2024 Alibaba Global Mathematics Competition

Question 1

A group of students took a trip to a city for vacation. There are 6 towers in the city positioned at A, B, C, D, E, F , respectively . Following a period of unrestricted mobility, every student found that he/she can only see the towers at A, B, C, D , and cannot see the towers at E and F.

(1) The positions of the students and the towers are regarded as points on the same plane, and no two of them are coincident.

(2) Any three points of A, B, C, D, E, F are not collinear.

(3) The only possibility to not see a tower is that the view is blocked by other towers. For example, if the position of a student, P, satisfies that P, A, B are collinear, and A lies on segment PB, then the student cannot see the tower at B.

Question: What is the maximum possible number of students in the group?

(A) 3 (B) 4 (C) 6 (D) 12

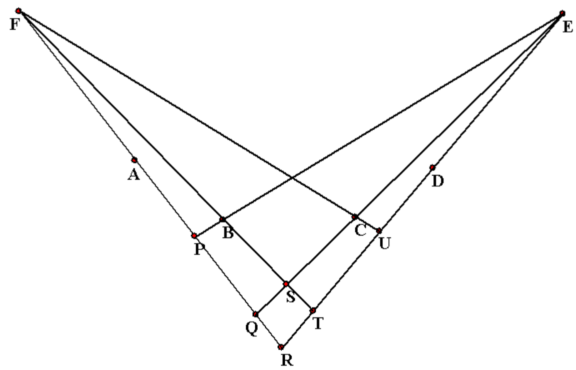
Answer 1

(C). Since no three positions of the towers are collinear, for any student, the view of E and F must be blocked by two different towers.

Choose 2 towers of the first four(Suppose they are the towers at A and B). We suppose the view of a student is blocked by these two towers, then the position of this student is either the intersection of extensions of EA and FB, or the intersection of extensions of FA and EB.

However, if the extensions of EA and FB intersectthen $AEFB$ is a convex quadrilateral, which means the intersection of EB and FA lies in these two segments. So, there are at most 1 student whose view is blocked by the towers at A and B.

Since there are 6 ways to choose 2 towers among the first four, the number of students is not greater than 6. The following figure is an example of 6 students (P, Q, R, S, T, U are the positions of the students).


F
E
A
D
B
C
U
P
S
Q
T
R

Question 2

In a fighter aircraft game, a player has 2 initial points to start with. In the game, the points are deducted continuously and linearly in time (at the rate of 1 point per unit time). When the game starts, after every independent exp(1) distributed time[1], an enemy plane appears on the screen. The player immediately takes action and can shoot it down instantly, or be shot down instantly. For each enemy plane the player shoots down, he/she can get 1.5 points, and has the option to choose to end the game immediately, or to wait until the next enemy plane to appear (but cannot quit in between). The process repeats. The level of difficulty increases with time: for the nth enemy plane that appears, the probability of shooting it down is $(0.85)^n$, and the probability of being shot down is $1 - (0.85)^n$, independently over time. Also, the game is automatically over whenever the player is shot down or all the points have been consumed, whichever comes first.

Question part:

(1) The player decides to finish the game after he/she shoots down the k-th enemy plane (if he/she is still in the game then). Suppose the player can keep all the points he/she accumulated at the end of the game, then in order to maximise his/her expected number points at the end of the game, which of the following is the optimal k?

 (A) 1.

 (B) 2.

 (C) 3.

 (D) 4.

(2) Suppose the player loses all the points accumulated if he/she is shot down. Then there is still an optimal time for him/her to stop the game in order to maximise the expected number of points. In this situation, which of the following is the closest to the expected number of points of the player when the game finishes (either finishing at his/her optimal choice, or the number points reaches 0, whichever comes first).

 (A) 2.

 (B) 4.

 (C) 6.

 (D) 8.

[1] exp(1) is the exponential distribution with mean 1

Answer 2

Enemy aircrafts appear according to a Poisson point process with rate 1. At any time, the player spends 1 point for each unit time period. After shooting down the kth enemy plane, he/she obtains 1.5 points. After that, for each unit period, the expected profit from shooting down the enemy aircraft is $1.5 \times (0.85)^n$.

(1) In this case, the expected loss from being shot down by an enemy aircraft is 0. Then we choose the largest k such that $1.5 \times (0.85)^n > 1$, that is, $n = 2$. The answer is (B).

