PRIMES Math Problem Set

PRIMES 2017

Due December 1, 2016

Dear PRIMES applicant:

This is the PRIMES 2017 Math Problem Set. Please send us your solutions as part of your PRIMES application by **December 1, 2016**. For complete rules, see `http://web.mit.edu/primes/apply.shtml`

Note that this set contains two parts: "General Math problems" and "Advanced Math." Please solve as many problems as you can in both parts.

You can type the solutions or write them up by hand and then scan them. Please attach your solutions to the application as a PDF file. The name of the attached file must start with your last name, for example, "`smith-solutions`". Include your full name in the heading of the file.

Please write not only answers, but also proofs (and partial solutions / results / ideas if you cannot completely solve the problem). Besides the admission process, your solutions will be used to decide which projects would be most suitable for you if you are accepted to PRIMES.

You are allowed to use any resources to solve these problems, *except other people's help.* This means that you can use calculators, computers, books, and the Internet. However, if you consult books or Internet sites, please give us a reference.

Note that posting these problems on problem-solving websites before the application deadline is strictly forbidden! Applicants who do so will be disqualified, and their parents and recommenders will be notified.

Note that some of these problems are tricky. We recommend that you do not leave them for the last day. Instead, think about them, on and off, over some time, perhaps several days. We encourage you to apply if you can solve at least 50% of the problems.

We note, however, that there will be many factors in the admission decision besides your solutions of these problems.

Enjoy!

General Math Problems

Problem G1. A positive multiple of 11 is *good* if it does not contain any even digits in its decimal representation.

(a) Find the number of good integers less than 1000.

(b) Determine the largest such good integer.

(c) Fix $b \geq 2$ an even integer. Find the number of positive integers less than b^3 which are divisible by $b+1$ and do not contain any even digits in their base b representation. (This is the natural generalization of part (a) with 10 replaced by b.)

Solution. First we address the first two parts. The answer to (a) is $5+10 = 15$, and (b) is 979.

The case of two digits is easy: the only multiples are 11, 33, 55, 77, 99.

Now consider three-digit numbers. Suppose the number is $\overline{abc}$. Then we must have $a-b+c$ is a multiple of 11 (by taking modulo 11). Moreover, this number is odd since each of a, b, c are odd. Since $-11 < a-b+c < 33$, we conclude that

$$11 \text{ divides } 100a + 10b + c \iff a + c = 11 + b.$$

In particular, given odd a and c there exists a (unique) suitable b exactly when $a + c \geq 12$. It is easy to see there are exactly 10 such pairs. This concludes the calculation. As for (b) we may take $(a, c) = (9, 9)$ to get $b = 7$.

Explicitly, the set of 15 numbers is: 11, 33, 55, 77, 99, 319, 517, 539, 715, 737, 759, 913, 935, 957, 979.

For part (c) the answer is $\frac{\frac{b}{2}(\frac{b}{2}+1)}{2}$ by the same argument as in (a). $\square$

Problem G2. A fair six-sided die whose sides are labelled 1, 4, 9, 16, 25, 36 is rolled repeatedly until the sum of the rolled numbers is nonzero and either even or a multiple of 3.

(a) Compute the probability that when we stop, the sum is odd.

(b) Find the expected value of the number of rolls it takes until stopping.

Partial marks may be awarded for approximate answers obtained by computer simulation. (This is also a good way to check your answer!)

Solution. This is a typical "Markov-chain" problem. Call an outcome "good" if we have arrived at an odd number which is a multiple of 3, and "bad" otherwise. Let a and b denote the probability of of reaching a good outcome assuming the current running sum is $1 \pmod 6$ or $5 \pmod 6$, respectively.

Then, we have that

$$\begin{aligned} a &= \frac{0+b+0+b+0+a}{6} \\ b &= \frac{0+1+0+1+0+b}{6} \end{aligned}$$

where the numerator of each equation corresponds to the state we arrive at when each of 1^2, 2^2, …, 6^2 is rolled, in that order. Thus solving, we obtain $b = \frac{2}{5}$, and hence $a = \frac{4}{25}$.

