Caijun Qin
02/03/2022

P1

Let T denote the total number of points in the game from both players A and B combined once a game ends with a clear winner.

Scenario 1: A and B are tied in points.
Scenario 2: A or B have a 1-point lead over the other player.
Scenario 3: A or B have a 2-point lead over the other player.
Scenario 4: A or B have more than 2 points more than the other player.

Given that a game ended due to a clear winner as recognized by the rules, scenarios 1 and 2 are not considered when computing T by not satisfying the game rules. In theory, both scenarios 3 and 4 satisfy the rules to declare a winner. However, scenario 3 is a prerequisite for scenario 4. Once scenario 3 occurs, the game immediately ends with a specific value for T before scenario 4 can happen. Therefore, a game with a winner always ends with $T = (k+2) + k = 2k + 2$ points, where k is the number of points of the losing player (player with lower points). In simple words, the winning player always has exactly 2 more points than the losing player once the game ends with a clear winner.

$T = 2k + 2 = 2(k+1) = 2n$ is divisible by 2 because T can be expressed as a product of 2 and another integer. By definition of an even integer, T is also an even integer. Note that $k \geq 0$, because no player can possibly have negative points on any step of the game.

Regardless of who wins, $2k$ trials are played before the winning 2-point lead occurs at the end. In each pair of trials e.g. (trial 1, trial 2), (trial 3, trial 4), …, either A can win on the earlier trial followed by B or vice versa. This means the pre-win $2k$ trials can have 2^k unique historical sequences of how points are earned by A and B, regardless of who wins at the end.

Consider these denotations to be used in expressing probability.
1) n can also be expressed with k, being $n = k + 1$.

2) Let $T(A)$ and $T(B)$ denote the number of points under players A and B, respectively, that together sum to T at game ending.

$$P(T = 2n) = P\big((T(A) = k+2, T(B) = k) \cup (T(A) = k, T(B) = k+2)\big)$$
$$= 2^k[p^{n+1}(1-p)^{n-1} + p^{n-1}(1-p)^{n+1}] = 2^{n-1}[p^{n+1}(1-p)^{n-1} + p^{n-1}(1-p)^{n+1}]$$

Player A winning is simply the first term in the above probability after distribution of 2^{n-1}.
$P(A\ wins) = P(T(A) = k+2, T(B) = k) = 2^{n-1}p^{n+1}(1-p)^{n-1}$

P2

Base case:

Caijun Qin
02/03/2022

$P(\cap_{i=1}^{n=2} E_i) = P(E_1 \cap E_2) = P(E_1)\frac{P(E_1 \cap E_2)}{P(E_1)} = P(E_1)P(E_2|E_1)$

Induction:
Assume
$P\left(\cap_{i=1}^{k} E_i\right) = P((E_1 \cap E_2 \cap \cdots \cap E_{k-1}) \cap E_k)$
$= \left[P(E_1)P(E_2|E_1) \cdots P\left(E_{k-1} \middle| \cap_{i=1}^{k-1} E_i\right)\right] P\left(E_k \middle| \cap_{i=1}^{k-1} E_i\right)$

Let Z denote $\cap_{i=1}^{k} E_i$ as a short-hand way.
$P\left(\cap_{i=1}^{k+1} E_i\right) = P(Z \cap E_{k+1}) = P(Z)\frac{P(Z \cap E_{k+1})}{P(Z)} = [P(E_1)P(E_2|E_1) \cdots P(E_k|Z)]P(E_{k+1}|Z)$

P3 The sample space is the set of all possible samples of k balls, regardless of color distribution. This assumes that each ball is unique (like having an ID), meaning a sample of i white balls and $k - i$ black balls does not necessarily equate to the same outcome as a second sample with the same color counts. A specific event, such as 3 white balls and 9 black balls for $k = 12$, contains all possible samples of 12 unique balls with 3:9 white-to-black ratio. Therefore, $|\Omega| = \binom{n+m}{k}$.

