Hari Probability Playing Cards

Bored Boy

June 2024

I pick cards randomly and solve them.

A $\clubsuit$ What is the expected value of a randomly drawn card from a standard deck? If you draw two cards randomly, what is the expected value of their sum?
Solution. The EV of a random card is $\boxed{7}$. Then, the EV of the sum of two random cards is (by LOE) $\boxed{14}$.

---

2 $\clubsuit$ On average, how many cards will you need to draw from a deck of cards until you draw the $A\clubsuit$?
Solution. With replacement, we have a 1 in 52 chance of drawing it at any time and the expectation of the associated bernoulli distribution is $\boxed{52}$.

Without replacement, an equal number of cards are expected to be on either side of the $A\clubsuit$ and so the expected number of cards to draw to draw the $A\clubsuit$ is

$$\frac{51}{2} + 1 = \boxed{\frac{53}{2}}$$

---

3 $\clubsuit$ 20 people are each holding a playing card. These cards are all shuffled and redistributed randomly. What is the expected number of people who get their original card back?
Solution. Let λ_i be an indicator variable for person i getting their card back. We seek

$$E(\sum_{i=1}^{20} \lambda_i$$

$$= 20\lambda_1$$

$$= 20\frac{1}{20}$$

$$= \boxed{1}.$$

---

4 $\clubsuit$
Solution.

---

5 $\clubsuit$ When inserting two jokers into the deck we drop the $5\clubsuit$ on the floor. We then shuffle the deck. What's the expcted number of cards between the two jokers?
Solution. There is an expected 17 cards before the first joker, $\boxed{17}$ between the first and second jokers, and 17 after the second joker; by symmetry, a card is equally likely to be in any three of these.

(To see this better, for any permutation of the 53 cards, let us view it as x cards followed by (say) the starry joker, followed by y cards followed by (say) the cloudy joker, followed by $51 - x - y$ cards. So it might look like, with X representing a permutation of size x, Y representing a permutation of size y, and $51 - X - Y$ representing a permutation of size $51 - x - y$,

$$X|J_s|Y|J_c|51 - X - Y$$

But we can then move the top two components to the bottom in a clearly bijective transformation,

$$Y|J_c|51 - X - Y|J_s|X$$

and again:

$$51 - X - Y|J_s|X|J_c|Y.$$

One can note that if we perform this transformation again, we do not get back the initial permutation (We get the initial permutation with J_s, J_c swapped.) But then applying it 3 more times gets back the initial permutation.

So we can partition all permutations into groups of size 6, and now clearly the average value of the first pile is 17 in each of these piles and thus must be 17 for the set of all permutations.)

---

6 ♣ You have a pile of a large number of cards (equal numbers of each suit.) How many cards on average do you have to draw before you have at least one of each suit?
Solution. Let A_n be the number of draws after drawing $n-1$ suits to draw a new suit. The time we seek is equal to $A_1 + \cdots + A_4$, so our expectation is

$$E(A_1 + \cdots + A_4) = E(A_1) + \cdots + E(A_4).$$

It takes one draw to get the first suit, $\frac{4}{3}$ for the second, $\frac{4}{2}$ for the third, and $\frac{4}{1}$ in expectation to get the fourth. We thus end up with an expected

$$\boxed{\frac{25}{3}}$$

draws to draw all 4 suits.

---

7 ♣ While showing off a fancy shuffling trick, you accidentally drop almost half the deck down a storm drain. You want to figure out which cards you still have, so you try to match up cards with the same rank and color (As with Ac, for example.) What is the expected number of matches found?
Solution. Every card has a partner. We want to find the expected number of pairs present in the top 34 cards of a deck.

Note the number of pairs present in the top 34 cards of a deck is

$$\frac{1}{2}\sum_{i=1}^{34}\lambda_i,$$

where λ_i is 1 if card i's partner is in the first 34 cards and 0 otherwise.

But then the expectation we seek is

$$\begin{aligned} E(\frac{1}{2}\sum_{i=1}^{34}\lambda_i) &= \frac{1}{2}(34)E(\lambda_1) \\ &= 17 * \frac{33}{51} \\ &= \boxed{11} \end{aligned}$$

---

8 ♣ Arrange the thirteen clubs in random order. How many times do you expect the sequence to change direction (increasing to decreasing or vice versa?)
Solution. This is equal to the expected number of elements that either exceed both their neighbors or are less than both of their neighbors. For any position besides the first or last, the likelihood of the card at that position satisfying these constraints is $\frac{2}{3}$. Then, the expected number of directional shifts is

$$\boxed{\frac{22}{3}.}$$

(This implicitly assigns an indicator variable to each internal position being a local extrema, then sums the expectations of these indicators.)

---

9 ♣ A very large deck of $2N+1$ card is cu at one of its $2N$ gaps randomly. On average, how many cards are in the smaller pile? What is the expected ratio of the size of the larger pile to the size of the smaller pile?

Solution. For the first problem, we can just case on the number of cuts in the smaller pile based on where the cut is:

$$E_N(X) = \frac{1+2+\cdots+N+N+(N-1)+\cdots+1}{2N} = \boxed{\frac{N+1}{2}}.$$

The expected value of the ratio of the size of the larger pile to the size of the smaller pile is (WLOG the cut is in the first half,)

$$\begin{aligned} &\frac{\sum_{i=1}^{N} \frac{2N-i}{i}}{N} \\ &= \frac{2N\sum_{i=1}^{N} \frac{1}{i} - N}{N} \\ &= \boxed{2H_N - 1}, \end{aligned}$$

and as $N \to \infty$ this approaches

$$\approx \boxed{2\log n + 0.154}$$

where we crudely approximate the Euler-Mascheroni constant.

---

10 ♣. Draw cards from a shuffled deck until you draw the Jc, Qc, and Kc. How many draws will it take on average?