Then, the answer to part (a) is

$$\frac{a+0+1+0+a+0}{6} = \frac{11}{50}.$$

Similarly, suppose we denote by x and y the expected number of steps assuming the current running sum is $1 \pmod 6$ or $5 \pmod 6$. By the same reasoning, we have

$$\begin{aligned} x &= 1 + \frac{0+y+0+y+0+x}{6} \\ y &= 1 + \frac{0+0+0+0+0+y}{6}. \end{aligned}$$

This gives $y = \frac{6}{5}$ and then $x = \frac{42}{25}$. So the answer to part (b) is

$$1 + \frac{x+0+0+0+x+0}{6} = \frac{39}{25}.$$

□

Problem G3. Let $\mathcal{H}$ be a hyperbola with center Z. Points A and B are selected on $\mathcal{H}$. Suppose that the tangents to $\mathcal{H}$ at points A and B intersect at a point C distinct from A, B, Z. Prove that line ZC passes through a point X in the interior of segment AB and determine the ratio AX/AB.

Solution. We claim ZC bisects AB, so the answer is $\frac{1}{2}$.

By taking a suitable affine transformation (which doesn't affect incidence or ratios) we may assume the asymptotes of $\mathcal{H}$ are perpendicular to each other. Now impose Cartesian coordinates so that $\mathcal{H}$ becomes $xy = 1$, in other words $Z = (0, 0)$. Let $A = (a, 1/a)$ and $B = (b, 1/b)$. Then the tangent at A has slope $-a^{-2}$ and hence has equation

$$y - \frac{1}{a} = -a^{-2}(x - a)$$

or

$$x + a^2 y = 2a.$$

Similarly, the tangent at B has equation $x + b^2 y = 2b$.

Solving, we deduce that the coordinates of C are

$$C = \left(\frac{2ab}{a+b}, \frac{2}{a+b} \right).$$

But the midpoint M of AB has coordinates $M = (\frac{a+b}{2}, \frac{a+b}{2ab})$, and it follows directly that Z, C, M are collinear. $\square$

Problem G4. Suppose $P(n)$ is a monic polynomial with integer coefficients for which $P(0) = 17$, and suppose distinct integers $a_1, \ldots, a_k$ satisfy $P(a_1) = \cdots = P(a_k) = 20$.

(a) Find the maximum possible value of k (over all P).

(b) Determine all P for which this maximum is achieved.

Solution. (a) The answer is $k = 3$.

To see this is maximal, note that $Q(n) = P(n) - 20$ is a monic polynomial with k distinct integer roots, so we must be able to write it as $Q(n) = (n - a_1)(n - a_2) \ldots (n - a_k) R(n)$ where $R(n)$ is some other integer coefficient monic polynomial. Putting $n = 0$ we get that $Q(0) = -3$ is the product of at least k distinct integers, so $k \le 3$ is immediate.

We now address part (b). WLOG, $a_1 < a_2 < a_3$. Now, we have $-3 = P(0) - 20 = -a_1 a_2 a_3 R(0)$, so $a_1 a_2 a_3 = 3$ or $a_1 a_2 a_3 = -3$. In the first case (a_1, a_2, a_3) must equal $(-3, -1, 1)$, so we have

$$P(n) = 20 + (n-1)(n+1)(n+3)R(n),$$

where R is any monic polynomial with $R(0) = 1$. In the second case (a_1, a_2, a_3) must equal $(-1, 1, 3)$, so we have

$$P(n) = 20 + (n-1)(n+1)(n-3)R(n),$$

where R is any monic polynomial with $R(0) = -1$. $\square$

Problem G5. A positive integer N is *nice* if all its decimal digits are 4 or 7.

(a) Find all nine-digit nice numbers which are divisible by 512.

(b) How many d-digit nice numbers are divisible by 512 for each d?

