

Unexpected Uses of Probability

Ravi Boppana
raviboppana@aol.com

July 14, 2005

1 The Probabilistic Method

The theme of these notes is problems whose statements have nothing to do with probability, but whose solutions involve probabilistic techniques. We will investigate examples from diverse branches of mathematics, mostly from olympiad contests.

We will focus on existence problems: proving that there exists a mathematical object with certain properties. Our main strategy, the probabilistic method, solves such problems in two steps:

- (i) Define a random object from an appropriate probability space.
- (ii) Show that the random object satisfies the desired properties with probability greater than 0.

Together, these two steps prove existence. The probabilistic method is a powerful tool for attacking existence problems. When faced with such a problem, consider applying the probabilistic method.

To learn more about the probabilistic method, you should read the “bible” of the field, the book by Alon and Spencer [1]. You can download the first four chapters of the first edition [11].

The remainder of these notes is divided as follows. Sections 2 through 6 provide many example problems from the fields of combinatorics, geometry, graph theory, algebra, and number theory. Section 7 provides a hint for solving each problem.

We will solve some of the problems during the talk. After the talk, try to solve some of the remaining problems. All of the problems can be solved with the probabilistic method. (Some can be solved in other ways too, but give the probabilistic method a try.) If you need a hint, take a peek at Section 7.

I have listed a reference for each problem. The reference I cite is rarely the original source. It is just a convenient place to find out more information.

2 Combinatorics

Problem 1 (IMC for University Students 2002 [6]) Two hundred students participated in a mathematical contest. They had 6 problems to solve. It is known that each problem was correctly solved by at least 120 participants. Prove that there must be two participants such that every problem was solved by at least one of these two students.

Problem 2 (IMO Shortlist 1987 [9]) Show that we can color the elements of the set $\{1, 2, \dots, 1987\}$ with 4 colors so that any arithmetic progression of ten terms, each in the set, is not monochromatic.

Problem 3 (Zarankiewicz) Show that there exists a partition of the set of positive integers into two classes such that neither class contains an infinite arithmetic progression, and neither class contains 3 consecutive integers.

Problem 4 (IMO 1987 [9]) Let $p_n(k)$ be the number of permutations of the set $\{1, \dots, n\}$, $n \geq 1$, which have exactly k fixed points. Prove that

$$\sum_{k=0}^n k p_n(k) = n!.$$

(Remark: A permutation f of a set S is a one-to-one mapping of S onto itself. An element i in S is called a fixed point of the permutation f if $f(i) = i$.)

Problem 5 With the same notation as in the previous problem, determine the value of

$$\sum_{k=0}^n k^2 p_n(k).$$

Problem 6 (IMO 1989 [9]) A permutation $(x_1, x_2, \dots, x_{2n})$ of the set $\{1, 2, \dots, 2n\}$, where n is a positive integer, is said to have property P if $|x_i - x_{i+1}| = n$ for at least one i in $\{1, 2, \dots, 2n - 1\}$. Show that, for each n , there are more permutations with property P than without.

Problem 7 (Kraft's inequality 1949 [3, Chap. 5]) Let C be a set of binary strings. (A binary string is a finite sequence of 0's and 1's). Say that C is a *prefix-free code* if no string in C is a prefix of another string in C .

(i) Show that if C is a prefix-free code, then

$$\sum_{x \in C} \frac{1}{2^{|x|}} \leq 1.$$

Here $|x|$ means the length of string x .

- (ii) Show that if C is a prefix-free code with $n \geq 1$ strings, then the average length of a string in C is at least $\log_2 n$.

Problem 8 (Sperner 1928 [1, Chap. 11]) Let \mathcal{F} be a family of subsets of $\{1, \dots, n\}$. Say that \mathcal{F} is an *antichain* if no set in \mathcal{F} is a subset of another set in \mathcal{F} .

- (i) Show that if \mathcal{F} is an antichain, then

$$\sum_{S \in \mathcal{F}} \frac{1}{\binom{n}{|S|}} \leq 1.$$

- (ii) Show that if \mathcal{F} is an antichain, then $|\mathcal{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$.