(2) In this case, suppose the player has t points after shooting down the $(n-1)$-th enemy plane. If he/she chooses to wait until the next plane to appear, then the expected number of points right after the appearance of the next plan (or he/she consumed all the points, whichever comes first) is

$$(0.85)^n \times \int_0^t \big(t + 1.5 - x\big)e^{-x}\mathrm{d}x = (0.85)^n \times \big(t + 0.5 \times (1 - e^{-t})\big)\,. \tag{1}$$

For $n = 1$, the value of (1) is larger than t for all $t \leq 2$, so the player should wait at least until the first enemy plane to appear. If he/she shoots down the first one and continues, then he/she has at least 1.5 points. The value of (1) with $n = 2$ is less than t for $t \geq 1.5$, so after the player shoots down the first enemy plane, it is better to quit at that moment than to "quit right after shooting down the second enemy plane". By Question 1, we know any optimal strategy necessarily stops at the appearance of the second plane, hence the strategy for the player should be "quit the game right after shooting down the first plane". The expected number of points is then the value of (1) at $n = 1$ and $t = 2$, which is approximately 2.067. The closest answer is (A).

Question 3

For a real number $T > 0$, a subset Γ of the Euclidean plane $\mathbb{R}^2$ is said to be T-dense if for every $v \in \mathbb{R}^2$, there exists $w \in \Gamma$ such that $\|v - w\| \leqslant T$. Let $A \in \mathrm{M}_2(\mathbb{Z})$ be a 2×2 integer matrix with $\det(A) \neq 0$.

(1) Assume that $\operatorname{tr}(A) = 0$. Prove that there exists $C > 0$ such that for every positive integer n, the set

$$A^n\mathbb{Z}^2 := \{A^n v : v \in \mathbb{Z}^2\}$$

is $C|\det(A)|^{n/2}$-dense.

(2) Assume that the characteristic polynomial of A is irreducible over the field of rational numbers. Prove the same result as in (1).

Remark: Here vectors in $\mathbb{R}^2$ and $\mathbb{Z}^2$ refer to column vectors, the inner product on $\mathbb{R}^2$ is the standard one, namely, $\langle v, w\rangle = v^t w$.

(Hint: In proving (2), one can use the following special case of Minkowski's Convex Body Theorem: Any closed parallelogram in $\mathbb{R}^2$ centered at the origin with area 4 must contain a nonzero vector in $\mathbb{Z}^2$.)

Answer 3

Denote $t = \operatorname{tr}(A)$, $d = \det(A)$. Then the characteristic polynomial of A is $\chi_A(\lambda) = \lambda^2 - t\lambda + d$. By the Cayley-Hamilton theorem, one has $\chi_A(A) = 0$, namely $A^2 = tA - dI_2$. We proceed in two steps.

Step 1. We first prove that under the condition of (1) or (2), for every positive integer n, there exist integers p_n, q_n, which are not both zero, and real numbers x_n, y_n in the interval $[-1, 1]$, such that

$$p_n A^{n+1} + q_n A^n = |d|^{n/2}(x_n A + y_n I_2), \tag{2}$$

and both sides of equation (2) are invertible matrices.

- Assume that the condition of (1) holds, namely $t = 0$. Then $A^2 = -dI_2$.
  - If n is even, then $A^n = (-d)^{n/2} I_2$, namely, equation (2) holds for $p_n = 0$, $q_n = 1$, $x_n = 0$, $y_n = \operatorname{sgn}((-d)^{n/2})$. In this case, both sides of equation (2) are invertible.
  - If n is odd, then $A^n = (-d)^{(n-1)/2} A$, namely, equation (2) holds for $p_n = 0$, $q_n = 1$, $x_n = \operatorname{sgn}((-d)^{(n-1)/2})|d|^{-1/2}$, $y_n = 0$. In this case, both sides of equation (2) are also invertible.

- Assume that the condition of (2) holds, namely, $\chi_A(\lambda)$ is irreducible over $\mathbb{Q}$. In view of $A^2 = tA - dI_2$, it follows that for every $n \geqslant 0$, there exist $a_n, b_n \in \mathbb{R}$ such that $A^n = a_n A + b_n I_2$. Moreover, since A and I_2 are linearly independent, a_n and b_n are uniquely determined by n. To get more information, notice that

$$a_{n+1}A + b_{n+1}I_2 = A^{n+1} = A(a_nA + b_nI_2) = a_n(tA - dI_2) + b_nA = (ta_n + b_n)A - da_nI_2.$$

This implies $\begin{pmatrix} a_{n+1} \\ b_{n+1} \end{pmatrix} = \begin{pmatrix} t & 1 \\ -d & 0 \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}$. Therefore, for $n \geqslant 1$ we have

$$\begin{pmatrix} a_{n+1} & a_n \\ b_{n+1} & b_n \end{pmatrix} = \begin{pmatrix} t & 1 \\ -d & 0 \end{pmatrix} \begin{pmatrix} a_n & a_{n-1} \\ b_n & b_{n-1} \end{pmatrix} = \cdots = \begin{pmatrix} t & 1 \\ -d & 0 \end{pmatrix}^n \begin{pmatrix} a_1 & a_0 \\ b_1 & b_0 \end{pmatrix} = \begin{pmatrix} t & 1 \\ -d & 0 \end{pmatrix}^n.$$

