as $n \to \infty$, and whp there are no bad pairs.

b) $whp$ $G^p$ is $k$-connected. Set $s = 2$ and $t = 0$. We have shown that whp any two vertices have at least $k$ common neighbors, and hence every separating set has size at least $k$. 

In a random tournament, whp for every choice of disjoint vertex sets $S$ and $T$ of sizes $s$ and $t$, there are at least $k$ vertices with an edge to every vertex of $S$ and from every vertex of $T$. The computations in part (a) apply here with $p = 1/2$. That is, a vertex succeeds for the pair $(S, T)$ if it has $s$ edges to $S$ and $t$ from $T$, so the probability of failure is $1 - 2^{-s+t}$. The remainder of the argument is the same.

14.3.3. Whp the first $\log_2(n\omega_n)$ vertices in $[n]$ form a dominating set in the graph $G^p$, where $p \in (0, 1)$ is fixed, $b = 1/(1 - p)$, and $(\omega)$ is any sequence tending to $\infty$. Let $k = \log_2(n\omega_n)$. We have specified vertices in $S$. A subsequent vertex fails to dominate with probability $(1 - p)^k$. Let $X$ be the number of undominated vertices; we have $E(X) = (n - k)(1 - p)^k$. It suffices to show $E(X) \to 0$. We compute $E(X) < n(1 - p)^k = nb^{-\log_2(n\omega_n)} = n(n\omega_n)^{-1} = 1/\omega_n \to 0$.

14.3.4. The smallest connected graph that is not balanced has five vertices and six edges, obtained from $K_4^-$ by adding a vertex of degree 1. Trees are balanced, with average degree $2(n - 1)/n$, since $2(n - 1)/k < 2(n - 1)/n$ when $k < n$. Also unicyclic graphs are balanced, since they have average degree 2, and any subgraph is unicyclic or is a forest (with average degree less than 2). To avoid these classes we need at least four vertices, but $K_4^-$ and $K_4$ are balanced, having average degrees 2.5 and 3 (other subgraphs are unicyclic or acyclic).

A graph $G$ obtained by adding a pendant edge to $K_5^-$ (at any vertex) has average degree 12/5, which equals 2.4, but the subgraph $K_4^-$ has average degree 2.5, so $G$ is not balanced. A 5-vertex graph with at most five edges is unicyclic or acyclic or $K_4^{-} + K_1$, so those are all balanced.

14.3.5. If $p$ is fixed and $k \in o(n/\ln n)$, then whp $G^p$ is $k$-connected. It suffices to show that whp any two vertices have $k$ common neighbors. Let $X$ be the number of pairs failing this. The probability that $u$ and $v$ have exactly $j$ common neighbors is $\binom{n}{j}p^j(1 - p)^{n - j}$. When $p$ is fixed, this probability increases with $j$ for $j < n/2$. Hence the value when $j < k - 1$ is bounded by the value when $j = k$, so we multiply that value by $1/k$ to have an upper bound. To bound $E(X)$, we also multiply by $\binom{n}{k}$.

To simplify the bound, use $\binom{n}{k}(n - k)^{2j} < n^k$ and $(p^j)^k < 1$ and $(1 - p^2)^{n - j} < (1 - p^2)^{n/2} < e^{-p^2n/2}$. Therefore, $E(X) \leq n^ke^{-p^2n/2} = e^\ln n - e^{-p^2n/2}$. Since $k \in o(n/\ln n)$ and $p$ is fixed, the positive term in the exponent grows more slowly than the negative term. Hence the exponent tends to $-\infty$, and $E(X) \to 0$. By Markov’s Inequality, the probability that every pair of vertices in $G^p$ has at least $k$ common neighbors tends to 1.

14.3.6. Threshold probability for the existence of cycles in the random graph $G(n, p)$.

If $pn \to 0$, then almost always $G$ has no cycles. Let $X$ be the number of cycles; we want $P(X = 0) \to 1$. By Markov’s Inequality, it suffices to show that $E(X) \to 0$. Each cycle with length $k$ has probability $p^k$ of occurring, and there are $n(n-1)/2k$ possible such cycles. We compute

$$E(X) = \sum_{k=3}^{\infty} \frac{n(n-1)}{2k} p^k < \sum_{k=1}^{\infty} (np)^k = \frac{np}{1 - np} \to 0,$$

$$P(X \leq k) = \sum_{j=0}^{k} \binom{n(n-1)}{2j} p^j(1 - p)^{n - 1 - j} \leq (k + 1)^{n^2k}/k!p^k.$$
we can view the coin flip to determine whether two vertices are adjacent as occurring at the time when we want to test that pair. In particular, we cannot determine whether $A_t$ occurs until after knowing the vertices eliminated in the earlier rounds.

At round $i$, we consider the $2k-2i+1$ pairs consisting of the least-indexed remaining vertex and one of the other remaining vertices. Event $A_t$ occurs only if all those pairs fail to be edges, which occurs with probability $(1/2)^{2k-2i+1}$. By the Union Bound, the probability that some round fails is at most $\frac{1}{2} \sum_{i=1}^{k} \left(\frac{1}{2}\right)^{k-i}$. Since the geometric sum is less than 4/3, the probability of failure for the algorithm is less than 2/3.

14.3.9. For fixed positive $\epsilon$, with $p = 1/n$, almost always $G$ has no component with more than $(1 + \epsilon)n/2$ vertices. It suffices to show that the probability of $G$ having at least $(1 + \epsilon)n/2$ edges tends to 0, since this big a component must have at least that many edges. Let $X$ be the number of edges. By Markov’s Inequality, $P(X \geq t) \leq E(X)/t$. With $E(X) \sim n/2$ and $t = (1 + \epsilon)n/2$, the ratio does not tend to 0, so this is not strong enough.

However, Chebyshev’s Inequality (Markov’s Inequality for $X$) is strong enough. We have $P(X - E(X))^2 \geq t^2) \leq E(X^2) - E(X)^2$.

Note that $X$ is a binomial random variable with $m$ trials and success probability $p$, where $m = \binom{n}{2}$ and $p = 1/n$. The trials (edge slots) are independent, so $E(X^2) = mp + m(m-1)p^2$. Also $E(X) = mp$, so $E(X^2) - E(X) = mp - mp^2 - mp(1-p)$.

Our upper bound on the probability becomes $\frac{n-1}{2} \left(1 - \frac{1}{n}\right) \frac{1}{(en/2)^2}$, which tends to 0.

14.3.10. The probability that no two among $s$ people with random birthdays have the same birthday within $t$-day years tends to 0 as $s \to \infty$ if $t \in o(s^2)$. Let $P(s, t)$ denote the stated probability.

Since $P(s, t) = \prod_{i=0}^{s-1} \left(1 - \frac{t}{i+1}\right)$, we have $-\ln P(s, t) = -\sum_{i=1}^{s-1} \ln(1-\frac{i}{t})$. Also $-\ln(1-x) > x$ for $x \in (0, 1)$. Hence

$$-\ln P(s, t) > \sum_{i=1}^{s-1} \frac{i}{t} > \frac{(s-1)s}{2t} \to \infty$$

as $s \to \infty$ when $t \in o(s^2)$. Thus $P(s, t) \to e^{-\infty} = 0$ as $s \to \infty$.

14.3.11. Given edge probability $p = \frac{1}{2} + \frac{1}{q}$ with $c < 1$, whp $G^p$ has at least $\frac{1}{2} \binom{n}{2}$ edges. Let $X$ be the number of edges, so $E(X) = \frac{1}{2} \binom{n}{2} + \frac{n-1}{2q-1}$. With $t = \frac{n-1}{2q-1}$, we want $P(|X - E(X)| \geq t) \to 0$. Since $X$ is a binomial random variable, $E(X^2) - E(X)^2 = \left(\binom{n}{2}\right)p(1-p)/t^2$. Since $p(1-p) \sim \frac{1}{4}$, the bound is asymptotic to $\frac{1}{2} \left(\frac{2}{n}\right)^{n-2}$, which is asymptotic to $\frac{1}{2} n^{2c-2}$.

When $p = \frac{1}{2} - \frac{1}{n}$ with $c < 1$, whp $G^p$ has at most $\frac{1}{2} \binom{n}{2}$ edges. The computation is essentially the same as above.

When $p$ is in the interval $\frac{1}{2} \pm \frac{1}{n}$ with $c > 1$, the expected number of edges is asymptotic to $\frac{1}{2} \binom{n}{2}$, and the asymptotic probabilities of having at least $\frac{1}{2} \binom{n}{2}$ or at most $\frac{1}{2} \binom{n}{2}$ edges are both positive.

14.3.12. For $p = \sqrt{(c/n) \ln n}$, when $c \geq 2$ whp every edge in $G^p$ lies in a triangle, and when $c < 2$ whp that property fails, as long as $c$ is constant.

Let $X$ be the number of edges of $G$ not belonging to triangles. There are $\binom{n}{2}$ possible edges. A pair $(x, y)$ contributes 1 to $X$ if it is an edge and not in a triangle. The probability of this is $p(1 - p^{2n-2}) < \frac{1}{2} n^{2e-n(2c-1)}$.

Let $p = \sqrt{(c/n) \ln n}$ with $c \geq 2$. Now $p^2(n-2) \sim c \ln n$, so the bound on $P(X)$ is asymptotic to $\frac{1}{2} n^{2c-2} \sqrt{(c/n) \ln n}$. With $c \geq 2$, we obtain $P(X) \to 0$, so whp every edge belongs to a triangle.

When $c < 2$, we have $P(X) \to \infty$, so we consider the Second Moment Method to prove the threshold. We aim to show $E(X^2) \to E(X)^2$. Since $X$ is a sum of indicator variables, $E(X^2)$ is the sum of $E(X)$ and the sum of probabilities of distinct edges appearing and not being in triangles. For each of the $6 \binom{n}{4}$ ordered pairs sharing an endpoint, the contribution is $p^2(1-p)(1-2p^2 + p^3)^{n-3}$. The contribution from each of the $6 \binom{n}{4}$ ordered pairs of two disjoint edges is $p^2(1-p)^22^{n-4}$. Since $6 \binom{n}{4} \sim \binom{n}{2}$, this total contribution is asymptotically $E(X)^2$. Thus it suffices to show that the contribution from pairs of edges sharing an endpoint has lower order.

With $p = \sqrt{(c/n) \ln n}$, we have $E(X^2) \sim (c/4) n^{3-2c} \ln n$. Asymptotically, the contribution from the pairs of edges sharing an endpoint is bounded by $n^3(c/n)(\ln n)^2 \sim (c/4) n^{3-2c} \ln n$, which is asymptotic to $n^{2-2c} \ln n$. Thus we need $2 - 2c < 3 - 2c$, which holds.

When $c$ is not constant, the approximations in the arguments fail. In particular, if $c = b/(n \ln n)$ with $b < 1$, then $p = b/n$ and there is positive probability of having no cycles at all and hence no triangles.

14.3.13. $n^{-1/p}$ is a threshold probability function for the appearance of $H$ as a subgraph of $G^p$, where $\rho$ is the maximum density of $H$. Let $F$ be a subgraph of $H$ with density $\rho$. Since $F$ is balanced, $c_n n^{-1/p}$ is a threshold function for the appearance of $F$. As in Theorem 14.3.18, when $c_n \to 0$ whp $G^p$ has no copy of $F$ and hence no copy of $H$.

Let $X$ count the copies of $H$ that appear. Let $k = |V(H)|$ and $l = |E(H)|$, and let $A$ be the number of automorphisms of $H$. Each copy of $H$ arises under $A$ of the $n(k)$ ways to map $V(H)$ into $[n]$, so there are $n(k)/A$
possible copies. Thus $\mathbb{E}(X) \sim n^{k^2}p^k/A$. Since $\rho \geq 1/k$, setting $p(n) = c_nn^{-1/\rho}$ yields $\mathbb{E}(X) \sim c_n^2n^{-k^2-1}/A \geq c_n^2/A$. Hence $\mathbb{E}(X) \to \infty$ when $c_n \to \infty$.

By the Second Moment Method, we need only prove $\mathbb{E}(X^2) \sim \mathbb{E}(X)^2$ when $c_n \to \infty$. This uses computations essentially the same as in Theorem 14.3.18. Again let $E_{HI}$ be the sum of the contributions to $\sum E(X,X')$ from all pairs $(H_i,H_j)$ (both being potential copies of $H$) such that $H_i \cap H_j = H'$. For each pair such that $H_i$ and $H_j$ share no edges, $\mathbb{P}(X,X' = 1) = p^{2l-s}$. Since we must choose vertices for each copy, the number of such pairs is bounded by $(n^6/A^2)^2$. This is an overcount, since many such choices share edges, but the total contribution to $\mathbb{E}(X^2)$ from these pairs is bounded by a quantity asymptotic to $\mathbb{E}(x)^2$.