Problem 9 (IMO 1998 [9]) In a competition, there are a contestants and b judges, where $b \geq 3$ is an odd integer. Each judge rates each contestant as either “pass” or “fail”. Suppose k is a number such that, for any two judges, their ratings coincide for at most k contestants. Prove that $k/a \geq (b-1)/(2b)$.

Problem 10 (Iberoamerican Olympiad 2001 [9]) Let S be a set with n elements. Let S_1, S_2, \dots, S_k be subsets of S (for $k \geq 2$), each with at least r elements. Prove that there exist i and j (with $1 \leq i < j \leq k$) such that the number of common elements of S_i and S_j is at least

$$r - \frac{nk}{4(k-1)}.$$

3 Geometry

Problem 11 (Engel [4, Chap. 4]) There are 650 points in a circle of radius 16. We have a washer (in other words, an annulus or ring) of inner radius 2 and outer radius 3. Prove that we can place the washer so that it covers at least 10 of the points.

Problem 12 (All Russian Olympiad 1961 [9]) There are 120 unit squares in a 20-by-25 rectangle. Prove that we can place a circle of unit diameter inside the rectangle without intersecting any of the squares.

Problem 13 (Engel [4, Chap. 4]) Let S be a subset of the unit interval with measure (total length) more than $1/2$. Prove that there exist two points of S that are exactly distance 0.1 apart.

Problem 14 (Engel [4, Chap. 4]) 12% of (the surface of) a sphere is painted black; the remainder is white. Prove that there is a rectangular solid with all white vertices.

Problem 15 (Engel [4, Chap. 4]) There are several circles of total circumference 10 inside a square of side length 1. Prove that there is a line that intersects at least 4 of the circles.

Problem 16 (Engel [4, Chap. 5]) A tramp has a coat of area 1 with 5 patches. Each patch has area at least $1/2$. Prove that two patches exist with common area at least $1/5$.

Problem 17 (IMO 1970 [9]) In a plane there are 100 points, no three of which are collinear. Consider all possible triangles having these points as vertices. Prove that no more than 70% of these triangles are acute-angled.

Problem 18 (IMO 1971 [9]) Prove that for every natural number m , there exists a finite set S of points in a plane with the following property: For every point A in S , there are exactly m points in S which are at unit distance from A .

Problem 19 (IMO 1989 [9]) Let n and k be positive integers and let S be a set of n points in the plane such that

- (i) No three points of S are collinear, and
- (ii) For any point P of S there are at least k points of S equidistant from P .

Prove that:

$$k \leq \frac{1}{2} + \sqrt{2n}.$$

4 Graph Theory

Problem 20 (Szele 1943 [1, Chap. 2]) In a (round-robin) *tournament*, every player plays one game with every other player. A *Hamiltonian path* of the tournament is an ordering of the players from left to right so that every player (except the last) beat the player immediately to its right. Let n be a positive integer. Show that there is a tournament with n players that has at least $n!/2^{n-1}$ Hamiltonian paths.

Problem 21 (Erdős 1963 [1, Chap. 1]) Let k be a positive integer. Say that a (round-robin) tournament is *k-unrankable* if for every set of k players, there is another player who beat all of them. Show that there is a tournament with at least k players that is *k-unrankable*.

Problem 22 (Turán 1941 [1, Chap. 6]) Let $G = (V, E)$ be a graph with m edges and n vertices. Let $d(v)$ be the degree (number of edges) of vertex v . An *independent set* is a set of vertices such that no two are adjacent (joined by an edge).

(i) Show that G contains an independent set of size at least

$$\sum_{v \in V} \frac{1}{d(v) + 1}.$$

(ii) Show that G contains an independent set of size at least $n^2/(2m + n)$.

Problem 23 (IMO Shortlist 1986 [9] and German Olympiad [4, Chap. 14])

Let S be a set of n points in space ($n \geq 3$). The segments joining these points are of distinct length, and r of these segments are colored red. Let m be the smallest integer for which $m \geq 2 \cdot r/n$. Prove that there always exists a path of m red segments with their lengths sorted increasingly.

