Quant Questions

Amruth Ayaan Gulawani

May 13, 2025

Contents

1 Introduction

This PDF contains a compilation of pretty much all the quant questions asked across different firms. Feel free to contact if you possess a better solution. The questions are split into three categories of difficulty and span over a wide area, namely; Combinatorics, Probability, Stochastic calculus and processes, Statistics and Regressions, Game Theory, Limit Theorems, Differential Equations, and many more!

An honest recommendation :

1. Spend 5-10 minutes on the easy questions
2. Spend 10-15 minutes on the medium questions
3. Spend a good half hour on the hard questions

Since the questions are diverse, I also recommend going through some old notes from your previous semesters to avoid losing your mind :D. I have also included the Code/Algorithmic solution repository to selected problems which require simulation/brute forcing. See here.

You'll also find a couple of insightful findings and links to materials, at real random intervals ;D

Some solutions have been inspired from the Mathematics Stack Exchange (MSE) community answers and Stackoverflow community.

Here is my MSE profile

2 Easy Questions

1. You and your friend are playing a game where you each pick an integer from 1 to 10 inclusive and add that to a running total. The person that states 50 is the winner. You get to go first, what number do you say first?

   Solution : Pick 6 as your first number, as 10+1=11 and 50-11=39, so you can always split your turns to win in two turns as a complement of the current running total. Hence the generic sequence would follow 6-17-28-39-50.

2. Let $x_0 = 0$ and let $x_1, \ldots, x_{10}$ satisfy that $\|x_i - x_{i-1}\| = 1$ for $1 \leq i \leq 10$, and $x_{10} = 4$. How many such sequences are there satisfying these conditions?

   Solution : Take a coin: take heads as $+1$ (or moving up) and tails as -1 (or moving down). So after 10 tosses, as we end up 4 positions upwards, we can say that our coin tosses resulted in 7 up tosses and 3 down tosses. This implies how we choose our down tosses matters: $\binom{10}{3} = 120$ ways. (OR) we can use a decision tree and brute force it and find a sequence of growth at 120 for 10 iterations.

3. Let x and y be vectors in R^n with angle $75°$ between them. Let A be an orthogonal $n \times n$ matrix. Find the angle between Ax and Ay.

   Solution : Since the scalar product and the norm do not change after multiplication with an orthonormal matrix the angle remains preserved. Observe the inner product as $x^T x$, hence

   $$(Ax)^T(Ay) = x^T A^T Ay$$

   and in orthogonal matrices $A^T A = AA^T = I_n$, hence inner product preserved and norm$(Ax) = \sqrt{(Ax)^T(Ax)} = \sqrt{x^T x} =$ norm(x), hence the norm also remains preserved. Apply inner product cosine rule. Hence answer is $75°$.

4. How many values of n are there such that a regular $n$gon (e.g $n = 5$ is a pentagon) has interior angles that (in degrees) are integer-valued (no decimals)?

   Solution : Since the interior angle of a regular polygon is given by $\frac{180\times(n-2)}{n} \rightarrow 180 - \frac{360}{n}$, find number of factors of 360 and apply $n \geq 3$, you'll get $24 - 2 = 22$ factors

5. The correlation between X and Y is 0.56. What is the correlation between $8X$ and $Y + 7$?

   Solution :

   $$\rho_{aX+b,\,cY+d} = \frac{\text{Cov}(aX+b,\,cY+d)}{\sigma_{aX+b}\cdot\sigma_{cY+d}} = \frac{ac\cdot\text{Cov}(X,Y)}{|a|\cdot\sigma_X\cdot|c|\cdot\sigma_Y} = \frac{ac}{|a||c|}\cdot\rho_{X,Y} = \text{sign}(a)\cdot\text{sign}(c)\cdot\rho_{X,Y}$$

   Therefore, on plugging in a, b, c, d in above correlation formula, we obtain that the correlation remains unaffected by the linear transformation and we can obtain the same correlation of 0.56.

6. What is the total number of paths from $(0, 0, 0)$ to $(3, 4, 5)$ in 3D space, given that we can move one unit to the right, forward, or up at each step?

   Solution : The total amounts of rights/forwards/ups you make is 3+4+5=12, so now the total number of paths becomes $\frac{12!}{3!4!5!} = 27720$.

7. You are playing a 2D game where your character is trapped within a 6×6 grid. Your character starts at $(1, 1)$ and can only move up and right. How many possible paths can your character take to get to $(6, 6)$?

