The Probabilistic Method

Espen Slettnes

Berkeley Math Circle

Oct 2, 2019

The probabilistic method is a a very powerful tool in combinatorics. The basic gist of it is to show that

(a) a statement E *can* be true by showing that its probability greater than 0. In turn,

(b) a common subtype of (a) is to show that E can be true is to show that the probability of its negative $(\neg E)$ is less than 1,

(c) a common subtype of (b) is to show that X can be at least or at most a by showing that $\mathrm{E}[X] \geq a$ or $\mathrm{E}[X] \leq a$, respectively, and

(d) a common subtype of (c) is to show that it is possible for $|X|$ to be at least or at most $a > 0$ by showing that $\mathrm{E}[X] = 0$ and $\mathrm{Var}(X) \geq a^2$ or $\mathrm{Var}(X) \leq a^2$, respectively.

It's best to get a feel for it by just seeing examples.

Example 1. (Russia 1999) In a certain school, every boy likes at least one girl. Prove that we can find a set S of at least half the students in the school such that each boy in S likes an odd number of girls in S.

Walkthrough:

(a) Flip a coin for every girl to determine whether she goes in S or not. What is the expected number of girls in S?

(b) Put every boy who likes an odd number of girls in S into S. What is the expected number of boys in S?

(c) What is the expected size of S? Conclude.

Example 2. (Erdős) A set of nonzero integers A is sum-free if there is no $a_1, a_2, a_3 \in A$ with $a_1 + a_2 = a_3$.

Show that any set of nonzero integers B contains a sum-free subset A with $|A| > |B|/3$.

Walkthrough:

(a) Find a prime $p \equiv 2 \pmod 3$, where the elements of B are distinct $\pmod p$. Let $p = 3n - 1$ and $\overline{B}$ be the set of residues $b \pmod p$ for all $b \in B$.

(b) Find a sum-free set A' of residues $\pmod p$ of size n. (Hint: One possible answer is a particular set of n consecutive residues.)

(c) Let $kA' = \{ka \mid a \in A'\}$ for any nonzero residue k. Show that $A_k = kA' \cap \overline{B}$ is sum-free.

(d) Find the expected value of $|A_k|$ for random k and conclude.

Example 3. (Sperner) If S is a subset of $2^{[n]}$ with no two $A, B \in S$ satisfying $A \subset B$,

$$|S| \leq \binom{n}{\lfloor n/2 \rfloor}.$$

($2^{[n]}$ is shorthand for the set of subsets of $\{1, \ldots, n\}$.)

Walkthrough:

(a) Consider a random permutation of $1, 2, \ldots, n$. For any set $A \in S$, let E_A be the event that the first $|A|$ terms of the permutation are the elements of A. What is the probability of E_A?

(b) Show that $E_A \wedge E_B$ never happens, and thus that $1 \geq \sum_{A \in S} \frac{1}{\binom{n}{|A|}}$.

(c) Conclude.

Example 4. (Bollobás Theorem).

Let $A_1, A_2, \ldots, A_n$ and $B_1, B_2, \ldots, B_n$ be distinct finite sets such that $A_i \cap B_i = \emptyset$ for all i, but $A_i \cap B_j \neq \emptyset$ for all $i \neq j$. Then

$$\sum_{i=1}^{n} \frac{1}{\binom{|A_i|+|B_i|}{|A_i|}} \leq 1.$$

Walkthrough:

(a) Consider a random permutation of the elements of $\bigcup_{i=1}^{n} A_i \cup B_i$. For each $i \in [n]$, let E_i be the event that all elements of A_i come before all elements of B_i. What is the probability of E_i?

(b) Show that $E_i \wedge E_j$ never happens.

(c) Conclude.

Example 5. (Erdős-Ko-Rado) Let $k \leq \frac{n}{2}$ and let S be a set of subsets of $[n] = \{1, \ldots, n\}$, each of size k. If $\forall A, B \in S,\ A \cap B \neq \emptyset$, then $|S|$ is at most $\binom{n-1}{k-1}$.

