2 Linearity of expectations

Let $X = c_1X_1 + \cdots + c_nX_n$ where $X_1, \dots, X_n$ are random variables, and $c_1, \dots, c_n$ constants. Then

$$\mathbb{E}[X] = c_1\mathbb{E}[X_1] + \cdots + c_n\mathbb{E}[X_n]$$

Note: this identity does not require any assumption of independence. On the other hand, generally $\mathbb{E}[XY] \neq \mathbb{E}[X]\mathbb{E}[Y]$ unless X and Y are uncorrelated (Independent random variables are always uncorrelated)

Here is a simple question with a simple solution (there are also much more involved solutions via enumerations, but linearity of expectations nearly trivializes the problem).

Question 2.0.1. What is the average number of fixed points of a random permutation of $[n]$ chosen uniformly at random?

Let X_i be the event that i is fixed. Then $\mathbb{E}[X_i] = 1/n$. So the expected number of fixed points is $\mathbb{E}[X_1 + \cdots + X_n] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] = 1$

2.1 Hamiltonian paths in tournaments

Important observation for proving existence: With positive probability, $X \geq \mathbb{E}[X]$ (likewise for $X \leq \mathbb{E}[X]$)

A **tournament** is a directed complete graph.

Theorem 2.1.1 (Szele 1943)**.** There is a tournament on n vertices with at least $n!2^{-(n-1)}$ Hamiltonian paths

Proof. Let X be the number of Hamiltonian paths in a random tournament.

For every permutation σ of $[n]$, one has the directed path $\sigma(1) \to \sigma(2) \to \cdots \to \sigma(n)$ with probability 2^{-n+1}.

Let X be the number of σ satisfying the above. $\mathbb{E}X = n!2^{-n+1}$. □

This was considered the first use of the probabilistic method. Szele conjectured that the maximum number of Hamiltonian paths in a tournament on n players is $n!/(2-o(1))^n$. This was proved by Alon (1990) using the Minc–Brégman theorem on permanents (we will see this later in the course when discussing the entropy method).

2.2 Sum-free set

A subset A in an abelian group is **sum-free** if there do not exist $a, b, c \in A$ with $a+b=c$.

Does every n-element set contain a large sum-free set?

Theorem 2.2.1 (Erdős 1965)**.** Every set of n nonzero integers contains a sum-free subset of size $\geq n/3$.

Proof. Let $A \subset \mathbb{Z} \setminus \{0\}$ with $|A| = n$. For $\theta \in [0,1]$, let

$$A_\theta := \{a \in A : \{a\theta\} \in (1/3, 2/3)\}$$

where $\{\cdot\}$ denotes fractional part. Then A_θ is sum-free since $(1/3, 2/3)$ is sum-free in $\mathbb{R}/\mathbb{Z}$.

For θ uniformly chosen at random, $\{a\theta\}$ is also uniformly random in $[0,1]$, so $\mathbb{P}(a \in A_\theta) = 1/3$. By linearity of expectations, $\mathbb{E}|A_\theta| = n/3$. □

Remark 2.2.2. Alon and Kleitman (1990) noted that one can improve the bound to $\geq (n+1)/3$ by noting that $|A_\theta| = 0$ for $\theta \approx 0$.

Bourgain (1997) improved it to $\geq (n+2)/3$ via a difficult Fourier analytic argument. This is currently the best bound known.

Eberhard, Green, and Manners (2014) showed that there exist n-element sets of integers whose largest sum-free subset has size $(1/3 + o(1))n$.

It remains an open problem to prove $\geq (n + \omega(n))/3$ for some function $\omega(n) \to \infty$

2.3 Turán's theorem and independent sets

Question 2.3.1. What is the maximum number of edges in an n-vertex K_k-free graph?

Taking the complement of a graph changes its independent sets to cliques and vice versa. So the problem is equivalent to one about graphs without large independent sets.

The following result, due to Caro (1979) and Wei (1981), shows that a graph with small degrees much contain large independent sets. The probabilistic method proof shown here is due to Alon and Spencer.

Theorem 2.3.2 (Caro 1979, Wei 1981)**.** Every graph G contains an independent set of size at least

$$\sum_{v \in V(G)} \frac{1}{d_v + 1},$$

where d_v is the degree of vertex v.