   Solution : The same as problem 6, $\frac{12!}{6!6!} = 924$.

8. Two distinct prime integers between 1 and 20, inclusive, are selected uniformly at random. Find the probability their sum is even.

   Solution : Just our sample space is tiny, list out all the primes := 2,3,5,7,11,13,17,19

So, now total probability is $\binom{8}{2}$ and there is only one case of picking out two odds to make an even, $\binom{7}{2}$, hence its as simple as doing $\frac{\binom{7}{2}}{\binom{8}{2}} = \frac{21}{28}$.

Some useful Fibonacci Sequence properties:

(a) $\sum_{i=1}^{n} F_i = F_{n+2} - 1$
(b) $\sum_{i=0}^{n-1} F_{2i+1} = F_{2n}$
(c) $\sum_{i=1}^{n} F_{2i} = F_{2n-1} - 1$
(d) $\sum_{i=1}^{n} F_i^2 = F_n F_{n+1}$
(e) $F_n = F_{n-1} + F_{n-2}$
(f) $\lim_{n\to\infty} \frac{F_{n+m}}{F_n} = \phi^m$

9. Let F_n be the nth Fibonacci number. Compute $\frac{F_1+F_2+F_3+\cdots+F_{300}}{F_3+F_6+F_9+\cdots+F_{300}}$.

Solution : Rewrite the numerator using the recurrence relation of Fib Sequence, to get the value 2

10. There is a game where an integer is picked uniformly between 2 and 10 inclusive. You win n if the integer is prime and lose $\frac{n}{2}$ if its composite (where n is the integer picked). How much would you pay to play this game? If you don't wish to play, enter 0.

Solution : Now this was a really tricky question! Simply list out the primes and composites as P=2,3,5,7 and C=4,6,8,9,10, Notice there are 9 numbers in total, each coming off an uniform distribution of $\frac{1}{9}$, hence the total gain you make by getting a prime is $(2+3+5+7)/9$ and total loss you make by getting a composite is $(4+6+8+9+10)/9$, Notice that you lose more than you make. So should you pay to play this game? No! hence you ought to pay none or 0.

11. Instead of the *Moment Generating Function* (MGF), for non-negative integer-valued discrete random variables X, we can define the *Probability Generating Function* (PGF), which is given by

$$p(z) = \mathbb{E}[z^X] \quad \text{for } |z| \leq 1.$$

Find the PGF of $X \sim \text{Geom}(p)$. Evaluate this function for $z = \frac{1}{3}$ and $p = \frac{2}{3}$.

Solution : Use LOTUS rule;

If X is a **discrete** random variable whose pmf is $f_X(x)$, then

$$\mathbb{E}[g(X)] = \sum_x g(x) f_X(x).$$

If X is a **continuous** random variable whose pdf is $f_X(x)$, then

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx.$$

The answer should turn out to be 0.25.

12. Let $X, Y \in unif(0,1)$ IID. Find $P\left[XY > \frac{1}{2}\right]$

Solution : We are given $X, Y \sim \text{Unif}(0,1)$ i.i.d. We want to compute:

$$\mathbb{P}(XY > \tfrac{1}{2}) = \iint_{xy>1/2} dx\, dy$$

Since the pair (X, Y) lies in the unit square $[0,1]^2$, we consider:

$$= \int_{1/2}^{1} \int_{1/(2y)}^{1} dx\, dy = \int_{1/2}^{1} \left(1 - \frac{1}{2y}\right) dy = \int_{1/2}^{1} 1\, dy - \int_{1/2}^{1} \frac{1}{2y} dy \approx 0.15$$

13. A computer is interpreting an 8bit binary string, which consists of 8 characters that are either 1 or 0. How many such strings begin with 1 or end with the two characters 00?

Solution : Let S be the event the string starts with 1 and O be the event it ends with 00. We want $|S \cup O| = |S| + |O| - |S \cap O|$. We will end up with 128+64-32=160 strings.

3 Medium Questions

1. Find two real numbers x and y composed of only 1s in their decimal expansion such that $xy = x + y$. What is xy?

   Solution : Rewrite it as $x = \frac{y}{y-1}$, then use trial and error to get $y = 11$, $x = \frac{11}{10} \Rightarrow xy = \frac{121}{10}$.

