



Jerry Qin

Personal Website and Blog

Blog

About

Categories

Projects

Contact



© Jerry Qin 2021

A Working List of Probability Questions: Week 1

In Preparation For Quantitative Trading Interviews

June 21, 2020 | Updated June 28, 2020

This series is mostly for my own self-documentation purposes as I prepare for recruiting at the end of summer. I'll be updating this page with 5 new problems/brain teasers and solutions every day to keep myself accountable. These problems will come from a variety of sources, including *A Practical Guide to Quantitative Finance Interviews*, Math Stack Exchange, and UC Berkeley's EECS 126 [course notes](#). You can skip to my week one review [here](#) or navigate to a specific day using the legend below. Thanks for tuning in!

[Day 1](#) | [Day 2](#) | [Day 3](#) | [Day 4](#) | [Day 5](#) | [Day 6](#) | [Day 7](#)

Day One: June 21

Today's problems are relatively simple as I spent a while figuring out the LaTeX formatting on websites.

Problem 1. What is the expected number of tosses to get three consecutive heads when flipping a fair coin?

Solution. Let $\mathbb{E}[3]$ denote the expected number of rolls for 3 consecutive heads. If we start rolling and roll a tail first, which has probability $\frac{1}{2}$, then the expected number of tosses will be $\mathbb{E}[3] + 1$. If we roll a head then a tail, which has probability $\frac{1}{4}$, then the expected number will be $\mathbb{E}[3] + 2$. If we roll 2 heads then a tail with probability $\frac{1}{8}$ then expected value is $\mathbb{E}[3] + 3$. Finally if we roll 3 heads consecutively, which has probability $\frac{1}{8}$, then $\mathbb{E}[3] = 3$. Combining these together, we have

$$\mathbb{E}[3] = \frac{1}{2}(\mathbb{E}[3] + 1) + \frac{1}{4}(\mathbb{E}[3] + 2) + \frac{1}{8}(\mathbb{E}[3] + 3) + \frac{1}{8}(3)$$

Solving this equation gives $\mathbb{E}[3] = 14$. We can extend this to find the expected number of rolls for n consecutive heads to get $\mathbb{E}[n]$:

$$\mathbb{E}[n] = \frac{1}{2}(\mathbb{E}[n] + 1) + \frac{1}{4}(\mathbb{E}[n] + 2) + \cdots + \frac{1}{2^n}(\mathbb{E}[n] + n) + \frac{1}{2^n}(n)$$

Solving for this in general gives us the formula $\mathbb{E}[n] = 2(2^n - 1)$. A number of other solutions of this type of problem can be found [here](#).

Problem 2. Suppose there are five pirates with 100 pieces of gold who are voting on splitting up the gold. The most senior pirate will propose a way to distribute the gold and requires 50% of the votes in order to pass his proposal. If he receives less than 50% of the votes then he will be killed and the process will begin with the next most senior pirate. How will the gold coins be divided?

Solution. Let's label each pirate in ascending order based on seniority, so pirate 5 will be most senior and 1 will be least senior. Building this up from the two pirate case: Pirate 2 can safely propose he take all 100 gold pieces as he constitutes 50% of the votes. Then Pirate 1 receives nothing.

Moving on to the three pirate case—Pirate 3 knows that if his plan is voted down, then Pirate 2 can take all 100 coins and Pirate 1 will receive nothing. All he needs to do is give Pirate 1 one coin to secure his vote. Adding a fourth pirate, Pirate 4 knows that if his plan

is voted down, then Pirate 2 will get nothing. Then all he needs to do is propose to offer Pirate 2 one coin and keep the remaining 99. Now considering all five pirates, Pirate 5 knows that if his plan is voted down, Pirate 1 and Pirate 3 will receive nothing. Then all he needs to do is offer 1 gold coin each to Pirate 1 and Pirate 3, keeping the 98 coins.

We can extend this to any number n pirates (so long as $n \leq 200$). If n is odd, the most senior pirate will offer Pirates 1, 3, \dots , $n - 2$ each one coin and keep the remainder for himself. If n is even, then he will offer Pirates 2, 4, \dots , $n - 2$ one coin each and keep the remainder.

Problem 3. For a best of 3 set tennis match, would you bet on it finishing in two sets or three sets given that probability of winning a set is constant?

Solution. Let $p = P(\text{player 1 wins})$ and $q = P(\text{player 2 wins})$. To win in two sets, we can have player 1 or player 2 win two consecutive sets, giving a total probability of $p^2 + q^2$. In order for the match to end in three sets, one player must win 2 sets (assume that's player 1) and the other one set. This corresponds with a probability of $2pq$. Note that $(p - q)^2 \geq 0$, implying that $p^2 + q^2 \geq 2pq$, so I would bet on the match finishing in two sets.

Problem 4. You break a stick of length 1 at two points uniformly. What is the probability that the three resulting pieces can form a triangle?

Solution. A triangle is possible to form so long as one piece is not longer than the sum of the other two (one side $> \frac{1}{2}$) and the difference between two pieces is not longer than the third. Consider the midpoint M of the original stick: if both cuts are made on the same side of M (probability $\frac{1}{2}$), then one side will have length greater than $\frac{1}{2}$ and a triangle is not possible. Let's say we break the stick at two points x and y with $x < y$ and x and y are on different side of M . There is probability $\frac{1}{2}$ that x is farther left on its half of the stick than y is right on its half or the vice versa. In this case, we can't form a triangle either because the difference between two sides will be longer than the third. Then only $(1 - \frac{1}{2}) \cdot (1 - \frac{1}{2}) = \frac{1}{4}$ of breaks can form a triangle (my justification of this seems a bit shady and I'll work on a more analytical solution another time).