Problem 24 (USAMO 1985 [9]) There are $n \geq 2$ people at a party. Prove that there are two people such that, of the remaining $n - 2$ people, there are at least $\lfloor n/2 \rfloor - 1$ of them, each of whom knows both or else knows neither of the two. Assume that “knowing” is a symmetric relation, and that $\lfloor x \rfloor$ denotes the greatest integer less than or equal to x .

Problem 25 (Asian Pacific Olympiad 1989 [9]) Let S be a set consisting of m pairs (a, b) of positive integers with the property that $1 \leq a < b \leq n$. Show that there are at least

$$4m \cdot \frac{(m - \frac{n^2}{4})}{3n}$$

triples (a, b, c) such that (a, b) , (a, c) , and (b, c) belong to S .

Problem 26 (USAMO 1995 [9]) Suppose that in a certain society, each pair of persons can be classified as either *amicable* or *hostile*. We shall say that each member of an amicable pair is a *friend* of the other, and each member of a hostile pair is a *foe* of the other. Suppose that the society has n persons and q amicable pairs, and that for every set of three persons, at least one pair is hostile. Prove that there is at least one member of the society whose foes include $q(1 - 4q/n^2)$ or fewer amicable pairs.

5 Algebra

Problem 27 (Bay Area Math Olympiad 2004 [2]) Suppose one is given n real numbers, not all zero, such that their sum is zero. Prove that one can label these numbers a_1, a_2, \dots, a_n in such a manner that

$$a_1 a_2 + a_2 a_3 + \dots + a_{n-1} a_n + a_n a_1 < 0.$$

Problem 28 (Rudin [8] and Chinese Olympiad 1986 [10]) Let z_1, \dots, z_n be complex numbers. Show that there is a set $S \subseteq \{1, \dots, n\}$ such that

$$\left| \sum_{j \in S} z_j \right| \geq \frac{1}{\pi} \sum_{j=1}^n |z_j|.$$

Problem 29 Let z_1, \dots, z_n be complex numbers on the unit circle. Show that there is a semicircular arc of the unit circle that contains at least

$$\frac{n + \left| \sum_{j=1}^n z_j \right|}{2}$$

of the numbers.

Problem 30 (Romania 2004) Let z_1, \dots, z_n be complex numbers. Prove that there exist numbers a_1, \dots, a_n , each ± 1 , such that

$$\left| \sum_{j=1}^n a_j z_j \right|^2 \leq \sum_{j=1}^n |z_j|^2.$$

Problem 31 (Weierstrass Approximation Theorem 1885 [1, Chap. 7])

Let F be a continuous function on $[0, 1]$. Let $\epsilon > 0$. Show that there is a polynomial G such that for each p in $[0, 1]$, we have $|F(p) - G(p)| \leq \epsilon$.

Problem 32 (Putnam 1947 [9]) Given $P(z) = z^2 + az + b$, a quadratic polynomial of the complex variable z with complex coefficients a, b . Suppose that $|P(z)| = 1$ for every z such that $|z| = 1$. Prove that $a = b = 0$.

Problem 33 (Math. Tripos 1929 [5, Chap. 6] and Putnam Feb. 1958 [9])
If a_0, a_1, \dots, a_n are real numbers satisfying

$$\frac{a_0}{1} + \frac{a_1}{2} + \dots + \frac{a_n}{n+1} = 0,$$

show that the equation $a_0 + a_1x + a_2x^2 + \dots + a_nx^n = 0$ has at least one real root.

Problem 34 Let p and q be nonnegative numbers that add up to 1. Let m and n be nonnegative integers. Prove that

$$(1 - p^m)^n + (1 - q^n)^m \geq 1.$$

6 Number Theory

Problem 35 (IMO Shortlist 1991 [9]) Let A be a set of n residues mod n^2 . Show that there is a set B of n residues mod n^2 such that at least half of the residues mod n^2 can be written as $a + b$ with a in A and b in B .