Solution. The three cards partition the deck into 4 piles of cards, each of which have the same expected size. The expected size is thus $\frac{49}{3}$, and then the expected number of draws among the first three piles is $\frac{147}{4}$. Then, the expected number of draws to get all three of these cards is $\boxed{\frac{159}{4}}$

J ♣ We have $2N$ cards: 1 through N of both clubs and spades. Remove the 1♣ and a random ♠. Remove a pair of cards: the 2♣ and the associated ♠ (the 2♠ if you can find it, or else a random ♠). Continue similarly (with the 3♣, 4♣ , etc.) until all cards are paired and removed. What's the chance N♣ and N♠ form a pair? On average, how many pairs have matching rank?

Solution. For the first problem, note that if player 1 goes to the u of spades, then everything up to $u-1$ of spades gets filled in. Then, if we thereafter have the u of clubs map to the v of spades for $v \neq u$, everything from u to $v-1$ is filled in.

Let's call x_1 the rank of the card that the 1♣ is paired with, and x_2 the rank of the card that the x_1♣ is paired with, and so on. Note that when x_τ equals 1 or n, the cards from $x_\tau+1$ to $n-1$ ♣ get paired with their corresponding ♠ card. Then, n♣ will get paired with n♠ if $x_\tau = 1$ and with 1♠ if $x_\tau = n$. These events occur with equal likelihood, so the solution to the first question is

$$\boxed{\frac{1}{2}.}$$

For the second part, let us have for each x♣ an indicator variable λ_x equal to 1 if it is paired with itself and 0 otherwise. We seek $E(\sum_{i=1}^{N} \lambda_i) = \sum_{i=1}^{N} E(\lambda_i)$.

The probability $\lambda_i = 1$ is the probability that the sequence x_j equals 1 or $i+1, \ldots, n$ before it equals i (the sequence must displace i♣ before it fixes i♣ by either overshooting it or equalling 1.) The first value among

$$\{1, i, i+1, \ldots, n\}$$

to be taken on by the sequence is equally likely to be any of these values by symmetry, so the probability it is exactly i is

$$\frac{1}{n-i+2}$$

and so the expecation of λ_i is

$$1 - \frac{1}{n-i+2}$$

except for $i = 1$, where the sequence cannot equal "1" before "1" and the expectation of λ_1 is

$$\frac{1}{n}.$$

Then,

$$\sum_{i=1}^{n} E(\lambda_i)$$

$$= \frac{1}{n} + \sum_{i=2}^{n} 1 - \frac{1}{n-i+2}$$

$$= \frac{1}{n} + \sum_{i=0}^{n-2} 1 - \frac{1}{n-i}$$

$$= \frac{1}{n} + n - 1 - \sum_{i=0}^{n-2} \frac{1}{n-i} = \frac{1}{n} + n - 1 - \frac{1}{n} - \sum_{i=2}^{n-1} \frac{1}{i}$$

$$= \boxed{n - H_{n-1}}$$

Q ♣
Solution.

K ♣ You and I are playing with a strange deck of cards that consists of the 1 through N ♣ s. I remove r cards from the deck. Without knowing which r I removed, how large of a sum-free subset can you promise me you can find in your cards?
Solution. Our answer is

$$\boxed{f(N,R) = \lfloor \frac{N-R}{2} \rfloor + 1},$$

with $f(N,R) : R \geq N = 0$.

Firstly, note the solution holds for $r = 0$; just take all the odd numbers (n odd) or the last $\frac{n+2}{2}$ numbers (n even). We can't do better as whatever the largest number is (say U,) we can only pick one of $1, U-1$, $2, U-2$, etc. which leads to at most our expression.

We now attempt to proceed inductively. Note that $f(n,r) \geq f(n-1,r)$: just ignore the nth number and draw from the mapping from $|s| = r$ to $|u| \subseteq s, |u| = f(n-1,r)$ with u sumfree guaranteed by induction. Also, $f(n,r) \leq f(n-1,r-1)$: if the last number is removed, we can do only as well as in the $f(n-1,r-1)$ case.

The latter rule directly gives $f(N,R) \leq f(N,R) = \lfloor \frac{N-R}{2} \rfloor + 1$

for nontrivial pairs (N,R.)

Finally, proceed by (R, then N) induction on the problem. Suppose our formula first does not hold for some (N', R'). Then note that $N' - R'$ must be even, as in the odd case we directly get the result from our lower and upper bounds via induction.

I do not know how to show this inductive step.

A ♢ A standard deck of cards is shuffled. What is the chance that the third card from the top is either a diamond or an ace?
Solution.

$$\boxed{4}13$$

2 $\diamondsuit$
Solution.

---

3 $\diamondsuit$
Solution.

---

4 $\diamondsuit$
Solution.

---

5 $\diamondsuit$ I have all the diamonds. You randomly choose one of the cards I have without looking at it. I look at my cards and then place 11 of them face up. The 5 of diamonds is not among them. You and I both want the 5 of diamonds. Should you trade the card you initially chose for my remaining card?
Solution. Let us first make a short note; we assume (as in monty hall: see https://en.wikipedia.org/wiki/Monty_Hall_problem) the game show host is restricted to randomly choosing 11 of the cards from the cards both not chosen by me and not containing the 5$\diamondsuit$ to show me. The second constaint can be shown to be optimal; if the game show host is allowed to show you your own card when it doesn't contain the 5$\diamondsuit$, then their odds vastly improve. I will not solve for the nash equilibrium of this updated game and simply assume this is a monty hall clone.

In the monty hall case, (by conditional probability and letting A be our picked card, B be the other unpicked card,) the probability B is the 5 is

$$\frac{\frac{1}{13}}{\frac{1}{13}+\frac{1}{13}\frac{1}{12}} = \boxed{\frac{12}{13}},$$

where we note that conditioning on B being wrong (and the fact that A is right,) the probability B is the unflipped card is $\frac{1}{13}$. One important thing to note is that we cannot simply say "since A could not be picked, we gain no new information about A and so B "absorbs" the probabilities of all flipped cards" which is a common misconception; see my CMIMC TCS problem at haridesikan.com to see why this is the case. The uniform prior distribution is a special case.

---

6 $\diamondsuit$ You are given a deck of cards that is equally likely to be perfectly shuffled or brand new. You draw the first three cards and they are in order. How much more likely is the deck to be brand new?
Solution. "How much more likely" could either refer to, with p denoting the probability the deck is perfectly shuffled, $p-0.5$ or $p-(1-p)$ (or some other interpretation I've missed.) I will simply return p, and you can trivially process the result.