2. Suppose you start with the 2×2 identity matrix I_2. At each step, select one of the four elements of the matrix uniformly at random. If the element you select is a 1, change it to 0. If the element you select is a 0, change it to 1. Find the expected number of steps needed to obtain a singular matrix (i.e. the determinant is 0).

   Solution : We begin with the 2×2 identity matrix:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

At each step, we randomly select one of the four elements. If the selected element is a 1, we flip it to 0; if it is a 0, we flip it to 1. The process continues until we reach a **singular matrix** (i.e., a matrix with determinant zero).

Our goal is to compute the expected number of steps required to reach a singular matrix starting from the identity matrix.

Each matrix is a 2×2 binary matrix. Since each entry is either 0 or 1, the total number of possible matrices is:

$$2^4 = 16$$

Each matrix can be represented by a 4-bit binary string corresponding to the flattened form of the matrix. We will now analyze the states and the transitions between them.

A matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is singular if its determinant is zero:

$$\det(A) = ad - bc = 0$$

By checking all 16 matrices, we find that the following matrices are singular:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Thus, there are 10 singular matrices and 6 non-singular matrices.

We will model the matrix-flipping process as a Markov chain, where the states correspond to the different binary matrices. The transitions occur by flipping one of the four entries in the matrix.

For each non-singular matrix, there are 4 possible state transitions corresponding to flipping each of the 4 elements. For each transition, the probability of moving to a different state is $\frac{1}{4}$. If the matrix is singular, we remain in that state because it is absorbing.

Let us define the set of non-singular states. Since there are 6 non-singular matrices, we will construct the transition matrix P for these 6 matrices. These matrices will transition with probability $\frac{1}{4}$ between each other and will not transition to singular matrices.

Let the non-singular matrices be represented by states 1 to 6, and the singular matrices by states 7 to 16. The transition matrix P is partitioned as follows:

$$P = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix}$$

Where: - Q is the 6×6 matrix of transitions between the non-singular states. - R is the 6×10 matrix of transitions from non-singular to singular states (which will be zero, because non-singular states do not transition directly to singular ones). - I is a 10×10 identity matrix representing the absorbing singular states.

We now calculate the **Fundamental Matrix** N, which is given by:

$$N = (I - Q)^{-1}$$

Where I is the identity matrix of size 6×6 and Q is the submatrix of P representing the transitions between the non-singular states.

After finding the fundamental matrix N, the expected number of steps to absorption from any of the non-singular states can be found by:

$$\mathbf{t} = N \cdot \mathbf{1}$$

where $\mathbf{1}$ is a column vector of ones. The result gives us the expected number of steps to reach a singular matrix starting from any non-singular state.

The expected number of steps to reach a singular matrix from the identity matrix (the starting non-singular state) is calculated to be approximately:

$$\boxed{1.714}$$

We have modeled the process as a Markov chain and used the concept of the fundamental matrix to compute the expected number of steps needed to reach a singular matrix from the identity matrix. The expected value is approximately 1.714 steps.

3. You are given an undirected graph with 10 nodes. From every node, you are able to access any other node (including itself), all with an equal probability of $\frac{1}{10}$. What is the expected number of steps to reach all nodes at least once (rounded to the nearest step)?

   Solution : This is the classic Coupon Collector's Problem. Therefore, for 10 nodes, we are required to calculate the first 10 harmonic terms,

   $E[x] = \underset{n \to 10}{nH_n} = \lfloor 10 \sum_{k=1}^{10} \frac{1}{k} \rfloor = 29$

4. Calculate the sum of the digits from 1 to 1 million, inclusive? For example, the sum of the digits of 46 is 10.

   Solution : If we start looking at the sum of the digits of all numbers between 0 and 999,999, start by adding leading 0's to pad all numbers to 6 digits.

   Consider the first digit. If the first digit is a 1, there are 10 choices for each of the remaining 5 digits. So, there are 105 numbers in that range that have 1 as a first digit. The same can also be said of any other number as the first digit. So, if you add the first digit of all those numbers together, you get (1+2+3+4+5+6+7+8+9)×105=4,500,000. The same is true for the remaining digits. So, the sum of all the digits of all numbers between 0 and 999,999 is 4,500,000×6=27,000,000. We counted 0, which we didn't really want, but the sum of the digits of 0 is 0, so that doesn't affect our sum. That leaves just 1,000,000 which has a digit sum of 1, making our final total 27,000,001.