Problem 5. Consider a game where there are two blue balls and two red balls in a bag. Each ball is removed from the bag without replacement and if you guess the color of the ball correctly you earn a dollar. How much would you pay to play this game?

Solution. On your first draw you have a $\frac{2}{4}$ chance of guessing correctly. On the second draw, you have a $\frac{2}{3}$ chance of guessing correctly so long as you guess the color opposite

of the first ball. By the third ball, either you have selected two balls of the same color with probability $\frac{1}{3}$ or you selected two balls of different color with probability $\frac{2}{3}$. In the first case you are guaranteed to know the colors of the next ball and in the second you have a 50% chance of guessing the next ball correctly. Finally, the probability of guessing correctly on the last ball is 1. Then calculating the expected value of the game from our individual draws:

$$\mathbb{E}[x] = \left(\frac{2}{4} \cdot 1\right) + \left(\frac{2}{3} \cdot 1\right) + \left(\frac{1}{3} \cdot 1 + \frac{2}{3} \cdot \frac{1}{2} \cdot 1\right) + 1 = \frac{17}{6}$$

So I would pay \$ $\frac{17}{6}$ to play this game.

Day Two: June 22

Today's problems focus on stochastic processes and Markov chains. A couple of these problems come from David Aldous's STAT 150 page. I also came across a pretty sick [compilation](#) of different resources this morning.

Problem 1. A calculating machine uses the digits 0 and 1 and transmits one of these digits through several stages. At every stage, there is a probability p that the digit will be changed when it leaves and probability $q = 1 - p$ that it will not. Suppose the machine starts at state 0. What is the probability that the machine will display the correct digit after 2 stages?

Solution. We can identify the Markov chain for this problem as below:

$$\begin{array}{cc} & \begin{matrix} 0 & 1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \end{matrix} & \begin{pmatrix} q & p \\ p & q \end{pmatrix} \end{array}$$

which produces the transition matrix

$$P = \begin{pmatrix} q & p \\ p & q \end{pmatrix}$$

We are interested in finding the probability that the state remains in state 0 after two time periods, which is indexed by the $(1, 1)$ -entry of P^2 . Then,

$$P^2 = \begin{pmatrix} q & p \\ p & q \end{pmatrix} \begin{pmatrix} q & p \\ p & q \end{pmatrix} = \begin{pmatrix} q^2 + p^2 & 2qp \\ 2pq & p^2 + q^2 \end{pmatrix}$$

and our desired probability is $q^2 + p^2$.

Problem 2. A transmitter is sending binary code (+ and - signals) that must pass through three relay signals before being sent on to the receiver. At each relay station, there is a 25% chance that the signal will be reversed. Suppose + signals make up 60% of the messages sent. If a + signal is received, what is the probability that a + was sent?

Solution. This was a problem in the notes of my statistics class last semester and was solved using a tree diagram. I decided to take a stochastic approach:

Let A be the event that a + was transmitted and B be the event that + is received after it passes through the third relay. We're interested in finding $P(A | B)$, which by Bayes' Theorem can be written as

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)} = \frac{P(B | A)P(A)}{P(B | A)P(A) + P(B | A^c)P(A^c)}$$

Note that A^c is the event that a - was transmitted. Let's consider the Markov chain for this problem:

	+	-
+	0.75	0.25
-	0.25	0.75

giving us the transition matrix

$$P = \begin{pmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{pmatrix} \rightarrow P^3 = \begin{pmatrix} 0.5625 & 0.4365 \\ 0.4365 & 0.5625 \end{pmatrix}$$

We can see that $P(B | A) = P_{1,1}^3 = 0.5625$ and $P(B | A^c) = P_{2,1}^3 = 0.4365$. Then, we have

$$P(A | B) = \frac{0.5625 \cdot 0.6}{0.5625 \cdot 0.6 + 0.4365 \cdot 0.4} = 0.6585$$

Problem 3. Consider a Markov chain with two states: 1 and 2. If $p_{1,2} = a$ and $p_{2,1} = b$. For which values of a and b do we have an absorbing Markov chain?

Solution. An absorbing Markov chain is one in which every state can reach a state where, once entered, cannot be left. Our transition matrix looks something like this:

$$P = \begin{pmatrix} ? & a \\ b & ? \end{pmatrix}$$

a and b are the probabilities of state 1 changing to 2 and vice versa. In order for a state to remain constant between time periods, we must have at least one of a or b be 0 so that upon reaching one of the states, we will leave it.

Problem 4. A Markov chain with the state space $\{1, 2, 3\}$ has the transition matrix

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0 & 1/2 & 1/3 \\ 0 & 0 & 1 \end{pmatrix}$$

Starting from each state, find the expected time until absorption occurs.

Solution. Clearly, absorption happens at state 3. Let $\mathbb{E}[i]$ denote the expected time to reach state 3 from $i = 1, 2, 3$. Then, we have

$$\begin{aligned} \mathbb{E}[1] &= 1 + \frac{1}{3}\mathbb{E}[1] + \frac{1}{3}\mathbb{E}[2] + \frac{1}{3}\mathbb{E}[3] \\ \mathbb{E}[2] &= 1 + \frac{1}{2}\mathbb{E}[2] + \frac{1}{2}\mathbb{E}[3] \\ \mathbb{E}[3] &= 0 \end{aligned}$$

Walking through the case of $\mathbb{E}[1]$, we get the recursion in the following manner: if we go forward one time period, then there is a $\frac{1}{3}$ chance that we are still at state 1 and our expected time is now $1 + \mathbb{E}[1]$. Similarly, we have a $\frac{1}{3}$ chance of being at state 2 or 3 respectively, which brings us to $\mathbb{E}[1] = 1 + \frac{1}{3}\mathbb{E}[1] + \frac{1}{3}\mathbb{E}[2] + \frac{1}{3}\mathbb{E}[3]$.

$$\text{Solving these equations gives us } \mathbb{E}[i] = \begin{cases} 5/2 & \text{if } i = 1 \\ 2 & \text{if } i = 2 \\ 0 & \text{if } i = 3 \end{cases}$$

Problem 5. A group of quants wants to determine their average salary but want to do so such that no person has access to anyone else's salary. Can they accomplish this, and, if so, how?