Problem 36 (USA Team Selection Test 2001 [7]) For a set S , let $|S|$ denote the number of elements in S . Let A be a set of positive integers with $|A| = 2001$. Prove that there exists a set B such that

- (i) $B \subseteq A$;

(ii) $|B| \geq 668$;

(iii) for any $u, v \in B$ (not necessarily distinct), $u + v \notin B$.

Problem 37 (Erdős 1965 [1, Chap. 1]) A set of numbers is *sum-free* if the sum of every two numbers (possibly equal) in the set is not in the set. Let A be a set of $n \geq 1$ nonzero integers. Show that there is a sum-free subset of A with more than $n/3$ elements.

Problem 38 (Erdős 1956 [1, Chap. 8]) If S is a set of nonnegative integers, and n is a nonnegative integer, define $R_S(n)$ to be the number of ways that n can be represented as the sum of two distinct integers in S . Show that there is a set S such that for all $n > 1$, we have

$$1 \leq R_S(n) \leq 1000 \ln n.$$

7 Hints

Below are brief hints for solving all of the problems via the probabilistic method. Most of the hints suggest at a minimum how to define the random object.

Hint 1 Let v and w be two random students (with replacement). What is the probability that neither v nor w solved Problem 1? Problem 2? And so on until Problem 6.

Hint 2 Consider a random 4-coloring of 1..1987. What is the probability that a particular progression of length 10 is monochromatic? How many arithmetic progressions of length 10 are there in 1..1987?

Hint 3 Color each odd integer red or blue randomly. Color each even integer the opposite color of the preceding odd integer.

Hint 4 Let f be a random permutation of 1.. n . What is the probability that 1 is a fixed point? 2? And so on up to n . Add these probabilities to get the expected number of fixed points of f . Do you see how that result is essentially the same as the given problem?

Hint 5 Let f be a random permutation of 1.. n . Let k be the number of fixed points of f . The preceding solution showed that $E[k] = 1$. Next show that $E[k(k-1)] = 1$.

Hint 6 Let x be a random permutation of 1.. $2n$. Let A_i be the event “ $|x_i - x_{i+1}| = n$ ”. What is the probability of A_i ? What is the probability of $A_i \cap A_j$? Now use PIE (the Principle of Inclusion-Exclusion) to bound from below the probability that some A_i is true.

Hint 7

- (i) Let r be a random infinite string. What is the probability that a particular string is a prefix of r ? What is the probability that some string in C is a prefix of r ?
- (ii) Apply the AM-GM (arithmetic-geometric mean) inequality to the first part.

Hint 8

- (i) A *chain* is a family of sets such that for each pair of sets, one is a subset of the other. A *maximal chain* is a chain with one subset of each size from 0 to n . Let \mathcal{C} be a random maximal chain. Note that \mathcal{C} and \mathcal{F} intersect in at most one element. What is the probability that some set in \mathcal{F} is in \mathcal{C} ?
- (ii) Use the first part. What is the largest binomial coefficient in a row of Pascal's triangle?

Hint 9 Let C be a random contestant. Let p be the number of judges that pass C , and f be the number of judges that fail C . Use the inequality $\binom{p}{2} + \binom{f}{2} \geq (b-1)^2/4$. Take expected value of both sides. Can you interpret $E[\binom{p}{2}]$ in terms of the original problem? How about $E[\binom{f}{2}]$?

Hint 10 Let v be a random element of S . Let d be the number of indices i such that S_i contains v . Use the inequality $\binom{d}{2} \geq (k-1)d/2 - k^2/8$. Take expected value of both sides. Can you interpret $E[d]$ in terms of the original problem? How about $E[\binom{d}{2}]$?

Hint 11 Choose a random point in a concentric circle of radius 19. Show that the expected number of points covered by a washer placed at that point is greater than 9.

Hint 12 Consider the inner 19-by-24 rectangle that is concentric with the large rectangle. Place the center of the circle at a random point in the 19-by-24 rectangle. What is the probability that the circle overlaps with a particular unit square?