   Extra! A really cool problem on number theory

5. How would you calculate the number of digits in n^m or $n!$?

   Solution : In both cases, we apply the approximation as follows,

$$N_1 = \lfloor m \times \log_{10}(n) \rfloor + 1$$

$$N_2 = \lfloor \log_{10}(n!) \rfloor + 1 \iff log_{10}(n!) = \sum_{k=1}^{n} \log_{10}(k) \implies \lfloor \sum_{k=1}^{n} \log_{10}(k) \rfloor + 1$$

6. Two cylinders of radius 1 intersect at right angles to one another. Furthermore, their central axes also intersect. Find the volume of the enclosed region.

   Solution : This is a Steinmetz solid.

   For a bicylinder, the volume is $\frac{16r^3}{3} \underset{r=1}{\implies} \frac{16}{3}$

7. Evaluate $\sum_{k=0}^{\infty} \frac{F_k}{10^{k+1}}$ where F_k is kth Fibonacci number and $F_0 = 0$ and $F_1 = 1$

   Solution : You can either proceed with generating functions or solve the recurrence relation for the sequence and then plug it in to obtain geometric progressions in both cases. The result should be $\frac{1}{89}$

8. There is a 3×3 grid of light bulbs, where each of the light bulbs is turned on with probability $\frac{1}{2}$. Find the probability that no two adjacent light bulbs (grid cells that share a common side) are powered on.

   Solution : Consider the matrix representation of all the cases in which no two adjacent bulbs remain turned on,

   '1' is for bulb on and '0' is for bulb off,

$$\begin{bmatrix}1&0&1\\0&1&0\\1&0&1\end{bmatrix} \quad \begin{bmatrix}0&1&0\\1&0&1\\0&1&0\end{bmatrix} \quad \begin{bmatrix}1&0&0\\0&0&1\\1&0&0\end{bmatrix} \quad \begin{bmatrix}1&0&0\\0&0&1\\0&0&0\end{bmatrix} \quad \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$$

   The first matrix brings on a total of 2^5 combinations, second matrix brings on a total of 2^4 combinations, third matrix comes with 4 rotations of symmetry, fourth matrix comes with 8 rotations of symmetry and fifth matrix gives us 4 rotations of symmetry. This results in 32+16+4+4+8 = 64 matrices and we need to subtract 1 case which is the no bulbs are on, hence our solution is $\frac{63}{2^9} = \frac{63}{512}$

9. Let $A = [1, 2, ..., 2023, 1^2, 2^2, ..., 2023^2]$. What is the median of A?

   Solution : We are given the set:

$$A = \{1, 2, \ldots, 2023\} \cup \{1^2, 2^2, \ldots, 2023^2\}$$

   This means:

$$A = \{1, 2, 3, \ldots, 2023\} \cup \{1, 4, 9, \ldots, 4092929\}$$

   So A contains:

   - 2023 elements from the first set,
   - 2023 elements from the second set.

   Therefore, the total number of elements in A is:

$$2023 + 2023 = 4046$$

   Since the number of elements is even, the median is the average of the 2023rd and 2024th elements of the sorted multiset A.

   Define:

$$L_1 = \{1, 2, \ldots, 2023\}$$
$$L_2 = \{1, 4, 9, \ldots, 4092929\}$$

   We now define a function $f(x)$, which gives the number of elements in A less than or equal to x:

$$f(x) = \min(x, 2023) + \lfloor \sqrt{x} \rfloor$$

   Here:

- $\min(x, 2023)$: number of elements $\leq x$ from L_1,
- $\lfloor \sqrt{x} \rfloor$: number of perfect squares $\leq x$ from L_2.

We need to find:

$$a_{2023} = \min\{x \mid f(x) \geq 2023\}$$
$$a_{2024} = \min\{x \mid f(x) \geq 2024\}$$

Let us test $x = 1979$:

$$f(1979) = 1979 + \left\lfloor \sqrt{1979} \right\rfloor = 1979 + 44 = 2023$$

Now test $x = 1980$:

$$f(1980) = 1980 + \left\lfloor \sqrt{1980} \right\rfloor = 1980 + 44 = 2024$$

So,

$$a_{2023} = 1979, \quad a_{2024} = 1980$$

$$\text{Median} = \frac{a_{2023} + a_{2024}}{2} = \frac{1979 + 1980}{2} = \boxed{1979.5}$$

This answer was heavily inspired from a CompSci Algorithms textbook, I recommend checking out the algorithmic solution in the repo!