Walkthrough:

(a) Take a random permutation of $[n]$. Say that A is an *arc* of this permutation if it is a set of consecutive terms, possibly wrapping around. For example, given the permutation $(3, 4, 8, 2, 7, 5, 1, 6)$, the arcs of size 4 are $\{2, 3, 4, 8\}$, $\{2, 4, 7, 8\}$, $\{2, 5, 7, 8\}$, $\{1, 2, 5, 7\}$, $\{1, 5, 6, 7\}$, $\{1, 3, 5, 6\}$, $\{1, 3, 4, 6\}$, and $\{3, 4, 6, 8\}$. Show that the intersection of all arcs in S is non-empty.

(b) What is the maximum possible number of arcs in S?

(c) Compute the probability that a given $A \in S$ is an arc of $[n]$. In terms of $|S|, n,$ and $k,$ what is the expected number of arcs in S?

(d) Conclude.

Graph Theory

Example 6. Show that a graph G with m edges has a bipartite subgraph with $\frac{m}{2}$ edges.

Walkthrough:

(a) Flip a coin on every vertex and define a corresponding bipartite subgraph.

(b) Show that the expected number of edges in the subgraph is $\frac{m}{2}$, and conclude.

Example 7. A *tournament* is a directed graph with exactly one directed edge between any two vertices.

Show that for any $n \in \mathbb{N}$, there is a tournament on n vertices with more than $\frac{n!}{2^{n-1}}$ Hamiltonian paths.

Walkthrough:

(a) Take a random tournament. What is the chance that a given sequence of n players is a Hamiltonian path? Conclude.

Example 8. (Erdős)

Let $R(s)$ denote the Ramsey number of s, i.e., the smallest integer n for which, when one colors the edges of K_n either red or blue, there must be a monochromatic K_s.

Show that $R(s) > 2^{s/2}$ for $s \geq 3$.

Walkthrough:

(a) Let $n = \lfloor 2^{s/2} \rfloor$. Showing that $R(s) > n$ is showing that there existing a coloring of K_n with no monochromatic K_s. Randomly color each edge of K_n red or blue. What is the probability that a given set of s vertices forms a monochromatic K_s?

(b) Show that it suffices to show $\binom{n}{s} < 2^{\binom{s}{2}-1}$, and verify this is true by showing that $\binom{n}{s} < \frac{n^s}{2^s} < 2^{\binom{s}{2}-1}$.

Example 9. (Caro-Wei). An independent set of a graph G is a set of vertices between which there are no edges. The independence number $\alpha(G)$ is the size of the maximal independent set.

Show that

$$\alpha(G) \geq \sum_{v \in G} \frac{1}{1+d(v)}$$

where $d(v)$ is the degree of v.

Walkthrough:

(a) Permute the vertices of G randomly. Let S be the set of vertices which are not connected to any vertex before itself in the permutation. Show that S is independent.

(b) Find the probability in terms of $d(v)$ that v is in S.

(c) Find the expected value of $|S|$ and conclude.

Example 10. (The Crossing Lemma). The crossing number of a graph G, denoted $\operatorname{cr}(G)$, is the minimum number of crossings with which you can draw the graph in 2 dimensions. (By definition, the crossing number of a planar graph is 0.)

Show that if the average degree of a graph G is at least 8, $\operatorname{cr}(G) \geq \dfrac{m^3}{64n^2}$ where n and m, respectively, are the number of vertices and edges in G.

Walkthrough:

(a) Show that $\operatorname{cr}(G) \geq m - 3n$ without using the fact that the average degree of G is at least 8.

(Hint: In fact, $\operatorname{cr}(G) \geq m - (3n-6)$.)