Solution. The critical idea is to show that for every $k \geq 1$, there is a unique k-digit nice number N_k (with digits 4 or 7) which is divisible by 2^k. When $k = 1$ this is clear. For $k > 1$, note that by taking modulo 2^{k-1} the last $k - 1$ digits are fixed, say A, and so we need exactly one of the numbers

$$10^{k-1} \cdot 4 + A, \quad 10^{k-1} \cdot 7 + A$$

to be divisible by 2^k. Taking modulo 2^k these numbers equal $A \pmod{2^k}$ and $A + 2^{k-1} \pmod{2^{k-1}}$, respectively. Since $A \equiv 0 \pmod{2^{k-1}}$, both uniqueness and existence follow. One can apply the content of this proof to then recover the number above.

As the proof is constructive, we may apply it to obtain the unique answer to part (a):

$$N_9 = 444{,}447{,}744 = 2^9 \times 868{,}062.$$

Now note that N_8 and N_7 differ from N_9 by $4 \cdot 10^8$ and $44 \cdot 10^7$, respectively, so it follows that N_7 and N_8 are also divisible by $2^9 = 512$. On the other hand we see $N_1, \ldots, N_6$ are not divisible by $2^9 = 512$. For $d > 9$, note that the any new digits no longer matter and hence the number of nice numbers divisible by 2^9 is exactly 2^{d-9}.

So the answer to (b) is

$$\begin{cases} 0 & 1 \leq d \leq 5 \\ 1 & 6 \leq d \leq 8 \\ 2^{d-9} & d \geq 9. \end{cases}$$

$\square$

Problem G6. A sequence $x_1, x_2, \ldots$ is defined by $x_1 = 10$ and

$$x_n = 3n \cdot x_{n-1} + n! - 3^n \cdot (n^2 - 2)$$

for integers $n \geq 2$. Derive a closed form for x_n (not involving Σ summation).

Solution. The answer is

$$x_n = \frac{3^n - 1}{2} \cdot n! + 3^n(n+2).$$

This is easily proved by induction on n. The base case $n = 1$ is immediate. For the inductive step,

$$\begin{aligned}
x_n &= 3n \cdot x_{n-1} + n! - 3^n \cdot (n^2 - 2) \\
&= 3n \left(\frac{3^{n-1} - 1}{2} \cdot (n-1)! + 3^{n-1}(n+1) \right) + n! - 3^n \cdot (n^2 - 2) \\
&= \left(3 \cdot \frac{3^{n-1} - 1}{2} + 1 \right) n! + 3^n \left(n(n+1) - (n^2 - 2) \right) \\
&= \frac{3^n - 1}{2} \cdot n! + 3^n \left(n + 2 \right)
\end{aligned}$$

as desired.

The way to discover the answer is to set $y_n = x_n / 3^n n!$ and write an explicit formula for $y_n - y_{n-1}$. Then we can compute y_n by summation, establishing the telescopic nature of the sum. □

Problem G7. Let ABC be a triangle, let a, b, c be the lengths of its sides opposite to A, B, C respectively, and let h_A, h_B, and h_C be the lengths of the altitudes from A, B, and C. Suppose that

$$\sqrt{a + h_B} + \sqrt{b + h_C} + \sqrt{c + h_A} = \sqrt{a + h_C} + \sqrt{b + h_A} + \sqrt{c + h_B}. \quad (1)$$

(a) Show that $(a + h_B)(b + h_C)(c + h_A) = (a + h_C)(b + h_A)(c + h_B)$.

(b) Prove that the three terms on the left-hand side of (1) are obtained by a permutation of the three terms on the right-hand side of (1). (Possible hint: consider polynomials with the three terms as roots.)

(c) Show that triangle ABC is isosceles.

Solution. Assume without loss of generality that ABC has area $1/2$, so $ah_A = bh_B = ch_C = 1$ (this can be achieved by rescaling). Then for part (a) note that

$$\begin{aligned}
(a + h_B)(b + h_C)(c + h_A) &= (a + 1/b)(b + 1/c)(c + 1/a) \\
&= abc + 1/abc + a + b + c + 1/a + 1/b + 1/c.
\end{aligned}$$

Now we show that the triples

$$\left(\sqrt{a+h_B}, \sqrt{b+h_C}, \sqrt{c+h_A}\right) \quad \text{and} \quad \left(\sqrt{a+h_C}, \sqrt{b+h_A}, \sqrt{c+h_B}\right)$$

are permutations of one another (part (b)). We already saw they have the same product. But the triples also have the same sum, and sum of squares. It follows, say by Vieta formula, that they are the roots of the same polynomial, hence are a permutation of each other.