Hint 13 Let x be a point chosen uniformly at random from the unit interval. Define y to be $x + 0.1$ if $\lfloor 10x \rfloor$ is even, and $x - 0.1$ if $\lfloor 10x \rfloor$ is odd. Note that y also has a uniform distribution on the unit interval. Also, the distance from x to y is 0.1. What is the probability that x is not in S ? How about y ?

Hint 14 Fix a coordinate system whose origin is the center of the sphere. Let p be a random point on the sphere. Consider the 7 reflections of p along all the axes. What is the probability that p is black? What about each of its reflections? [By the way, I don't know whether the result is true for a cube instead of a rectangular solid. Also, can the constant 12% be made significantly larger?]

Hint 15 Fix one side of the square, and choose a random point on that side. Consider the line that passes through that point and is perpendicular to that side. Show that the expected number of circles that the line passes through is more than 3.

Hint 16 Choose a random point on the coat. Let d be the number of patches that the point belongs to. Use the inequality $\binom{d}{2} \geq 2d - 3$. Take expected value of both sides. What is the expected value of d in terms of the original problem? What about the expected value of $\binom{d}{2}$?

Hint 17 First prove that among four points, the probability that a triangle is acute is at most 75%. Using that result, prove that among five points, the probability that a triangle is acute is at most 70%. Finally use that result to solve the problem.

Hint 18 Choose m random vectors on the unit circle. Let S be the set of sums of each of the 2^m subsets of the vectors.

Hint 19 For each point in S , fix one circle centered at that point that contains at least k other points in S . Let P be a random point in S . Let d be the number of the circles that contain P . Note that the expected value of d is at least k . Note that any two circles intersect in at most two points. Hence, show that the expected value of $\binom{d}{2}$ is at most $2(n - 1)$. Finally use the inequality $E[\binom{d}{2}] \geq \binom{E[d]}{2}$.

Hint 20 Let T be a random tournament with n players. (The outcome of each game is random and mutually independent of other games.) What is the expected number of Hamiltonian paths of T ?

Hint 21 Let n be a sufficiently large integer (say 3^k). Let T be a random tournament with n players. [By the way, there are explicit constructions known for unrankable tournaments, but they required advanced concepts such as quadratic residues and Riemann's hypothesis for finite fields.]

Hint 22

- (i) Consider a random ordering of the n vertices from left to right. Let I be the set of vertices whose neighbors are all to the right of it. Note that I is an independent set. What is the expected size of I ?
- (ii) Apply the AM–HM (arithmetic-harmonic mean) inequality to the expression in the first part.

Hint 23 Translation into graph theory: Given a graph with n vertices and r edges, whose edges are labeled $1, 2, \dots, r$, show that there exists a path of at least $2r/n$ edges with their labels in increasing order. Let's place a "hiker" at

a random vertex, chosen uniformly from the n vertices. In Step 1, look at the edge labeled 1; if at one of the endpoints of that edge, the hiker will traverse that edge. Similarly for Steps 2, 3, \dots , r . Note that after each step, the hiker is still uniformly distributed. What is the expected number of edges that the hiker traversed?

Hint 24 Translation into graph theory: Given a graph with $n \geq 2$ vertices, show that there are two vertices v and w such that at least $\lfloor n/2 \rfloor - 1$ of the other vertices are joined to both or neither of v and w . Let v and w be two vertices chosen at random (without replacement). Let x be an arbitrary vertex. Can you bound from above the probability that x is joined to exactly one of v and w ?

Hint 25 Translation into graph theory: Show that a graph with n vertices and m edges has at least $m(4m - n^2)/(3n)$ triangles. By the QM-AM (quadratic-arithmetic mean) inequality, the sum of the squares of the degrees is at least $4m^2/n$. Let $\{v, w\}$ be a random edge of the graph. Show that the expected number of triangles that contain v and w is at least $4m/n - n$.

Hint 26 Translation into graph theory: Show that in a triangle-free graph with n vertices and q edges, there is a vertex v such that removing v and its neighbors leaves at most $q(1 - 4q/n^2)$ edges. Again, by the QM-AM inequality, the sum of the squares of the degrees is at least $4q^2/n$. Let v be a random vertex. Show that the expected number of edges lost when v and its neighbors are removed is at least $4q^2/n^2$.