10. The sides and height of a triangle are four consecutive whole numbers. What is the area of the triangle?

    Solution : Use the Heron formula and plug in a-1,a,a+1 and compare with alternate formula for area of triangle and then trial error to find the sides and height as 12,13,14 and 15

11. If $X \in unif(0, \pi)$, find $\cos(E[X|\sin(X)]$

    Solution :

    $$E[X|A] = \int_x P[X = x|A]\,dx$$

    So, you'll need to figure out that $\sin(x)$ is injective in the given interval, thereby making $\cos(E[X|\sin(X)]) \to \cos(E[X])$. Apply the expectation formula for uniform probability distribution and you'll get 0.

12. Consider the 30×30 matrix A with $A_{ij} = 30$ for $1 \leq i \leq 30$ and $A_{ij} = -1$ for $i \neq j$. Let $\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}$ be the vector of distinct eigenvalues $\lambda_1 \neq \lambda_2$ of A and $g = \begin{bmatrix} g_1 \\ g_2 \end{bmatrix}$ be the vector of geometric multiplicities corresponding to λ_1 and λ_2, respectively. Find $||\lambda|| + ||g||^2$.

    Solution : A fact from Linear Algebra states that if M is $n \times n$ with eigenvalues (repeated according to multiplicity) $\{\lambda_i\}_{i=1}^n$, then the eigenvalues of $M + kI_n$ is $\{\lambda_i + k\}_{i=1}^n$. Note that we can write $A = 1_{30} + 291_{30}$, where 1_{30} is the 30×30 matrix of all ones.

    So, here's the detailed solution for the above problem,

    The matrix A can be expressed as:

    $$A = 29I + J$$

    where:

    - I is the 30×30 identity matrix.
    - J is the 30×30 matrix with all entries equal to 1.

    The matrix J has eigenvalues:

    - 30 with multiplicity 1 (corresponding to the eigenvector $\mathbf{e} = [1, 1, \ldots, 1]^T$).
    - 0 with multiplicity 29 (since $\text{rank}(J) = 1$, the nullity is 29).

Since $A = 29I + J$, its eigenvalues are:

$$\lambda_1 = 29 + 30 = 59 \quad \text{(multiplicity 1)},$$

$$\lambda_2 = 29 + 0 = 29 \quad \text{(multiplicity 29)}.$$

- For $\lambda_1 = 59$: The eigenspace is 1-dimensional, so $g_1 = 1$.
- For $\lambda_2 = 29$: The eigenspace is 29-dimensional, so $g_2 = 29$.

$$\lambda = \begin{bmatrix} 59 \\ 29 \end{bmatrix}, \quad g = \begin{bmatrix} 1 \\ 29 \end{bmatrix}.$$

$$\lambda + g = \begin{bmatrix} 59 + 1 \\ 29 + 29 \end{bmatrix} = \begin{bmatrix} 60 \\ 58 \end{bmatrix}.$$

$$\|\lambda + g\|^2 = 60^2 + 58^2 = 3600 + 3364 = 6964.$$

$$\boxed{6964}$$

4 Hard Questions

1. You generate a uniformly random number in the interval $(0, 1)$. You can generate additional random numbers as many times as you want for a fee of 0.02 per generation. This decision can be made with the information of all of the previous values that have been generated. Your payout is the maximum of all the numbers you generate. Under the optimal strategy, find your expected payout of this game.

   Solution :

   We sequentially draw numbers $X_i \sim \text{Uniform}(0, 1)$. Each sample costs \$0.02 (except the first). The goal is to maximize:

   $$\mathbb{E}[\max(X_1, \ldots, X_n)] - 0.02 \cdot (n - 1)$$

   Let the strategy be: continue sampling only if the current maximum is less than a threshold t, and stop otherwise.