Using Bayes' rule yields

$$P(\text{new given 3 in order}) = \frac{0.5}{0.5+0.5*\frac{1}{52*51*50}}$$
$$= \boxed{\frac{132600}{132601}}$$

---

7 $\diamondsuit$
Solution.

---

8 $\diamondsuit$ We are playing with a random deck. I remove one $\diamondsuit$ and some non $\diamondsuit$ but I don't tell you how many! I secretly flip a fair coin; on heads I return the missing $\diamondsuit$, on tails I return one of the other missing cards. You shufle the deck and draw one card; it's a $\diamondsuit$. What's the probability my coin came up heads?

Solution. We assume that the number of nondiamonds removed is uniformly chosen from 0 to 39 (1 to 39 gives the same result.) Then $P(H|d) =$

$$\frac{\frac{1}{2}\frac{1}{40}\sum_{i=0}^{39}\frac{13}{52-i}}{\frac{1}{2}\frac{1}{40}\sum_{i=0}^{39}\frac{13}{52-i}++\frac{1}{2}\frac{1}{40}\sum_{i=0}^{39}\frac{12}{52-i}}$$

$$= \boxed{\frac{13}{25}}.$$

Note whether we begin at $i = 0$ or 1 doesn't matter, so the 2 interpretations of the problem give the same result.

---

9 $\diamondsuit$ Which of the following is most likely when cards are drawn with replacement from a shuffled deck: at least 1 diamond from 4 cards, at least 2 diamonds from 8 cards, or at least 3 diamonds from 12 cards?
Solution. My intuition is as follows. Let's imagine these as continuous distributions and consider the condition for $4n$ draws. We want that there are more than n diamonds. Intuitively, since the step size is 1 card, this is somewhat analogous to the probability that $S_n > \frac{n}{4} - C$ for some fixed constant C (representing the constant that is 1 card.) If we now imagine that we can "pseudo-apply" CLT here, the sum of N is distributed as $N(\frac{n}{4}, \sigma^2 n)$. We now seek to maximize $\int_{-C}^{0} N(0, \sigma^2 n)$ or more simply, $\int_0^C N(0, n)$. This is clearly maximized at smaller n.

Now, let's compute.

The first is $1 - \frac{1}{4^4}$. The second is $1 - \frac{1}{4^8} - 8 * \frac{3}{4^8}$ The third is $1 - \frac{1}{4^12} - 12 * \frac{3}{4^12} - 66 * \frac{9}{4^12}$.

So the third is most likely to have at least the average number of diamonds. My intuition was all wrong.

TODO: Explain why my intuition was wrong.

---

10 $\diamondsuit$
Solution.

J $\diamondsuit$ Consider the following game: place the $A\diamondsuit - 6\diamondsuit$ and the $A\spadesuit - 6\spadesuit$ in a bag. You randomly draw cards one at a time from the bag without replacement, putting aside matching cards as soon as they appear in you hand. The game ends and you lose if you ever hold three cards, no two of which match. What is the chance you win?
Solution. We proceed by seeking a recurrence. Call P_i the probability you win with the A-i of two suits and the same game; we seek P_6. Note $P_1 = P_2 = 1$.

Now, if we have k pairs, we must draw some card (WLOG an A.) Now, we have $n - 1$ cards left.

If we draw the other Ace ($\frac{1}{2n-1}$), our probability of success is P_{n-1}. Otherwise, say WLOG we hold an A and a 2 (with probability $\frac{2n-2}{2n-1}$. Then, we MUST draw either an A or 2 the next turn ($\frac{2}{2n-2}$,) at which point (WLOG we drew the A) we hold a 2 in our hand and the 3-n of both suits remain in the deck. But again, note this is P_{n-1}, as having $n - 2$ pairs in the deck and a single pair split between the deck and your hand is equivalent to having $n - 1$ pairs in the deck; the latter state always transitions to the former in 1 move.

So we have

$$P_n = \frac{1}{2n-1}P_{n-1} + \frac{2n-2}{2n-1}\frac{2}{2n-2}P_{n-1}$$

$$= \frac{3}{2n-1}P_{n-1},$$

$$= \frac{3^{n-1}2^n n!}{(2n)!},$$

and plugging in 6 yields $\boxed{\frac{9}{385}}$.

Q $\diamondsuit$ Take a random permutation of [13]. Given the longest increasing subsequence has length 10, find the probability the $K\diamondsuit$ is the last card.
Solution. After much work, I searched OEIS to see if my pattern (number of permutations of $[n]$ with longest increasing subsequence of length $n-3$) was recognized, which is basically giving up. It was recognized, but not with any closed form – there is no closed form. Further investigation shows that only asymptotics are available.

Searching $OEIS$ yields
https://oeis.org/A245666 https://oeis.org/A245665
which together yield

$$\boxed{\frac{153341}{275705}}$$

I do not plan on doing bashy counting on this problem. I may look at asymptotics, but the problem definitely should have been phrased that way to begin.

K $\diamondsuit$
Solution.

A $\heartsuit$ A tiny airplane has gotten stuck inside an inflatable ball and is flying around randomly. When you stop and look, what is the probability the ball is closer to the center of the ball than to the outside?
Solution. We seek the ratio between the volume of the half-unit ball and the volume of the unit ball. Note that the volume is proportional to r^3, and we simply get

$$\frac{\frac{1}{2}^3}{1} = \boxed{\frac{1}{8}}$$

---

2 $\heartsuit$ Two planes are travelling at 60mph and 180mph respectively. They both stop independently after 0-60 seconds (uniformly at random.) What is the probability that the planes travel less than a mile total in sum?
Solution. This is the same as adding one Unif(0,1) and a Unif(0,3) and asking the likelihood the sum is less than 1. The area of the triangle bounded by y=0, x=0, and y+x=1 is $\frac{1}{2}$ and the total area of the space of possibilities is 3, for a likelihood of $\frac{\frac{1}{2}}{3} = \boxed{\frac{1}{6}}$.