In particular, $\det \begin{pmatrix} a_{n+1} & a_n \\ b_{n+1} & b_n \end{pmatrix} = d^n$. Consider the following closed parallelogram centered at the origin:

$$\Delta_n := \left\{ |d|^{n/2} \begin{pmatrix} a_{n+1} & a_n \\ b_{n+1} & b_n \end{pmatrix}^{-1} \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in [-1, 1] \right\}.$$

Since the matrix $|d|^{n/2} \begin{pmatrix} a_{n+1} & a_n \\ b_{n+1} & b_n \end{pmatrix}^{-1}$ has determinant ± 1, the area of Δ_n is 4. By Minkowski's Convex Body Theorem, Δ_n contains a nonzero integer vector, namely, there exist $p_n, q_n \in \mathbb{Z}$ that are not all zero and $x_n, y_n \in [-1, 1]$ such that

$$\begin{pmatrix} a_{n+1} & a_n \\ b_{n+1} & b_n \end{pmatrix} \begin{pmatrix} p_n \\ q_n \end{pmatrix} = |d|^{n/2} \begin{pmatrix} x_n \\ y_n \end{pmatrix}.$$

This implies that

$$\begin{aligned} p_nA^{n+1} + q_nA^n &= p_n(a_{n+1}A + b_{n+1}I_2) + q_n(a_nA + b_nI_2) \\ &= (p_na_{n+1} + q_na_n)A + (p_nb_{n+1} + q_nb_n)I_2 \\ &= |d|^{n/2}(x_nA + y_nI_2). \end{aligned}$$

Thus equation (2) holds. On the other hand, since $\chi_A(\lambda)$ is irreducible over $\mathbb{Q}$, A has no eigenvalue in $\mathbb{Q}$. It follows that $p_nA + q_nI_2$ is invertible, and hence both sides of equation (2) are invertible.

Step 2. Suppose $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. We prove that

$$C := \left((|a_{11}| + 1)^2 + a_{12}^2 + a_{21}^2 + (|a_{22}| + 1)^2 \right)^{1/2}$$

satisfies the requirement. First, it is straightforward to verify that for any $x, y \in [-1, 1]$ and $u \in \mathbb{R}^2$, we have $\|(xA + yI_2)u\| \leqslant C\|u\|$. Let $n \geqslant 1$, $v \in \mathbb{R}^2$. Denote

$$v' = |d|^{-n/2}(x_nA + y_nI_2)^{-1}v.$$

Let $w' \in \mathbb{Z}^2$ be such that $\|v' - w'\| \leqslant 1$, and let

$$w = |d|^{n/2}(x_n A + y_n I_2)w' = A^n(p_n A + q_n I_2)w' \in A^n\mathbb{Z}^2.$$

Then

$$\|v - w\| = |d|^{n/2}\|(x_n A + y_n I_2)(v' - w')\| \leqslant C|d|^{n/2}.$$

This completes the proof. $\square$

Question 4

Let $d \geq 0$ be an integer and V be a $(2d+1)$-dimensional complex linear space with a basis

$$\{v_1, v_2, \cdots, v_{2d+1}\}.$$

For an integer j $(0 \leq j \leq \frac{d}{2})$, write U_j for the subspace generated by

$$v_{2j+1}, v_{2j+3}, \cdots, v_{2d-2j+1}.$$

Define a linear transformation $f : V \to V$ by

$$f(v_i) = \frac{(i-1)(2d+2-i)}{2} v_{i-1} + \frac{1}{2} v_{i+1}, \ 1 \leq i \leq 2d+1.$$

Here we put $v_0 = v_{2d+2} = 0$.

(1) Show that eigenvalues of f are $-d, -d+1, \cdots, d$.

(2) Write W for the sum of eigenspaces of f of eigenvalues $-d+2k$ $(0 \leq k \leq d)$. Find the dimension of $W \cap U_0$.

(3) For any integer j $(1 \leq j \leq \frac{d}{2})$, find the dimension of $W \cap U_j$.

Answer 4

In $M_2(\mathbb{C})$, let

$$h_0 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad y_0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Then,

$$[h_0, x_0] = h_0 x_0 - x_0 h_0 = 2x_0,$$
$$[h_0, y_0] = h_0 y_0 - y_0 h_0 = -2y_0,$$
$$[x_0, y_0] = x_0 y_0 - y_0 x_0 = h_0.$$

Put

$$T = \exp(\frac{\pi}{4}(x_0 - y_0)) = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}.$$

Then, $T(x_0 + y_0)T^{-1} = h$.