Proof. Consider a random ordering (permutation) of the vertices. Let I be the set of vertices that appear before all of its neighbors. Then I is an independent set.

For each $v \in V$, $\mathbb{P}(v \in I) = \frac{1}{1+d_v}$ (this is the probability that v appears first among $\{v\} \cup N(v)$). Thus $\mathbb{E}|I| = \sum_{v \in V(G)} \frac{1}{d_v+1}$. Thus with positive probability, $|I|$ is at least this expectation. □

Remark 2.3.3. Equality occurs if G is a disjoint union of cliques.

Remark 2.3.4 (Derandomization). Here is an alternative "greedy algorithm" proof of the Caro–Wei inequality.

Permute the vertices in non-increasing order of their degree.

And then greedily construct an independent set: at each step, take the first available vertex (in this order) and then discarding all its neighbors.

If each vertex v is assigned weight $1/(d_v + 1)$, then the total weight removed at each step is at most 1. Thus there must be at least $\sum_v 1/(d_v + 1)$ steps.

Taking the complement

Corollary 2.3.5. Every n-vertex graph G contains a clique of size at least $\sum_{v \in V(G)} \frac{1}{n-d_v}$.

Note that equality is attained when G is multipartite.

Now let us answer the earlier question about maximizing the number of edges in a K_{r+1}-free graph.

The **Turán graph** $T_{n,r}$ is the complete multipartite graph formed by partitioning n vertices into r parts with sizes as equal as possible (differing by at most 1).

Easy to see that $T_{n,r}$ is K_{r+1}-free.

Turán's theorem (1941) tells us that $T_{n,r}$ indeed maximizes the number of edges among n-vertex K_{r+1}-free graphs.

We will prove a slightly weaker statement, below, which is tight when n is divisible by r.

Theorem 2.3.6. (Turán's 1941) Every n-vertex K_{r+1}-free graph has $\leq \left(1 - \frac{1}{r}\right) \frac{n^2}{2}$ edges.

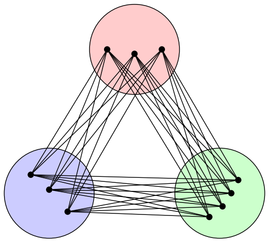

Figure 3: The Turán graph $T_{10,3}$.

Proof. Since G is K_{r+1}-free, by Corollary 2.3.5, letting $\overline{d}$ be average degree and $m = n\overline{d}/2$ be the number of edges, we see that the size $\omega(G)$ of the largest clique of G satisfies

$$r \geq \omega(G) \geq \sum_{v \in V} \frac{1}{n - d_v} \geq \frac{n}{n - \overline{d}} = \frac{n}{n - 2m/n}.$$

Rearranging gives $m \leq \left(1 - \frac{1}{r}\right) \frac{n^2}{2}$. □

Remark 2.3.7. By a careful refinement of the above argument, we can deduce Turán's theorem that $T_{n,r}$ maximizes the number of edges in a n-vertex K_{r+1}-free graph, by noting that $\sum_{v \in V} \frac{1}{n-d_v}$ is minimized over fixed $\sum_v d_v$ when the degrees are nearly equal.

2.4 Crossing number inequality

Consider drawings of graphs on a plane using continuous curves as edges.

The **crossing number** $\mathrm{cr}(G)$ is the minimum number of crossings in a drawing of G.

A graph is **planar** if $\mathrm{cr}(G) = 0$.

$K_{3,3}$ and K_5 are non-planar; furthermore, the following famous theorem characterizes these two graphs as the only obstructions to planarity

Kuratowski's theorem (1930): every non-planar graph contains a subgraph that is topologically homeomorphic to $K_{3,3}$ or K_5

(Also related: Wagner's theorem (1937) says that a graph is planar if and only if it does not have $K_{3,3}$ or K_5 as a minor. It is not too hard to show that Wagner's theorem and Kuratowski's theorem are equivalent)

Question 2.4.1. What is the minimum possible number of crossings that a drawing of:

- K_n? (Hill's conjecture)
- $K_{n,n}$? (Zarankiewicz conjecture; Turán's brick factory problem)
- a graph on n vertices and $n^2/100$ edges?