Now we prove (c). If any altitudes or side lengths are equal, we're done. So the only case left to consider is if

$$a+h_B = b+h_A, \qquad b+h_C = c+h_B, \qquad c+h_A = a+h_C$$

and the symmetric case. In this situation, we get $a-1/a = b-1/b = c-1/c$, but in that case, the triangle is isosceles. $\square$

Advanced Math Problems

Problem M1. For every positive integer n set

$$a_n = 1^{-2} + 2^{-2} + \cdots + n^{-2}.$$

(a) Prove that the infinite sum

$$\sum_{n \geq 2} \frac{1}{n^2 a_n a_{n-1}}$$

is convergent.

(b) Determine its value.

Solution. This is a telescoping sum: the partial sums are equal to

$$\sum_{n=2}^{N} \frac{1}{n^2 a_n a_{n-1}} = \sum_{n=2}^{N} \left(\frac{1}{a_{n-1}} - \frac{1}{a_n} \right) = \frac{1}{a_1} - \frac{1}{a_N}.$$

Note that $\lim_N a_N = \frac{\pi^2}{6}$. So convergence is immediate and we obtain the answer $1 - \frac{6}{\pi^2}$. □

Problem M2. Let $n \geq 1$ be a positive integer. An $n \times n$ matrix M is generated as follows: for each $1 \leq i, j \leq n$, we randomly and independently write either i or j in the (i,j)th entry, each with probability $\frac{1}{2}$. Let E_n be the expected value of $\det M$.

(a) Compute E_2.

(b) For which values of n do we have $E_n \geq 0$?

Solution. For (a), the answer is $-1/4$, and for (b) the answer is all $n \neq 2$.

In fact we claim that

$$E_n = \begin{cases} 1 & n = 1 \\ -1/4 & n = 2 \\ 0 & n \geq 3. \end{cases}$$

The determinant can be written as

$$\det M = \sum_{\pi} \prod_{i=1}^{n} \left(\frac{i + \pi(i)}{2} + \frac{i - \pi(i)}{2} \cdot \varepsilon_{i,\pi(i)} \right)$$

over permutations π on $\{1,\dots,n\}$, where $\varepsilon_{i,j} \in \{-1,1\}$ at random. If we imagine expanding the sum, then we at once see that any term with an $\varepsilon_{i,j}$ has expected value zero, since all the $\varepsilon_{i,j}$'s are independent.

Thus by linearity of expectation we can drop all the $\varepsilon_{i,j}$'s, and obtain

$$\mathbb{E}[\det M] = \sum_{\pi} \prod_{i=1}^{n} \left(\frac{i+\pi(i)}{2} \right)$$

This is the determinant of the matrix N whose (i,j)th entry is $\frac{1}{2}(i+j)$.

Having thus reduced the problem we remark that

- For $n=1$, we have $N = \begin{bmatrix} 1 \end{bmatrix}$ so $\det N = 1$.
- For $n=2$, we have $N = \begin{bmatrix} 1 & 3/2 \\ 3/2 & 2 \end{bmatrix}$ so $\det N = -\frac{1}{4}$.
- For $n \geq 3$, the second row is the average of the first and third, and consequently the determinant vanishes.

$\square$

Problem M3. Let $k \geq 1$ be a positive integer. Find an $\varepsilon > 0$ in terms of k, as large as you can, such that the following statement is true.

> Consider a family $\mathscr{F}$ of subsets of $S = \{1,2,\dots,N\}$, where $N > k$ is an integer. If for any $T \subseteq S$ with $1 \leq |T| \leq k$ we have
>
> $$\frac{1}{2} - \varepsilon < \frac{|\{X \in \mathscr{F} : |X \cap T| \text{ is odd}\}|}{|\mathscr{F}|} < \frac{1}{2} + \varepsilon,$$
>
> then some set in $\mathscr{F}$ contains none of $\{1,\dots,k\}$.