There are $\binom{n}{i}$ possible combinations of i white balls. Similarly, there are $\binom{m}{k-i}$ possible combinations of $k - i$ black balls. Therefore, the number of unique samples of size k containing i white balls is simply the product $\binom{n}{i}\binom{m}{k-i}$. The probability of an event specified by this specific distribution of colors is the previous product over the sample space size for samples of k balls.

$$P(X = i) = \frac{\binom{n}{i}\binom{m}{k-i}}{\binom{n+m}{k}}$$

P4

Choose 13 cards out of deck of 52 cards such that exactly 1 of 4 aces is included.
$P(E_1) = \frac{\binom{4}{1}\binom{48}{12}}{\binom{52}{13}} = \frac{4*69668534468}{635013559600} = 0.438847539$
Choose 13 cards out of remaining 39 cards such that exactly 1 of 3 remaining aces is included.
$P(E_2|E_1) = \frac{\binom{3}{1}\binom{36}{12}}{\binom{39}{13}} = \frac{3*1251677700}{8122425444} = 0.46230441$
Choose 13 cards out of remaining 26 cards such that exactly 1 of 2 remaining aces is included.
$P(E_3|E_1 \cap E_2) = \frac{\binom{2}{1}\binom{24}{12}}{\binom{26}{13}} = \frac{2*2704156}{10400600} = 0.52$
Choose final 13 cards out of remaining 13 cards such that exactly 1 of 1 remaining aces (last ace not yet chosen) is included.
$P(E_4|E_1 \cap E_2 \cap E_3) = \frac{\binom{1}{1}\binom{12}{12}}{\binom{13}{13}} = \frac{1*1}{1} = 1$

Caijun Qin
02/03/2022

$P(E_1) * P(E_2|E_1) * P(E_3|E_1 \cap E_2) * P(E_4|E_1 \cap E_2 \cap E_3) = P(E_1 \cap E_2 \cap E_3 \cap E_4) = 0.438847539 * 0.46230441 * 0.52 * 1 = 0.1054981994$

P5

Let *A*, *B*, and *C*, denote the event that the chosen coin is two-headed, fair, and comes up heads 75% of the time, respectively. Let *H* denote the event of getting heads, and let *T* denote the event of getting tails.

$$P(A|H) = \frac{P(H|A)P(A)}{P(H)} = \frac{1*\frac{1}{3}}{P(H|A)P(A)+P(H|B)P(B)+P(H|C)P(C)} = \frac{\frac{1}{3}}{1*\frac{1}{3}+\frac{1}{2}*\frac{1}{3}+\frac{3}{4}*\frac{1}{3}} = \frac{\frac{1}{3}}{\frac{1}{3}+\frac{1}{6}+\frac{1}{4}} = \frac{\frac{1}{3}}{\frac{3}{4}} = \frac{4}{9}$$

P6

When using quicksort on the sequence of *n* numbers $A_0 = \{x_1, x_2, \ldots, x_n\}$, let the sorted version of that sequence be denoted as $S_0 = \{x_{(1)}, x_{(2)}, \ldots, x_{(n)}\}$.

Let *X* be the number of comparisons made throughout quicksort, computed below.

$X = \Sigma_{i=1}^{n-1} \Sigma_{j=i+1}^{n} I(i,j)$

where $I(i,j) = \begin{cases} 1, & if\ x_{(i)}\ and\ x_{(j)}\ are\ compared \\ 0, & otherwise \end{cases}$

The maximum number of times any two elements, $x_{(i)}$ and $x_{(j)}$, can be compared is once, and that only happens when one of the elements is a pivot. After that pivot's particular subarray becomes partitioned, the pivot will never be compared to another number for the rest of sorting. If $x_{(i)}$ and $x_{(j)}$ are from the same subarray but become separated into two different partitions, these two elements are never compared at all during sorting. These two possible paths for each pair of elements explain the usage of a binary indicator variable *I* to indicate whether a comparison occurs or not for each pair of unique elements.

The expression $X = \Sigma_{i=1}^{n-1} \Sigma_{j=i+1}^{n} I(i,j)$ is structured after the worst-case scenario of the original sequence A_0 already being sorted and iteratively choosing the pivot from left to right. We need to examine the expectation of every possible comparison that may take place during sorting. In this scenario, every possible comparison occurring implies $I(i,j) = 1$ for the entire sorting process, leading the sum of the first $n-1$ positive integers, $\frac{n(n-1)}{2}$.