It thus suffices to show that the contribution from pairs sharing at least one edge has lower order. When $H'$ has $s$ edges, $H_i \cup H_j$ has $2l-s$ edges, so $\mathbb{P}(X,X = 1) = p^{2l-s}$. Also $H'$ has $r$ vertices, where $r \geq 2$. Since $H'$ has maximum density $p$, we have $s/r \leq p$, so $2l-s \geq 2l-pr$. This means $\mathbb{P}(X,X' = 1) \leq p^{2l-pr}$. To bound the number of such pairs, we pick at least two vertices for the second subgraph from the set used by the first subgraph. Hence the contribution for the remaining pairs is bounded by

$$\frac{1}{2}(\sum_{r=2}^k n(n-k)(k-r)p^{2l-pr}$$

Since $k$ is fixed, this is bounded by $O\left(\sum_{r=2}^k n^{2k-r}p^{2l-pr}\right)$, which is $\mathbb{E}(X)^2O\left(\sum_{r=2}^k n^{-r}p^{2l-pr}\right)$. With $p = c_nn^{-1/\rho}$, the summation becomes $\sum_{r=2}^k c_n^{2k-r}$. With $c_n \to \infty$ and $r \geq 2$, all $k-2$ terms in this sum tend to 1, so the total contribution from these pairs is $o(\mathbb{E}(X)^2)$ as desired.

### 14.3.14. In almost every tournament, every vertex is a king.

Every vertex is a king when every vertex be reachable from every other vertex by a path of length at most 2. Let $X$ be the number of ordered pairs of vertices where both fails. Failing to reach $y$ from $x$ by a path of length at most 2 requires that for $w \notin \{x,y\}$, the edges $xw$ and $wy$ do not both have that specified orientation. Hence the probability that the ordered pair $(x,y)$ fails is bounded above by $(3/4)^{n-2}$ (orienting $xy$ toward $x$ yields another factor of 1/2, but that is not important).

Since there are $n(n-1)$ ordered pairs, $\mathbb{E}(X) \leq n(n-1)(3/4)^{n-2}$. Since this bound tends to 0 as $n \to \infty$, Markov’s Inequality implies that almost every tournament has no bad pairs and thus has every vertex being a king.

### 14.3.15. Transitive subtournaments.

a) For $e > 0$, almost every tournament has no transitive subtournament with more than $2\log n + (1+\epsilon)$ vertices. Let $X$ be the number of transitive subtournaments of order $k$. On each set of $k$ vertices, a transitive tournament corresponds to an ordering of the vertices. Hence $\mathbb{E}(X) = \binom{n}{k}k!2^{-\binom{k}{2}}$.

By Markov’s Inequality, it suffices to show that $\mathbb{E}(X) \to 0$ (or $\log \mathbb{E}(X) \to -\infty$) when $k > (1+\epsilon)2\log n$. We have $\mathbb{E}(X) \leq n^k2^{-\binom{k}{2}}$, which we write as $\log \mathbb{E}(X) \leq k\log n -(k-1)/2$. If $k > 1+\epsilon+2\log n$, then $\log \mathbb{E}(X) < -ke/2$, and the bound tends to $-\infty$ as $n \to \infty$.

Comment: Every $n$-vertex tournament $T$ has a transitive subtournament with $\log n$ vertices. We prove by induction on $n$ that $T$ has a transitive subtournament with at least $1 + \lfloor \log n \rfloor$ vertices. This is trivial for $n=1$. When $n > 1$, some vertex $x$ has outdegree at least $\lceil (n-1)/2 \rceil$, which equals $\lfloor n/2 \rfloor$. In the subtournament induced by $N^+(x)$, the induction hypothesis yields a transitive subtournament of order at least $1 + \lfloor \log(\lfloor n/2 \rfloor) \rfloor$, which equals $\lfloor \log n \rfloor$. Adding $x$ produces a transitive subtournament of order $1 + \lfloor \log n \rfloor$ in $T$.

b) $R(k,k) > 2(k-2)^2$. Generate a random tournament $T$ with vertex set $[n]$. Convert $T$ to a 2-edge-coloring of $K_n$ by coloring each edge oriented from lower to higher endpoint red and each edge oriented from higher to lower endpoint blue. If the coloring has a monochromatic copy of $K_k$, then $T$ has a transitive subtournament with $k$ vertices (the converse is false). When $k > 2+\log n$, the transitive subtournament does not exist. Hence $\log n$ has no monochromatic $K_k$, yielding $R(k,k) > n$. Solving for $n$ in terms of $k$ yields $R(k,k) > 2(k-2)^2/k$.

### 14.3.16. If $k = \log n - (2 + \epsilon)\log \log n$, then almost every tournament, every set of $k$ vertices has a common successor. The probability that a $k$-set has no common successor is $(1 - 2^{-k})^{n-k}$, since each vertex outside the set must not be a common successor. Let $X$ be the number of $k$-sets with no common successor; we have $\mathbb{E}(X) \sim \binom{n}{k}(1 - 2^{-k})^{n-k}$. An upper bound on $\mathbb{E}(X)$ is \((\frac{2}{3})^k e^{2^{-k}(n-k)}\). If this bound tends to 0 for some choice of $k$ in terms of $n$, then the claim holds for this choice of $k$.

We seek $k$ such that \((\frac{2}{3})^k e^{2^{-k}(n-k)}\) grows more slowly than $e^{2^{-k}(n-k)}$. Taking natural logarithms, we want $k(\ln n + 1 - \ln k) < 2^{-k}(n-k)$. Factoring out $k$ from the second factor on the right and $\ln n$ from the second factor on the left and then taking base-2 logarithms, we want

$$\log k + \log \ln n + (1 - \frac{\ln k}{\ln n}) - k < \log n + \log(1 - \frac{\epsilon}{n}).$$

Roughly speaking, we want $k + \log k < \log n - \log \ln n$. That is, $k$ should be enough less than $\log n$ that adding $\log k$ still keeps the value less than $\log n - \log \ln n$. Converting from $\log n$ to $\log \ln n$ on the right only introduces an additive constant, since the $\ln n$ is inside $\log$. The $\epsilon$ in the definition of $k$ is more than enough to take care of that.

Setting $k$ as specified above yields $\mathbb{E}(X) \to 0$, and the property almost always holds.
14.3.17. Given \( p = (1 - \varepsilon) \ln n/n, \) whp \( G^p \) has at least \((1 - \varepsilon)n^c\) isolated vertices, where \( c \) is any positive constant. As in the proof of the threshold function for the disappearance of isolated vertices (Theorem 14.3.17), the expected number of isolated vertices is \( n^c\). Chebyshev’s Inequality bounds the probability that \( X \) deviates much from its expected value. In particular, letting \( X \) be the number of isolated vertices, \( X < (1 - \varepsilon)n^c \) requires \( X \) to deviate by at least \( cn^c \) from its expectation, so we set \( t = c\mathbb{E}(X) = cn^c \) in Chebyshev’s Inequality. This yields

\[
\mathbb{P}[X < (1 - \varepsilon)n^c] < c^{-2} \left( \frac{\mathbb{E}(X^2)}{n^c} - 1 \right).
\]

The computation in Theorem 14.3.17 shows \( \mathbb{E}(X^2) \sim \mathbb{E}(X)^2 \), and hence \( \mathbb{P}(X \geq (1 - \varepsilon)n^c) \to 1 \) as \( n \to \infty \), for any positive constant \( c \).

14.3.18. For fixed \( p \), with \( b = 1/p \), a threshold for the maximum \( k \) such that whp every \( k \)-set in \( V(G^p) \) has a common neighbor is \( c \log_b n \) with \( c = 1 \). That is, setting \( k = c \log_b n \), when \( c < 1 \) whp \( G^p \) has no \( k \)-set with no common neighbor, and when \( c > 1 \) whp \( G^p \) has such a \( k \)-set.

Say that a vertex subset is bad if its elements have no common neighbor. Let \( X \) be the number of bad \( k \)-sets. A fixed \( k \)-set \( S \) is bad with probability \((1 - p)^{kn-k} \), since it is contained in the neighborhood of a fixed vertex outside \( S \) with probability \((1 - p)^{kn-k} \). Thus \( \mathbb{E}(X) = \sum_{k=0}^{\infty} \mathbb{P}(X = k) = \sum_{k=0}^{\infty} (1 - p)^{kn-k} \).

Note that \( \mathbb{E}(X) < \left( \frac{1}{1-p} \right)^k e^{-p(n-k)} \). Writing \( k = c \log_b n \) yields \( p^k = n^{-c} \). Since \( k \in \mathbb{N}(n) \), we can write \( \mathbb{E}(X) < n^c e^{-n^{1-c}} \). Hence \( \ln \mathbb{E}(X) < k \ln n - n^{1-c}/2 \). If \( c < 1 \), then the bound tends to \(-\infty\) and \( \mathbb{E}(X) \to 0 \). By Markov’s Inequality, \( \mathbb{P}(X = 0) \to 1 \), and almost every \( G^p \) has no bad \( k \)-set.

When \( c > 1 \), we need an asymptotic expression for \( \mathbb{E}(X) \) to apply the Second Moment Method. With \( k \in \mathbb{N}(n) \), the factor \( n-k \) is asymptotic to \( n \). For \( (1 - p)^{kn-k} \) to be asymptotic to \( e^{-p^k n} \), we need \( np^{2k} \to 0 \). That is, \( n^{1-2c} \to 0 \), which is valid when \( c > 1/2 \).

Hence \( \mathbb{E}(X) \sim \left( \frac{1}{1-p} \right)^k e^{-n^{1-c}} \to \infty \) when \( c > 1 \). When \( k \)-sets \( S \) and \( T \) are disjoint, the events corresponding to them being bad are independent, since they are determined by disjoint sets of pairs of vertices. Since the probability of an individual \( k \)-set being bad tends to 1, and the probability of overlapping \( k \)-sets being bad is at most 1, to prove \( \mathbb{E}(X^2) \sim \mathbb{E}(X)^2 \) it suffices to show that the number of pairs of overlapping \( k \)-sets has smaller order than the number of pairs of disjoint \( k \)-sets.

To form two disjoint \( k \)-sets, we pick \( 2k \) elements and partition them; there are \( 2 \binom{n}{k} \) ways to do this, asymptotic to \( \frac{1}{2} \binom{n}{k} \) since \( k \in \mathbb{N}(n) \). The number of pairs with \( r \) common elements is \( \frac{1}{2} \binom{k}{r} \binom{n-k}{r} \). The ratio to the number of disjoint pairs is asymptotic to \( \frac{2^{r}(r)!}{(1-r)(r)!} \). This ratio is bounded by \((k^2/n)^r\). The ratio is largest when \( r = 1 \), and there are \( k \) possible values of \( r \), so the proof is completed by observing that \( k^2/n \to 0 \).

14.3.19. The property \( Q_k \): for every choice of disjoint vertex sets \( S \) and \( T \) of size \( k \), there is an edge with endpoints in \( S \) and \( T \).

a) Candidate for a threshold. For \( G \in \mathbb{G}(n, p) \), let \( X \) count the pairs \( S, T \in \binom{\frac{n}{2}}{\frac{k}{2}} \) with \( [S, T] = \emptyset \). For any such pair, the indicator variable has probability \((1 - p)^{k^2} \) of being 1. Hence \( \mathbb{E}(X) = (\binom{k}{2} \binom{n-k}{2})^{1/2} \), where \( q = 1 - p \). Since \( k \) is fixed, \( \mathbb{E}(X) \sim n^{2k}q^{k^2}/(k^2) \). If \( p \) is chosen so that \( q = o(n^{2k}) \), then \( \mathbb{E}(X) \to 0 \) and almost every graph satisfies \( Q_k \), by Markov’s Inequality. If \( q \) is larger, so that \( n^{2k}q^k \to \infty \), then \( \mathbb{E}(X) \to \infty \).

Hence \( q = cn^{2k} \) will be thresholded a function for \( Q_k \) if \( n^{2k}q^k \to \infty \), guarantees that almost no graph in \( \mathbb{G}(n, p) \) satisfies \( Q_k \). The Second Moment Method suffices if \( \mathbb{E}(X^2) \sim \mathbb{E}(X)^2 \). Since \( X \) is a counting variable, \( \mathbb{E}(X^2) = \mathbb{E}(X) + 2 \sum_{i<j} \mathbb{P}(X_i = X_j = 1) \), with the \( X_i \) indexed by the pairs of disjoint \( S, T \). Since \( k \) is fixed, the number of pairs \([S, T], (S', T')\) with \( S, T, S', T' \) pairwise disjoint is asymptotic to \( n^k/(k!)^2 \), and the probability that the variables for two such pairs both have value 1 is \( q^{k^2} \). Thus these terms contribute asymptotically \( \mathbb{E}(X^2) \to \mathbb{E}(X)^2 \), and it suffices to show that the total contribution from all other terms has lower order.