   Let V denote the expected value when following the optimal policy. At each step:

   $$V = \max\left(x, \mathbb{E}[\max(x, X)] - 0.02\right), \quad X \sim \text{Uniform}(0, 1)$$

   The expectation becomes:

   $$\mathbb{E}[\max(x, X)] = x \cdot x + \int_x^1 y\,dy = x^2 + \frac{1 - x^2}{2} = \frac{1 + x^2}{2}$$

   $$V(x) = \max\left(x, \frac{1 + x^2}{2} - 0.02\right)$$

   Solve for x where:

   $$x = \frac{1 + x^2}{2} - 0.02 \Rightarrow x^2 - 2x + 1.04 = 0 \Rightarrow (x - 1)^2 = -0.04$$

   No real solution. Instead, we must find numerically the t such that:

   $$\mathbb{E}[\max(t, X)] = t + 0.02$$

   Using numerical iteration, we find the threshold $t \approx 0.82$, and the corresponding expected reward:

   $$\boxed{\mathbb{E}[\text{optimal reward}] \approx 0.82}$$

   Alternate explanation here

2. Evaluate $\sum_{k \in S} \frac{1}{k^3}$ where S is the set of all positive integers such that $\frac{1}{k}$ has a terminating decimal expansion.

   Solution : Best way is to trial error this, start by writing down a few positive integers, 1,2,3 and so on and observe their reciprocal's decimal expansions and that's when you'll notice a pattern that only numbers of the form $2^m 5^n$ have a terminating decimal expansion.

   So we're required to find $\sum_{k \in S} \frac{1}{k^3}$, where S contains all the numbers of the form $2^m 5^n$,

   $$\sum_{a=0}^{\infty} \sum_{b=0}^{\infty} \frac{1}{(2^a 5^b)^3} = \frac{250}{217}$$

   You're curious why this was categorized as hard, well it's identifying the set S, which a lot miss out on.

3. Consider all subsets of $S = \{1, 2, 3, \ldots, 30\}$ so that each pair of numbers in that subset are coprime. Find the subset $\Omega \in S$ satisfying the previous condition whose elements have the largest sum. What is the sum of all the elements in Ω?

   Solution : Just to be clear, coprime numbers, also known as relatively prime numbers, are two or more integers that have no common factors other than 1.

   Notice the phrase "each pair of numbers in that subset" are coprime, some may misinterpret it as making subsets of even sizes, but that's not the case.

   Off topic, but the total number of subsets for n elements is 2^n, so if you're planning on writing down all the subsets, you'll have to write 2^{30} subsets.

   And it does not end at just finding the candidate subsets, you're apparently expected to find that one particular subset whose elements have the largest sum, like who made this diabolical question?!

   How do you even go about this?!

   To ensure all of the elements are coprime, the easiest way to make a candidate set would be the collection of all prime numbers at most 30. However, we also want to include 1, as 1 is coprime to every positive integer. Therefore, a candidate set would be $\{1, 2, 3, 5, 7, 11, 13, 17, 19, 23, 29\}$. This set has sum 130. Can you do better by combining and substituting different values? Yeah maybe?

   Our candidate subset will definitely involve 1, so we can work our way backwards from the largest number, 30? but that will put a constraint of not being able to put larger even numbers in the subset, 29? works, $(1, 29)$, now put 23? wait, seems too boring.

   How about we improve our previously obtained subset? if we removed 2,3,5 and put 30, that would cause the obtained sum to go from 130-2-3-5+30=150, we brought an improvement of 20 points, our subset now is $\{1, 7, 11, 13, 17, 19, 23, 29, 30\}$.

   Ok, I'll stop here, and restart this process,

   Back to the original problem I started with the set with only primes, $\{1, 2, 3, 5, 7, 11, 13, 17, 19, 23, 29\}$.

   Now I needed to remove and add elements to improve the sum from 130 and upwards. So I started with putting in 30 and remove 2,3,5 to obtain the improved sum of 150; $\{1, 7, 11, 13, 17, 19, 23, 29, 30\}$. Now I'm restricted to add any even numbers and also any number of the form 3^k and 5^k. This implies 30 was a bad choice, so back to our original subset and sum, but this time I remove 2 and 5 and add 20, $\{1, 3, 7, 11, 13, 17, 19, 20, 23, 29\}$,

   Ah yes, if I removed 3 and put 27, that does improve the sum, $\{1, 7, 11, 13, 17, 19, 20, 23, 27, 29\}$,

   Also, I can remove 7 and put in 28 but removed 20, $\{1, 11, 13, 15, 17, 19, 23, 27, 28, 29\}$,

   Also, we can put 25 instead of 15, and we end on the candidate subset, $\{1, 11, 13, 17, 19, 23, 25, 27, 28, 29\}$, sum = 193.

4. You and your opponent each draw a number from $U(0, 1)$ without revealing it. You each have the option to redraw and keep that value instead, and the other person doesn't know which choice they made. The person with the higher number wins. The optimal strategy is of the form reroll if the number is less than k. Find k to 3 decimal places.