The following result, due to Ajtai–Chvátal–Newborn–Szemerédi (1982) and Leighton (1984), lower bounds the number of crossings for graphs with many edges.

Theorem 2.4.2 (Crossing number inequality)**.** In a graph $G = (V, E)$, if $|E| \geq 4|V|$, then

$$\operatorname{cr}(G) \gtrsim \frac{|E|^3}{|V|^2}$$

Corollary 2.4.3. In a graph $G = (V, E)$, if $|E| \gtrsim |V|^2$, then $\operatorname{cr}(G) \gtrsim |V|^4$.

Proof. Recall **Euler's formula:** $v - e + f = 2$ for every connected planar graph

For every connected planar graph with at least one cycle, $3|F| \leq 2|E|$ since every face is adjacent to ≥ 3 edges, whereas every edge is adjacent to exactly 2 faces. Plugging into Euler, $|E| \leq 3|V| - 6$.

Thus $|E| \leq 3|V|$ for all planar graphs. Hence $\operatorname{cr}(G) > 0$ whenever $|E| > 3|V|$.

By deleting one edge for each crossing, we get a planar graph, so $|E| - \operatorname{cr}(G) > 3|V|$, i.e.,

$$\operatorname{cr}(G) \geq |E| - 3|V|$$

This is a "cheap bound" that we will boost using the probabilistic method.

For graphs with $|E| = \Theta(n^2)$, this gives $\operatorname{cr}(G) \gtrsim n^2$. This not a great bound. We will use the probabilistic method to boost this bound.

Let $p \in [0, 1]$ to be decided. Let $G' = (V', E')$ be obtained from G by randomly keeping each vertex with probability p. Then

$$\operatorname{cr}(G') \geq |E'| - 3|V'|$$

So

$$\mathbb{E}\operatorname{cr}(G') \geq \mathbb{E}|E'| - 3\mathbb{E}|V'|$$

We have $\mathbb{E}\,\mathrm{cr}(G') \le p^4\,\mathrm{cr}(G)$, $\mathbb{E}|E'| = p^2|E|$ and $\mathbb{E}|V'| = p\mathbb{E}|V|$. So

$$p^4\,\mathrm{cr}(G) \ge p^2|E| - 3p|V|.$$

Thus

$$\mathrm{cr}(G) \ge p^{-2}|E| - 3p^{-3}|V|.$$

Setting $p \in [0,1]$ so that $4p^{-3}|V| = p^{-2}|E|$, we obtain $\mathrm{cr}(G) \gtrsim |E|^3 / |V|^2$. □

2.4.1 Application to incidence geometry

Question 2.4.4. What is the maximum number of incidences between n distinct points and n distinct lines on a plane?

Let $\mathcal{P}$ be a set of points and $\mathcal{L}$ a set of lines. Denote the number of incidences by

$$I(\mathcal{P},\mathcal{L}) := |\{(p,\ell) \in \mathcal{P}\times\mathcal{L} : p \in \ell\}|$$

Example: n points and n lines:

$$\mathcal{P} = [k]\times[2k^2] \quad \text{and} \quad \mathcal{L} = \{y = mx + b : m \in [k], b \in [k^2]\}$$

Every line contains k points from $\mathcal{P}$. Taking $3k^3 \approx n$ gives $k^4 = \Theta(n^{4/3})$ incidences.

Can we do better?

No. The following foundational theorem in incidence geometry implies that one has $O(n^{4/3})$ incidences between n points and n lines.

Theorem 2.4.5 (Szemerédi–Trotter 1983)**.** Given a set $\mathcal{P}$ of points and $\mathcal{L}$ of lines in $\mathbb{R}^2$,

$$I(\mathcal{P},\mathcal{L}) \lesssim |\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} + |\mathcal{P}| + |\mathcal{L}|.$$

We will show how to prove the Szemerédi–Trotter theorem using the crossing number inequality. This proof is due to Székely (1997).