The steps for doing so are like those in the threshold for the appearance of balanced graphs. In fact, since “not \( Q_k \)” is the event \( K_{k,k} \in G \) and \( K_{k,k} \) is strictly balanced, this is a special case of Theorem 14.3.18.

b) For \( k = c \log n \) with \( c > 2 \), almost every graph \((p = 1/2) \) has property \( Q_k \). Again, let \( X \) be the number of pairs \( S, T \in \binom{n}{k} \) such that \( [S, T] = \emptyset \). The approximation \( \binom{n}{k} \sim n^k k! \) is valid when \( k = c \log n \). Hence \( \mathbb{E}(X) \sim (n^2(1-p)^k k!) \). With \( k = c \log n \), we have \( n^2(1/2)^k = n^{2-c} \). Hence \( \mathbb{E}(X) \to 0 \) when \( c > 2 \). By Markov’s Inequality, almost surely there are no bad pairs, and \( Q_k \) holds. From the complementary event, we conclude that almost every graph fails to contain \( K_{k,k} \) when \( k = c \log n \) with \( c > 2 \).

14.3.20. Whp, the maximum length of a constant string among \( n \) flips of a fair coin is \( (1 + o(1)) \ln n \). Let \( k = (1 + \varepsilon) \ln n \) with \( \varepsilon \) constant. We show that when \( \varepsilon < 0 \) whp there is a constant string of length \( k \), while when \( \varepsilon > 0 \) whp there is no such string.

Proof 1. Let \( X \) count the positions that begin strings of \( k \) heads, so \( X = \sum X_i \), where \( X_i = 1 \) if such a strings begins at position \( i \). We have \( \mathbb{E}(X) = \frac{n-k+1}{2^{k-1}} \binom{n}{k} \). For \( k = (1 + \varepsilon) \ln n \), \( \mathbb{E}(X) \) is asymptotic to \( n^{-\varepsilon} \), which approaches 0 when \( \varepsilon > 0 \) and \( \infty \) when \( \varepsilon < 0 \). (Including strings of tails merely doubles \( \mathbb{E}(X) \).) Hence \( \mathbb{P}(X = 0) \to 1 \) when \( \varepsilon > 0 \), and by the Second Moment Method \( \mathbb{E}(X^2) \sim \mathbb{E}(X)^2 \) imply \( \mathbb{P}(X = 0) \to 0 \) when \( \varepsilon < 0 \).
other than $i$ is sent (probability $1 - 1/n$), or $i$ is sent and fails (probability $p/n$). Thus the failure probability at each time is $1 - (1 - p)/n$. With $q = 1 - p$, we have $E(X_i) = (1 - q/n)^m < e^{-qm}/n$. Let $m = cn ln n$, so $E(X_i) < n^{-c}$, and $E(X) < n^{-1-c}$. If $c > 1/q$, then $E(X) → 0$ and almost surely every message arrives.

If $c < 1/q$, then $E(X) → 0$. To obtain $P(X = 0) → 0$ so that almost surely some message fails to arrive, we use the Second Moment Method. We need $E(X) ∼ n^{1-c}$. The approximation $(1 - x)^m$ is valid when $mx^2 → 0$, so we want $a^2m/n^2 → 0$. Since $m/n^2 = (c ln n)/n$ and $q < 1$, the approximation is valid (even if $q$ is a function of $n$).

As usual $E(X^2) = E(X) + 2 ∑_{i<j} P(X_i X_j = 1)$. Let $A$ be the event that $i$ and $j$ never arrive. This has probability $1 - 2/n + p^2/n$ at each time, so $P(A) = (1 - 2q/n)^m ∼ e^{-2qm/n}$. With $2$ such events, $E(X^2) - E(X) ∼ n^{2-2c} ∼ E(X)^2$, and the Second Moment Method applies.

**14.3.23.** The value $n ln n + (1 + ε)n ln n$ is a threshold for the number of trials needed to obtain $n$ equally likely outcomes more than once each. For $ε > 0$ it almost surely happens, and for $ε < 0$ it almost surely does not.

For $m$ trials, let $X$ count the outcomes occurring at most once. The probability that outcome $j$ occurs at most once is $(1 - p)^m + mp(1 - p)^{m-1}$, where $p = 1/n$. Let $b = [1 - p + mp (1 - p)^{m-1}]$, by linearity, $E(X) = nb$.

With $m/n = n ln n + (1 + ε)n ln n$ and $p = 1/n$, where $ε$ is constant, we have $(1 - p)^{m-1} < e^{-mp} = 1/(n ln n)^{1+ε}$. Also $1 - p + mp ∼ mp ∼ ln n$. Thus $E(X) = nb ∼ nmpe^{-mp} ∼ (n ln n)^{1+ε}$.

If $ε > 0$, then $E(X) → 0$, and almost always every outcome occurs more than once. If $ε < 0$, then $E(X) → ∞$. By the Second Moment Method, it suffices to show $E(X^2) ∼ E(X)^2$.

Letting $X_i$ be the indicator variable for failing to get outcome $i$, we have $E(X^2) = E(X) + n(n-1)E(X_i X_j)$. The probability of failure for both outcome $i$ and outcome $j$ is

$$(1 - 2p)^m + 2mp(1 - 2p)^{m-1} + m(m - 1)p^2(1 - 2p)^{m-2}.$$ 

Constant powers of $(1 - 2p)$ tend to 1, while $(1 - 2p)^m ∼ e^{-2mp}$. Also $mp → ∞$, so the third term in the expression above is dominant, and $E(X_i X_j) ∼ (mp)^2e^{-2mp}$. Thus $E(X^2) ∼ (nmp)^2e^{-2mp}$. Since $E(X) ∼ nmpe^{-mp}$, we have $E(X^2) ∼ E(X)^2$, as desired.

**14.3.24.** If $Q$ is the property that each inverse image under a random function $f$ from $[m]$ to $[n]$ has size more than $k$, then $m(n) = n ln n + kn ln n$ is a (fairly narrow!) threshold function for $Q$. Let $X$ be the number of points in $[n]$ hit at most $k$ times. For each $r ∈ [n]$, the probability that $f^{-1}(r)$ has size $j$ is $n!/(j!)((1 - p)^{m-j})$, where $p = 1/n$. The probability that $f^{-1}(r)$ has
size at most \( k \) is the summation of this up to \( j = k \). Let this probability be \( b \); thus \( \mathbb{E}(X) = nb \).

We claim that the contribution to \( b \) from terms with \( j < k \) is of lower order than the term when \( j = k \). Let \( a = (1 - p)^{m-k+1} \). We bound the sum by a multiple of a geometric sum. If \( mp \to \infty \), this yields

\[
b = \sum_{j=0}^{k-1} \binom{m}{j} p^j (1-p)^{m-j} \leq \sum_{j=0}^{k-1} (mp)^j = a \left( \frac{mp}{mp-1} \right)^k \sim a(mp)^k
\]

On the other hand, \( (\binom{m}{k}) p^k (1-p)^{m-k} \) is bounded below by a constant times \( a(mp)^k \). Hence \( mp \to \infty \) and \( k \) constant yields \( b \sim \binom{m}{k} p^k (1-p)^{m-k} \).

We want to choose \( m(n) \) so that \( nb \) approaches 0 or \( \infty \), depending on the choice of a parameter in \( m(n) \). Since \( k \) is constant, \( (1-p)^k \to 1 \) and the binomial coefficient in the top term is asymptotic to \( m^k / k! \). Thus \( b \sim \frac{1}{k!} m^k p^k (1-p)^m \). Also \( np^2 \to 0 \), so \( 1-p \) is asymptotic to \( e^{-p} \).

With \( m(n) = n \ln n + cn \ln n \), we have \( mp = n \ln n + cn \ln n - \ln n \) and \( em^p = n \ln n \). We now compute

\[
\mathbb{E}(X) = nb \sim \frac{(mp)^k}{k! e^{mp}} \sim n \frac{(ln n)^k}{k! (ln n)^k} = \frac{1}{k!} (ln n)^{k-c}.
\]

If \( c < k + \epsilon \), then \( \mathbb{E}(X) \to 0 \), and almost always every target point is hit more than \( k \) times. If \( c < k - \epsilon \), then \( \mathbb{E}(X) \to \infty \). The Second Moment Method then will imply that almost always some target point is hit at most \( k \) times if we prove that \( \mathbb{E}(X^2) \sim \mathbb{E}(X^2) \).

Let \( X = \sum_{r=1}^n X_r \), where \( X_r \) is the event that \( |f^{-1}(r)| \leq k \). The probability that \( |f^{-1}(r)| = i \) and \( |f^{-1}(s)| = j \) is \( \binom{m}{i,j,m-i-j} p^i (1-2p)^{m-i-j} \), from the multinomial distribution. Thus sum of this over \( i,j \) both at most \( k \) equals \( \mathbb{E}(X, X_s) \). Again because \( m \) grows while \( k \) is fixed, the sum is asymptotic to the single term with \( i = j = k \). The multinomial coefficient is asymptotic to \( m^k / (k! k!) \). With the other approximations as above, we have

\[
\mathbb{E}(X^2) = \mathbb{E}(X) + \sum_{r<s} \mathbb{E}(X_r, X_s) \sim \mathbb{E}(X) + n^2 b^2 \sim \mathbb{E}(X^2).
\]

**Comment:** The special case with \( k = 0 \) strengthens the Coupon Collector Problem. Not only does the expected number of trials to obtain all prizes equal \( n \ln n \), but also the probability is concentrated around the expectation. If we perform only \( n \ln n - \epsilon \ln n \) trials, then almost never will we get all the prizes, but if we perform \( n \ln n + \epsilon \ln n \) trials, then almost always we will get all the prizes.

**14.3.25. If trials have \( n \) equally likely outcomes, independently, then after \( n(\ln n + x) \) trials the number of outcomes that have not occurred is Poisson distributed with mean \( e^{-x} \).** Let \( X \) count the outcomes not seen in \( m \) trials. Let \( X_i \) be the indicator variable for the \( i \)th outcome not being seen. Since the trials are independent and uniform, \( \mathbb{P}(X_i = 1) = (1 - \frac{1}{m})^m \) for any set of \( r \) indices. Hence the binomial moment \( \mathbb{E}(X^k) \) equals \( \binom{m}{k} (1 - \frac{1}{m})^m \).

For fixed \( r \), we have \( (\binom{m}{k}) p^k (1-p)^{m-k} \) to \( \infty \). Also, with \( m = n(\ln n + x) \) we have \( m(r/n)^2 \to 0 \), and hence \( (1-r/n)^m \to e^{-r/n} \). Thus \( \mathbb{E}(X) \sim \frac{1}{r!} n e^{-r/(\ln n + x)} = e^{-rx} / r! \). By the Poisson Paradigm (Convergence of Moments Method), \( X \) has the claimed distribution.

**14.3.26. For a real constant \( c \), the asymptotic probability that a graph drawn from \( G(n, c/n) \) has no triangle is \( e^{-c^3/6} \).** Let \( X \) be the number of triangles appearing; \( X \) is the sum of \( \binom{n}{3} \) indicator variables that have probability \( p^3 \) of being 1 , where \( p = c/n \). Thus \( \mathbb{E}(X) = \binom{n}{3} = n^3 / 3^3 / c^3 / 6 \). For fixed \( r \), there are asymptotically \( n^r (r!3!) \) choices of \( r \) variables for vertex-disjoint triangles, each having probability \( p^r \) of all occurring.

Sets that share some edges have higher probability of occurring, but the number of them has smaller order in \( n \). The total of these terms, the number of which is bounded in terms of \( r \), is bounded by \( o(n^3) \). Thus \( S_r(X) \) is asymptotic to \( \frac{1}{r!} (\mathbb{E}(X))^r \). By the Poisson Paradigm (Convergence of Moments Method), the distribution of the number of triangles is Poisson with mean asymptotic to \( c^3 / 6 \). Hence the probability of avoiding all triangles is asymptotic to \( e^{-c^3/6} \).

**14.3.27. In \( G(n, p) \) with \( p = 1 - \frac{k \ln n + x}{n} \), the probability of having a dominating set of size \( k \) is asymptotic to \( 1 - e^{-kx/k!} \).** Let \( X \) be the number of dominating sets of size \( k \).