Solution. Taking $\varepsilon = \frac{1}{2^{k+1}-2}$ is sufficient.

It is more natural to phrase the problem in terms of binary strings. To this end, we consider each set in $\mathscr{F}$ as a binary string of length N (with 1 indicating membership as usual).

We will prove that for any k indices $i_1, \dots, i_k$, there is some string in $\mathscr{F}$ with all zeros in these positions.

Let A be the set of all length k binary strings. For $a \in A$, let $p(a)$ be the proportion of strings $x \in \mathscr{F}$ with $x_{i_1} \dots x_{i_k} = a$. Then for every string $w \neq 0$ of length k, we have

$$\sum_{\substack{a \in A \\ w \cdot a = 0}} p(a) \geq \frac{1}{2} - \varepsilon$$

where $w \cdot a$ is dot product modulo 2. Now we sum this over all $2^k - 1$ choices of w to get

$$\sum_{w \neq 0} \sum_{\substack{a \in A \\ w \cdot a = 0}} p(a) \geq (2^k - 1)\left(\frac{1}{2} - \varepsilon\right).$$

We can add in the $w = 0$ term to get

$$\sum_{w} \sum_{\substack{a \in A \\ w \cdot a = 0}} p(a) \geq 1 + (2^k - 1)\left(\frac{1}{2} - \varepsilon\right).$$

Now note that a given string $\mathbf{0} \neq a \in A$ appears as a $p(a)$ for half the choices of w, i.e. it appears 2^{k-1} times. On the other hand the all-zero string $a = \mathbf{0}$ appears every time. So we deduce that

$$2^k p(\mathbf{0}) + \sum_{0 \neq a \in A} 2^{k-1} p(a) \geq 1 + (2^k - 1)\left(\frac{1}{2} - \varepsilon\right)$$

or

$$2^k p(\mathbf{0}) + 2^{k-1}\left(1 - p(\mathbf{0})\right) \geq \left(\frac{1}{2} + \varepsilon\right) + 2^k\left(\frac{1}{2} - \varepsilon\right).$$

This implies:

$$p(\mathbf{0}) \geq 2^{-k} - \frac{2^k - 1}{2^{k-1}}\varepsilon = 2^{-k} - 2\left(1 - 2^{-k}\right)\varepsilon.$$

So we are done as long as

$$\varepsilon < \frac{2^{-k}}{2(1 - 2^{-k})} = \frac{1}{2^{k+1} - 2}$$

since in that case $p(\mathbf{0}) > 0$. But this follows from the choice of ε. $\square$

Problem M4. There are $n \geq 3$ married couples attending a daily couples therapy group. Each attendee is assigned to one of two round tables, so that no one sits at the same table with his/her spouse. The order of seating at each table remains fixed once and for all.

Initially, s of the attendees have contracted a contagious disease. For each person P, consider P's two neighbors at the table, as well as P's spouse (who, of course, sits at the other table). Each day, if at least two of these three people are sick, then P gets sick too, and remains sick forever.

Eventually everyone gets sick. Across all possible seating arrangements, what is the smallest possible value of s?

Solution. The answer is $s = \left\lceil \frac{1}{2}(n+1) \right\rceil$.

Let G be the 3-regular graph connecting P to P's neighbors and spouse. At any point, consider the quantity

$$X = \#\text{sick} + \#\left\{\text{edges } e \in G \text{ touching exactly one sick person}\right\}.$$

Observe that when a person catches the disease, the quantity X either decreases or remains the same. Moreover, immediately before the final person gets sick, we have $X = 2n+2$. On the other hand, we initially have $X \le 4s$. Thus $s \ge \left\lceil \frac{1}{2}(n+1) \right\rceil$.

In fact, this is sharp. For n even, we can start with every other person at the first table sick, and any (one) person at the second table sick. For $n = 2k+1$ odd, a similar construction works if we make persons in positions $1, 3, \dots, 2k-1$ at the first table sick as well as spouse of the person in position $2k+1$. □

Problem M5. Consider an $n \times n$ binary matrix T (all entries are either 0 or 1). Assume at most $0.01n^2$ of the entries of T are zero.