---

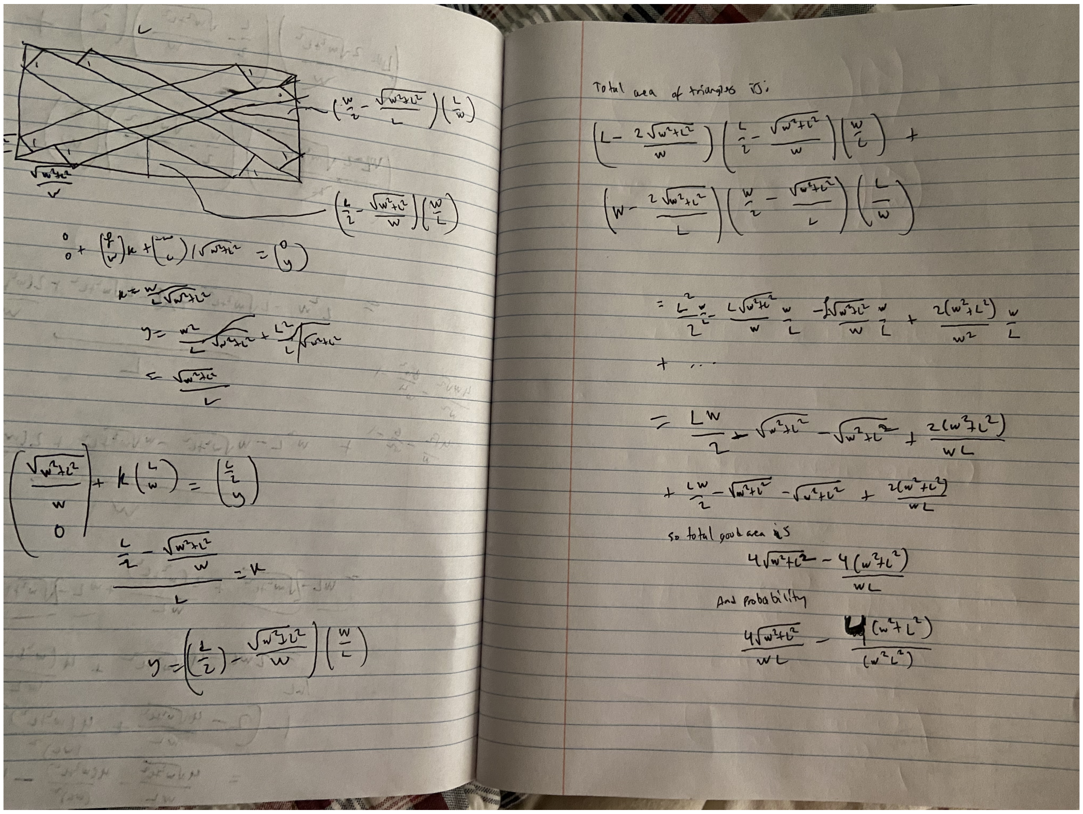
3 $\heartsuit$
Solution.

---

4 $\heartsuit$
Solution.

---

5 $\heartsuit$ A unit circle's center is randomly chosen within the interior of a rectangle of dimensions l, w. What is the probability it intersects one of the diagonals?
Solution. The actual problem includes a physical interpretation which is too complicated to work out, so I simplify the problem to allow the circle to have parts outside the rectangle.



I am no technological geometer, and I humbly offer my scratch work. My solution ends up being

$$\boxed{\frac{4}{WL}(\sqrt{W^2+L^2}-\frac{W}{L}-\frac{L}{W})}$$

---

6 ♡
Solution.

---

7 ♡ N cyclists are biking around a circular track at a rate of 1 rpm. They start in uniformly random locations and random orientations. Just before a pair would collide, they instantly reverse direction. What is the expected number of times all the cyclists reverse direction each minute?
Solution. Cyclists reverse direction on a collision, so we seek the expected number of collisions. The primary observation is to note that each cyclist can be viewed as simply passing through opposing cyclists – when a cyclist A and their bumper B reverse directions, if we label A 'B' and B 'A' then the situation looks like the cyclists passed through each other. Then it is clear that in this pass through version, each cyclist passes through $\frac{N-1}{2}$ of the other cyclists in expectation (the ones facing the opposite direction.) They do this twice per revolution, so each cyclist does their little phasing trick $N-1$ times per revolution in expectation. Finally, note the question asks "number of times cyclists reverse direction each minute" so depending on interpretation we can answer $\boxed{N(N-1)}$ (every instance of a turn around, counting a phase through twice for the perspective of each cycler being called a "time") or $\boxed{\frac{(N)(N-1)}{2}}$ if we take the meaning of time in the literal sense; as in periods of time of length 0, counting each phase-through once.

---

8 ♡ N cyclists are biking around a circular track at a rate of 1 rpm. They start in uniformly random locations and random orientations. Just before a pair would collide, they instantly reverse direction. What is the probability all cyclists are where they started after a minute?

Solution. I quite liked this problem. Read the solution to 7♡ to understand a bit of machinery surrounding this problem; namely, that if we seek to analyze properties of this system that don't depend on us tracking the identities of cyclists, we can just act as if the cyclists phase through each other. We call this the "phasing lemma".

However, the property we seek does depend on cyclists' identities. We make some observations anyways using the phasing lemma anyways.

Note that by phasing lemma, the ending configuration of cyclists ignoring their labels is equivalent to the initial configuration.

Now, also note that any cyclist in the actual configuration (not the configuration equivalent-by-phasing-lemma to our real configuration) can never cross through another cyclist. That is, if a cyclist starts between two others, they will end between those two cyclists. Formally, suppose the cyclists are initially numbered $C_0, \ldots, C_{n-1}$ counterclockwise and starting from the rightmost point on the circle – that is, by increasing theta if we imagine the cyclists initial positions to be written in polar form. Then if cyclist C_0 ends up where C_i started, then cyclist C_j must end up where cyclist $C_{j+i \mod N}$ started.

If all cyclists are facing the same way, then it is trivial that they all end up at their starting position. Now, suppose that all cyclists do not face the same way.