In \( G(n, p) \), the probability that a particular \( k \)-set \( S \) of vertices is a dominating set is the probability that every vertex outside \( S \) has at least one neighbor in \( S \). This probability is \( 1 - (1-p)^{n-k} \). When \( p \to 1 \), constant powers of \( 1-p \) tend to 0, and hence the probability is asymptotic to \( 1 - (1-p)^{n-k} \). Note also that \( n(k \ln n + x)^2 \to 0 \). Hence \( X \) is a sum of \( \binom{n}{k} \) indicator variables \( X_i \), such that

\[
\mathbb{P}(X_i = 1) \sim [1 - (1-p)^{k}]^n = (1 - \frac{k \ln n + x}{n})^n \sim e^{-k \ln n + x} = e^{-kx}.
\]

We conclude \( \mathbb{E}(X) \to e^{-kx} / k! \). Proving \( S_r(X) \to \mathbb{E}(X)^r / r! \) would show that \( X \) is Poisson distributed with mean \( e^{-kx} / k! \), and hence the probability of having a dominating set of \( k \) would be \( 1 - \mathbb{P}(X = 0) \), equal to \( 1 - e^{-\mathbb{E}(X)} \).

**14.3.28. **Not complete*** Sharp threshold for diameter at most 2.** Given \( G \) from \( G(n, p) \), let \( X \) be the number of pairs of nonadjacent vertices having no common neighbor. Note that \( \text{diam}(G) \leq 2 \) if and only if
X = 0. Note that X = \sum X_i, where X_i is the indicator variable for the
ith vertex pair being nonadjacent and having no common neighbor.
As in Example 14.3.12, E(X) = \binom{n}{2}(1 - p)(1 - p)^{n-2}, and p small
enough so that np^4 \to 0 implies E(X) \sim \frac{1}{2} n^2 e^{-np^2}. With p = \left(\frac{\ln n}{n}\right)^{1/2},
we have E(X) \sim \frac{1}{2} n^2 e^{-\epsilon n}. If c = 2 + \epsilon, then \Pr(X = 0) \to 1 and whp diam(G) \leq 2.
Proving E(X^2)/E(X)^2 \to 1 when c = 2 - \epsilon establishes \left(\frac{\ln n}{n}\right)^{1/2} as
a threshold probability function for the property diam(G) \leq 2, by the
Second Moment Method. Recall that E(X^2) = E(X)^2 + \sum_{i \neq j} E(X_i X_j).
When X_i and X_j refer to disjoint pairs, the events E(X_i X_j = 1) and
E(X_i = 1) and E(X_j = 1) are almost independent. We have \Pr(X_i X_j = 1) = (1 - p)^2 \binom{n}{2} \binom{2n - 4}{n - 2} (1 - 4p^2(1 - p)^2 - 4p^2(1 - p) - p^4),
where the last factor is the probability that neither pair finds a common neighbor in the other pair and tends to 1. Hence \Pr(X_i X_j = 1) \to \Pr(X_i = 1)^2.
To obtain a sharp threshold with an asymptotic probability e^{-\epsilon} for
this property, we need to prove S_r(X) \to \mu'/r! for the rth binomial moment
S_r(X), defined to be \sum X_i X_j \cdots X_k, where the sum is taken over all
r-sets of the indicator variables summing to X.
More generally, for any fixed r we have \Pr(X_i \cdots X_k) \to
14.3.29. ***Not complete*** For a discrete random variable X, if E(e^{tX})
is finite for some interval of t around 0, then knowing all the moments of
a distribution of X is equivalent to knowing the distribution. Let I be an
open interval containing 0 such that E(e^{tX}) is finite for t \in I. With p_k =
\Pr(X = k),
\[
E(e^{tX}) = \sum_{k \geq 0} p_k e^{tk} = \sum_{k \geq 0} \sum_{n \geq 0} p_k \frac{(tk)^n}{n!} = \sum_{n \geq 0} \sum_{k \geq 0} p_k k^n \frac{t^n}{n!} = \sum_{n \geq 0} \sum_{k \geq 0} E(X^n) \frac{t^n}{n!}.
\]
Here, we have used that all terms in the summation are positive, so that
the sum converges absolutely and terms can be summed in any order.
If X and Y are nonnegative integer-valued with the same moments,
and both E(e^{tX}) and E(e^{tY}) are finite for t \in D, then the above computation yields
E(e^{tX}) = E(e^{tY}). Letting also q_k = \Pr(Y = k), we obtain
\[
\sum_{n \geq 0} \sum_{k \geq 0} \left(\frac{p_k - q_k}{k} k^n\right) \frac{t^n}{n!} = 0.
\]

When this is 0 for an interval of t around 0, how do we conclude p_k = q_k
for all k?

Comment: For continuous random variables, having finite moments
is not enough for the moments to determine the distribution. See N.I.
Akhiezer, The classical moment problem and some related questions in

\[\text{Section 14.3: Moments and Thresholds} \]

14.3.30. Formula for \Pr(X = k) in terms of binomial moments. For X =
\sum_{i=1}^m X_i, let S_r(X) be the rth binomial moment of X. Given S_r(X) =
\sum_{j=r}^{m} \binom{m}{j} \Pr(X = j), we compute
\[
\sum_{r=k}^{m} (-1)^{r-k} \binom{r}{k} S_r(X) = \sum_{r=k}^{m} (-1)^{r-k} \binom{r}{k} \sum_{j=r}^{m} \binom{m}{j} \Pr(X = j) = \sum_{j=k}^{m} \Pr(X = j) \sum_{r=k}^{m} (-1)^{r-k} \binom{r}{k} \binom{m}{j} = \Pr(X = k).
\]

14.3.31. \sum_{i=0}^{a} (-1)^i \binom{a}{i} = \begin{cases} 1 & \text{if } b = 0 \\ 0 & \text{if } 0 < b < a \\ (-1)^b \frac{b-1}{a} & \text{if } b \geq a \end{cases}
If b = 0, then all terms are 0 except the term l = 0, which equals 1.
If 0 < b < a, then by the Binomial Theorem the value is (1 - 1)^b,
which equals 0 since a > 0.
If b > a, then the induction step for a proof by induction on b (basis
b = a) is
\[
\sum_{i=0}^{a} (-1)^i \binom{a}{i} = \sum_{i=0}^{a} (-1)^i \binom{a-1}{i} \binom{b-1}{l} = (-1)^a \binom{b-2}{a} - \sum_{i=0}^{a-1} (-1)^i \binom{b-2}{a-1} \binom{b-2}{a-1} = (-1)^a \binom{b-1}{a}
\]

14.3.32. ***Solution not yet written*** Let p = c/n for constant c.
(a) Prove for s < n(20c^2) that whp no set of s vertices in G^p induces
at least 2s edges.
(b) Prove that whp \Delta(G) \leq \ln n / \ln n.
(c) Prove that whp the vertices of degree at least \frac{2\ln n - 3\ln \ln n}{3\ln n} form an independent set.

14.3.33. The probability that a random d-regular graph generated using
the pairing model has no triangle tends to e^{-\left(d-1\right)^6 \frac{m}{n}}. Let X be the number
of triangles in the multigraph G generated by the pairing model.
When two vertices are selected from each of three groups, they can produce
three pairings that correspond to a triangle in eight ways. Since

we can generate the pairing by iteratively matching an arbitrary vertex to a random unmatched vertex, the probability of generating three such specified pairs is \( \frac{1}{n}\frac{1}{n-1}\frac{1}{n-2}\). There are \( \binom{n}{3} \) ways to choose two vertices from each of the three groups, and there are \( \binom{n}{3} \) ways to choose three groups. Hence \( \mathbb{E}(X) \sim \frac{8}{6\sqrt{\pi}}(d-1)^{3/2} n^{-1} \ln n \rightarrow (d-1)^{3/2}/6. \)

For fixed \( r \), it then suffices to show \( S_r(\lambda) \rightarrow \lambda /r! \), where \( \lambda = (d-1)^{3/2}. \) Pick \( r \) of the indicator variables for triples of vertices that generate a triangle. The resulting \( 3r \) pairs all occur asymptotically with probability \( \prod_{j=1}^{n} \frac{1}{3n-2-r} \) (when multiple triples use a particular pair, the probability is different, but the frequency of this has lower order. Because this probability is asymptotic to \( \left( \frac{3}{1-1/(dn)} \right)^{3/2} \) and we can generate the \( r \) triples in any order, \( S_r(\lambda) \rightarrow \lambda /r! \), as desired.

**14.3.34.** The expected number of cycles in a random 2-regular graph generated by the pairing model is asymptotic to \( \frac{1}{2} \ln n \). We generate a random pairing among \( 2n \) points grouped in sets of size 2. We can do this by picking any desired point to initiate the \( i \)th pair, as long as the mate of this point is chosen randomly from the remaining unmatched \( 2n-2(i-1)-1 \) points. To initiate the \( i \)th pair, choose the other point in the 2-set containing the point that completed the previous pair, if that point is available; otherwise choose any point. In this way, the collapsed 2-regular graph being generated consists of some cycles and one path being extended.

There is exactly one point whose completion of the \( i \)th pair completes a cycle. Hence the probability that a cycle is completed on the \( i \)th step is exactly \( 1/(2n-2i+1) \). The expected number of cycles completed is the sum of these probabilities, which equals \( \sum_{i=1}^{n} 1/(2j-1) \). Asymptotically, alternate terms are taken in the sum for \( \ln n \), and the value of the sum tends to \( \frac{1}{2} \ln n \).

In Theorem 3.1.20, the generating function for permutations of \( [n] \) by number of cycles is shown to be \( x^{(n)} \), the rising factorial. Dividing by \( n! \) yields a generating function for probabilities of the various numbers of cycles. To find the expected number of cycles, we differentiate \( A(x) = (1/n!) \sum a_k x^k \) and set \( x = 1 \). The derivative of \( x^{(n)} \) is \( x^{(n)} \sum_{i=1}^{n} 1/(x+i-1) \). With \( x = 1 \), we obtain \( n/(n!/n! \sum_{i=1}^{n} 1/i) \), which is asymptotic to \( \ln n \).

Actually, the method here can also be used to generate random permutations. At each step, we can choose the image of any desired element uniformly from among the remaining elements that have not yet been chosen as an image. At the \( i \)th stage, we choose the image of the element most recently chosen as an image, unless the previous step completed a cycle. At the \( i \)th step, the probability of completing a cycle is \( 1/(n-i+1) \), since only one of the remaining images will do so. Summed over \( i \), the expected number of cycles is again the harmonic number \( \sum_{i=1}^{n} 1/j \), asymptotic to \( \ln n \).

Less obvious would be a direct argument for the factor of 2 between the expected number of cycles in a permutation of \([n]\) and the expected number of cycles in a 2-regular labeled \( n \)-vertex graph.

**14.3.35.** The number of \( d \)-regular graphs with vertex set \([n]\) is asymptotic to \( n^{\sqrt{2e((d-1)/d)})^2} \). Within the pairing model that generates \( d \)-regular multigraphs, the fraction of sample points that correspond to \( d \)-regular graphs (no loops or multiple edges), the probability that a multigraph generated in this model is simple, which by Theorem 14.3.29 is asymptotic to \( e^{((d-1)/d)})^2} \).

The number of sample points in the model is the number of pairings of \( dn \) points. As computed in Chapter 1, the number of pairings of \( t \) points is \( t!/(2t)!2^{t/2} \). This is the size of the sample space \( S \).

Each \( d \)-regular graph with vertex set \([n]\) arises \( (d^n)^2 \) times in the pairing model. To see this, for each vertex \( v \) we have a list of \( d \) neighbors among \([n]\), in order. The edges to those neighbors depart from the \( d \) vertices of the group assigned to \( v \) in some order; there are \( d! \) choices for that order. These orderings at each of the \( n \) vertices are independently chosen.

The answer is thus \( e^{((d-1)/d)})^2} |S|/(d^n)^2 \). Using Stirling’s Approximation, we have \( t!/(2t)!2^{t/2} \sim t^{1/2} e^{-t/2} \). Setting \( t = dn \) completes the proof.

**14.3.36.** In the \( d \)-regular pairing model on \( n \) vertices, the number of vertices in cycles of length at most \( \log_{d-1} \) is \( o(n) \), where \( \omega_n \) is any sequence tending to \( \infty \). Let \( \ell = \log_{d-1} \omega_n \).

As in the evolutionary version of \( G(n,p) \), the symmetry in the pairing model permits an evolutionary version in which the first endpoint of the next pair revealed can be any unpaired vertex in any group. In particular, we can grow the multigraph by distance from a given group \( u \).

Let \( t_i \) be the number of vertices reached within distance \( i \) from \( u \) in the resulting multigraph. Always \( t_i \leq 1 + d \sum_{j=0}^{i-1} (d-1)^j \); thus \( t_i \leq 2(d-1)^i \). Vertex \( u \) lies on a cycle of length at most \( \ell \) if and only if for some \( i \) with \( i \leq \ell/2 \) one of the vertices found when generating vertices at distance \( i \) has already been found. Since at most \( t_i \) vertices have already been found, the probability that a given edge is bad is at most \( t_i/n \). Since at most \( t_i \) edges will be generated, the expected number of new edges that are bad is at most \( t_i^2/n \).

