Probabilistic Method

"To prove that a structure with certain desired properties exists, one define an appropriate probability space of structures and then shows that the desired properties hold in this space with positive probability." - *The Probabilistic Method* by Noga Alon and Joel Spencer

There are two general steps to using the probabilistic method:

1. Define an appropriate probability space.
2. Calculate the probability of getting the desired object.

You can then reason from there about the existence of the desired object using that probability.

If the probability of finding a valid structure over all possibilities is greater than 0, then such an object exists.

If the probability of finding a valid structure over all possibilities is less than 1, then there exists an object that does not have those properties.

If you calculate the expected value of some random variable X, then we can say that with positive probability, there exists $X' \geq E[X]$ and likewise $X' \leq E[X]$.

As a silly example, what is the average number of legs per person? Some people have no legs or one leg (and much more rarely have three). So the average amount of legs $E[\textit{number of legs}]$ is slightly less than 2. The probabilistic method is just saying that because $1 < E[\textit{number of legs}] < 2$, there must exist someone with less than two legs.

Below are some tools you can use:

Theorem 1 (Union Bound)**.** For a countable set of events $A_1, A_2, \ldots, A_n$, we have

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i)$$

Independence: We say that an event A is **independent** of a *set of events* $\{B_1, \ldots, B_m\}$ if for all disjoint S, $T \subseteq [m]$, we have

$$\Pr\left[A \;\middle|\; \bigcap_{i \in S} B_i \cap \bigcap_{j \in T} \bar{B}_j\right] = \Pr[A]$$

Theorem 2 (Lovász Local Lemma)**.** Let A_i, …, A_N be events having dependency graph with maximum degree d. If $\max_i \Pr[A_i] \leqslant p$ and $ep(d+1) \leqslant 1$, then

$$\Pr\left[\bigcap i \bar{A}_i\right] > 0$$

Linearity of Expectation Problems

All these problems can be solved by decomposing the random variable you're interested in into indicator variables.

Problem 1: You take a 52-card deck, with 26 red cards and 26 black cards, shuffle it, and lay the cards out one-by-one in a row. What's the expected number of "monochromatic blocks", or groups of cards that are all the same color? For example, RRBRBBBBR has 5 blocks.

Problem 2: You have a bag initially containing n red marbles and m blue marbles. You pull marbles out of the bag randomly one by one, without replacement, until there are none left. What's the expected number of times you pull a red ball out immediately after pulling a blue ball out?

Problem 3: You and some friends enjoy playing a card game where you take n standard 52-card decks, shuffle them, and divide them up equally among $2n$ players, so that each player has a 26-card hand. We want to find the value of n that makes the game the most "fun", where we'll use the *variance* of the number of trump cards in your hand as a proxy for fun. Trump cards are simply one given suit, for example, in one game, spades might be the trump suit. What value of n maximizes this variance? **Hint:** Use $\text{Var } X = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.

Problem 4: Suppose we have n cars all driving on a single lane one-way road. Each driver picks a speed uniformly at random, and then attempts to drive at this speed. However, if the car in front of them is going slower than this, they will slow down to match that speed. As $t \to \infty$, "clumps" of cars will form all going the same speed. What is the expected number of such clumps?

Probabilistic Method Problems

Problem 5: (should be straightforward) Given subsets $S_1, \ldots, S_m \subset S$ such that $|S_i| = n$ for all i, is it possible to two-color each element of S such that none of the S_i are monochromatic (e.g. all red or all blue)?

Show that one can do such a coloring when $m < 2^{n-1}$.

Problem 6: (should be straightforward) We call a **tournament** a directed complete graph. Note that we can interpret vertices as players and directed edge (a, b) or $a \to b$ to mean that player a beats player b. All players play one another in this competition, so the "graph" is complete. Now recall that a Hamiltonian path is a path that visits all vertices exactly once. You can interpret this as a ranking of players where each person beats the next person in the path. Prove that there is a **tournament** on n vertices with at least $\frac{n!}{2^{n-1}}$ Hamiltonian paths.

Problem 7: Suppose $n \geqslant 4$ and let H be an n-uniform hypergraph with at most $4^{n-1}/3^n$ edges. Prove that there is a coloring of the vertices of H by four colors so that in every edge all four colors are represented.

Problem 8: Prove that there is an absolute constant $c > 0$ with the following property: let A be an $n \times n$ matrix with pariwise distinct entries. Then there is a permutation of the rows of A

so that no column in the permuted matrix contains an increasing sub-sequence of length at least $c\sqrt{n}$.

Problem 9: A k-SAT formula is an expression such as

$$(x_1 \text{ OR } x_4 \text{ OR } \bar{x}_6) \text{ AND } (x_1 \text{ OR } \bar{x}_2 \text{ OR } x_5)$$

where the variables x_i take values true or false, $\bar{x}_i$ means not x_i, and k variables or their negation are OR'd together in each clause. A formula is satisfiable if there is an assignment of values to the variables making the expression true.

a.) Using Markov's inequality, show that any k-SAT formula with fewer than 2^k clauses is satisfiable.

b.) Using the Lovasz Local Lemma, show that any k-SAT formula in which no variable lies in more than $2^k/(ek)$ clauses is satisfiable.

Problem 10: Using LLL, show that it is possible to color the edges of the complete n-vertex graph K_n with $k = \lceil 3\sqrt{n} \rceil$ colors so that no triangle is monochromatic.

Problem 11: Prove that any 10 (arbitrary) points in the 2-D plane can be covered by disjoint coins of radius 1, that is the coins cannot physically overlap. (If this seems "obvious", consider increasing the number 10 to say 100. Now can we always cover any 100 points with coins?)