   Solution : It seems obvious that any optimal strategy for a player is to redraw if the first round is less than some value t and not redraw if the first round is greater than t.

   If you do this then the density and cdf of your result after making your choice,

   when $0 \le x \le t$, has $f_t(x) = t$ and $F_t(x) = tx$, as you must have had drawn less than t in the first draw and drew again

   when $t \le x \le 1$, has $f_t(x) = t + 1$ and $F_t(x) = tx + x - t$, as you may have drawn less than t in the first draw and drawn again or you may have drawn more than t in the first draw and not drawn again

   Suppose Player A adopts this strategy with cutoff t and Player B adopts this strategy with cutoff s. Then the probability of Player B winning is

$$\int_0^1 F_t(x) f_s(x)\, dx$$

which when $0 < t < s < 1$ is $\int_0^t (tx)s\,dx + \int_t^s (tx + x - t)s\,dx + \int_s^1 (tx + x - t)(s+1)\,dx$ and comes out at $\frac{1}{2}(1 - t + s + st + st^2 - s^2 - s^2 t)$ and is maximised when $s = \frac{t^2+t+1}{2t+2}$ so long as that is between t and 1 and when $0 < s < t < 1$ is $\int_0^s (tx)s\,dx + \int_s^t (tx)(s+1)\,dx + \int_s^1 (tx + x - t)(s+1)\,dx$ and comes out at $\frac{1}{2}(1 - t + t^2 + s - st + st^2 - s^2 t)$ and is maximised when $s = \frac{t^2-t+1}{2t}$ so long as that is between 0 and t.

An unbeatable strategy comes when Player A uses a cutoff where $\frac{t^2+t+1}{2t+2} = t = \frac{t^2-t+1}{2t}$, i.e. when $t = \frac{\sqrt{5}-1}{2} \approx 0.618$.

5. There are 100 noodles in your bowl of ramen. You take the ends of two noodles uniformly at random and connect the two, putting the connected noodle back into the bowl and continuing until there are no ends left to connect. On average, how many circles will you create? Round to the nearest whole number.

   Solution : Okay crazy! Start with base case of 1 noodle, you have two ends join them and you get a circle, the expected number of circles = 1. Now take 2 noodles, and you get 4 endpoints namely (a,b,c,d) and now these four can be paired up in taking 2 from 4 choices which gives us 6 choices and let (a,b) and (c,d) be the endpoints of the two noodles. Then, we can take both form a circle individually, after consecutive steps (being swapped), or they can form one large circle if a-c and b-d and then join them as a whole.

   Alright, that sets up the expected number of circles as $\frac{1}{6} \times 2 + \frac{1}{5} \times 2 + \frac{4}{6} \times 1 = 1 + \frac{1}{3}$.

   You can then inductively prove that, $E_n = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1}$, we were required to find $\lfloor E_{100} \rfloor = 3$

6. You have a bag of 90 beans which all weigh the same, except for one that is slightly heavier or lighter than the others. What is the minimum number of times a balance scale must be used to find the abnormal bean?

   Solution : With each weighing, a balance scale gives you 3 possible outcomes:

   Left = Right $\rightarrow$ both sides normal

   Left > Right $\rightarrow$ either left has heavy or right has light

   Left < Right $\rightarrow$ either left has light or right has heavy

   Off topic, my dumb mind thought of this "Can it not be as straight forward as arrange those 90 beans in a row and best case is the bad bean is 3rd so after first two and measured, then when you remove one bean put the 3rd you see the scale deflect, so thats like 2 times you used the scale? isnt this minimum?", why this thought is absurdly wrong is an exercise left to the reader (aha, I always wanted to do this!)

   Back to our problem, we obviously need to group some beans and then identify which group is defective and then apply brute force on that group, very similar to divide and conquer stratergy.

   If you have N beans, one of which is faulty, and that fault can be either heavier or lighter. there are 2 possibilities: 2N

   Also, our ternary tree provides us with 3^k options, so we are required to minimise our TST traversals, as $3^k \geq 2N \implies 3^k \geq 180 \implies k = 5$

   It will take 5 tries to find the defective bean.

7. Let C_n count the number of cycles in a random permutation of $\{1, 2, \ldots, 2n\}$ that are larger than n in length. Compute $\lim_{n\to\infty} E[C_n]$.