Then, let us call D_{ri} the time the cyclist C_i spends moving clockwise and D_{li} the time the same cyclist spends moving counterclockwise. We have that

$$D_{ri} \equiv D_{li} \mod 1, D_{ri} + D_{li} = 1$$

(when our problem requirement holds, the cyclist ends where they starts and moves the entire minute) and $D_{ri}, D_{li} > 0$ (every cyclist makes at least one collision, and there is no case where a cyclist experiences two collisions simultaneously. Every cyclist makes at least one collision as follows: find the cyclist C' that supposedly does not make a collision, note that another cyclist starts facing the opposite direction, and then by the phasing lemma some cyclist will cross the real path of C'.) These together imply that

$$D_{ri} = D_{li} \forall i.$$

Now, this means that the average over all i of $D_{ri} - D_{li}$ is 0. But now by phasing lemma, this average is also $R - L$, where L is the number of cyclists initially facing counterclockwise and R is the number of cyclists initially facing clockwise. So we have that $L = R$ is a necessary condition for our result to hold when all cyclists are not initially monodirectional.

But now finally, we show this condition is also sufficient.

AFSOC that $L = R$ but not all of the $D_{ri} - D_{li}$ are 0.

Thm: Of the $D_{ri} - D_{li}$ terms, *they all have the same sign.*

Recall from earlier that the cyclists' ending positions are fixed by a single cyclist's ending position. Then, (suppose) C_0 ended up where C_i was. Then C_k ends up where $C_{k+i \mod n}$ ended up $\forall k$. Let x_k be the initial position (theta divided by 2π) at which C_k is located. Then note that C_k ends at x_{k+i} and starts at x_k, and so we either have that $D_{lk} - D_{rk} = x_i - x_0$ or $D_{rk} - D_{lk} = x_0 - x_i + 1$.

Now, for the proof. We proceed via contradiction (remember, this is nested within the outer contradiction that not al the $D_{ri} - D_{li}$ terms are 0.) Suppose not. Then, note there exist $C_a, C_{a+1 \mod n}$ two consecutive cyclists for which $D_{r(a+1)} > D_{l(a+1)}, D_{ra} < D_{la}$. That is, the first of the two went more counterclockwise to its destination and the second went more clockwise. We are guaranteed to have two such consecutive cyclists "pointing towards each other on average" via an easy contradiction argument (Hint – you are also guaranteed to have two pointing away from each other.) C_{a+1} must pass or at least arrive at C_a's starting position, and the same holds in reverse. It is now very evident (imagine the circle as now a line; neither of these points is permitted to get to their destination by moving in the other direction by construction) that both cyclists cannot arrive at the other's starting point (which is necessarily on the way to their destination by phasing lemma) while maintaining that the first is to the clockwise direction of the second. This is a contradiction, which shows that all the $D_{ri} - D_{li}$ terms have the same sign.

But now *this* is a contradiction as if all these terms have the same sign, their average cannot be 0. So finally, all the $D_{ri} - D_{li}$ terms are 0, and the initial configuration having an equal number of counterclockwise-facing and clockwise-facing cyclists to start is a sufficient condition to ensure all cyclists end up where they started.

Finally, this evaluates out to

$$\boxed{\frac{1}{2^{n-1}} \text{ if n odd and } \frac{1 + \binom{n-1}{\frac{n}{2}}}{2^{n-1}} \text{ if n even.}}$$

---

9 ♡ Cars of length 2 are parking on a street of length 100, with parking meters evenly spaced every length 1. Each minute, a car arrives and parks at a uniformly randomly open spot, aligning its front bumper with a parking meter. When no more cars can possibly fit, roughly what fraction of parking meters are unoccupied by a car?

Solution. Let F_n be the answer to the above question, with "100" replaced with n and "fraction" replaced with "number". So the answer is $\frac{F_100}{100}$. Let's case on the position on the first car that parks, just looking to find F_n. If it parks at the very left (with probability $\frac{1}{n-1}$, we are left with an expected F_{n-2} spaces unoccupied. If it parks one space over, we are left with $1+F_{n-3}$. In particular, if the first car's left endpoint is placed i spaces in (with i in $[0, n-2]$), then we are left with $F_i + F_{n-2-i}$. This occurs for each i with probability $\frac{1}{n-1}$, so

$$\begin{aligned} F_n &= \frac{1}{n-1}\sum_{i=0}^{n-2} F_i + F_{n-2-i} \\ &= \frac{2}{n-1}\sum_{i=0}^{n-2} F_i. \end{aligned}$$

Now, note

$$F_{n-1} = \frac{2}{n-2}\sum_{i=0}^{n-3} F_i,$$

so

$$F_n = \frac{2\sum\limits_{i=0}^{n-3} F_i}{n-1} + \frac{2F_{n-2}}{n-1} = \frac{(n-2)F_{n-1} + 2F_{n-2}}{n-1}.$$

At this point we may note that if we were more clever it may have been possible to derive this recurrence directly, but alas I could not.

Now, we look at the generating function for this recurrence, denoted $F(x)$. Here, we assume familiarity with generating fucntions and differentiating them in the interest of space.

$$\begin{aligned} F(x) &= F_0x^0 + F_1x^1 + \sum_{i=2}^{\infty} F_ix^i \\ &= x + \sum_{i=2}^{\infty} \frac{i-2}{i-1}F_{i-1}x^i + 2\sum_{i=2}^{\infty} \frac{F_{i-2}}{i-1}x^i \\ &= x + \sum_{i=2}^{\infty} F_{i-1}x^i - \sum_{i=2}^{\infty} \frac{1}{i-1}F_{i-1}x^i + 2\sum_{i=2}^{\infty} \frac{F_{i-2}}{i-1}x^i \\ &= x + xF(x) - \sum_{i=2}^{\infty} \frac{1}{i-1}F_{i-1}x^i + 2\sum_{i=2}^{\infty} \frac{F_{i-2}}{i-1}x^i \end{aligned}$$

We can look at the third term as follows: let (we henceforth denote functions just by the letter, with F denoting $F(x)$ and similar)

$$\begin{aligned} G &= \sum_{i=2}^{\infty} \frac{1}{i-1}F_{i-1}x^i \\ &= x(\frac{G}{x})', \\ G &= x\int \frac{F}{x}dx. \end{aligned}$$