(1)Define $h, x, y : V \to V$ as follows:

$$x(v_i) = v_{i+1}, \ y(v_i) = (i-1)(2d+2-i)v_{i-1}, \ h(v_i) = 2(i-d-1)v_i.$$

Then,

$$[h, x] = 2x, \quad [h, y] = -2y, \quad [x, y] = h.$$

Define $\phi : \mathfrak{sl}(2,\mathbb{C}) \to \mathfrak{gl}(V)$ by

$$\phi(h_0) = h,\ \phi(x_0) = x,\ \phi(y_0) = y.$$

Then, ϕ is an Lie algebra homomorphism. Again, put $S = \exp(\frac{\pi}{4}(x-y))$, we have $S(x+y)S^{-1} = h$. Thus, $f = \frac{1}{2}(x+y)$ is conjugate to $\frac{1}{2}h$. Then, eigenvalues of f are the same as that of $\frac{1}{2}h$, which are $-d, -d+1, \cdots, d$.

(2)Put $A = \exp(\pi \mathbf{i} f)$ and $B = \exp(\frac{\pi \mathbf{i}}{2} h)$. Then, W is the $(-1)^d$-eigenspace of A and U_0 is $(-1)^d$-eigenspace of B. The above homomorphism ϕ is the differential of a Lie group homomorphism $\Phi : \mathrm{SL}(2,\mathbb{C}) \to \mathrm{GL}(V)$. In $\mathrm{SL}(2,\mathbb{C})$, we have

$$A_0 := \exp(\pi \mathbf{i} f_0) = \begin{pmatrix} 0 & \mathbf{i} \\ \mathbf{i} & 0 \end{pmatrix}$$

and

$$B_0 := \exp(\frac{\pi}{2}\mathbf{i} h_0) = \begin{pmatrix} \mathbf{i} & 0 \\ 0 & -\mathbf{i} \end{pmatrix}.$$

Note that

$$A_0^2 = B_0^2 = A_0 B_0 A_0^{-1} B_0^{-1} = -I \text{ and } A_0 \sim B_0 \sim A_0 B_0.$$

Since $\Phi(-I) = 1$, then A and B are commuting involutions in $\mathrm{GL}(V)$ and $A \sim B \sim AB$. When d is even, we have

$$\dim W \cap U_0 = \frac{1}{2}(3 \dim V^A - \dim V) = \frac{1}{2}(d+2).$$

When d is odd, we have

$$\dim W \cap U_0 = \frac{1}{2}(\dim V - \dim V^A) = \frac{1}{2}(d+1).$$

(3)We have $U_{\lfloor \frac{d+2}{2} \rfloor} = 0$. Thus, $\dim W \cap U_{\lfloor \frac{d+2}{2} \rfloor} = 0$. By (2), $\dim W \cap U_0 = \lfloor \frac{d+2}{2} \rfloor$. It suffices to show that: for any j $(0 \leq j \leq \frac{d}{2})$, we have

$$\dim W \cap U_j - \dim W \cap U_{j+1} \leq 1.$$

Then, it follows that

$$\dim W \cap U_j = \lfloor \frac{d+2}{2} \rfloor - j.$$

As $\mathrm{Ad}(A_0)x_0 = y_0$ and $\mathrm{Ad}(A_0)y_0 = x_0$, we get $\mathrm{Ad}(A)x = y$ and $\mathrm{Ad}(A)y = x$. We also have $\mathrm{Ad}(A_0)^{-1}h_0 = -h_0$. Thus, $\mathrm{Ad}(A)^{-1}h = -h$. Put $u_0 = v_{d+1}$, which is a generator of the 0-eigenspace of h. As $\mathrm{Ad}(A)^{-1}h = -h$, we have $h(Au_0) = -A(hu_0) = 0$. Then, Au_0 is also a 0-eigenvector of h. Thus, $Au_0 = tu_0$, $t \neq 0$. As $A^2 = 1$, we must have $t = \pm 1$. For any $j \leq \frac{d}{2}$, $v_{2d+1-2j} = x^{d+1-2j}u_0$ and v_{2j+1} is proportional to $y^{d+1-2j}u_0$. Then, The action of A switches v_{2j+1} and $c_j v_{2d+1-2j}$ for some non-zero constant c_j. Thus, $\dim W \cap U_j - \dim W \cap U_{j+1} \leq 1$.

Question 5

For any convex polyhedron V in $\mathbb{R}^3$ symmetric about a point, prove that we can find an ellipsoid E, containg V in its interior, and the surface area of E does not exceed 3 times that of V.

Answer 5

We can prove the following general results:

(1)Let V be a nonempty convex open set in $\mathbb{R}^3$ which is symmetric about the origine,among all the ellipsoids E that are symmetric about the origine and contains V in its interior, there exists a unique E_0 that attains the minimal volume.