(a) Find a constant $c > 0$, as large as you can, such that: for integers $m \ge 2$, the trace of T^m is at least $(cn)^m$.

(b) Does any $c < 1$ work?

Solution. We first address part (a) and prove $c = 0.7$ works. An index $1 \le j \le n$ is called *big* if both the jth row or jth column have fewer than $0.1n$ zeros. Note that at most $0.2n$ of the j's are not big.

We now claim that

> Let $m \ge 2$. If i and j both big, the (i,j)th entry of T^m is at least $(0.7n)^{m-1}$.

This follows by induction in the following way. For $m = 2$ the entry is actually at least $0.8n$, since the ith row and jth column altogether have at most $0.2n$ zeros. For the inductive step, we have

$$(T^{m+1})_{ij} = \sum_k (T^m)_{ik} \cdot T_{kj}$$

and we note that there are at most $0.1n$ indices j with $T_{kj} = 0$, and $0.2n$ indices j which are not big. For the remaining $0.7n$ indices j we may use the induction hypothesis.

To complete part (a), note that the trace of T^m is at least $0.8n \cdot (0.7n)^{m-1}$ by the induction hypothesis.

The answer to part (b) is negative. We show $c = 0.99$ fails. Suppose T is a 10×10 matrix such that the $(1, 2)$th entry of T is zero and all other entries are 1. A straightforward calculation shows that the characteristic polynomial of this matrix T is equal to

$$-\lambda^{10} + 10\lambda^9 - \lambda^8$$

which has nonzero roots $5 \pm 2\sqrt{6}$. Thus $\operatorname{Tr} T^m = (5 + 2\sqrt{6})^m + (5 - 2\sqrt{6})^m$. Since $5 + 2\sqrt{6} \approx 9.8990 < 0.99 \cdot 10$, this resolves the problem. $\square$

Problem M6. You're given a simple graph G with n vertices, and want to develop an algorithm which either finds a 4-cycle or proves that none exists. For example, a naïve algorithm taking $O(n^4)$ runtime would be to simply brute-force search all $\binom{n}{4}$ possible sets of four vertices.

(a) Exhibit an algorithm with the best runtime you can find.

(b) Give the best lower bounds you can on the runtime of any such algorithm.

Solution. We provide an algorithm which solves the problem in $O(n^2)$ time. This is asymptotically optimal, because it is linear in the input size.

Impose an arbitrary ordering of the vertices, so that we may refer to them as $1, \dots, n$. Looping through vertices v, we record all pairs (w_1, w_2) of vertices such that vw_1 and vw_2 are edges, and $w_1 < w_2$. We stop if we ever see any pair twice; note that this is equivalent to a 4-cycle.

The key insight is that there are only $\binom{n}{2}$ possible pairs, so this algorithm terminates in at most $\binom{n}{2}$ steps! (A priori one might expect this algorithm to take $O(n^3)$ time, but this is not so.) $\square$

Open-ended problem. Research in mathematics is different from problem solving. One of the main differences is that you can ask your own questions. This problem is unusual. You do not need to solve anything, you have to invent questions and show your vision.

We start with a famous problem:

> You have a segment $[0, 1]$. You choose two points on this segment at random. They divide the segment into three smaller segments. What is the probability that the three smaller segments can be the sides of a triangle?

Imagine that you are a mentor and want to create a research project related to this problem. Your task is to invent up to five questions based on this problem.

Let us give you an example:

Question. The problem is equivalent to asking the probability that none of the three smaller segments exceed $1/2$ in length. A new question might be: Given x, what is the probability that none of the three smaller segments exceed x in length?

This question is a rather trivial generalization, which has an easy solution. Your goal is to invent something more challenging and stimulating. This problem will be judged separately based on how interesting your questions are.

To test how good your question is, it's a good idea to think about it a little bit, consider the simplest examples, etc. If this leads you to solving the question completely, then the question was too easy, and you should think how to make it harder! Please discuss the insight gained this way for each question.