Similarly, let

$$H = \sum_{i=2}^{\infty} \frac{F_{i-2}}{i-1}x^i,$$

$$(\frac{H}{x})' = \sum_{i=2}^{\infty} F_{i-2}x^{i-2} = F,$$

so

$$H = x \int F dx.$$

Then, going back to our initial expression,

$$F = x + xF - x \int \frac{F}{x} dx + 2x \int F dx,$$
$$\frac{F}{x} = 1 + F - \int \frac{F}{x} dx + 2 \int F dx,$$

and differentiating yields

$$\frac{xF' - F}{x^2} = F' - \frac{F}{x} + 2F,$$
$$xF' - F = x^2 F' - xF + 2x^2 F,$$
$$\implies \frac{F'}{F} = \frac{2x^2 - x + 1}{x - x^2},$$
$$\implies -\frac{dF}{F} = \frac{2x^2 - x + 1}{x^2 - x} dx.$$

Integrating and long dividing yields

$$-\log F + C = \int (2 + \frac{x+1}{x^2 - x}) dx,$$

and partial fractions yields

$$-\log F + C = \int (2 + \frac{2}{x-1} - \frac{1}{x}) dx.$$
$$\implies \log F + C = \log x - 2\log(x-1) - 2x,$$

$$\implies CF = \frac{e^{-2x} x}{(x-1)^2},$$

or (multiplying together the Taylor series of the components and convolving,)

$$F = C \sum_{i=0}^{\infty} x^i \sum_{j=0}^{i-1} \frac{(-2)^j}{j!} (i - j).$$

Then $F_i = C \sum_{j=0}^{i-1} \frac{(-2)^j}{j!} (i-j)$, and we can quickly see that plugging in $i = 1$ that $C = 1$. So

$$F_i = \sum_{j=0}^{i-1} \frac{(-2)^j}{j!} (i - j).$$

In particular, our solution is $F_{100}/100 =$

$$\frac{1}{100} \sum_{j=0}^{99} \frac{(-2)^j}{j!} (100 - j).$$

Let's instead consider the solution for the proportion $P_i = F_i/i$:

$$P_i = \frac{1}{i} \sum_{j=0}^{i-1} \frac{(-2)^j}{j!} (i - j),$$

$$= \sum_{j=0}^{i-1} \frac{(-2)^j}{j!} - \frac{1}{i} \sum_{j=0}^{i-1} \frac{(-2)^j}{(j-1)!},$$

and we may note that the terms of both summations start getting quite small in magnitude quickly. So we can extend the summations to infinity without adding too much,

$$\approx \sum_{j=0}^{\infty} \frac{(-2)^j}{j!} - \frac{1}{i}\sum_{j=0}^{\infty} \frac{(-2)^j}{(j-1)!},$$

$$\approx e^{-2} - \frac{1}{i}(1 + -2e^{-2})$$

$$\approx \boxed{e^{-2} - \frac{1}{100}(1 - 2e^{-2})}.$$

---

10 ♡
Solution.

J ♡
Solution.

Q ♡
Solution.

K ♡
Solution.

A ♠ Two crabs are facing each other on the beach. Each second they randomly walk one step to the right or one step to the left. After 6 seconds, what is the probability they are still across from each other?
Solution. This is

$$\frac{\sum_{i=0}^{6} \binom{6}{i}^2}{2^{12}},$$

and here we will include a non-combinatorial way to derive

$$\sum_{i=0}^{n} \binom{n}{i}^2$$

for the reader.

Note that this is the coefficient of x^n in $(x+1)^{2n} = \sum_{i=0}^{n} \binom{n}{i} x^i \sum_{i=0}^{n} \binom{n}{i} x^{n-i}$. But

$$(x+1)^{2n} = \sum_{i=0}^{2n} \binom{2n}{i} x^i,$$

and the coefficient of x^n is $\binom{2n}{n}$.

With that aside out of the way, we end up with

$$= \binom{12}{6}/2^{12} = \boxed{\frac{231}{1028}}$$

---

2 ♠ In a random binary sequence, which is more likely to come first? 00 or 01? What is the expected time to see 00? 01? What about 001?

Solution. Unless we instantly roll BB, we are guaranteed to roll RB first. So BB comes first 1/4 of the time, and $\boxed{RB}$ is most likely to arrive first.

To see BB we have a simple states problem. We solve:

$$E = \frac{1}{4}(2) + \frac{1}{2}(E+1) + \frac{1}{4}(E+2).$$

(These are the cases: BB, (R), BR.) This yields $\boxed{E(t_{BB}) = 6}$ as the average time to see BB.

To see RB we have

$$E = \frac{1}{2}(E+1) + \frac{1}{2}(3),$$

which yields $\boxed{E(t_{RB} = 4}$ as the average time to see BB. (We can also see this as the average time to see a red, which is 2, summed with the average time to see a blue after, which is also 2, by LOE.)

For BBR, we have

$$E = \frac{1}{2}(E+1) + \frac{1}{4}(E+2) + \frac{1}{4}(4),$$

and so

$$\boxed{E(t_{BBR}) = 8}$$

Putting this here: https://martingalemeasure.wordpress.com/2014/02/02/monkey-typing-abracadabra-14/

I read this 2 years ago, didn't fully understand it, and have been looking for it ever since. Pretty important moment for me.

---

3 ♠
Solution.

---

4 ♠ A bunny cannot decide where to take a nap. There is a rose bush 4 steps to the left and sunflowers 8 steps to the right. The bunny is 25% to step left and 75% to step right. How many steps will it take in expectation until the bunny can get some rest?

Solution. Approaching this problem with states sounds annoying. Instead, we solve the more general problem via function equation / martingale techniques. First, we phrase the more general problem:

A random walk with probability p of moving right begins at point a on the number line,

and terminates upon reaching 0 or L. Find the expected number of steps taken.

Our FE is

$$f(x) = 1 + pf(x+1) + (1-p)f(x-1)$$

subject to $f(0) = f(L) = 0$.

I spent a good amount of time working on this FE, and a quick write-up cannot do it justice. (To be fair, time spent on an FE is infinitely more enjoyable than time spent on a states problem.)