Solution. The quants can do this by having the first quant taking a random number and adding his salary to it and passing that number onto the second quant. Then each subsequent quant will add their salary to the sum and tell the number to the next quant.

Finally, the first quant removes the random number they chose and divides by the number of total quants for the average salary.

Day Three: June 23

I threw in some statistics questions for today's problems in addition to the usual bunch.

Problem 1. There are N lions and 1 sheep in a field. All the lions really want to eat the sheep, but the problem is that if a lion eats a sheep, it becomes a sheep. A lion would rather stay a lion than be eaten by another lion. There is no other way for a lion to die than to become a sheep and then be eaten. When is it safe for any lion to eat?

Solution. Let's consider this similarly to Problem 2 of Day 1. If there is one lion, then it can happily eat the sheep. If there are two lions, no lion would eat the sheep for fear of getting eaten by the other lion. If there are three lions, then a lion eating the sheep would reduce us to the two lion scenario, in which no lion would eat the sheep. We can continue on this train of thought for any number N and conclude that the sheep will be eaten when N is odd. I actually think this problem is actually quite similar to the game of Nim or other impartial combinatorial games in the sense that we're thinking in terms of a recursion.

Problem 2. You draw cards from a deck without replacement. What is the expected number of cards you draw before you get an ace?

Solution. I had to refer to Stack Exchange for this question. You can find the solution [here](#). I was initially tempted to use μ_x for a negative binomial distribution but I later realized that cards are drawn without replacement. Let D be the r.v. that represents the number of draws before we draw an ace and assign each of the non-ace cards in the deck a number between 1 and 48. Then we can define the random indicator variable $X_i = 1$ if the card with label i was drawn before any ace and 0 otherwise. Then, we have that

$$D = X_1 + X_2 + \cdots + X_{48}$$

From linearity of expectation, we have

$$\mathbb{E}[D] = \mathbb{E}[X_1 + X_2 + \cdots + X_{48}] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_{48}]$$

Let's consider $P(X_1 = 1)$. If we look card 1 along with the 4 aces, it's pretty simple to see that $P(X_1 = 1) = \frac{1}{5}$. It follows that $\mathbb{E}[X_1] = \frac{1}{5}$. Symmetry guarantees that all the

X_i have the same distribution so $P(X_i = 1) = P(X_j = 1)$ for all $i, j \in [1, \dots, 48]$. Then each $\mathbb{E}[X_i] = \frac{1}{5}$ and $\mathbb{E}[D] = \frac{48}{5}$.

Problem 3. You roll a die up to three times. You can decide to stop and choose the number on the die (where the number is your payoff) during each roll. What's your strategy?

Solution. Trading firms often ask these types of questions, where you have to devise an optimal strategy to a given game with uncertainty. Work backwards from the end of the game, we know that the expected value of the last roll is $\frac{7}{2}$. Then, for our second to last roll, it makes sense to reroll if we rolled a 1, 2, 3, and just keep the money and quit if we rolled a 4, 5, 6. Then, the expected value of our last two rolls is $\frac{1}{6}(4 + 5 + 6) + \frac{1}{2} \cdot \frac{7}{2} = 4.25$ (it should make sense why we multiply each EV by $\frac{1}{2}$). Using similar logic to before, it makes sense to reroll if we rolled a 1, 2, 3, 4 and keep the money if we rolled a 5 or 6. Then, our strategy is to roll once if we roll a 5 or a 6, reroll otherwise. When rerolling, we'll keep anything 4 and beyond and reroll our final time otherwise.

Problem 4. A hat contains 100 coins, 99 of them are guaranteed to be fair and 1 that has a $\frac{1}{2}$ chance to be double-headed. A coin is randomly chosen at random and flipped 7 times. If it landed heads each time what is the probability that one of the coins is double-headed?

Solution. Let's use Bayes' Theorem for this question. Let A be the event that one of the 100 coins is double-headed and B be the event that a randomly chose coin lands heads on all 7 flips. We're interested in finding $P(A | B)$. Then,

- $P(A) = \frac{1}{100}$
- $P(A^c) = \frac{99}{100}$
- $P(B | A) = P(\text{all head flips given unfair coin}) = \frac{1}{2^7}$
- $P(B | A^c) = P(\text{all head flips given no unfair coin}) = \frac{1}{2^7}$

The third bullet point can be solved for by breaking down the components of the probability of rolling all 7 heads in the case of having a guaranteed rigged coin, which is $\frac{1}{100} \cdot 1 + \frac{1}{2^7} \cdot \frac{99}{100}$. Then we have

$$P(A | B) = \frac{\left(\frac{1}{100} + \frac{1}{2^7} \cdot \frac{99}{100}\right) \cdot \frac{1}{2^7}}{\left(\frac{1}{100} + \frac{1}{2^7} \cdot \frac{99}{100}\right) \frac{1}{2^7} + \frac{1}{2^7} \cdot \frac{1}{2^7}} = 0.3194$$

So if we roll 7 heads in a row, there is a 32% chance that there is an unfair coin in the bunch.

Problem 5. Let $X, Y \sim N(0, 1)$ be i.i.d. random variables. What is the probability that $X > 3Y$?