An ellipsoid in $\mathbb{R}^3$ that is symmetric about the origine is in 1-1 correspondance with (strictly) positive quadratic forms of three variables. We denote the set of positive quadratic forms of three variables by Q_+. For any $q \in Q_+$, the (possibly degenerate) ellipsoid it corresponds is $\{q(x,x) = 1\}$.

It is important to notice that the convex set V determines a unique *norm* N on $\mathbb{R}^3$(such that $V = \{x, N(x) < 1\}$). Then the set of all positive quadratic forms that correspond to (possibly degenerate) ellipsoids E containing V is

$$K = \{q \in Q_+, 0 \leqslant q(x,x) \leqslant (N(x))^2, \forall x \in \mathbb{R}^3\}.$$

It is easy to observe:

- K is not empty;
- K as a subset of $\mathbb{R}^{3\times 3}$(equipped with the topology determined by euclidean metric) is a bounded closed set, therefore compact;
- K is convex (if $0 \leqslant q_1(x,x), q_2(x,x) \leqslant (N(x))^2$, then for any $\lambda \in [0,1]$, $0 \leqslant \lambda q_1(x,x) + (1-\lambda)q_2(x,x) \leqslant (N(x))^2$).

We define $v(q) =$ the volume of$\{q(x,x) \leqslant 1\}$, it is a continuous function on K (the inverse of the product of the three eigenvalues, which is *not* bounded above). By the compactness of K, there exists q_N such that $v(q_N)$ attains the minimum value.

Now suppose that there exists another q' such that $v(q') = v(q_N)$, we want to show that $q' = q_N$. For this purpose, we need the identity

$$\iiint e^{-q(x,x)} d^3x = \frac{3}{2} v(q) \int_0^\infty t^{1/2} e^{-t} dt,$$

and the convexity of exponential functions, then we get

$$I(q) := \iiint e^{\frac{-q'(x,x) - q_N(x,x)}{2}} d^3x \leqslant \frac{1}{2} \left(\iiint e^{-q'(x,x)} d^3x + \iiint e^{-q_N(x,x)} d^3x \right).$$

So there exists a unique E_0(that corresponds to q_N), such that the volume attains minimum.

(2) Morevoer, we have $E_0 \subset \overline{\sqrt{3}V}$.

On the ellipsoid E_0, we pick a point a where $N(x)$ attains maximum. It is easy to observe

- the tangent plane Π of E_0 at a has only one common point (a) with E_0;
- we define $H = \{y, q_N(a, y) = 0\}$, then $\Pi = a + H$;
- for any $y \in H$, $t \in \mathbb{R}$, define $\varphi(t) = N(a + ty)$, by the general poperties of norms, it is a convex function that satisfies $\varphi(t) \leqslant N(a)\sqrt{1 + t^2 q_N(y, y)}$;
- therefore, the minimum of φ is attained at $t = 0$. If, for any $x \in \mathbb{R}^3$, we denote by $\pi_H(x)$ its orthogonal projection to H (with respect to q_N), and $\pi_a(x) = x - \pi_H(x)$, then we have $N(\pi_a(x)) \leqslant N(x)$;
- for any $\varepsilon, \delta \in (0, 1)$, define

  $$q(x, x) = (1 + \varepsilon)q_N(\pi_a(x), \pi_a(x)) + (1 - \delta)q_N(\pi_H(x), \pi_H(x)),$$

  we have $q \in Q_+$, and $I(q) = (1 + \varepsilon)^{-1/2}(1 - \delta)^{-1} I(q_N)$;
- for any $x \in \mathbb{R}^3$, we have $q_N(\pi_a(x), \pi_a(x)) + q_N(\pi_H(x), \pi_H(x) = q_N(x, x) \leqslant N^2(x)$, so

  $$\begin{aligned} q(x, x) =& (1 - \delta)(q_N(\pi_a(x), \pi_a(x)) + q_N(\pi_H(x), \pi_H(x)) + (\delta + \varepsilon)q_N(\pi_a(x), \pi_a(x)) \\ \leqslant& N^2(x) - \left[\delta - \frac{\delta + \varepsilon}{N^2(a)}\right] N^2(x). \end{aligned}$$

- if $N^2(a) > 3$, then we can choose certain $\varepsilon, \delta \in (0, 1)$ such that $\left[\delta - \frac{\delta + \varepsilon}{N^2(a)}\right] \geqslant 0$ and $(1 + \varepsilon)^{-1/2}(1 - \delta)^{-1} < 1$, that means $I(q) < I(q_N)$, contradiction.

Hence, $N^2(a) < 3$, i.e., $E_0 \subset \sqrt{3}V$.

(1)+(2) is a theorem of Fritz John dated back in 1940's.