Over the entire process, the expected number of bad edges is bounded by \( \sum_{j=0}^{\ell/2-1} t_j^2/n \), which is \( O((d-1)^\ell/n) \). With \( t \) as chosen, this expected number of bad edges is \( O(1/\omega_n) \). Considering all vertices, the expected number of
vertices for which a bad edges completing a short cycle is found is $O(n/\omega_n)$. By Markov’s Inequality, the probability that the number of vertices lying in cycles this short is $O(n/\omega_n)$ tends to 1. That is, whp there are $o(n)$ vertices in short cycles.

14.3.37. For $r = c \log_b n$ with $c > 2$, almost every graph has no $r$-clique. Generate a random graph with constant edge probability $p = 1/b$. Let $X$ be the number of $r$-cliques; $X$ is the sum of $\binom{n}{r}$ indicator variables for the possible cliques. Since each such clique requires the presence of $\binom{r}{2}$ edges to exist, linearity of the expectation yields $\binom{n}{r}p^\binom{r}{2}$ as the expected number of $r$-cliques. By Markov’s Inequality, it suffices to show $\mathbb{E}(X) \to 0$. For $r = c \log_b n$, we can approximate $\binom{n}{r}$ by $(ne/r)^r$. Thus we can approximate $\mathbb{E}(X)$ by $\left(\frac{pe}{r}\right)^{r/2}$. Since $p^{r/2} = n^{-c/2}$, we have $\mathbb{E}(X) \sim \left[en^{-c/2}/\log n\right]^r$. This tends to 0 when $c > 2$.

For $c < 2$, almost every graph has an $r$-clique (sketch). The expression above yields $\mathbb{E}(X) \to \infty$ when $c < 2$. By the Second Moment Method, it suffices to show $\mathbb{E}(X^2) \sim \mathbb{E}(X)^2$. (If this holds, then the application of Markov’s Inequality to $\mathbb{E}[(X - \mathbb{E}(X))^2]$ yields $P(X = 0) \to 0$.

Let $X_1, \ldots, X_{\binom{n}{2}}$ be the indicator variables for the presence of the $r$-cliques. To explore $\mathbb{E}(X^2)$, we use $X^2 = \sum_{i \neq j} X_i X_j$. We want to show that the expectation is asymptotic to $[\mathbb{E}(X)]^2$, which is asymptotic to $[en^{-c/2}/\log n]^2$. We aim to discard lower-order terms.

When two $r$-cliques share at most one vertex, the corresponding $r$-cliques occur independently, and then $\mathbb{E}(X_i X_j) = P(X_i = X_j = 1) = \mathbb{E}(X_i)\mathbb{E}(X_j)$. The number of ordered pairs of completely disjoint cliques is $\binom{n}{r, r, n-2r}$; each occurs with probability $p^{2\binom{r}{2}}$. With $r = c \log n$, we have $\binom{n}{r, r, n-2r} \sim \binom{n}{r}^2$. Thus it suffices to show that the sum of the remaining terms has lower order.

The number of ordered pairs of cliques with one common vertex is $\binom{n}{1, r-1, r-1, n-2r+1}$. These again have probability $p^{2\binom{r}{2}}$ of common occurrence, but with $r = c \log n$ we have $\binom{n}{1, r-1, r-1, n-2r+1} = o\left(\binom{n}{r, r, n-2r}\right)$, so we can ignore these. To complete the computation, it suffices to consider the pairs sharing at least one edge. The contribution from these with an overlap of $s$ vertices is $\binom{n}{r, r, r-s, n-2r+2s}p^{2\binom{r}{2}}$. It suffices to show that the ratio of each such term to $\binom{n}{r, r, n-2r}p^{2\binom{r}{2}}$ goes to zero sufficiently fast.


a) Ore’s Theorem has asymptotic strength 0. Ore’s Condition is that $d(x) + d(y) \geq n(G)$ whenever $xy \notin E(G)$. A pair $\{x, y\}$ is bad if $xy \notin E(G)$ and at most $n - 1$ of the $2n - 4$ possible edges from $\{x, y\}$ to other vertices occur. The probability is 1/2 times the probability of at most $n - 1$ heads in $2n - 4$ coin flips. The latter probability $p$ exceeds 1/2 by $\frac{1}{2} \left(\frac{n}{n-2}\right)2^{-(2n-4)} + \left(\frac{n}{n-1}\right)2^{-(2n-4)}$. Thus

$$\frac{1}{2} < p < \frac{1}{2} + \frac{3}{n} \left(\frac{n-1}{n-2}\right)2^{-(2n-4)} \sim \frac{1}{2} + \frac{1}{\log n}.$$ 

Let $X$ be the number of pairs where Ore’s Condition fails. There are $\binom{n}{2}$ pairs, so $\mathbb{E}(X) = \frac{1}{4} \binom{n}{2} + O(n^{3/2}) \sim n^2/8$. We want $P(X = 0) \to 0$. Since $\mathbb{E}(X) \to \infty$, the Second Moment Method makes $\mathbb{E}(X^2) \sim \mathbb{E}(X)^2$ sufficient.

Letting $X_i$ and $X_j$ denote the indicator variables for the $i$th and $j$th pairs, $\mathbb{E}(X^2) = \mathbb{E}(X) + \sum_{i \neq j} P(X_i X_j = 1)$. If the two pairs are disjoint, say $\{x, y\}$ and $\{x', y'\}$, then for $X_i X_j = 1$ we must have at most $n - 1$ of the $2n - 8$ edges from $\{x, y\}$ to the remaining $n - 4$ vertices, and similarly for $\{x', y'\}$, and those two events are independent. Some outcomes with between $n - 1$ and $n - 4$ of these edges (for each event) are also forbidden when edges arise between $\{x, y\}$ and $\{x', y'\}$, but these have lower order.

Considering the difference in probability from 1/2 as above (for forbidding $xy$ and $x'y'$), we find that $P(X_i X_j = 1) = \frac{1}{16} + O(n^{-1/2})$ when $\{x, y\}$ and $\{x', y'\}$ are disjoint. The number of such ordered pairs $\{i, j\}$ is $\binom{n}{2}^2$, which is asymptotic to $n^4/4$. When the pairs are not disjoint, the probability is a bit higher, but the number of such events is $O(n^3)$. Thus $\mathbb{E}(X^2) \sim \mathbb{E}(X) + \frac{1}{16} n^4/4 + O(n^{7/2}) \sim \mathbb{E}(X)^2$.

It now suffices to show that almost every graph has a spanning cycle. There are $\frac{1}{2}(n - 1)!$ possible cycles, and each occurs with probability $2^{-n}$, so $\mathbb{E}(X) \to \infty$, where $X$ is the number of spanning cycles. The Second Moment Method can be used to show that almost always $X \geq 1$, since the fraction of pairs of spanning cycles that are edge-disjoint tends to 0. Alternatively, part (b) guarantees that almost all graphs are Hamiltonian.

b) The Chvátal–Erdős Theorem ($\alpha(G) \geq \alpha(G)$ implies Hamiltonian) has asymptotic strength 1. Let $X$ be the number of independent $r$-sets (see Theorem 14.3.34). An $r$-set is independent with probability $2^{-\binom{r}{2}}$, so $\mathbb{E}(X) = \binom{n}{2} 2^{-\binom{r}{2}}$. Note that $\binom{r}{2} < (ne/r)^r$. If $r \to \infty$ and $\frac{ne}{2^{(r-1)/2}} < 1$, then $\mathbb{E}(X) \to 0$, which yields $P(X = 0) \to 1$, and hence $\alpha(G) < r$ almost always. Thus it suffices to have $\log n - \frac{1}{2} \log r + \log e < (r - 1)/2$. This holds when $r = 2 \log n$ (and also with $r$ slightly larger than $2 \log n - 2 \log \log n$).

It suffices to show that almost always every two vertices have more than $k$ common neighbors, where $k = 2 \log n$. Let $X$ be the number of bad pairs. When $X$ has a binomial distribution and $k$ is small relative to the expectation, $P(X \leq k) \leq c \mathbb{P}(X = k)$ for some constant $c$. Thus $\mathbb{E}(X) \leq \binom{n}{2} c \binom{n-2-k}{k} 2^{-(n-k)2} < c(n/3)^n$. The bound tends to 0 since $n^k(3/4)^n \to 0$ when $k = 2 \log n$. Section 14.3: Moments and Thresholds
14.3.39. The interval number and boxicity of the n-vertex random graph \((G^{1/2})\) are each at least \(n/(4 \lg n)\). Here the boxicity of a graph \(G\) is the least number of interval graphs whose intersection is \(G\), and the interval number of \(G\) is the least \(t\) such that \(G\) is the intersection graph of subsets of \(\mathbb{R}\) composed of at most \(t\) intervals.

In a representation of \(G\) as the intersection of \(t\) interval graphs, or in a representation as the intersection graph of unions of \(t\) intervals, we must specify \(nt\) intervals. Each vertex is assigned \(2t\) endpoints of intervals. For each problem, we show that the number of possible representations is a vanishing fraction of all the graphs when \(t\) is too small.

Consider first the interval number. What matters in generating edges is only the order of the endpoints, and we may assume that the signed endpoints are distinct. Each vertex will be assigned \(2t\) endpoints for its intervals from a list of \(2nt\) endpoints. This can be done in \(\binom{2nt}{2t, \ldots, 2t}\) ways. Using Stirling’s Approximation, the logarithm (base 2) of this multinomial coefficient is asymptotic to \((2nt + \frac{1}{2}) \log n - (n - \frac{1}{2}) \log (4\pi t)\). Thus when \(t = \frac{1}{2} n/\lg n\), the logarithm of the number of graphs with vertex set \([n]\) whose interval number is bounded by \(t\) is less than \(\frac{1}{2} n^2 - n \lg n + O(n \lg^2 n)\). The logarithm of the total number of graphs is \(\binom{n}{2} = \frac{1}{2} n^2 - n/2\). Thus the fraction of the graphs with vertex set \([n]\) that have interval number at most \(t\) tends to 0.

For the number of graphs with vertex set \([n]\) that are the intersection of \(t\) interval graphs, the asymptotic computation is similar. For one interval graph, we can produce a representation in \(\binom{2n}{2, \ldots, 2}\) ways, and we raise this to the power \(t\). The logarithm is asymptotic to

\[
\log t \left((2n + 1/2)(1 + \log n) - 2n \log e - n\right).
\]

Since \(\log e > 1\), this value is less than \(2nt \lg n - nt\). With \(t = \frac{1}{2} n/\lg n\), the value is less than \(\frac{1}{2} n^2 - \frac{1}{4} n^2/\lg n\). Thus the fraction of the graphs with vertex set \([n]\) that have boxicity at most \(t\) tends to 0.

14.3.40. If \(p \in (0, 1)\) and \(k_1, \ldots, k_r\) are nonnegative integers summing to \(m\), then \(\prod_{i=1}^{r} \left(1 - (1 - p)^{k_i}\right) \leq \left(1 - (1 - p)^{m/r}\right)^r\). We assume that \(k_1, \ldots, k_r\) are positive real numbers: we don’t need integrality, and the inequality is clear if some \(k_i\) is 0. Let \(f(x) = \ln(1 - (1 - p)^x)\). Using calculus, we find

\[
f''(x) = -(1 - p)^x \left(\frac{\ln(1 - p)}{(1 - (1 - p)^x)^2}\right).
\]

Since \(f''(x) < 0\), the function \(f\) is concave for \(x > 0\). Concavity yields

\[
\frac{1}{r} \sum_{i=1}^{r} f(k_i) \leq f\left(\frac{1}{r} \sum_{i=1}^{r} k_i\right) = f(m/r).
\]

Thus \(\sum_{i=1}^{r} f(k_i) \leq r f(m/r)\), and exponentiating completes the proof.

14.3.41. Given the statements in Example 14.3.31 about the behavior of vertex degrees in \(G(n, \frac{1}{2})\), the isomorphism algorithm there runs in time \(O(n^2)\) for \(n\)-vertex graphs and works whp, assuming also that \(n\) numbers can be sorted using \(O(n \log n)\) pairwise comparisons. By “whp”, we mean that the probability of success tends to 1 when run on graphs \(G\) and \(H\) such that \(G\) is drawn from \(G(n, 1/2)\) and an adversary then chooses \(H\). Otherwise, if both \(G\) and \(H\) are random (or even if \(H\) is arbitrary but \(G\) is drawn from \(G(n, 1/2)\) independently of \(H\)), then an algorithm works whp simply by returning “not isomorphic” without looking at the graphs.