Solution. This is a pretty simple problem that I'm using as statistics review. Let's define $Z = X - 3Y$. Then we simply need to find $P(Z > 0)$. We know that Z is normally distributed as it's a linear combination of standard normal variables so we only need to find the mean and standard deviation in order to find $P(Z > 0)$. Note that

$$\mathbb{E}[Z] = \mathbb{E}[X] - 3\mathbb{E}[Y] = 0$$

Also,

$$\sigma(Z) = \sigma(X - 3Y) = \sigma(X) + (-3)^2\sigma(Y) = 10$$

Note that since Z is normally distributed about 0, we technically don't even need $\sigma(Z)$ since $P(Z > 0) = \frac{1}{2}$ for $Z \sim N(0, \sigma)$. Then, $P(X > 3Y) = \frac{1}{2}$.

Day Four: June 24

I didn't get to this until pretty late this evening so I did some relatively simple questions for today.

Problem 1. In unprofitable times corporations sometimes suspend dividend payments. Suppose that after a dividend has been paid the next one will be paid with probability 0.9, while after a dividend is suspended the next one will be suspended with probability 0.6. In the long run what is the fraction of dividends that will be paid?

Solution. This is another Markov chain problem so we'll tackle it like we have before. We can call the two states of this chain State 1: "dividend paid" and State 2: "dividend suspended". We're provided that $p_{1,1} = 0.9$ and $p_{2,2} = 0.6$, which implies that $p_{1,2} = 0.1$ and $p_{2,1} = 0.4$. This gives us the transition matrix

$$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \end{pmatrix}$$

Let π be the stationary distribution for our Markov chain. In the steady state,

$$\begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix} \begin{pmatrix} 0.9 & 0.1 \\ 0.6 & 0.4 \end{pmatrix}$$

Then, $0.9\pi_1 + 0.6\pi_2 = \pi_1$ and $0.1\pi_1 + 0.4\pi_2 = \pi_2$. Note also the $\pi_1 + \pi_2 = 1$ be definition. Then, solving this system of linear equations gives us that $\pi_1 = 0.8$ and $\pi_2 = 0.2$ and

$$\pi = \begin{bmatrix} \pi_1 & \pi_2 \end{bmatrix} = \begin{bmatrix} 0.8 & 0.2 \end{bmatrix}$$

so 80% of dividends will be paid.

Problem 2. If you repeatedly flip a fair coin. What is the expected number of flips you need to do in order to get a head immediately followed by a tail?

Solution. We know that we'll end with HT so let's think about the tosses we make before this happens. If we roll a heads on our first try, then we have to roll only heads until we roll our final HT . For this reason, if there are m tosses before HT , then it has to be of the form $T^k H^{m-k}$ where $k = 0, 1, \dots, m$. Then let X be a random variable that describes the number of tosses until we get T . We can flip T on our first try ($p = \frac{1}{2}$), or we could flip heads first and we'd effectively start the process over again ($p = \frac{1}{2}$). Then,

$$\mathbb{E}[X] = \frac{1}{2} \cdot 1 + \frac{1}{2}(1 + \mathbb{E}[X])$$

Solving this gives us $\mathbb{E}[X] = 2$. Now define Y is the random variable that describes the number of tosses until we get our first HT . We can either roll a tails on our first try ($p = \frac{1}{2}$) and making it so that we have to start over again, or we can roll heads on our first try ($p = \frac{1}{2}$), and all we need to do now is roll at T , whose EV we just modeled with X . Then,

$$\mathbb{E}[Y] = \frac{1}{2}(1 + \mathbb{E}[Y]) + \frac{1}{2}(1 + \mathbb{E}[X])$$

We can plug in $\mathbb{E}[X] = 2$ and find that $\mathbb{E}[Y] = 4$! Other solutions to this can be found [here](https://jerryqin.com/posts/a-working-list-of-probability-problems).

Problem 3. Let X be an r.v. such that $\mathbb{E}[X^2] < \infty$. What constant c minimizes $\mathbb{E}[(X - c)^2]$?

Solution. I thought that this would be a difficult problem but then I realized that it was extremely simple. Taking the derivative with respect to c gives us

$$\begin{aligned}\frac{d}{dc}\mathbb{E}[(X - c)^2] &= \frac{d}{dc}\mathbb{E}[X^2 - 2cX + c^2] = \mathbb{E}[-2X + 2c] \\ &= -2\mathbb{E}[X] + 2c\end{aligned}$$

Setting this equal to 0 makes it pretty clear that $c = \mathbb{E}[X]$ minimizes the expression. You could also solve for this by finding the X that minimizes the expression by differentiating with respect to X and get the same conclusion but the question specified finding a c value. Also, it never hurts to plug this back in and make sure that it is actually a minimum.

Problem 4. What is the expected number of rolls of a fair die needed to get all six numbers?

Solution. This was a nice brush-up on geometric distributions. It takes one roll in order to roll out first number. After our first roll, the number of rolls until a second number is rolled is geometrically distributed with success probability $p = \frac{5}{6}$. Since $\mu_X = \frac{1}{p}$ for a geometric distribution, the expected number of rolls until a second number is rolled is $\frac{5}{6}$. Continuing in this fashion, the number of rolls before each additional unique number geometrically distributed with probability $p = \frac{n}{6}$ where n is the number of remaining numbers to be rolled. Then,

$$\mathbb{E}[6] = \sum_{i=1}^6 \frac{6}{i} = 14.7$$

We can extend this further and say that for an n -sided die,

$$\mathbb{E}[n] = n \sum_{i=1}^n \frac{1}{i}$$

Problem 5. You are on a game show, and there are 3 doors. Two of the doors conceal something worthless, and one door conceals a valuable prize. The game show host, Monty Hall, knows where the prize is. He lets you pick a door, then he opens one of the remaining two doors to reveal something worthless. He then offers you the chance to switch doors. Should you?