Trivial bound: $I(\mathcal{P},\mathcal{L}) \le |\mathcal{P}||\mathcal{L}|$

Using that every pair of points determine at most one line, and counting triples $(p,p',\ell) \in \mathcal{P}\times\mathcal{P}\times\mathcal{L}$ with $p \ne p'$ and $p,p' \in \ell$, this is $\le |\mathcal{P}|^2$ and

$$\ge \sum_{\ell\in\mathcal{L}} |\mathcal{P}\cap\ell|(|\mathcal{P}\cap\ell| - 1) \ge |I(\mathcal{P},\mathcal{L})|^2/|\mathcal{L}| - |I(\mathcal{P},\mathcal{L})|$$

Combining we get

$$I(\mathcal{P}, \mathcal{L}) \lesssim |\mathcal{P}||\mathcal{L}|^{1/2} + |\mathcal{L}|$$

By point-line duality, also

$$I(\mathcal{P}, \mathcal{L}) \lesssim |\mathcal{L}||\mathcal{P}|^{1/2} + |\mathcal{P}|$$

This gives $n^{3/2}$ for n points and n lines. Can we do better? Note that this is tight for planes over finite fields. Need to use topology of Euclidean space.

Proof of Szemerédi–Trotter theorem. Assume that there are no lines with < 2 incidences (otherwise remove such lines repeatedly until this is the same; we remove $\leq |\mathcal{L}|$ incidences this way).

Draw a graph based on incidences. Vertices are point in $\mathcal{P}$ and edges join consecutive points of $\mathcal{P}$ on a given line of $\mathcal{L}$.

A line with k incidences gives $k - 1 \geq k/2$ edges, so the total number of edges is $\leq |I(\mathcal{P}, \mathcal{L})|/2$.

There are at most $|\mathcal{L}|^2$ crossings. So by crossing number inequality

$$|\mathcal{L}|^2 \geq \operatorname{cr}(G) \gtrsim \frac{|E|^3}{|V|^2} \gtrsim \frac{|I(\mathcal{P}, \mathcal{L})|^3}{|\mathcal{P}|^2} \quad \text{if } |I(\mathcal{P}, \mathcal{L})| \geq 8|\mathcal{P}|.$$

So $I(\mathcal{P}, \mathcal{L}) \lesssim |\mathcal{P}|^{2/3}|\mathcal{L}|^{2/3} + |\mathcal{P}|$. Remember to add $|\mathcal{L}|$ to the bound from the first step of the proof (removing lines with < 2 incidences). □

2.5 Dense packing of spheres in high dimensions

Question 2.5.1. What is the maximum density of a packing of non-overlapping unit balls in $\mathbb{R}^n$ for large n?

Here the **density** is fraction of volume occupied (fraction of the box $[-n, n]^d$ as $n \to \infty$)

Let Δ_n denote the supremum of unit ball packing densities in $\mathbb{R}^n$

Exact maximum only solved in dimension $1, 2, 3, 8, 24$. Dimensions 8 and 24 were only solved recently (see this Quanta magazine story). Dimensions 8 and 24 are special because of the existences of highly symmetric lattices (E_8 lattice in dimension 8 and Leech lattice in dimension 24).

What are examples of dense packings?

We can add balls greedily. Any *maximal* packing has density $\geq 2^{-n}$. Doubling the ball radius would cover space

What about lattices? $\mathbb{Z}^n$ has sphere packing density $\operatorname{vol}(B(1/2)) = \frac{\pi^{n/2}}{(n/2)!2^n} < n^{-cn}$.

Best upper bound: Kabatiansky–Levenshtein (1978): $\Delta_n \le 2^{-(0.599\cdots+o(1))n}$

Existence of a dense lattice? (Optimal lattices known in dimensions 1–8 and 24)

We will use the probabilistic method to show that a random lattice has high density.

How does one pick a random lattice?

A **lattice** the $\mathbb{Z}$-span of of its basis vectors $v_1, \dots, v_n$. It's covolume (volume of its fundamental domain) is given by $|\det(v_1|v_2|\cdots|v_n)|$.

So every matrix in $\mathrm{SL}_n(\mathbb{R})$ corresponds to a unimodular lattice (i.e., covolume 1).

Every lattice can be represented in different ways by picking a different basis (e.g., $\{v_1 + v_2, v_2\}$). The matrices $A, A' \in \mathrm{SL}_n(\mathbb{R})$ represent the same lattice iff $A' = AU$ for some $U \in \mathrm{SL}_n(\mathbb{Z})$.