Now, for our convex polyhedron V, the the theorem tells that for the corresponding ellipsoid E_0, we have $\frac{1}{\sqrt{3}}E_0 \subset V$. Hence, by the monotonicity of surface area of convex bodies in $\mathbb{R}^3$(this can be obtained via Cauchy's surface area formula, or, as we deal only a smooth ellipsoid and a polyhedron, it can be proved directly), we have

$$\text{Area of } \frac{1}{\sqrt{3}}E_0 \leqslant \text{Area of } \partial V,$$

So

$$\text{Area of } E_0 \leqslant 3 \times \text{Area of } \partial V.$$

Remark. Of course the constant 3 here is NOT optimal.

Question 6

(1) Consider tossing a coin, where the probability of getting a head is 1/3. Independently toss the coin for n times, and let X_n be the number of head we get. Find the following limit

$$\lim_{n\to\infty} P(X_n \text{ is even})$$

where $P(\cdot)$ denotes the probability.

(2) Suppose there are five boxes. Each time we independently put a ball in one of them chosen at random. We denote by p_i the probability that Box i is chosen at each trial. Let $X_n^{(i)}$ be the number of balls in Box i after this process repeated for n times. Find the following limit

$$\lim_{n\to\infty} P\left(X_{2n}^{(i)},\ i = 1, 2, \cdots, 5,\ \text{are all even}\right)$$

[Note: We thank Mr. Cai, Qiao from Beijing Mingcheng Academy as this question is inspired by communication with him during Fall, 2023. He is NOT informed by any means of this competition or the adoption of this question]

Answer 6

1. Conditioning on the result of the first toss:

   - If the first toss gives us a head, then we will need the total number of head to be odd in the second to the nth tosses.
   - Otherwise, we will need the total number of head to be even in the second to the nth tosses.

   Let $p_n = P(X_n \text{ is even})$, by total probability theorem we have

   $$p_n = \frac{1}{3} \times (1 - p_{n-1}) + \frac{2}{3} \times p_{n-1} = \frac{1}{3} + \frac{p_{n-1}}{3}$$

   thus

   $$\left(p_n - \frac{1}{2}\right) = \frac{1}{3} \times \left(p_{n-1} - \frac{1}{2}\right)$$

   which gives a contraction mapping that converges to 1/2.

2. First, let us consider the case of only boxes. Let $p_n(q)$ be the probability of getting an even number of heads in n tosses with the head probability to be q. Using the same

calculation as in Part (1) we immediately have for all $q \neq 1/2$, $\lim_{n\to\infty} p_n(q) = 1/2$. And when $q = 1/2$, one may also have

$$p_n(\frac{1}{2}) = \frac{1}{2} \times [1 - p_{n-1}(\frac{1}{2})] + \frac{1}{2} \times p_{n-1}(\frac{1}{2}) \equiv \frac{1}{2}$$

Now for the case of three boxes, say the probability of each is p_1, p_2, p_3. Denote by (X_1, X_2, X_3) the number of balls in each of them after $2n$ trials. By definition, (X_1, X_2, X_3) satisfies a multinomial distribution with parameter $(2n, p_1, p_2, p_3)$, while $X_1 \sim Binomial(2n, p_1)$. At the same time, given $X_1 = n_1$, the conditioned distribution of X_2 is given by $Binomial(n - n_1, \frac{p_2}{p_2+p_3})$.

Now let $A_{2n}^{(3)}$ be the event that (X_1, X_2, X_3) are all even. Again by total probability theorem:

$$P(A_{2n}^{(3)}) = \sum_{k=0}^{n} P(X_1 = 2k) p_{2n-2k}\left(\frac{p_2}{p_2+p_3}\right) \tag{3}$$

Note that we have shown that $\lim_{m\to\infty} p_m\left(\frac{p_2}{p_2+p_3}\right) = 1/2$Thus for all $\epsilon > 0$, there exists M_1 such that

$$p_m\left(\frac{p_2}{p_2+p_3}\right) \in (1/2 - \epsilon, 1/2 + \epsilon),\ \forall m \geq 2M_1$$

Meanwhile, there exists N_1 such that

$$p_m\left(p_1\right) \in (1/2 - \epsilon, 1/2 + \epsilon),\ \forall m \geq 2N_1$$