Solution. This is the Monty Hall Problem. Lets call the doors A, B, C . If you chose A , you have $p = \frac{1}{3}$ of guessing the prize, and $p = \frac{2}{3}$ of guessing incorrectly. This probability of A being correct does not change after the host offers to let you switch after he uncovers one of the other doors (lets say B). In allowing you to switch doors, the host is offering you the option to bet on or against whether your first guess was right with $p = \frac{1}{3}$ or that

it was wrong $p = \frac{2}{3}$. In essence, when you choose door A there a $\frac{2}{3}$ chance that the prize door is either between doors B or C . This doesn't change when you open up door B to see nothing in it—instead it now means that there is a $\frac{2}{3}$ chance that door C holds the prize so its better for you to switch.

Day Five: June 25

Today's problems focus on betting games and review of probability distributions.

Problem 1. You play a game where you toss two fair coins in the air and win \$1 each time you toss. However, if you have tossed 2 heads at least once, and 2 tails at least once, you lose all your winnings and the game ends. You can also choose to stop playing at anytime. What's your strategy?

Solution. I pulled this problem from stack exchange so you can find some solutions [here](#). Solution 1 on this page derives the solution in a cleaner manner than the solution my friend and I came up with for this.

It takes at least 2 tosses in order to get HH and TT . Then, after we roll our first bad event we have probability $\frac{1}{4}$ to lose everything and probability $\frac{3}{4}$ to win an additional dollar. Coin tosses are independent so each k^{th} toss we make after our first bad toss, we have probability $(\frac{3}{4})^k$ of gaining an additional k dollars and $1 - (\frac{3}{4})^k$ of losing the whole thing. So if we get our first bad toss on toss n (by which we have acquired n dollars), our expected value for $n + k$ is $\mathbb{E}[n + k] = (\frac{3}{4})^k(n + k)$. To find the optimal time to stop tossing, we need to find the boundary of $\mathbb{E}[n + k] \leq \mathbb{E}[n + k + 1]$. That is,

$$\begin{aligned} \left(\frac{3}{4}\right)^k(n + k) &\leq \left(\frac{3}{4}\right)^{k+1}(n + k + 1) \\ 4(n + k) &\leq 3(n + k + 1) \\ 4n + 4k &\leq 3n + 3k + 3 \\ n + k &\leq 3 \end{aligned}$$

Note that it's not always optimal to stop at a total number of rolls ≤ 3 . The whole premise of this question revolves around the first bad toss. So if $n < 3$ (our first bad toss is on toss 1 or 2), we should keep tossing until we've tossed 3 times. Otherwise, we should just take our earnings and end the game.

Problem 2. A room of 100 people put their business cards in a hat, then each person randomly draws a business card. What's the expected number of people who draw their own business card?

Solution. Each person has a $\frac{1}{100}$ chance of drawing their particular name, assuming that everyone draws at the same time. Let X_i be a Bernoulli r.v. with success being person i drawing their own card. Then, $P(X_i = 1) = \frac{1}{n}$ from our first observation. It follows that

$$\mathbb{E}[X_i] = \sum_{x \in \mathbb{R}} xp(x) = \sum_{x=0}^1 xp(x) = 0 \cdot \frac{99}{100} + 1 \cdot \frac{1}{100} = \frac{1}{100}$$

Then,

$$\mathbb{E}[X_1 + \cdots + X_{100}] = \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_{100}] = n\left(\frac{1}{n}\right) = 1$$

so we expect 1 person to draw their own name. Note that 100 is not a special number or anything—if we replace every instance of 100 with n in our argument, we observe that our EV is still 1 person.

Problem 3. There are 100 noodles in a soup bowl you are eating. You are blindfolded and told to take the ends of some noodles and connect them (each end is chose with equal probability). You do this until there are no more free ends. What is the expected number of circles created in this fashion?

Solution. This is actually be solved in a stochastic manner. Let $f(n)$ be a random variable that describes the number of circles that are made from connecting noodles. Starting from $n = 1$ noodle, you must connect the two ends together, forming $\mathbb{E}[f(1)] = 1$ circle. For $n = 2$ noodles, there are 4 ends and you must make 2 connections, forming $\binom{4}{2} = 6$ possible connections. Of these, 2 connections will connect the same ends of the same noodle together, giving us 1 noodle and 1 circle. The other four combinations give us a single noodle. Notice that we're only doing so much as to connect two of the ends, which brings us to a $n = 1$ case. Then the expected number of circles for $n = 2$ is

$$\mathbb{E}[f(2)] = \frac{2}{3}(1 + \mathbb{E}[f(1)]) + \frac{4}{6}\mathbb{E}[f(1)]$$

Solving this gives us that $\mathbb{E}[f(2)] = \frac{4}{3} = 1 + \frac{1}{3}$. Moving on to $n = 3$, we have $\binom{6}{2} = 15$ choices for connections. Of these, 3 choices will yield 1 circle and 2 noodles and the other 12 will simply yield 2 noodles. Then, we have

$$\mathbb{E}[f(3)] = \frac{3}{15}(\mathbb{E}[1 + f(2)]) + \frac{12}{15}\mathbb{E}[f(2)]$$

Solving this yields $\mathbb{E}[f(3)] = \frac{23}{15} = 1 + \frac{1}{3} + \frac{1}{5}$. There's a pretty clear pattern that's starting to appear. If we extend it to n noodles in general, we have that

$$\mathbb{E}[f(n)] = 1 + \frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{2n-1} = \sum_{k=1}^n \frac{1}{2k-1}$$

Plugging $n = 100$ gives us $\mathbb{E}[100] \approx 3.28$ (I cheated a bit and used a calculator to find this sum).