So the space of unimodular lattices is $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$, which has a finite Haar measure (even though this space not compact), so can normalize to a probability measure.

We can pick a **random unimodular lattice** in $\mathbb{R}^n$ by picking a random point in $\mathrm{SL}_n(\mathbb{R})/\mathrm{SL}_n(\mathbb{Z})$ according to its Haar probability measure.

The following classic result of Siegel acts as like a linearity of expectations statement for random lattices.

Theorem 2.5.2 (Siegel mean value theorem)**.** Let L be the random lattice in $\mathbb{R}^n$ as above and $S \subset \mathbb{R}^n$. Then

$$\mathbb{E}|S \cap L \setminus \{0\}| = \lambda_{\mathrm{Leb}}(S)$$

Proof sketch.

1. $\mu(S) = \mathbb{E}|S \cap L \setminus \{0\}|$ defines a measure on $\mathbb{R}^n$ (it is additive by linearity of expectations)
2. This measure is invariant under $\mathrm{SL}_n(\mathbb{R})$ action (since the random lattice is choosen with respect to Haar measure)
3. Every $\mathrm{SL}_n(\mathbb{R})$-invariant measure on $\mathbb{R}^n$ is a constant multiple of the Lebesgue measure.
4. By considering a large ball S, deduce that $c = 1$. □

Theorem 2.5.3 (Minkowski 1905)**.** For every n, there exist a lattice sphere packing in $\mathbb{R}^n$ with density $\ge 2^{-n}$.

Proof. Let S be a ball of volume 1 (think $1 - \epsilon$ for arbitrarily small $\epsilon > 0$ if you like) centered at the origin. By the Siegel mean value theorem, the random lattice is has expected 1 nonzero lattice point in S, so with positive probability it has no nonzero lattice point in S. Putting a copy of $\frac{1}{2}S$ (volume 2^{-n}) at each lattice point then gives a lattice packing of density $\geq 2^{-n}$ □

Here is a factor 2 improvement. Take S to be a ball of volume 2. Note that the number of nonzero lattice points in S must be even (if $x \in S$ then $-x \in S$). So same argument gives lattice packing of density $\geq 2^{-n+1}$.

The above improvement uses 2-fold symmetry of $\mathbb{R}^n$. Can we do better by introducing more symmetry?

Historically, a bunch of improvements of the form $\geq cn2^{-n}$ for a sequence of improving constants $c > 0$

Venkatesh (2012) showed that one can get a lattice with a k-fold symmetry by building it using two copies of the cyclotomic lattice $\mathbb{Z}[\omega]$ where $\omega = e^{2\pi/k}$. Every lattice of this form has k-fold symmetry by multiplication by ω.

Skipping details, one can extend the earlier idea to choose a random unimodular lattice in in dimension $n = 2\phi(k)$ with k-fold length-preserving symmetry (without fixed points). An extension of Siegel mean value theorem also holds in this case.

By apply same argument with S being a ball of volume k, we get a a lattice packing of density $\geq k2^{-n}$ in $\mathbb{R}^n$. This bound can be optimized (in term of asymptotics along a subsequence of n) by taking primorial $k = p_1 p_2 \cdots p_m$ where $p_1 < p_2 < \cdots$ are the prime numbers. This gives the current best known bound:

Theorem 2.5.4 (Venkatesh 2012)**.** For infinitely many n, there exists a lattice sphere packing in $\mathbb{R}^n$ of density

$$\geq (e^{-\gamma} - o(1)) n \log \log n 2^{-n}.$$

Here $\gamma = 0.577\ldots$ is Euler's constant.

Open problem 2.5.5. Do there exist lattices (or sphere packings) in $\mathbb{R}^n$ with density $\geq (c + o(1))^n$ for some constant $c > 1/2$?

2.6 Unbalancing lights

Theorem 2.6.1. Let $a_{ij} = \pm 1$ for all $i, j \in [n]$. There exists $x_i, y_j \in \{-1, 1\}$ for all $i, j \in [n]$ such that

$$\sum_{i,j=1}^{n} a_{ij} x_i y_j \geq \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{3/2}$$

Interpretation: $n \times n$ array of lights. Can flip rows and columns. Want to turn on as many lights as possible.