When \(r = \lfloor 3 \lg n \rfloor\), the cited 1981 result by Bollobás implies that the \(r\) vertices of \(G\) with highest degrees \(v\) are uniquely identified. We then get a unique identification of all the vertices if no two vertices \(v_i\) and \(v_j\) in \(G\) have the same neighborhoods within the resulting set \(S\) of \(r\) vertices. A particular neighbor occurs for \(v_i\) with probability \(2^{-r}\). Given a uniform distribution over \(2^r\) options, two random samples agree with probability \(2^{-r}\). Over \(n - r\) samples, two of them agree with probability at most \((\frac{n-r}{2})^{2^{-r}}\), by the Union Bound. Since \(\frac{1}{2} n^2 2^{-3 \log n} = \frac{1}{4} n^{-1} \to 0\), whp there is no agreement, and the neighborhoods in \(S\) are distinct.

We can then check whether \(H\) is isomorphic to \(G\) by applying the same procedure to \(H\) and comparing the adjacency matrices. If \(G\) and \(H\) are isomorphic, then the adjacency matrices will match. If they are not isomorphic, then the adjacency matrices cannot match.

The vertex degrees can be found from the adjacency matrix (unpermuted) in \(O(n^2)\) time, then sorted in \(O(n \log n)\) time. Since \(S\) has size \(3 \lg n\), the neighborhoods within \(S\) can also be found in \(O(n \log n)\) time. Viewing them as binary numbers, we may allow another factor of \(\log n\) since they will be binary numbers with \(3 \lg n\) bits, so we can sort them in \(O(n \log(n^2))\) time. We do this also for \(H\) and then compare the adjacency matrices in \(O(n^2)\) time.

14.3.42. Perfect matchings in bipartite graphs. Let \(G\) be a random subgraph of \(K_{n,n}\), with parts \(A\) and \(B\) and independent edge probability \(1 + e) \frac{\ln n}{n}\). Say that \(S\) fails if \(|N(S)| < |S|\). By Hall’s Theorem, \(G\) has a perfect matching if and only if no set fails.

a) If \(\varepsilon < 0\), then whp \(G\) has no perfect matching. The vertices in \(B\) are isolated with probability \((1 - p)^n\), independently. Hence the probability that there is no isolated vertex in \(B\) is \((1 - (1 - p)^n)^n\). With \(p = 1 + \varepsilon) \frac{\ln n}{n}\), we have \((1 - p)^n \sim n^{-1+\varepsilon}\). When \(\varepsilon > 0\), but when \(\varepsilon < 0\) the probability of having no isolated vertex in \(B\) is asymptotic to \(1 - (1/n^{1+\varepsilon})^n\). Since

\[
n \ln [1 - (1/n^{1+\varepsilon})] = -n^{-\varepsilon} \to -\infty,
\]

the probability of having no isolated vertex in \(B\) tends to 0.
Chapter 14: The Probabilistic Method

Section 14.4: Concentration Inequalities

Curiously, the expected number of perfect matchings is
\[ \frac{-p^2}{(n/\log n)^2} + \frac{n}{(n/\log n)^2} \]

when \( p \to \infty \). Similarly, the expected number of spanning cycles in the binomial model \( G(n, p) \) tends to \( \infty \) for an edge probability slightly below the threshold for the disappearance of isolated vertices.

b) If \( S \) is a minimal failing set in \( G \), then \( |N(S)| = |S| - 1 \) and \( G[S \cup N(S)] \) is connected. Since neighbors of a vertex in one part are in the other part, a minimal failing set is contained in one part. If \( |N(S)| < |S| - 1 \), then the set obtained by deleting any element of \( S \) is a failing set. If \( G[S \cup N(S)] \) is not connected, then the subset of \( S \) in one of the components is a failing set. Hence minimality of \( S \) yields the claims.

c) If \( G \) has no perfect matching, then \( A \) or \( B \) contains a failing set with at most \( \lceil n/2 \rceil \) elements. Let \( S \) be a minimal failing set; by symmetry we may assume \( A \subseteq S \). Let \( T = B - N(S) \); note that \( N(T) \subseteq A - S \). Since \( |N(S)| < |S| \) and \( |B| = |A| \), we have \( |N(T)| < |T| \), so \( T \) is also a failing set.

If \( |S| > \lceil n/2 \rceil \), then \( |N(S)| \geq |S| - 1 \geq \lfloor n/2 \rfloor \), by part (b). This implies \( |T| \leq n - \lceil n/2 \rceil = \lfloor n/2 \rfloor \). Since \( T \) is a failing set, there is a minimal failing set contained in \( T \), and its size is at most \( \lfloor n/2 \rfloor \).

d) For edge probability \( p \) with \( p = (1 + \varepsilon)\ln n/n \) and \( \varepsilon > 0 \), whp \( G \) has a perfect matching. Write \( 1 + \varepsilon = c \) with \( c > 1 \). First consider failing sets of size 1 and 2. With \( Y \) counting the isolated vertices, \( E(Y) = 2n(1 - p)^n \leq 2n e^{-\varepsilon e} \to 0 \), so whp there is no failing set of size 1. A minimal failing set of size 2 consists of two vertices of degree 1 and a common neighbor; let \( Z \) count these. After choosing the pair, there are \( n \) choice for the common neighbor, two edges to ensure, and \( 2(n - 1) \) edges to forbid, so

\[ \mathbb{E}(Z) = 2(n^2)p(1 - p)^2 = n^2 \left( \frac{\ln n}{n} \right)^2 e^{-2p}\ln(n-1) = (c \ln n)^2 n^{-1 - 2c} e^{2p} \to 0. \]

In general, \( K_{r,s} \) has \( r^{-1}s^{-1} \) spanning trees. Hence for a minimal failing set \( S \) of size \( s \), the subgraph of \( K_{n,n} \) induced by \( S \cup N(S) \) has \( (s - 1)^{-1}s^{-2} \) spanning trees. To show that the expected number of minimal failing subsets tends to 0, we show \( \mathbb{E}(X) \to 0 \), where \( X \) is the number of spanning trees in induced subgraphs of the form \( G[S \cup N(S)] \) with \( S \) a minimal failing subset of \( A \) having size between 3 and \( \lfloor n/2 \rfloor \). (The same computation holds for failing sets in \( B \).)

Each such tree has \( 2s - 2 \) edges and occurs with probability \( p^{2s-2} \). Importantly, we must also forbid the \( s(n-s+1) \) possible edges from \( S \) to \( B - N(S) \). We compute

\[ \mathbb{E}(X) < \sum_{s=3}^{\lfloor n/2 \rfloor} \binom{n}{s} \left( \frac{n}{s-1} \right)^{s-2} (s-1)^{-1} p^{2s-2} \left( 1 - p \right)^{s(n-s+1)} \]

\[ < \sum_{s=3}^{\lfloor n/2 \rfloor} \left( \frac{ne}{s} \right)^{s-2} (s-1)^{-1} p^{2s-2} e^{-n e (n/2)} \]

\[ = \sum_{s=3}^{\lfloor n/2 \rfloor} \left( \frac{ne}{s} \right)^{s-2} e^{-\frac{c \ln n}{n}} < \sum_{s=1}^{\infty} \left( \frac{\ln n}{n} \right)^2 \to 0. \]

Comment: The small sets are eliminated so that the geometric sum in the bound starts with a term that tends to 0. Another proof first eliminates the failing sets of size at most 2 to show that in \( G[S \cup N(S)] \) each vertex of \( N(S) \) has degree at least 2. This again requires \( 2s - 2 \) edges in \( G[S \cup N(S)] \), and then the rest of the computation is similar.

14.3.43. Existence of expanders of linear size. An \( (n, \alpha, \beta, d) \)-expander is a bipartite graph \( G \subseteq K_{n,n} \) with \( |A| = |B| = n \), \( \Delta(G) \leq d \), and \( |N(S)| \geq \beta |S| \) whenever \( |S| \leq an \).

a) If \( X \) is the size of the union of \( dk \)-subsets of \( [n] \) chosen at random, then \( P(X \leq l) \leq (\frac{1}{l})^{dl/k} \). If \( X \leq l \), then all the \( k \)-sets are confined to one \( l \)-set. By multiplying the probability of this occurrence for a particular \( l \) by \( \binom{n}{k} \), we obtain a loose upper bound (the events for distinct \( l \)-sets are not disjoint). For a particular \( l \)-set, the probability it contains any selected \( k \)-set in the sequence is \( \left( \frac{1}{l} \right)^{k} = \prod_{i=0}^{k-1} \left( \frac{l-i}{l} \right) \leq (l/n)^{k} \).

b) If \( \alpha \beta < 1 \), then there is a constant \( d \) such that, for all \( n \) sufficiently large, an \( (n, \alpha, \beta, d) \)-expander exists. We generate bipartite graphs by choosing \( d \) random complete matchings between \( A \) and \( B \), discarding extra copies of edges. This yields a simple graph with maximum degree at most \( d \). We show that in this probability distribution the probability that the resulting graph is not an expander is less than 1. Let \( S \) be a violated set if \( |N(S)| \leq \beta |S| \). Let \( E \) be the event that a violated set exists; we bound \( P(E) \) by a quantity that is less than 1 when \( n \) is sufficiently large. The \( d \) random matchings provide \( d \) random \( k \)-sets as neighbors of \( S \) when \( |S| = k \). By (a), we have

\[ P(E) < \sum_{k=1}^{an} \binom{n}{k} \left( \frac{\beta k}{n} \right)^{kd} = \sum_{k=1}^{an} \left( \frac{ne}{k} \right)^{k} \left( \frac{ne}{\beta k} \right)^{kd} \]

\[ = \sum_{k=1}^{an} \left( e^{1+\beta \beta} (\beta k/n)^{d-\beta-1} \right)^{k} < \sum_{k=1}^{an} \left( e^{1+\beta \beta} (\alpha \beta)^{d-\beta-1} \right)^{k}. \]

If \( \alpha \beta < 1 \), then we can choose \( d \) to make the constant ratio in the geometric series as small as desired. It suffices to choose \( d \) so that \( e^{1+\beta \beta} (\alpha \beta)^{d-\beta-1} < 1/2 \), or \( d > 1 + \beta - (1 + \beta + \ln 2\beta) / \ln(\alpha \beta) \).
14.4. CONCENTRATION INEQUALITIES

14.4.1. Being right at least half the time. We want to place a bound on the probability of having too many failures. Let $X$ count the failures. The failure probability is $1/4$. To have $.99$ probability of being correct at least half the time among $n$ trials, we want $\Pr(X - n/4 \geq n/4) \leq .01$. By the Chernoff Bound, $\Pr(X - n/4 \geq n/4) \leq e^{-2nt^2} = e^{-n/8}$. Hence it suffices to have $e^{-n} \leq 10^{-16}$, or $n \geq 16 \ln 10$. Thus $n \geq 37$ suffices.

14.4.2. Comparison of Chebyshev and Chernoff bounds on the tail probability $\Pr(|X - E(X)| \geq \lambda \sqrt{\text{Var}(X)})$ when $X$ is distributed as Bin$(n, p)$. Chebyshev’s Inequality in Example 14.4.3 yields $\Pr(|X - E(X)| \geq \lambda \sqrt{\text{Var}(X)}) \leq \lambda^{-2}$. The Chernoff Bound is $\Pr(X - np \geq nt) \leq e^{-2nt^2}$. To compare them, we set $nt = \lambda \sqrt{\text{Var}(X)} = \lambda \sqrt{np(1-p)}$. Thus $nt^2 = (nt)^2/n = \lambda^2 p(1-p)$.

Hence the Chernoff Bound improves the upper bound from $\lambda^{-2}$ to $e^{-2\lambda^2 p(1-p)}$, which simplifies to $e^{-\lambda^2/2}$ when $p = 1/2$. Sample values illustrate the improvement for a few standard deviations from the expected value $n/2$.

<table>
<thead>
<tr>
<th>$\lambda$</th>
<th>Chebyshev</th>
<th>Chernoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.000</td>
<td>.607</td>
</tr>
<tr>
<td>2</td>
<td>.250</td>
<td>.135</td>
</tr>
<tr>
<td>3</td>
<td>.111</td>
<td>.011</td>
</tr>
</tbody>
</table>

14.4.3. If $p$ is fixed and $\epsilon > 0$, then whp $G^p$ has connectivity at least $(1-\epsilon)p^2n$. It suffices to show that almost every $G^p$ has the property that any two vertices have at least $k$ common neighbors, where $k = c p^2(n-2)$ with $c < 1$. Let $X$ denote the number of common neighbors for two fixed vertices: $E(X) = p^2(n-2)$. The distribution is binomial with success probability $p^2$.