Problem 4. Two archers shoot at a target. The distance of each shot from the center of the target is uniformly distributed from 0 to 1, independently of the other shot. What is the PDF (probability density function) of the distance from the center for the losing shot?

Solution. Let X be the distance from the center of the shot by Archer 1 and Y be that of Archer 2. Let $Z = \max\{X, Y\}$ be the distance from the center of the losing shot. As X and Y are uniformly distributed over $[0, 1]$, we have that for all $z \in [0, 1]$,

$$P(X \leq z) = P(Y \leq z) = \frac{z - 0}{1 - 0} = z$$

Then, as X and Y are independent,

$$\begin{aligned} F_Z(z) &= P(\max\{X, Y\} \leq z) = P(X \leq z, Y \leq Z) \\ &= P(X \leq z)P(Y \leq z) = z^2 \end{aligned}$$

When we differentiate we then get the pdf of Z

$$f_Z(z) = \begin{cases} 2z & \text{if } z \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

This is an interesting problem that I haven't seen before. I've seen another form of the question that asks for the pdf of the winning shot, which is a bit more tricky as Z would be defined as $\min\{X, Y\}$ and it wouldn't be clear that we could use the $P(X, Y \leq z)$ since the minimum requires only one of them to satisfy the condition.

Problem 5. Let X_1, \dots, X_n be independent random variables uniformly distributed on $[0, 1]$. Find $P(X_1 + \cdots + X_n \leq 1)$?

Solution. Like Problem 3, I'm going to try to start from the $n = 1$ case and try to build up a solution. Clearly, $P(X \leq 1) = 1$ by definition. $P(X_1 + X_2 \leq 1)$ is the probability

that X_1 and X_2 , distributed over the unit square, lie on or below the line that separates the lower left and upper right triangles so $P(X_1 + X_2 \leq 1) = \frac{1}{2} = \frac{1}{2!}$. Moving on to $n = 3$, $P(X_1 + X_2 + X_3)$ describes the probability that the set of points in the unit cube all lie in the bottom-left-front triangular pyramid (probability $\frac{1}{6}) = \frac{1}{3!}$. Again, a bit of a pattern is showing up: it seems like $P(X_1 + \dots + X_n) = \frac{1}{n!}$. We can confirm this with induction of a more general case, that for $0 \leq t \leq 1$

$$P(X_1 + \dots + X_n \leq t) = \frac{t^n}{n!}$$

For simplicity sake, I'll call $S_n := X_1 + X_2 + \dots + X_n$. We've already proved the base case earlier in our solution. Then suppose our assumption holds for n , then for $n + 1$:

$$\begin{aligned} P(S_{n+1} \leq t) &= \int_0^1 P(S_n + x \leq t) f(x) dx = \int_0^t \frac{(t-x)^n}{n!} dx \\ &= \frac{t^{n+1}}{(n+1)!} \end{aligned}$$

and thus our claim is proved by induction. This last part does require a bit of massaging but should be clear after applying a bit of calculus/analysis. Then, it follows that $P(S_n \leq 1) = \frac{1}{n!}$

Day Six: June 26

Most of the questions today come from *A Practical Guide to Quantitative Finance Interviews* and past interview questions.

Problem 1. Bond A will default next year with $p = 0.5$ and bond B will default next year with $p = 0.3$. What is the range of probability that at least one bond defaults and what is the range of their correlation?

Solution. The upper bound on the the probability of at least one bond defaulting takes place when A defaulting and B defaulting are mutually exclusive, in which case,

$$P(A \text{ or } B \text{ defaults}) = P(A) + P(B) = 0.5 + 0.3 = 0.8$$

For the lower bound, we consider the case when A and B are completely dependent events. Then,

$$P(A \text{ or } B \text{ defaults}) = \max\{A, B\} = 0.5$$

Then $P(A \text{ or } B \text{ defaults}) \in [0.5, 0.8]$. To calculate the correlation between A and B , let I_A and I_B be indicator variables for the events that A and B default and let ρ be their correlation. We have that $\mathbb{E}[I_A] = 0.5$ and $\mathbb{E}[I_B] = 0.3$. Since A and B are Bernoulli random variables, we have that $\sigma_{I_A} = p_A(1 - p_A) = 0.5^2 = 0.25$. Similarly, $\sigma_{I_B} = 0.21$. We have that

$$\begin{aligned} P(A \text{ or } B \text{ defaults}) &= \mathbb{E}[I_A] + \mathbb{E}[I_B] - \mathbb{E}[I_A I_B] \\ &= \mathbb{E}[I_A] + \mathbb{E}[I_B] - (\mathbb{E}[I_A]\mathbb{E}[I_B] - \text{Cov}(I_A, I_B)) \\ &= 0.5 + 0.3 - (0.5 \cdot 0.3 - \rho \sigma_{I_A} \sigma_{I_B}) \\ &= 0.65 - \frac{\sqrt{0.21}}{2\rho} \end{aligned}$$

Plugging in the bounds of $P(A \text{ or } B \text{ defaults})$, we have $\rho \in \left[-\sqrt{\frac{3}{7}}, \sqrt{\frac{3}{7}}\right]$.

Problem 2. What is the expected number of dice rolls until we get two 6's in a row?