First, we note that this is almost the gamblers ruin FE, with martingale-solution

$$(\frac{1-p}{p})^x.$$

A quick look reveals that this is not a solution (duh). The boundary conditions, for one, are doomed. We need to subtract 1, and a natural way of doing this is to add ax such that $pa(x+1) + (1-p)a(x-1) = -1 + ax$ (we isolate the '1' in our FE and try to add a linear term to get rid of it) or

$$(2p-1)(a) = -1 \implies a = \frac{-1}{2p-1}.$$

Now, we have a better attempt at a solution:

$$(\frac{1-p}{p})^x - \frac{1}{2p-1}x.$$

Importantly, it is not clear how to use the boundary conditions. Our "solution" doesn't even reference L! I was stumped for a while until my friend mentioned another functional equation – the same FE, but with only the boundary condition at x=0 – and that it was "underconstrained". This led to me realize that subtracting 1 from our existing solution was a "valid solution" for some L – that is,

$$(\frac{1-p}{p})^x - \frac{1}{2p-1}x - 1$$

is a martingale and satisfies $f(0) = 0$, and probably satisfies $f(x) = 0$ for some $x \neq 0$. Now, this x may not be an integer, so in the strictest sense this might be a solution with no physical interpretation for an invalid L. But this naturally leads to the question of how to incorporate our "free variable". A bit of thinking yields

$$f(x) = c(L)((\frac{1-p}{p})^x - 1) - \frac{1}{2p-1}x.$$

This is basically unfindable without first underconstraining the problem, as $c(L)$ is a complicated function in L. Indeed, we have that

$$c(L) = \frac{1 + \frac{L}{2p-1}}{(\frac{1-p}{p})^L - 1},$$

and our final solution is

$$f(x) = \frac{1 + \frac{L}{2p-1}}{(\frac{1-p}{p})^L - 1}((\frac{1-p}{p})^x - 1) - \frac{1}{2p-1}x.$$

Finally, lets look at our particular case. We have $L = 12$ and $p = \frac{3}{4}$ for a solution of

$$\boxed{16\frac{4593}{6643}}$$

TODO: check this

---

5 ♠ A beaver is looking for materials for her dam. There are a wide mountain and a wide river 20 steps apart and parallel. The beaver randomly steps in 1 of the 4 cardinal direction on her search. If a step would have taken her up the mountain, she instead goes towards the river. If she is at the mountain, how many steps can she expect it to take to return to the river?

Solution. Let the river be to the north at $y = 20$ and let us start at $y = 0$. Let us do away with the mountain and instead just have two rivers 40 steps apart at $y = 20, y = -20$. (That is, this new problem is exactly the same as the problem with a mountain and a river, and the implicit bijection is hopefully evident; every step starting or ending below y=0 in our new problem is mirrored in the initial problem.) This is to do away with boundary conditions at $y = 0$.

Clearly, the expected time to one of the rivers is only dependent on our current y value. We have a system of states equations for states $E(y)$ for all y in $[-19, 19]$:

$$E(y) = 1 + \frac{1}{2}E(y) + \frac{1}{4}E(y+1) + \frac{1}{4}E(y-1),$$

and now we apply the classical technique in such situations where when we have large symmetric states equations: we treat it as a functional equation in y instead and searching for $E(y)$. (If we find a function satisfying the functional equation and that $E(-20) = E(20) = 0$, then by definition the function will simultaneously satisfy all states equations.)

Now,

$$\frac{1}{2}E(y) = 1 + \frac{1}{4}E(y+1) + \frac{1}{4}E(y-1),$$
$$\implies E(y) = 2 + \frac{1}{2}E(y+1) + \frac{1}{2}E(y-1).$$

One may recall that the simple random walk has martingale $Y^2 - t$, and corresponding solution to functional equation $-y^2$. In fact, $-2y^2$ is a satisfactory solution to this new functional equation. We can add any constant and keep our solution a solution, so we add 800 so the function is 0 at $20, -20$ to yield

$$E(y) = 800 - 2y^2 \implies E(0) = \boxed{800}.$$

---

6 ♠ A chipmunk is stuffing its mouth with acorns. It starts with 6 and every second, it adds a nut with probability $\frac{2}{3}$ and drops one with probability $\frac{1}{3}$. What is the probability it stuffs its mouth with 36 acorns before it drops all of its acorns?
Solution.

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

is our martingale. Then, letting P be the probability of holding 36 acorns,

$$\frac{1}{2}^6 = P\frac{1}{2}^{36} + (1-P)\frac{1}{2}^0$$
$$\implies \frac{1}{2^6} = P\frac{1}{2^{36}} + 1 - P,$$

and

$$P = \frac{1 - \frac{1}{2^6}}{1 - \frac{1}{2^{36}}} = \boxed{\frac{2^{36} - 2^{30}}{2^{36} - 1}}.$$

---

7 ♠ A crow and a raven are each picking up coins off the street at the same pace. They are clever and keep track of the number of heads they picked up. They each pick up N coins total. What is the chance they never collected the same number of heads after starting? **Solution.**

Let the crow and raven pick up coins in order, so the Crow's first coin precedes the Raven's first coin precedes the Crow's second coin .... There are $2N$ coins picked up in this way.

The key observation is that we can view the sequence of 2N coins picked up as a random up-right walk of length $2N$, where we move up if either the current coin is heads and being picked up by raven or tails and being picked up by crow, and move right if the current coin is tails and picked by raven or heads and picked by crow. Note that this new random walk meets $y = x$ precisely when the number of heads by both birds is equal.

Now, we simply seek the likelihood that a random walk of this sort doesn't meet $y = x$. Such a walk must end at $(2N - A, A)$ for $A \in [0, 2N]$. We compute the probability that we have such a walk for $A \in [0, N-1]$ and then multiply by 2. (Note paths ending at [N,N] can't fulfill the criteria.)

First, we count the number of paths beginning with a right move that end at $(2N - A, A)$.