Proof. Choose $y_1, \dots, y_n$ randomly. And then choose x_i to make the i-th row sum nonnnegative. Let

$$R_i = \sum_{j=1}^{n} a_{ij} y_j \qquad \text{and} \qquad R = \sum_{i=1}^{n} |R_i|.$$

How is R_i distributed? Same distribution as $S_n = \epsilon_1 + \cdots + \epsilon_n$, a sum of n i.i.d. uniform $\{-1, 1\}$. And so for every i

$$\mathbb{E}[|R_i|] = \mathbb{E}[|S_n|] = \left(\sqrt{\frac{2}{\pi}} + o(1) \right) \sqrt{n},$$

e.g., by central limit theorem

$$\begin{aligned} \lim_{n \to \infty} \mathbb{E}\left[\frac{|S_n|}{\sqrt{n}} \right] &= \mathbb{E}[|X|] \qquad \text{where } X \sim \text{Normal}(0,1) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |x| e^{-x^2/2} \, dx = \sqrt{\frac{2}{\pi}} \end{aligned}$$

(one can also use binomial sum identities to compute exactly: $\mathbb{E}[|S_n|] = n 2^{1-n} \binom{n-1}{\lfloor (n-1)/2 \rfloor}$, though it is rather unnecessary to do so.) Thus

$$\mathbb{E}[R] = \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{3/2}.$$

Thus with positive probability, $R \geq \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{3/2}$. □

The next example is tricky. The proof will set up a probabilistic process where the parameters are not given explicitly. A compactness argument will show that a good choice of parameters exists.

Theorem 2.6.2. Let $V = V_1 \cup \cdots \cup V_k$, where $V_1, \dots, V_k$ are disjoint sets of size n. The edges of the complete k-uniform hypergraph on V are colored with red/blue. Suppose that every edge formed by taking one vertex from each $V_1, \dots, V_k$ is colored blue. Then there exists $S \subset V$ such that the number of red edges and blue edges in S differ by more than $c_k n^k$, where $c_k > 0$ is a constant.

Proof. Let's do this proof for $k = 3$. Proof easily generalizes to other k.

Let p_1, p_2, p_3 be real numbers to be decided. We are going to pick S randomly by including each vertex in V_i with probability p_i, independently. Let

$$a_{i,j,k} = \#\{\text{blue edges in } V_i \times V_j \times V_k\} - \#\{\text{red edges in } V_i \times V_j \times V_k\}.$$

Then

$$\mathbb{E}[\#\{\text{red edges in } S\} - \#\{\text{blue edges in } S\}]$$

equals to some polynomial

$$f(p_1, p_2, p_3) = \sum_{i \le j \le k} a_{i,j,k} p_i p_j p_k = n^3 p_1 p_2 p_3 + a_{1,1,1} p_1^3 + a_{1,1,2} p_1^2 p_2 + \cdots.$$

(note that $a_{1,2,3} = n^3$ by hypothesis). We would be done if we can find $p_1, p_2, p_3 \in [0,1]$ such that $|f(p_1, p_2, p_3)| > c$ for some constant $c > 0$ (not depending on the $a_{i,j,k}$'s). Note that $|a_{i,j,k}| \le n^3$. We are done after the following lemma

Lemma 2.6.3. Let P_k denote the set of polynomials $g(p_1, \dots, p_k)$ of degree k, whose coefficients have absolute value ≤ 1, and the coefficient of $p_1 p_2 \cdots p_k$ is 1. Then there is a constant $c_k > 0$ such that for all $g \in P_k$, there is some $p_1, \dots, p_k \in [0,1]$ with $|g(p_1, \dots, p_k)| \ge c$.

Proof of Lemma. Set $M(g) = \sup_{p_1, \dots, p_k \in [0,1]} |g(p_1, \dots, p_k)|$ (note that sup is achieved as max due to compactness). For $g \in P_k$, since g is nonzero (its coefficient of $p_1 p_2 \cdots p_k$ is 1), we have $M(g) > 0$. As P_k is compact and $M \colon P_k \to \mathbb{R}$ is continuous, M attains a minimum value $c = M(g) > 0$ for some $g \in P_k$. ■ □

18.226 Probabilistic Method in Combinatorics
Fall 2020