Solution. This is similar to our expected value of HT problem. Let X be a random variable that describes the number of throws to get two sixes in a row. There are three scenarios that we should consider:

Case 1: throwing two consecutive sixes ($p = \frac{1}{36}$)

Case 2: throwing a six on the first toss and a different number on the second ($p = \frac{5}{36}$)

Case 3: nothing throwing a six on the first toss ($p = \frac{5}{6}$)

In Case 1, we're done and our ev is 2. In Case 2, rolling a different number means we're exactly back where we started and our ev is $2 + \mathbb{E}[X]$ since we "wasted" our first two throws. In Case 3, we're back where we started and we've only used up the first throw, giving expectation of $1 + \mathbb{E}[X]$. Then,

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{36} \cdot 2 + \frac{5}{36}(2 + \mathbb{E}[X]) + \frac{5}{6}(1 + \mathbb{E}[X]) \\ &= \frac{42}{36} + \frac{35}{36}\mathbb{E}[X] \end{aligned}$$

Solving this gives us $\mathbb{E}[X] = 42$ so we expect to roll 42 times total in order to get our first string of consecutive 6's.

Problem 3. You flip a weighted coin with $P(\text{heads}) = 0.6$ four times. Every time that a heads is rolled A wins \$1 and every time a tails is flipped, B wins \$1. What is the expected value of a bet where if A wins overall, you get his winnings, if B wins overall you get 0.

Solution. This questions is from DRW's Glassdoor page. It's a bit unclear what the question means by expected value of a bet so I'll be solving for the expected value of the game. Additionally, since it doesn't state what happens if A and B tie, I'll be increasing the number of rolls to five. Note that A wins if we roll 3, 4, 5 heads and B wins if we roll 0, 1, 2 heads. Let X be a Bernoulli random variable with success being flipping heads ($p = 0.6$). Then, X follows a binomial distribution and we can calculate the expected value of the game as follows:

$$\begin{aligned}\mathbb{E}[X] &= \sum_{x \in \mathbb{R}} xp(x) = \sum_{x=3}^5 xb(x; 5, 0.6) = \sum_{x=3}^5 x \binom{5}{x} p^x (1-p)^{5-x} \\ &= 3 \binom{5}{3} (0.6)^3 (0.4)^2 + 4 \binom{5}{4} (0.6)^4 (0.4)^1 + 5 \binom{5}{5} (0.6)^5 \\ &\approx 2.23\end{aligned}$$

Then, the expected value of the game is 2.23, so it would make sense of us to play is so long as our cost of playing is less than \$2.23.

Problem 4. You're presented with two empty jars and 100 marbles on a table, 50 of which are white and the other 50 are black. You are to put all 100 of the marbles into the two jars in any way you choose. I then blindfold you and shake the jars up to ensure good mixing and rearrange the placing of the jars randomly. You can then request the left-hand jar or the right-hand jar. You get to choose exactly one jar and can withdraw at most one marble from the jar and you do not get a second chance if you are unhappy with your choice. How many of each color marble should you put in each jar to maximize the probability your blindfolded random draw obtains a white marble?

Solution. At first I thought that both jars needed to contain 50 marbles, in which case it didn't matter how we distributed the black and white marbles. Due to the mixing and the randomization of which of the two jars you're picking, you have $p = 0.5$ of picking either jar. Say you have x white marbles in one jar and $50 - x$ in another jar. Then, your probability of picking a white marble is $\frac{1}{2} \cdot \frac{x}{50} + \frac{1}{2} \cdot \frac{50-x}{50}$ which gives a 50% chance of picking a white marble no matter how you distribute them.

I then realized that we could distribute the marbles however we'd like, meaning that we could keep at little as 0 marbles in one jar and all 100 in the other jar. Instead of devising a function to describe how to optimize for this problem I decided to approach it with several cases. We know we have $p = \frac{1}{2}$ of choosing either jar. We want to maximize the chance of picking a white marble regardless of which jar you choose. This can be done by placing 1 white marble in the first jar, and the other 99 marbles in the second jar. This ensures that we have $p = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{49}{99} \approx 0.75$ of picking a white marble. Just testing out some other placements of marbles it's clear that this is the optimal configuration. Approaching the problem analytically, we simply need to find the maximum of the function

$$f(x, y) = \frac{1}{2} \cdot \frac{x}{y} + \frac{1}{2} \cdot \frac{50-x}{100-y}$$

where x is the number of white marbles we place in the first jar and y is the total number of marbles we place in the first jar.

Problem 5. I will spin a fair roulette wheel with five sections. Four of the five sections pay 1 and the other pays 5. If the cost is \$1.5 per spin should you play the game? What if you can only spin once and the payoffs to enter and cost per spint were multiplied by 1000?

Solution. This is a really simple EV problem. Let X be the payoff of playing the game. Then,

$$\mathbb{E}[X] = \frac{4}{5} \times 1 + \frac{1}{5} \times 5 = 1.8 > 1.5$$

When you have the option of playing as much as you like, it makes sense to play the game as the actual payoff converges to the expected payoff in the long run so we expect to make \$0.30 profit every 5 times we play.

If we can only spin once, our EV for the game doesn't change. However, this scenario guages how risk-averse we might be: 80% of the time we expect to lose 50 cents, and 20% of the time we expect to win \$3.50 beyond what we spent to play the game. Since the cost of the game is relatively low, I would probably still take the bet. However, when the numbers are scaled up by 1000, we either have a net loss of \$500 or net gain of \$3500. While the problems hasn't changed in nature, the risk of playing the game has increased dramatically, and depending on my initial amount of money, I would be much more cautious when playing.

Day Seven: June 27

Today's problems focus on betting games and expected value.

Problem 1. Suppose there are 11 pirates who are placing their treasure in a safe. As a democratic group of pirates, they have decided that any majority of the pirates should be able to open the safe. They decide to implement this by placing several locks on the safe, which all need to be opened to open the safe. Each lock can have multiple keys but each key can only open one lock. What is the smallest number of locks needed to secure the safe in this manner? In addition, how many keys should each pirate carry if each pirate can carry multiple keys?