$E[X] = E\left[\Sigma_{i=1}^{n-1} \Sigma_{j=i+1}^{n} I(i,j)\right] = \Sigma_{i=1}^{n-1} \Sigma_{j=i+1}^{n} E[I(i,j)]$

As $I(i,j)$ is an independent random variable, $E[I(i,j)] = (I(i,j) = 0) * P(I(i,j) = 0) + (I(i,j) = 1) * P(I(i,j) = 1) = 0 * P(I(i,j) = 0) + 1 * P(I(i,j) = 1) = P(I(i,j) = 1)$

Caijun Qin
02/03/2022

Note that $x_{(i)}$ and $x_{(j)}$ are compared if and only if one of the two elements are chosen as the pivot. If a pivot is chosen to be less than $x_{(i)}$ or greater than $x_{(j)}$, then the two elements will remain in the same partitioned subarray after splitting the current subarray but will not be compared during this recursion step. If the pivot happens to be in between $x_{(i)}$ and $x_{(j)}$, then the two values will be separated into different partitioned subarrays, never to be compared at all. Therefore, only when one of the two elements in a subarray $\{x_{(i)}, x_{(i+1)}, \ldots, x_{(j-1)}, x_{(j)}\}$ is chosen as pivot will they be compared to each other. That probability is given below by choosing the pivot from the two endpoints of a subarray of $j - i + 1$ elements.

$P(I(i,j) = 1) = \frac{2}{j-i+1}$

Plug this back as $E[X_{ij}]$.
$E[X] = \Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}\left(\frac{2}{j-i+1}\right) = \Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}\left(\frac{2}{x}\right) = 2\Sigma_{i=1}^{n-1}\Sigma_{j=i+1}^{n}\left(\frac{1}{x}\right)$
Note that the substitution $x = j - i + 1$ has been made to more tersely represent the number of elements in the subarray enclosed by $x_{(i)}$ and $x_{(j)}$.

The inner summation can be approximated with an integral for large n.
$\Sigma_{x=1}^{n}\left(\frac{1}{x}\right) \approx \int_1^n \frac{dx}{x} = \ln x + C|_1^n = \ln n$

$\therefore E[X] \approx 2\Sigma_{i=1}^{n-1} \ln n = 2(n-1)\ln n = 2n\ln n - 2\ln n \approx 2n\ln n$
$E[X] \approx 2n\log n$

Note that in many fields, log is symbolically synonymous with ln.

P7

(i) As a property of subsets, $A \subseteq B \Rightarrow A \cup B = B$.
To ensure disjoint subsets, let $C = B \setminus A$. Sets A and C are disjoint. The equality $A \cup C = B$ holds. Under the *countable additivity* property of measure, $\mu(A \cup C) = \mu(A) + \mu(C)$.
$\mu(B) = \mu(A \cup C) = \mu(A) + \mu(C)$
Since μ always maps to a nonnegative value i.e. range is $[0, \infty)$, $\mu(A) \geq 0$ and $\mu(C) \geq 0$.
$\mu(A) \leq \mu(A) + \mu(C) \because$ the sum of a nonnegative number (in this case $\mu(A)$) and another nonnegative (in this case $\mu(C)$) is always at least the first number $\mu(A)$. Expressed more simply, adding a number with another number that is at least zero (0) yields a sum that is at minimum equal to the first number.

Replacing the right-hand side of the inequality, $\mu(A) + \mu(C)$, with the equivalent $\mu(B)$, we have
$A \subseteq B \Rightarrow \mu(A) \leq \mu(B)$

Caijun Qin
02/03/2022

(ii) Given the premise $X \subseteq Y$ where X and Y are sets, we know that Y can be expressed as a union of disjoint subsets, namely $Y = (Y \setminus X) \cup X$. By *countable additivity* property of measure, we also know that $\mu(Y) = \mu\big((Y \setminus X) \cup X\big) = \mu(Y \setminus X) + \mu(X)$, since $(Y \setminus X)$ and X are disjoint sets for which the union is Y.

$\Rightarrow \mu(Y \setminus X) = \mu(Y) - \mu(X)$
We will use this later.

Given that $A_1 \subseteq A_2 \subseteq \cdots$, the union of all non-decreasing subsets given above can be re-expressed as a union of disjoint subsets of A_n.

$A = \cup_n A_n = B_1 \cup (\cup_n B_n)$ where $B_i = A_i \setminus A_{i-1}$ for $i = 2,3,\dots,n$ and $B_1 = A_1$. Note that subsets $B_1, B_2, \dots$ are disjoint.