Moreover, for the fixed M_1recall that X_1 has a binomial districution with parameters $(2n, p_1)$. By Chebyshev's inequality, for the aforementioned $\epsilon > 0$there is some N_2such that for all $n \geq N_2$, we have

$$P(X_1 \geq 2n - 2M_1) < \epsilon$$

Letting $N = \max\{N_1, N_2\}$by(3) we have

$$\begin{aligned}
P(A_{2n}^{(3)}) &= \sum_{k=0}^{n} P(X_1 = 2k) p_{2n-2k}\left(\frac{p_2}{p_2+p_3}\right) \\
&\leq \sum_{k=0}^{n-M} P(X_1 = 2k) p_{2n-2k}\left(\frac{p_2}{p_2+p_3}\right) + P(X_1 \geq 2n - 2M) \\
&\leq \left(\frac{1}{2} + \epsilon\right) \sum_{k=0}^{n-M} P(X_1 = 2k) + \epsilon \\
&\leq \left(\frac{1}{2} + \epsilon\right) p_{2n}\left(p_1\right) + \epsilon \leq \left(\frac{1}{2} + \epsilon\right)^2 + \epsilon
\end{aligned}$$

and

$$
\begin{aligned}
P(A_{2n}^{(3)}) &= \sum_{k=0}^{n} P(X_1 = 2k) p_{2n-2k} \left(\frac{p_2}{p_2 + p_3} \right) \\
&\geq \sum_{k=0}^{n-M} P(X_1 = 2k) p_{2n-2k} \left(\frac{p_2}{p_2 + p_3} \right) \\
&\geq \left(\frac{1}{2} - \epsilon \right) \sum_{k=0}^{n-M} P(X_1 = 2k) \\
&\geq \left(\frac{1}{2} - \epsilon \right) \left[p_{2n} \left(p_1 \right) - P(X_1 \geq 2n - 2M) \right] \geq \left(\frac{1}{2} - \epsilon \right) \left(\frac{1}{2} - 2\epsilon \right)
\end{aligned}
$$

Note that ϵ is arbitrary, we have $P(A_{2n}^{(3)}) \to 1/4$. Using exactly the same argument, we can increase the number of boxes recursively and get $P(A_{2n}^{(5)}) \to 2^{-4}$. □

Question 7

Consider a music box with a circular track and a flowering tree at a point on the track. When the music box is turned on, it plays music and the track rotates clockwise at a constant speed.

You can place two tokens on the track which represent lovers. Let's call them Red and Green. When Red and Green are not under the tree, they move separately on the track. When a token reaches the tree, it waits under the tree for a period of time. During this time, if the other token also reaches the tree, then the two tokens meet and will then move together along the track, never to be separated again. Otherwise, when the waiting time is over, the two tokens will continue to move along the track separately.

Consider the mathematical model of this music box. We parameterize the circular track as a circle with length 1, and assume that the tokens and the tree can be represented as points on the circle. Specifically, we use $X(t) \in [0, 1]$ and $Y(t) \in [0, 1]$ to represent the positions of Red and Green on the track at time t, respectively, while the position of the tree is $\phi = 1$, or equivalently, $\phi = 0$.

When they have not reached the tree yet (see the left figure), their positions change according to

$$\frac{d}{dt}X(t) = 1, \quad \frac{d}{dt}Y(t) = 1.$$

Suppose at time t_0, Green reaches the tree (see the middle figure), i.e., $Y(t_0) = 1$, it waits for at most

$$\tau = K(X(t_0))$$

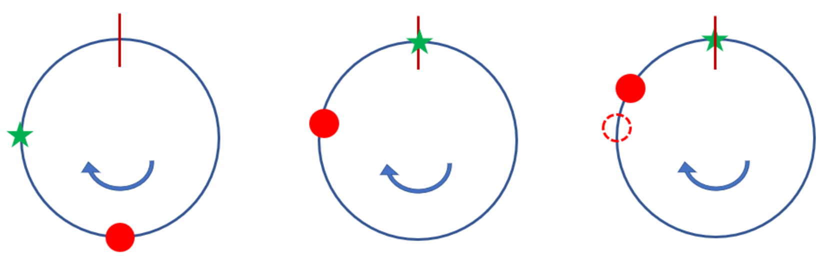
time. In other words, the maximum waiting time depends on Red's position at that moment.

During the waiting period, Green stays still, while Red continues to move. If at some time $t^* \in (t_0, t_0 + \tau]$ during the waiting period, Red also reaches the tree, i.e., $X(t^*) = 1$, then the two pieces meet. If the waiting time ends (see the right figure) and Red has not reached the tree yet, they will continue to move, and their positions become

$$X(t_0 + \tau) = X(t_0) + \tau, \quad Y(t_0 + \tau) = 0.$$

Note that although Green's coordinate is reset, its position on the circle remains the same.

If at some point Red reaches the tree, it also waits according to the same rule, with the maximum waiting time depending on Green's position at that moment. Clearly, the fate of Red and Green depends on the form of the maximum waiting time function $K(\phi)$.



(1) We define $f : \mathbb{R} \to \mathbb{R}$ to be a smooth function that satisfies

$$f' > 0, \quad f'' < 0, \quad f(0) = 0, \quad f(1) = 1.$$

Let ε be a sufficiently small positive constant. We define the waiting time function as

$$K(\phi) = f^{-1}\big(f(\phi) + \epsilon\big) - \phi.$$

Prove that with the unique exception, regardless of the initial distance between Red and Green, they will eventually meet.