(b) On every vertex, flip a coin that comes up with heads with probabilty p. Let H be the subgraph formed by all vertices where the coin landed on a head. What is the expected number of vertices in H? Number of edges? Crossing number (just an upper bound)?

(c) Conclude that $p^4 \operatorname{cr}(G) \geq p^2 m - 3np$.

(d) Show that $p = 4n/m$ is less than or equal to 1. Conclude.

Example 11. (Noga Alon)

The *minimum degree* of a graph G, denoted $\delta(G)$, is the minimum degree of any vertex $v \in V(G)$.

A set $X \subseteq V(G)$ *covers* a vertex v if $v \in X$ or v is adjacent to some vertex in X. A *dominating set* S is a any such vertex set that covers G. The *minimum dominating set* is the smallest such subset, and the *domination number*, denoted $\gamma(G)$, is the size of this minimum dominating set.

Show that $\gamma(G) \leq n \dfrac{1 + \ln(\delta(G)+1)}{\delta(G)+1}$.

Walkthrough:

(a) Take a vertex to be in X with probability p. What is the expected size of X?

(b) What is the expected number of vertices *not* dominated by X?

(c) Show that $\gamma(G) \leq np + ne^{-p(\delta+1)}$.

(d) Show that $p = \frac{\ln(\delta+1)}{\delta+1}$ is between 0 and 1, inclusive, and conclude.

Example 12. The *chromatic number* $\chi(G)$ of a graph G is the smallest number of colors required to color its vertices so that no two adjacent vertices have the same color.

If each vertex is assigned a list of possible colors, the *list chromatic number* $\operatorname{ch}(G)$ is the minimum number of colors needed per list to guarantee that there is a coloring of the graph where each vertex uses a color from its list. Note that $\operatorname{ch}(G) \geq \chi(G)$.

Show that $\operatorname{ch}(G) \leq \chi(G) \ln(n)$, where $n = |V(G)|$, the number of vertices in G.

Walkthrough:

(a) In the optimal coloring, for $1 \leq i \leq \chi(G)$, let A_i be the set of vertices that are colored i.

For every $1 \leq v \leq n$, let L_v be a list of (at least) $\chi(G)\ln(n)$ colors. Assign each list color at random to some A_i. Let C_i be the set of colors assigned to A_i.

Show that it suffices for all $v \in A_i$ to satisfy $L_v \cap C_i \neq \emptyset$.

(b) What is the probability that $L_v \cap C_i = \emptyset$?

(c) Show that it suffices that $e^{|L_v|/\chi(G)} \geq n$. (Hint: $1 + \epsilon < e^{1+\epsilon}$ for $\epsilon \neq 0$.)

(d) Conclude.

Olympiad Examples

Example 13. (IMO shortlist 2006, C3) Let S be a finite set of points in the plane such that no three of them are on a line. For each convex polygon P whose vertices are in S, let $a(P)$ be the number of vertices of P, and let $b(P)$ be the number of points of S which are outside P. A line segment, a point, and the empty set are considered as convex polygons of 2, 1, and 0 vertices respectively. Prove that for every real number x

$$\sum_P x^{a(P)}(1-x)^{b(P)} = 1,$$

where the sum is taken over all convex polygons with vertices in S.

Walkthrough:

(a) Show that it suffices to prove that the equation holds for $0 \leq x \leq 1$.

(b) On every point, flip a coin that lands on heads with probability x. Interpret $x^{a(P)}(1-x)^{b(P)}$ as a probability.

(c) Show that there is exactly one P for which the corresponding event happens. Conclude.

Example 14. (IMO Shortlist 1999, C4) Let A be a set of N residues $\pmod{N^2}$. Prove that there exists a set B of of N residues $\pmod{N^2}$ such that $A+B = \{a+b | a \in A, b \in B\}$ contains at least half of all the residues $\pmod{N^2}$.

Walkthrough:

(a) Choose B at random. What is the probability that a given residue is not in $A+B$?

(b) Conclude.