By the Chernoff Bound, $\Pr(X - E(X)) \leq -(n-2)t) \leq e^{-2(n-2)t^2}$. To bound $\Pr(X \leq k)$, we set $t = (1-c)p^2$, which is constant. Thus the bound tends to 0 exponentially quickly in $n$. By the Union Bound, the probability that any of the $\binom{n}{2}$ pairs fails to have this many common neighbors also tends to 0.

14.4.4. If $p > \sqrt{\ln n}/n$ for $c > 2$, then whp $G^p$ has diameter at most 2. Given any two vertices, let $X$ be their number of common neighbors in $G^p$. This variable has a binomial distribution with success probability $p^2$. We want to bound the probability of $X = 0$. This is the probability that $X \leq \mu - \mu$. Since we are interested in $p < 1/4$, we use the alternative form of the Chernoff Bound on lower tails, $\Pr(x \leq \mu - s) \leq e^{-s^2/2\mu}$. With $s = \mu$, the bound is $e^{-3/2\mu}$. Hence $\Pr(X = 0) \leq e^{-(n-2)p^2/2} \sim n^{-c^2/2}$. Since we must consider all $\binom{n}{2}$ pairs of vertices, using the Union Bound yields $\Pr(\text{diam}(G^p) > 2) \leq \binom{n}{2}n^{-c^2/2}$. When $c > 2$, we have the desired behavior.

(Comment: The actual threshold obtained by using the exact value of $\Pr(X = 0)$, is $(1+\epsilon)\ln n/n$, obtained in ??).

14.4.5. Whp the random graph has minimum degree at least $\frac{n}{2} - \sqrt{cn \ln n}$, for any constant $c$ greater than $1/2$. The degree of a fixed vertex is a binomial random variable with $n - 1$ trials and success probability $1/2$. We use the Chernoff Bound to bound the probability of the event $X \leq d$. By the Union Bound, multiplying by $n$ gives a bound on the probability $\delta(G^p) \leq d$, even though the events for distinct vertices having small degree are not independent. It suffices to choose $d$ so that the resulting bound tends to 0.

By the Chernoff Bound, $\Pr(X - (n-1)/2 \leq -(n-1)t) \leq e^{-(2n-1)t^2}$. To simplify the bound, let $t = \sqrt{(c \ln n)/(n-1)}$. The bound becomes $e^{-2\ln n}$, which equals $n^{-2c}$. If $c > 1/2$, then $n$ times this bound still tends to 0, as desired.

Hence we have $\Pr(\delta(G) \leq \frac{n}{2} - \sqrt{c(n-1)\ln n}) \to 0$ if $c > 1/2$. Changing $n - 1$ to $n$ yields a slightly weaker statement.

14.4.6. Whp the random graph has connectivity at least $\frac{n}{2} - \sqrt{\ln n}$, for any constant $c$ greater than $1$. The number of common neighbors of two fixed vertices is a binomial random variable $X$ with $n - 2$ trials and success probability $1/4$. We use the Chernoff Bound to bound the probability of the event $X \leq d$. By the Union Bound, multiplying by $n$ gives a bound on the probability $\kappa(G^p) \leq d$, even though the events for distinct pairs having few common neighbors are not independent. It suffices to choose $d$ so that the resulting bound tends to 0.

By the Chernoff Bound, $\Pr(X - (n-2)/4 \leq -(n-2)t) \leq e^{-(2n-2)t^2}$. To simplify the bound, let $t = \sqrt{(c \ln n)/(n-2)}$. The bound becomes $e^{-2\ln n}$, which equals $n^{-2c}$. If $c > 1$, then $\binom{n}{2}$ times this bound still tends to 0, as desired.

Hence we have $\Pr(\kappa(G) \leq \frac{n}{2} - \sqrt{c(n-2)\ln n}) \to 0$ if $c > 1$. Changing $n-2$ to $n$ affects only lower-order terms, so with $c > 1$ the claim still holds.

14.4.7. Almost always the random graph has no bipartite subgraph with more than $n^2/8 + n^{3/2}$ edges. For a fixed bipartition $(A, B)$ of the vertices, let $m = |A| |B|$, so $m \leq n^2/4$. The distribution of the number $X$ of edges respecting this bipartition is binomial with $m$ trials. Achieving $n^2/8 + n^{3/2}$ edges is impossible unless $m \geq n^2/8 + n^{3/2}$. It is easiest when $m$ is largest, equal to $n^2/4$.
In this case $n^2/8 + n^{3/2}$ is $n^{3/2}$ above the expectation. Since $n^{3/2} = \frac{1}{4}n^2(4n^{-1/2})$, we set $t = 4n^{1/2}$. By the Chernoff Bound, the probability of being this much above the expectation is bounded by $e^{-2(\alpha^3/4)(4n^{-1/2})}$, which equals $e^{-8n}$.

As we have noted, the probability of any bipartition capturing at least this many edges is no bigger than this. Since there are at most $2^n$ bipartitions to consider, by the Union Bound the probability of any of them capturing this many edges still tends to 0.

### 14.4.8. When $n$ balls are dropped uniformly at random into $n$ boxes, independently, whp not too many balls fall into a single box.

Let $X$ be the random variable counting the balls in the first box. The distribution of $X$ is binomial with $n$ trials and success probability 1/n, so $\mu = \mathbb{E}(X) = 1$.

- The probability that some box has more than $1 + 2\ln n$ balls is asymptotically at most $O(1/n)$. The alternative form of the Chernoff Bound is $\mathbb{P}(X - \mu > s) \leq e^{-s^2/(2\mu + s)}$. With $\mu = 1$, we compute $\mathbb{P}(X > 1 + 2\ln n) \leq e^{-4(\ln n)^2/(2+2\ln n)} = e^{-2(\ln n)(1-1/\ln n)} = e^{-2\ln n + O(1)} = O(n^{-2})$. Since there are $n$ boxes, we multiply by $n$. By the Union Bound the probability of having some box with more than $1 + 2\ln n$ balls is at most $O(n^{-1})$.

- Whp no box has more than $\ln_{\ln n}$ balls. The probability that a given box has at least $k$ balls is bounded by $\binom{n}{k}p^k(1-p)^{n-k}$, where $p = 1/n$. Using the upper bound $\binom{n}{k} \leq (ne/k)^k$, we have $\mathbb{P}(X \geq k) < (e/k)^k = e^{k\ln(e/k)}$.

Letting $k = (1+\epsilon)\ln n/\ln n$, we have $\ln(e/k) = -(1 - o(1))\ln k = -(1 + \epsilon - o(1))\ln n$. Hence $\mathbb{P}(X \geq (1+\epsilon)\ln n) < n^{-1-o(1)}$. Since there are $n$ boxes, we multiply by $n$. By the Union Bound the probability of having some box with more than $\ln_{\ln n}$ balls tends to 0.

### 14.4.9. There exists an $n$-vertex tournament such that $D_T(\sigma) < 2n^{3/2}/\ln n$ for every ranking $\sigma$, where $D_T(\sigma) = a - b$ with $a$ and $b$ being the number of vertices pairs where $\sigma$ agrees or disagrees with $T$, respectively. Consider a random tournament, ordering each pair with probability 1/2 in each direction. For a fixed ranking $\sigma$, let $X$ be the number of pairs on which the ranking agrees with the random tournament. The distribution of $X$ is binomial, with $\binom{n}{2}$ trials and success probability 1/2. By the Chernoff Bound, $\mathbb{P}(X - \frac{1}{2}\binom{n}{2} \geq \epsilon\binom{n}{2}) \leq e^{-n(n-1)\epsilon^2}$.

A tournament $T$ agrees with a ranking $\sigma$ on $\frac{1}{2}\binom{n}{2} + \frac{1}{2}D_T(\sigma)$ pairs. Hence we consider $t$ such that $\binom{n}{2} = \frac{1}{2}D_T(\sigma)$. Thus for application of the Chernoff Bound we set $t = n^{3/2}/\ln n/\binom{n}{2}$.

We obtain $\mathbb{P}(X \geq \frac{1}{2}\binom{n}{2} + n^{3/2}/\ln n < e^{-4n\ln n} = n^{-4n}$. There are $n! \approx n^ne^{-n}\sqrt{2\pi n}$, by the Union Bound the probability that some ordering agrees more than this often is less than 1. Hence there exists a tournament as an outcome of the experiment such that $D_T(\sigma) < 2n^{3/2}/\ln n$ for every ranking $\sigma$.

### 14.4.10. If the fraction of the population preferring $A$ to $B$ is $p$, and $X$ is the fraction preferring $A$ in a random sample of $n$ people, then $\mathbb{P}(|X - p| \leq \epsilon p) > 1 - \delta$ if $n \geq 8\ln(2\delta^{-1})/\epsilon^2$ (given $p > .25$). For $1 \leq i \leq n$, let $Y_i$ be the indicator random variable for event that the $i$th person in the poll prefers $A$ to $B$. We have $X = Y_i n$, where $Y = \sum Y_i$. Also $\mu = \mathbb{E}(Y) = np$, so $\mu = \mathbb{E}(Y) = np$. By the Chernoff Bound,

$$
\mathbb{P}(|X - p| \geq \epsilon p) = \mathbb{P}(|Y - \mu| \geq \epsilon \mu) \leq 2e^{-2n(\epsilon p)^2}.
$$

To ensure $\mathbb{P}(|X - p| \geq \epsilon p) > 1 - \delta$, we require $2e^{-2n(\epsilon p)^2} < \delta$. Solving for $n$ yields $n > \ln(2/\delta)/(2\epsilon^2)$. With $p > .25$, in terms of $\epsilon$ and $\delta$ alone it suffices to have $n > (8/\epsilon^2)\ln(2/\delta)$.

### 14.4.11. **Not complete** Chernoff–Hoeffding Bound:

For independent random variables $X_1, \ldots, X_n$ with $a_i \leq X_i \leq b_i$ and $X = \sum_{i=1}^n X_i$,

$$
\mathbb{P}(X - \mathbb{E}(X) \geq s) \leq e^{-2s^2/\sum_{i=1}^n (b_i - a_i)^2},
$$

We follow the proof of Theorem 14.4.6. Let $p_i = \mathbb{E}(X_i)$, and let $m = \mu + s$, where $\mu = \mathbb{E}(X)$. By Markov’s Inequality,

$$
\mathbb{P}(X \geq m) = \mathbb{P}(e^{mX} \leq e^{ms}) \leq \mathbb{E}(e^{mX})/e^{ms}.
$$

Independence of $X_1, \ldots, X_n$ yields $\mathbb{E}(e^{mX}) = \prod_i \mathbb{E}(e^{mX_i})$. For $a_i \leq X_i \leq b_i$, we have $e^{ux} \leq e^{(b_i - a_i)x} + e^{b_i - a_i}e^{ux}$, by the convexity of the exponential function. Hence $\mathbb{E}(e^{mX}) \leq \prod_i (e^{b_i - a_i} + e^{b_i - a_i}e^{ux})$. Applying the Arithmetic Mean / Geometric Mean Inequality to the product of these expectations yields

$$
\prod_{i=1}^n \mathbb{E}(e^{mX_i}) \leq \left[\sum_{i=1}^n \frac{1}{n}(e^{b_i - a_i} + e^{b_i - a_i}e^{ux})\right]^n.
$$

$$
\mathbb{P}(X \geq n(p + t)) \leq e^{-u(n(p + t))(1 - p + pe^t)^n},
$$

and after that the proof of the main statement is the same as that of ?? . In this case, the distribution need not be symmetric around the mean, but letting $Y_i = 1 - X_i$ and applying the bound on the upper tail for $\sum Y_i$ yields the bound on the lower tail for $X$.

### 14.4.12. A hypergraph with $m$ edges and $n$ vertices has discrepancy at most $\sqrt{2n \ln(2m)}$. Let $\alpha$ denote the desired bound, and let $J$ be a random coloring with colors $+1$ and $-1$. For each edge $A$, define an indicator variable $X_A$ such that $X_A = 1$ if the imbalance on $A$ exceeds $\alpha$. Let $X$ be the total number of such edges. Choosing $\alpha$ to make $\mathbb{E}(X) < 1$ guarantees discrepancy at most $\alpha$. 

The imbalance of $f$ on $A$ is the sum of $|A|$ variables taking values ±1 with probability 1/2. Let $Y' = (Y + |A|)/2$; this transformation yields an binomial random variable. By the Chernoff Bound,

$$\mathbb{P}(Y > \lambda \sqrt{|A|}) = \mathbb{P}(Y' - |A|/2 > (\lambda/2) \sqrt{|A|}) < e^{-\lambda^2/2}.$$

Setting $\lambda = \alpha/\sqrt{|A|}$ yields $\mathbb{P}(Y > \alpha) < e^{-\alpha^2/(2|A|)}$. Using $|A| \leq n$ and $\alpha = \sqrt{2n \ln(2m)}$, we have

$$\mathbb{E}(X_A) = \mathbb{P}(|Y| > \alpha) < 2e^{-\alpha^2/(2|A|)} \leq 2e^{-(2n \ln 2m)/(2n)} = 1/m.$$

Since there are $m$ edges, we have $\mathbb{E}(X) < 1$, as desired.