Under *countable additivity* property of measure:
$\mu(A) = \mu(\cup_n A_n) = \mu(B_1 \cup (\cup_{n=2}^{\infty} B_n)) = \mu(B_1) + \Sigma_{n=2}^{\infty}\mu(B_n) = \lim_{n\to\infty}[\mu(A_1) + \Sigma_{i=2}^{n}\mu(B_i)]$
$= \lim_{n\to\infty}[\mu(A_1) + \mu(A_2 \setminus A_1) + \mu(A_3 \setminus A_2) + \cdots]$
$= \lim_{n\to\infty}[\mu(A_1) + (\mu(A_2) - \mu(A_1)) + (\mu(A_3) - \mu(A_2)) + \cdots]$
$= \lim_{n\to\infty}\mu(A_n)$

$\therefore \mu(A) = \lim_{n\to\infty}\mu(A_n) = \mu(\cup_n A_n) \Rightarrow \mu(A_n) \uparrow \mu(A)$

(iii) Since $\mu(A_n) < \infty$ for some n, and $A_n \supseteq A_{n+1}$, we can conclude that $\mu(\cap_n A_n) < \infty$.
$A = A_1 \setminus \cap_n A_n$
$= A_1 \cap (\cap_n A_n)^C$ // equivalent expression for set difference
$= A_1 \cap (\cup_n A_n^C)$ // DeMorgan's laws
$= \cup_n (A_1 \cap A_n^C)$ // distributive property of intersection
$= \cup_n (A_1 \setminus A_n)$ // equivalent expression for set difference

Now, apply measure over A.
$\mu(A) = \mu(A_1 \setminus \cap_n A_n) = \mu(A_1) - \mu(\cap_n A_n)$ // measure of difference between set and subset, derived in earlier parts of problem

$= \lim_{n\to\infty}[\mu(A_1) - \mu(\cap_n A_n)] = \mu(A_1) - \lim_{n\to\infty}\mu(\cap_n A_n)$

From two lines earlier, we see that the equality can be expressed by $\mu(A) = \mu(A_1) - \mu(\cap_n A_n)$ while also by $\mu(A) = \mu(A_1) - \lim_{n\to\infty}\mu(\cap_n A_n)$, we can see that together:
$\mu(A_1) - \mu(\cap_n A_n) = \mu(A_1) - \lim_{n\to\infty}\mu(\cap_n A_n)$
$\Rightarrow \lim_{n\to\infty}\mu(\cap_n A_n) = \mu(\cap_n A_n)$

Caijun Qin
02/03/2022

Note that a property of subsets is that if $X \subseteq Y$, then $X = X \cap Y$.
By induction, $A_n = A_1 \cap A_2 \cap ... \cap A_n$ given that $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots \supseteq A_n$.

$\Rightarrow \lim_{n\to\infty} \mu(\cap_n A_n) = \lim_{n\to\infty} \mu(A_n) = \mu(\cap_n A_n) = \mu(A)$
$\therefore \mu(A) = \lim_{n\to\infty} \mu(A_n) = \mu(\cap_n A_n) \Rightarrow A_n \downarrow A$

(iv) A function that is additive, which a measure μ serves as an example, is also strongly additive.
Let a measure be defined by $\mu: \mathrm{F} \to \mathbb{R}^+$.
$\Rightarrow \forall A, B \in F: \mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B)$

If A and B are disjoint, then $\mu(A) + \mu(B) = \mu(A \cup B) + \mu(\emptyset) = \mu(A \cup B) + 0 = \mu(A \cup B)$.

Suppose there is a collection of possibly infinite sets $U = \{A_1, A_2, ...\}$. Then by extension of strongly additive functions for two sets as previously stated,
$\Sigma_n \mu(A_n) = \mu(\cup_n A_n) + \Sigma_{i<j} \mu\left(A_i \cap A_j\right)$

Let second term on the right-hand side be denoted by S. $S > 0$ if any non-empty intersections exist between elements of U, otherwise $S = 0$, indicating that all elements of U are mutually disjoint.

It is clear that $\mu(\cup_n A_n) \leq \mu(\cup_n A_n) + \Sigma_{i<j} \mu\left(A_i \cap A_j\right)$ holds.
Replace the right-hand side with a single summation to conclude
$\mu(\cup_n A_n) \leq \Sigma_n \mu(A_n)$