For some A, the number of paths of this form is

$$\binom{2N-1}{A}.$$

Now, we make the following observation: for all A in our desired range, there exists a bijection between paths to $(2N - A, A)$ passing through $y = x$ and paths to $(2N - A + 1, A - 1)$ (the term "paths" here refers to paths starting with a right move.) Just flip every step made in a path of the former variety before and including the first step onto $y = x$ and we end up with a path to $(2N - A + 1, A - 1)$ (up/right gets swapped.)

Then, we can view the number of paths to $(2N - A, A)$ for A in the desired range as

$$\sum_{A=0}^{N-1} A_A - B_A,$$

where A_A denotes the number of right-beginning paths to $(2N-A, A)$ and B_A denotes the number of paths of the above variety passing through $y = x$. But our bijection above says

$$B_A = A_{A-1},$$

and so the sum simply evaluates to

$$A_{N-1} - A_{-1} = \binom{2N-1}{N} - 0.$$

Doubling this yields

$$2\binom{2N-1}{N} = \frac{(2N-1)!(2N)}{(N!)(N-1)!(N)} = \binom{2N}{N},$$

and so the probability of such a path being valid is the number of such valid paths over the total number of possible paths,

$$= \boxed{\frac{\binom{2n}{n}}{4^n}}$$

---

8 ♠ B blue jays and C cardinals are sitting together in a tree. The birds take turns flying away. What is the likelihood that there is a time when the same number of blue jays and cardinals have flown away?
Solution. $t = 0$ doesn't count.

We look for the probability that there is no such time. We call the event that there is no such time 'N' (for NOT happened.). WLOG $B > C$. Note if a C goes first then there is no chance of N occurring. So it suffices to find the number of paths from $0, 0$ to (B, C) that stay strictly underneath the $B = C$ diagonal except for the beginning and then divide by $\binom{B+C}{B}$, and then finally take the complement.

If you look at the Wikipedia page for catalan numbers, you will see the now-standard argument to be used in these situations (search for 'bad paths'.) The idea is that the first step must be in the positive B direction. Then, we must stay below or on the $C = B - 1$ diagonal. There is a bijection between paths (starting at (1,0) that cross the $C = B - 1$ diagonal and touch the $C = B$ diagonal before arriving at (B,C) (therefore not being part of our event N) and paths from $(1, 0)$ to (C-1,B+1), by taking all up moves and right moves after the first incidence with the line $B = C$ and swapping them. So there are

$$\binom{B+C-1}{C-2}$$

bad paths that cross the diagonal,

$$\binom{B+C-1}{B-1}$$

paths in total that start at (1,0) and end at (B,C), and so there are

$$\binom{B+C-1}{B-1} - \binom{B+C-1}{C-2}$$

good paths. Finally, there are

$$\binom{B+C}{C} - (\binom{B+C-1}{B-1} - \binom{B+C-1}{C-2})$$

paths that cross the B=C diagonal at some point, and the desired probability is

$$\boxed{1 - \frac{\binom{B+C-1}{B-1} - \binom{B+C-1}{C-2}}{\binom{B+C}{C}}}$$

---

9 ♠
Solution.

---

10 ♠
Solution.

What is the probability a 2d random walk returns to the origin?

Solution. First, we find the expected number of returns to the origin.

Let R_n denote the event that the walk returns to the origin after $2n$ steps (possibly not for the first time.) Also, note that the walk can only return to the origin after an even number of steps.

Then of these 2n steps, we need 'a' lefts-and-rights and 'b' ups-and-downs such that $a + b = n$. In other words, we have the number of walks of length $2n$ that return to the origin after 2n steps is

$$\sum_{a+b=n} \binom{2n}{a, a, b, b},$$

where we use a multinomial coefficient to make clear that we pick a lefts, a rights, b ups, and b downs in our path.

$$= \sum_{a+b=n} \frac{(2n)!}{a!a!b!b!}.$$

$$= \sum_{a+b=n} \binom{2n}{n} \frac{n!^2}{a!a!b!b!},$$

motivated by the fact that this substitution matches each $a!b!$ term in the denominator with an $n!$ term in the numerator,

$$= \binom{2n}{n} \sum_{a+b=n} \binom{n}{a, b}^2,$$

and now the multinomial coefficient is a little silly.

$$= \binom{2n}{n} \sum_{i=0}^{n} \binom{n}{i}^2 = \binom{2n}{n}^2,$$

and the probability of return is

$$= \frac{\binom{2n}{n}^2}{4^{2n}}.$$

The expected number of returns is just (by LoE) the sum from n=1 to infinity of this term, which is

$$\sum_{n=1}^{\infty} \frac{\binom{2n}{n}^2}{4^{2n}}$$

Stirling's approximation yields

$$\sqrt{2\pi n}(\frac{n}{e})^n \leq n! \leq \sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{1}{12n}},$$

so

$$\sum_{n=1}^{\infty} \frac{((2n!)/(n!)^2)^2}{4^{2n}} \geq \sum_{n=1}^{\infty} \frac{(\sqrt{2\pi(2n)}(\frac{2n}{e})^{2n}/(\sqrt{2\pi n}(\frac{n}{e})^n e^{\frac{1}{12n}})^2)^2}{4^{2n}}$$

$$= \sum_{n=1}^{\infty} \frac{4\pi n 2^{4n}(\frac{n}{e})^{4n}}{4\pi^2 n^2 (\frac{n}{e})^{4n} e^{\frac{1}{3n}} 4^{2n}} = \sum_{i=1}^{n} \frac{1}{\pi n e^{\frac{1}{3n}}} \geq \sum_{i=1}^{n} \frac{1}{\pi n e^{\frac{1}{3}}},$$

which diverges, so the expected number of returns is infinite.

Then, let the probability of return be p. Then the expected number of returns E is

$$E = p(1 + E)$$

(with probability p, you return and then get E more returns in expectation. If you don't return, you get no returns,) or

$$p = \frac{E}{1+E}, E = \frac{p}{1-p}.$$

It's clear now that $\boxed{p = 1}$. The argument is direct when E is noninfinite; in any case, we can see that $p \in [0, 1]$, and $p \in [0, 1)$ results in a contradiction, justifying our intuition for the division by 0 or limit shenanigans a little more rigorously.

J ♠
Solution.

Q ♠
Solution.

K ♠
Solution.