(2) We consider an f function of the following form:

$$f(\phi) = \frac{1}{b} \ln\Big(1 + (e^b - 1)\phi\Big),$$

where $b > 0$ is a constant. When $b \ll 1$, $\varepsilon \ll 1$, please estimate the number of circles Red and Green have traveled before meeting.

Answer 7

(1) We consider the case when Green has finished waiting. Let Red's coordinate be ϕ at this time, and their coordinates in ascending order be $(0, \phi)$.

After a time of $1 - \phi$, Red reaches the tree, and their coordinates become $(1 - \phi, 1)$. Then Red's maximum waiting time is $\tau = K(1 - \phi)$.

If $1 - \phi + K(1 - \phi) \geq 1$, Red and Green meet during the waiting period.

If $1 - \phi + K(1 - \phi) < 1$, Green does not reach the tree during Red's waiting period, and their coordinates become $(1 - \phi + K(1 - \phi), 0)$. Rearranging in ascending order, their coordinates become $(0, 1 - \phi + K(1 - \phi))$, i.e., $(0, f^{-1}(f(1 - \phi) + \epsilon))$.

We define

$$h(\phi) = f^{-1}(f(1 - \phi) + \epsilon), \quad H(\phi) = h(h(\phi)).$$

Suppose Red and Green have not met yet, and we observe their coordinates in ascending order right after one of them finishes waiting. If their coordinates are $(0, \phi)$ after the n-th observation, then their coordinates after the $(n + 1)$-th observation are $(0, h(\phi))$, and after the $(n + 2)$-th observation are $(0, H(\phi))$. Note that every other observation corresponds to the same token waiting at the tree.

We first study the $h(\phi)$ function. To ensure they don't meet after one observation, we need $1 - \phi + K(1 - \phi) < 1$. It is easy to see that the domain of $h(\phi)$ is $(\delta, 1)$, where $\delta = 1 - f^{-1}(1 - \varepsilon)$.

Taking the derivative, we have

$$h'(\phi) = -(f^{-1})'(f(1-\phi)+\varepsilon)f'(1-\phi).$$

By the inverse function derivative rule, we know

$$(f^{-1})'(f(1-\phi)+\varepsilon) = \frac{1}{f'(f^{-1}(f(1-\phi)+\varepsilon))}.$$

Note that $f^{-1}(f(1-\phi)+\varepsilon) > 1-\phi$, then by the concavity of f, we have

$$h'(\phi) = -\frac{f'(1-\phi)}{f'(f^{-1}(f(1-\phi)+\varepsilon))} < -1.$$

Next, we study the $H(\phi)$ function. If they don't meet after two observations, then the domain is $(\delta, h^{-1}(\delta))$. Clearly, this domain is non-empty when ε is sufficiently small.

Taking the derivative, we have $H'(\phi) = h'(h(\phi))h'(\phi)$, so $H' > 1$.

We now study the fixed points and stability of H. First, it is easy to see that $h(\phi)$ has a unique fixed point ϕ^* in $(\delta, h^{-1}(\delta))$. Then naturally $H(\phi^*) = \phi^*$. From $H' > 1$, we can easily show that

$$H(\phi) > \phi, \quad \text{when} \quad \phi > \phi^*; \qquad H(\phi) < \phi, \quad \text{when} \quad \phi < \phi^*;$$

Thus, H has a unique unstable equilibrium point.

Therefore, unless the initial distance between the two tokens is ϕ^*, they will eventually meet.

(2) If we let $g = f^{-1}$, then it is easy to obtain

$$g(z) = \frac{e^{bz}-1}{e^b-1}.$$

We can directly calculate

$$h'(\phi) = -\frac{g'(f(1-\phi)+\varepsilon)}{g'(f(1-\phi))} = -e^{b\varepsilon}.$$

So we have $H'(\phi) = e^{2b\varepsilon}$, i.e., H is a linear function. Then from $H(\phi^*) = \phi^*$, we obtain

$$H(\phi) = e^{2b\varepsilon}(\phi - \phi^*) + \phi^*.$$

If we define

$$\phi_k = H^k(\phi_0), \quad \Delta_0 = |\phi_0 - \phi^*|, \quad \Delta_k = |\phi_k - \phi_*|,$$

then we can directly calculate

$$\Delta_k = e^{2b\varepsilon k}\Delta_0.$$

Note that at the $2k$-th observation, each token has made k revolutions, and when $\Delta_k = O(1)$, the two pieces will meet. Thus, we can estimate that at the meeting time,

$$k = O\left(\frac{1}{b\varepsilon}\ln\frac{1}{\Delta_0}\right).$$

The number of revolutions made by the tokens before meeting is also $O\left(\frac{1}{b\varepsilon}\ln\frac{1}{\Delta_0}\right)$.