Hint 27 Let a_1, \dots, a_n be a random permutation of the original numbers. What is the expected value of the left-hand side?

Hint 28 Let r be a random ray from the origin. Let S be the set of indices j such that z_j makes an acute angle with r . Look at the projection of each z_j onto r . What is the expected absolute value of each projection?

Hint 29 By rotating, we may assume that the sum of the vectors is a nonnegative real number. Let y_0 be a random real number between -1 and 1 . Let p be the point on the right half of the unit circle whose y -coordinate is y_0 . Place the semicircle so that its midpoint is at p . What is the probability that a specific complex number is on that semicircle? What is the expected number of our complex numbers that are on that semicircle?

Hint 30 Choose each a_j randomly and independently to be -1 or $+1$. What is the expected value of the left-hand side?

Hint 31 Because f is continuous on a bounded interval, it is uniformly continuous and bounded there. Let n be a sufficiently large integer. Let h be the number of heads when n coins are flipped, each with probability p of heads. Let

$G(p)$ be the expected value of $F(h/n)$. Note that G is a polynomial of degree n . Because h is likely to be near pn , the value of $G(p)$ will be near $F(p)$.

Hint 32 Let z be a random point on the unit circle. What is the expected value of $|P(z)|^2$?

Hint 33 Let x be a random point on the unit interval $[0,1]$. What is the expected value of the given polynomial at x ?

Hint 34 Consider a random m -by- n matrix in which each cell is filled with a 1 with probability p , and a 0 with probability q . What is the probability that every row has at least one 1? What is the probability that every column has at least one 0?

Hint 35 Let B be a set of n residues mod n^2 chosen randomly (with replacement). Can you bound from above the probability that a particular residue is not in $A + B$?

Hint 36 Let p be a large prime of the form $3k+2$. Let C be the set $k+1..2k+1$. Note that C is a sum-free set mod p . Let r be a random integer from 1 to $p-1$. Let B consist of those numbers a in A such that $ar \bmod p \in C$. What is the expected size of B ?

Hint 37 This problem is a generalization of the previous one and has the same proof. [By the way, it is not known whether the constant $1/3$ can be increased.]

Hint 38 Let S be a random set of positive integers such that x is in S with probability $10\sqrt{\ln x/x}$ (for $x \geq 1000$). Show that the expected value of $R_S(n)$ is on the order of $\ln n$. Show that $R_S(n)$ is likely to be near its expected value. [By the way, it is a great open problem to determine whether there exists a set S such that $R_S(n)$ is always between 1 and another constant. My guess is that there is no such set.]

References

- [1] Noga Alon and Joel Spencer (2000), *The Probabilistic Method* (Second Edition), Wiley. See [11] for the first four chapters of the first edition.
- [2] Bay Area Mathematical Olympiad, <http://mathcircle.berkeley.edu>.
- [3] Thomas M. Cover and Joy A. Thomas (1991), *Elements of Information Theory*, Wiley.
- [4] Arthur Engel (1998), *Problem-Solving Strategies*, Springer.
- [5] Godfrey H. Hardy (1952), *A Course of Pure Mathematics* (Tenth Edition), Cambridge University Press.

- [6] International Mathematics Competition for University Students,
<http://www.imc-math.org>.
- [7] Kiran S. Kedlaya, *USA/IMO Archives*,
<http://www.unl.edu/amc/a-activities/a7-problems/problemarchive.html>.
- [8] Walter Rudin (1986), *Real and Complex Analysis* (Third Edition),
McGraw-Hill.
- [9] John Scholes, *Kalva Maths Problems*, <http://www.kalva.demon.co.uk>.
- [10] Andy Liu (1998), *Chinese Mathematics Competitions and Olympiads 1981–1993*, Australian Mathematics Trust.
- [11] Joel Spencer, *Sample Chapters of The Probabilistic Method*,
<http://www.cs.nyu.edu/cs/faculty/spencer/nogabook>.