Example 15. (USAMO 2012, problem 2) A circle is divided into 432 congruent arcs by 432 points. The points are colored in four colors such that some 108 points are colored Red, some 108 points are colored Green, some 108 points are colored Blue, and the remaining 108 points are colored Yellow. Prove that one can choose three points of each color in such a way that the four triangles formed by the chosen points of the same color are congruent.

Walkthrough:

(a) Consider a random symmetry of the 432-gon formed by the points. How many are there? (Don't include the identity.)

(b) Color the red points that land on green points orange. What's the expected number of orange points? How many orange points are guaranteed to achievable?

(c) Pick another random symmetry, and color these orange points that land on blue points purple. What is the expected number of purple points? How many purple points are guaranteed to be achievable?

(d) Pick yet another random symmetry and color these purple points that land on yellow points pink. What is the expected number of pink points?

(e) Conclude.

Example 16. (USAMO 2012, #6) For integer $n \geq 2$, let $x_1, x_2, \ldots, x_n$ be real numbers satisfying

$$x_1 + x_2 + \ldots + x_n = 0, \qquad \text{and} \qquad x_1^2 + x_2^2 + \ldots + x_n^2 = 1.$$

For each subset $A \subseteq \{1, 2, \ldots, n\}$, define
$$S_A = \sum_{i \in A} x_i.$$
(If A is the empty set, then $S_A = 0$.)

Prove that for any positive number λ, the number of sets A satisfying $S_A \geq \lambda$ is at most $2^{n-3}/\lambda^2$. For which choices of $x_1, x_2, \ldots, x_n, \lambda$ does equality hold?

Walkthrough:

(a) Flip n coins and let $X_i = x_i$ if the i^{th} coin comes up heads, and $X_i = 0$ if it comes up tails. Let A be the set of indices of coins that came up heads. Write
$$\sum_{i=1}^{n} X_i = S_A.$$

(b) What is $\mathrm{E}[X_i]$? $\mathrm{E}[S_A]$? $\mathrm{Var}(X_i)$? $\mathrm{Var}(S_A)$?

(c) What does Chebyshev's inequality say when you plug in 2λ?

(d) Show that $\mathrm{P}(S_A \geq \lambda) = \mathrm{P}(-S_A \geq \lambda)$.

(e) Conclude that the inequality given in the problem holds.

(f) What does the equality case of Markov's inequality say about the equality case of Chebychev's inequality? What does the equality case of Chebychev's inequality say about the equality case for this problem?

(g) Conclude.

Example 17. (USAMO 2010, #6) A blackboard contains 68 pairs of nonzero integers. Suppose that for each positive integer k at most one of the pairs (k, k) and $(-k, -k)$ is written on the blackboard. A student erases some of the 136 integers, subject to the condition that no two erased integers may add to 0. The student then scores one point for each of the 68 pairs in which at least one integer is erased. Determine, with proof, the largest number N of points that the student can guarantee to score regardless of which 68 pairs have been written on the board.

Walkthrough:

(a) WLOG (why?), all pairs of the form (k, k) have positive entries. Let A be the set of such pairs. Find a strategy to erase at least $|A|$ such pairs.

(b) Flip a coin on all positive integers n, and erase all occurences of n if it lands heads or all occurences of $-n$ if it lands tails. In terms of $|A|$, what is a lower bound for the expected value of the number of erased pairs?

(c) What lower bound do you get on N from (a) and (b)?

(d) Let's try to do better. If instead, we put a bias on the coin so that it lands on heads with probability $p > 0.5$, what's the new bound on the expected score?

(e) What value of p maximizes the minimum bound over all possible sizes of A? What bound do you get for this p?

(f) What does this give as a lower bound for N?

(g) We now show achievability of this N. When is equality achieved in the bound from (d)? What does this suggest the equality case might look like for the problem?

(h) Play around until you find an equality case. Hint: Keep as much symmetry as possible.