14.4.13. The claim is this: For a hypergraph $H$ where every edge has size at least $r$ and intersects at most $k$ other edges, the discrepancy of $H$ is at most $\alpha$ if $k \leq \frac{\alpha}{\sqrt{2r}}$. With different names for the parameters, this claim is Exercise 5.1 on page 46 of Molloy–Reed [2002]. However, the claim is false. A counterexample is given by the $n$-vertex hypergraph $H$ in which the edges are the sets of $n/2$ vertices. Set $r = 1$, $k = 2^n$, $\alpha = \alpha/\sqrt{6(n+3) \ln 2}$. All hypotheses are satisfied, but the discrepancy is $n/2$, much bigger than $\sqrt{n}$. Hence we change the statement to assume that each edge has size at most $r$ and require $\alpha \geq 1$ to avoid parity issues. Since we are making this change, for simplicity we also say that edge edge intersects at most $k$ changes (including itself).

We may assume $k \geq 2$, since every hypergraph whose edges are disjoint has discrepancy at most 1. Randomly color $V(H)$ with colors in $\{1, -1\}$, letting $X(v)$ denote the color on vertex $v$. For an edge $e$, let $X_e = \sum_{v \in e} X(v)$; thus the discrepancy of the edge-coloring is $\max_{e \in E[H]} |X_e|$. To relate $X$ to a binomial distribution, let $Y(v) = (X(v) + 1)/2$ and $Y_e = \sum_{v \in e} Y(v) = (X_e + |e|)/2$. Let $m = |e|$. Now $Y$ is distributed as Bin$(m, 1/2)$ and $X = 2Y - m$. By the Chernoff Bound, with $\mu = \mathbb{E}(Y) = m/2$, we have

$$\mathbb{P}(|X_e| \geq \alpha) = \mathbb{P}(|2Y - m| \geq \alpha) = \mathbb{P}(|Y - \mu| \geq \alpha/2) \leq 2e^{-2(\alpha/2)^2/m} = 2e^{-\alpha^2/2m} \leq 2e^{-\alpha^2/2r}.$$

For each edge $e$, let $A_e$ be the event that $e$ has discrepancy larger than $\alpha$, and let $P = \mathbb{P}(A_e) \leq 2e^{-\alpha^2/2r}$. To guarantee the existence of a coloring such that no edge has discrepancy exceeding $\alpha$, it suffices by the symmetric Local Lemma to check $e^{pd} \leq 1$, where each bad event $A_e$ is mutually independent of all but $d$ events.

Since $A_e$ is mutually independent of all events associated with edges disjoint from $e$, we have $d \leq k$. We compute $e^{pd} \leq e^{2e^{-\alpha^2/2r}k}$. Since $e \leq 4$, having $k \leq e^{3/2}/8$ suffices. (Comment: The denominator 8 can be replaced by 2e. It was stated as 8 because Molloy and Reed use the stronger hypothesis $4pd \leq 1$ for the Local Lemma.)

14.4.14. Among edge-colorings $f$: $E(K_n) \to \{+1, -1\}$, some coloring $f$ has discrepancy at most $\sqrt{\ln 2(n^{3/2} + \sqrt{n} n^{1/2})}$, where the discrepancy of an edge-coloring $f$ is $\max_{e \in E[K_n]} |\sum_{v \in e} f(uv)|$. Thus, let $d(S) = |\sum_{v \in S} f(uv)|$.

Generate a coloring $f$ by choosing 1 or −1 randomly on each edge, independently. Let $S$ be a set of $k$ vertices. To relate $d(S)$ to a binomial distribution, let $Y_{uv} = \frac{1}{2}(f(uv) + 1)$ and $Y = \sum_{u \in S} Y_{uv}$, so that $d(S) = |2Y - m|$, where $m = \binom{k}{2}$. Note that $Y$ has distribution Bin$(m, 1/2)$. By the Chernoff Bound,

$$\mathbb{P}(d(S) \geq \alpha) = \mathbb{P}(|Y - m/2| \geq \alpha/2) \leq 2e^{-\alpha^2/(2m)} < 2e^{-\alpha^2/k} \leq 2e^{-\alpha^2/n^2}.$$

With $2^n$ choices for $S$, by the Union Bound the existence of a coloring with $\alpha < \alpha$ for all $S \subseteq V(K_n)$ follows from $2^n \cdot 2e^{-\alpha^2/n^2} \leq 1$. Therefore $\alpha \geq n \sqrt{(n + 1) \ln 2}$ suffices, and indeed $n \sqrt{(n + 1) \ln 2} \leq \sqrt{\ln 2(n^{3/2} + \sqrt{n} n^{1/2})}$.

14.4.15. ***Solution not yet written.*** In $n$ tosses of a fair coin, let $Y$ be the number of heads minus the number of tails. Compare the bounds on $\mathbb{P}(|Y| > \lambda \sqrt{n})$ using martingales and the simple Chernoff Bound.

14.4.16. ***Solution not yet written.*** Let $f$ be a random function from $[n]$ to $[n]$. Let $Y$ be the number of elements of $[n]$ that are missing from the image of $f$. Prove $\mathbb{P}(|Y - \mathbb{E}(Y)| \geq t) \leq e^{-t^2/(2n)}$.

14.4.17. The martingale for homogenizing triples in a graph. We start with $G_0$. At each step pick a set $S$ of three random vertices. If $S$ induces $f$ edges in $G_{i-1}$, then turn $S$ into a triangle with probability $j/3$ and into an independent set with probability $1 - j/3$, yielding $G_i$.

(a) The probability that the process ends with $K_n$ is $|E(G_0)|/\binom{n}{3}$. Note that in all cases $\mathbb{E}(|E(G_i)|) = |E(G_{i-1})|$. Letting $Y_i = |E(G_i)|$, we conclude that $\{Y_i: i \geq 0\}$ is a martingale, and always $Y_i - Y_{i-1} \leq 2$.

Since the expectation of each $Y_i$ at the start equals $Y_0$, the expected number of edges when the process ends is $|E(G_0)|$. The graph at the end can only be $K_n$ or $K_n$. Letting $p$ be the probability of ending with $K_n$, we have $p\binom{n}{3} + (1 - p) \cdot 0 = |E(G_0)|$, which yields $p = |E(G_0)|/\binom{n}{3}$.

(b) When $|E(G_0)| = \frac{1}{3} \binom{n}{3}$, the expected number of steps to finish the process is $\Omega(n^4)$. By the Hoeffding-Azuma Inequality, $\mathbb{P}(|Y_i - Y_0| \geq \alpha) \leq 2e^{-\alpha^2/8i}$. When $|E(G_0)| = \frac{1}{3} \binom{n}{3}$, the process ends when $\alpha = \frac{1}{3} \binom{n}{3} \sim n^2/4$. Since the square of this is $O(n^4)$, for any time $t$ that grows like $o(n^4)$ the probability of finishing the process within that time tends to 0.

If $t = n^4/128$, then the probability we finish before time $t$ is at most
2/e. With probability at least $1 - 2/e$, the time taken is at least $n^4/128$. So the expectation is at least $(1 - 2/e)n^4/128$, which is $\Omega(n^4)$.

14.4.23. The maximum length of an increasing sublist in a random permutation is (somewhat) highly concentrated around its expectation. Generate a random permutation of $[n]$ by iteratively choosing a random value $X_k$ in the interval $[0, 1]$, uniformly and independently. Since the choices are independent, the experiment is a cartesian product space. After the $k$th step, the order of the chosen values form a permutation of $[k]$, with position $i$ having value $\sigma(i)$ if $X_i$ is the $\sigma(i)$th smallest value among $X_1, \ldots, X_k$. Since the choices are uniform over the interval, at the $k$th step all permutations of $[k]$ are equally likely.

Let $f(X)$ be the maximum length of an increasing sublist in the final permutation. Let $(Y_0, \ldots, Y_n)$ be the Doob process with respect to $X$. That is, $Y_0 = \mathbb{E}(f(X))$, and $Y_k = \mathbb{E}(f(X) \mid X_1, \ldots, X_k)$ for $k \in [n]$. Let $A$ be the event defined by $\{X_j : j \neq k\}$. The value of $X_k$ ranges between 0 and 1. Wherever it lies, it may or may not belong to a longest increasing sublist. In particular, the values of $f$ on the outcomes in $A$ differ by at most 1.

Therefore, Lemma 14.4.23 applies to guarantee $|Y_k - Y_{k-1}| \leq 1$ for all $k$. We can therefore apply Azuma’s Inequality. With $Y_0 = \mathbb{E}(f(X))$ and $Y_n = f(X)$, we have $\mathbb{P}(|f(X) - \mathbb{E}(f(X))| \geq \lambda \sqrt{n}) \leq 2e^{-\lambda^2/2}$.

14.4.19. ***Solution not yet written.*** (Martingale Tail Inequality)
Let $Y$ be a martingale with respect to $X$ satisfying the Bounded Differences Condition for $c_1, \ldots, c_n$. Prove Theorem 14.4.19: For $t > 0$,
\[ \mathbb{P}(Y_n - Y_0 \geq t) \leq e^{-t^2/2 \sum c_i^2}. \] (Hoeffding [1963], Azuma [1967])

14.4.20. ***Solution not yet written.*** Generate two random binary lists of length $n$; the bits are chosen by unbiased coin flips, independently. Let $Y_n$ be the length of a longest common subsequence in the two lists (a common subsequence need not use the same positions in the two lists and need not appear in consecutive positions). Prove $\mathbb{P}(|Y_n - \mathbb{E}(Y_n)| \geq \lambda) \leq 2e^{-\lambda^2/8n}$.

14.4.21. ***Solution not yet written.*** Let $X$ be the number of triangles in the random $n$-vertex graph. By linearity, $\mathbb{E}(X) = \binom{n}{3}$. Prove $\mathbb{P}(|X - \mathbb{E}(X)| > \lambda n^2) \leq 2e^{-\lambda^2/2}$ for some constant $c$.

14.4.22. ***Solution not yet written.*** (○) Bin-packing. The numbers $a_1, \ldots, a_n$ are drawn uniformly and independently from the interval $[0, 1]$. They must be placed in bins, each having total capacity 1. Let $X$ be the number of bins needed. Use IBDI to prove that $\mathbb{P}(|X - \mathbb{E}(X)| \geq \lambda \sqrt{n}) \leq 2e^{-\lambda^2/2}$.

Section 14.4: Concentration Inequalities

14.4.23. Completion of the proof of Corollary 14.4.30. In the heart of the computation, we use that $f$ is $c$-Lipschitz. First we expand each of the relevant quantities by conditioning on the values of $X_i$. We use also the independence of $X_1, \ldots, X_m$. Let $X' = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$. Since $\sum_a \mathbb{P}[X_i = a'] = 1$,
\[ \mathbb{E}[f(X) \mid X_{(i-1)}] = \sum_a \mathbb{E}[f(X) \mid X', X_i = a] \mathbb{P}[X_i = a]. \]
\[ = \sum_a \sum_{a'} \mathbb{E}[f(X) \mid X', X_i = a'] \mathbb{P}[X_i = a'] \mathbb{P}[X_i = a]. \]

By first using $\sum_a \mathbb{P}[X_i = a] = 1$ and then conditioning,
\[ \mathbb{E}[f(X) \mid X_{(i)}] = \sum_a \mathbb{E}[f(X) \mid X', X_i = a] \mathbb{P}[X_i = a] \]
\[ = \sum_a \sum_{a'} \mathbb{E}[f(X) \mid X', X_i = a'] \mathbb{P}[X_i = a'] \mathbb{P}[X_i = a]. \]

Now we compute
\[ \mathbb{P}(|\mathbb{E}[f(X) \mid X_{(i)}] - \mathbb{E}[f(X) \mid X_{(i-1)}]|) \]
\[ = \left| \sum_a \sum_{a'} \mathbb{E}[f(X) \mid X', X_i = a'] - \mathbb{E}[f(X) \mid X', X_i = a] \right| \mathbb{P}[X_i = a'] \mathbb{P}[X_i = a] \]
\[ \leq \sum_a \sum_{a'} \mathbb{E}[f(X) \mid X', X_i = a'] - \mathbb{E}[f(X) \mid X', X_i = a] \mathbb{P}[X_i = a'] \mathbb{P}[X_i = a] \]
\[ \leq \sum_a \sum_{a'} c_i \mathbb{P}[X_i = a'] \mathbb{P}[X_i = a] = c_i. \]