   Solution : This is an interesting problem!

   I found a theorem stating that the expected number of cycles in a random permutation in a set of n-elements is the nth harmonic sum.

   We are only interested in cycles of length greater than n, so the formula should be:

$$E[C_n] = \sum_{k=n+1}^{2n} \frac{1}{k}$$

There can only be at most 1 cycle of length $k \geq n+1$ in a permutation of $2n$ items. Thus the expected number of such cycles is the probability that it occurs, which is the number of permutations which have such a cycle, divided by $(2n)!$.

As no permutation can have have a cycle of length k_1 and one of length k_2 where $k_1, k_2 \geq n+1$, we can just sum the number of permutations containing a cycle of length k from $n+1$ to $2n$, without worrying about overlaps.

For a given k, there are $\binom{2n}{k}$ choices of items to go in the cycle. Once these items are fixed, there are $(k-1)!$ possible cycles (pick an item, the cycle can take it to one of $k-1$ items, which can then be taken to one of $(k-2)$ items and so on ....). Finally, there are $(2n-k)!$ possible permutations for the remaining $2n-k$ items.

Thus the number of permutations containing a cycle of length k is:

$$\binom{2n}{k} \times (k-1)! \times (2n-k)! = \frac{(2n)!}{k}.$$

Summing from $k = n+1$ to $2n$ and dividing by $(2n)!$ gets us to;

$$E[C_n] = \sum_{k=n+1}^{2n} \frac{1}{k}.$$

Now we may rewrite this as

$$E[C_n] = \sum_{k=n+1}^{2n} \frac{1}{k/n} \frac{1}{n} = \sum_{x=1+\delta x,\ x/\delta x \in \mathbb{N}}^{2} \frac{1}{x} \delta x,$$

where $x = k/n$ and $\delta x = \frac{1}{n}$.

Thus:

$$\lim_{n\to\infty} E[C_n] = \int_1^2 \frac{1}{x} dx = \log(2) \approx 0.69314718056$$

8. Suppose that $X, Y \sim \text{Unif}(0,1)$ i.i.d. Compute the probability that $\left\lceil \frac{Y}{X} \right\rceil$ is a perfect square.

Solution :

The below is derived from the bivariate transformation.

The distribution of $Z = \frac{Y}{X}$ is given by:

$$f_Z(z) = \begin{cases} \frac{1}{2} & \text{if } 0 \leq z \leq 1, \\ \frac{1}{2z^2} & \text{if } z > 1, \\ 0 & \text{otherwise.} \end{cases}$$

The probability that $\left\lfloor \frac{Y}{X} \right\rfloor$ is a perfect square is:

$$\int_0^1 \frac{1}{2}\, dz + \sum_{n=1}^{\infty} \int_{n^2}^{n^2+1} \frac{1}{2z^2}\, dz = \frac{1}{2} + \frac{1}{2}\left(\sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{n^2+1} \right)$$

$$= \frac{1}{2} + \frac{1}{2}\left(\frac{\pi^2}{6} - \frac{\pi}{2}\coth(\pi) + \frac{1}{2} \right) = \frac{3}{4} + \frac{\pi^2}{12} - \frac{\pi}{4}\coth(\pi) \approx \boxed{0.78418}$$

Similarly, the probability that $\left\lceil \frac{Y}{X} \right\rceil$ is a perfect square is:

$$\int_0^1 \frac{1}{2}\, dz + \sum_{n=2}^{\infty} \int_{n^2-1}^{n^2} \frac{1}{2z^2}\, dz = \frac{1}{2} + \frac{1}{2}\left(\sum_{n=2}^{\infty} \frac{1}{n^2-1} - \sum_{n=2}^{\infty} \frac{1}{n^2} \right)$$

$$= \frac{1}{2} + \frac{1}{2}\left(\frac{3}{4} + \frac{1}{6}(6 - \pi^2)\right) = \frac{11}{8} - \frac{\pi^2}{12} \approx \boxed{0.55253}$$

9. Consider the set of 10 consecutive integers $\{1, 2, \ldots, 10\}$. How many subsets contain exactly 1 pair of consecutive integers?

   Solution : Very good question!

   I have the bit-manipulation approach,

   There are $2^{10} = 1024$ subsets.

   For each subset (represented by a 10-bit number), check how many consecutive pairs it has (check all pairs i and i+1), if it has exactly one pair, count that.

   Do check this out for more such questions!