Solution. Note that any combination of 5 pirates should not be able to open the safe. That is, there must be a unique lock for every combination of 5 pirates such that none of these pirates have the key to this lock. Then, there should be a total of $\binom{11}{5} = 462$ locks. Each of these locks should have 6 copies in order to maintain the majority criterion and these keys should be evenly distributed between all 11 pirates, giving us $\frac{462 \cdot 6}{11} = 252$ keys per pirate.

Problem 2. There is a population of ants that for each time period has $p = 0.2$ of dying out (E), $p = 0.5$ of staying the same (S), and $p = 0.3$ of doubling in size. Is the ant population likely to die long term?

Solution. This is a spinoff of a question regarding amoeba populations from *A Practical Guide*. Let's call $P(E)$ the probability that the ant population goes extinct. We can break this down into the different conditional probabilities: $P(E|S)$, $P(E|E)$, and $P(E|D)$. $P(E|S) = P(E)$ since the state is the same as the beginning state, $P(E|D) = P(E)^2$ (assuming each additional population of ants is independent to one another) as there are now two populations of ants, and $P(E|E) = 1$. Then,

$$P(E) = 0.5 \cdot P(E) + 0.3 \cdot P(E)^2 + 0.2 \cdot 1$$

Solving for this gives us that $P(E) = \frac{2}{3}$ so it is quite likely the ant population dies out. The question doesn't clarify what it means to double in size long term and E , D , and S are not collectively exhaustive events so it's quite tricky to calculate the other two outcomes. One could potentially use a Markov chain to simulate what happens in the future.

Problem 3. A gambler starts with an initial fortune of i dollars. On each successive game, the gambler wins \$1 with probability p , $0 < p < 1$, or loses \$1 with probability $q = 1 - p$. He will stop if he ever accumulates N dollars or loses all his money. What is the probability that he will end up with N dollars?

Solution. This is the gambler's ruin problem. Let P_i denote the probability that the gambler's fortune will reach N instead of 0 for each initial state $0 \leq i \leq N$. Note that $P_0 = 0$ and $P_N = 1$. The next after i is going to be $i + 1$ with probability p or $i - 1$ with probability q . Then, $P_i = pP_{i+1} + qP_{i-1}$. As $p + q = 1$, we have that

$$P_{i+1} - P_i = \frac{q}{p}(P_i - P_{i-1}) = \left(\frac{q}{p}\right)^2 (P_{i-1} - P_{i-2}) = \left(\frac{q}{p}\right)^i (P_1 - P_0)$$

We can then build an expression for P_i from the ground up starting from P_2 :

$$P_1 = pP_2 + qP_0 \rightarrow P_2 = \frac{1}{p}P_1 = \left[1 + \frac{q}{p}\right] P_1$$

$$P_2 = pP_3 + qP_1 \rightarrow P_3 = \frac{1}{q}P_2 - \frac{q}{p}P_1 = \left[1 + \frac{q}{p} + \left(\frac{q}{p}\right)^2\right] P_1$$

...

$P_i = \left[1 + \frac{q}{p} + \dots + \left(\frac{q}{p}\right)^{i-1}\right] P_1$ Note that the terms in the brackets is just a finite geometric sum. Applying this to P_N ,

$$P_N = 1 = \left[1 + \frac{q}{p} + \dots + \left(\frac{q}{p}\right)^{N-1}\right] P_1 = \begin{cases} \frac{1 - \left(\frac{q}{p}\right)^N}{1 - \frac{q}{p}} P_1 & \text{if } \frac{q}{p} \neq 1 \\ NP_1 & \text{if } \frac{q}{p} = 1 \end{cases}$$

We can manipulate this to solve for P_i through P_1 :

$$P_1 = \begin{cases} \frac{1 - \frac{q}{p}}{1 - \left(\frac{q}{p}\right)^N} P_1 & \text{if } \frac{q}{p} \neq 1 \\ \frac{1}{N} & \text{if } \frac{q}{p} = 1 \end{cases} \text{ and } P_i = \begin{cases} \frac{1 - \left(\frac{q}{p}\right)^i}{1 - \left(\frac{q}{p}\right)^N} P_1 & \text{if } \frac{q}{p} \neq \frac{1}{2} \\ \frac{i}{N} & \text{if } \frac{q}{p} = \frac{1}{2} \end{cases}$$

Given even odds of winning and losing each round, our probability of coming out with N dollars is proportional to our beginning state. What's interesting is that even if we decrease the odds of success slightly, say changing it from $\frac{1}{2}$ to $\frac{38}{80}$, there is a drastic decrease in the ability to come out with N dollars versus losing everything. I hope to include a Markov chain version of this problem with specific constraints and work through that sometime soon.

Problem 4. Suppose you roll a dice and are paid the value that you roll. If you roll a 5, 6 you can roll again and if you roll 1, 2, 3, 4 the game ends. How much would you pay to play this game?

Solution. Let X be the payoff of playing this game. There is $p = \frac{2}{3}$ that you roll a losing number and your payoff is what you rolled the first round. In this case, our expected value

is the average of $\{1, 2, 3, 4\} = \frac{5}{2}$. Also, there is a $\frac{1}{3}$ chance where we roll a 5 or a 6 and have the chance to roll again, bringing us back to our initial state. Then, we have

$$\mathbb{E}[X] = \frac{2}{3} \cdot \frac{5}{2} + \frac{1}{3} \left(\frac{11}{2} + \mathbb{E}[X] \right)$$

Solving this gives us $\mathbb{E}[X] = \frac{21}{4}$. Then I would pay up to \$5.25 to play this game.

Problem 5. Your recruiter loads two bullets separated by one chamber into a 6-chamber revolver. He then spins it, points it at his head and shoots but survives since he fired an empty chamber. You're given the gun and must fire at your head but you may choose to spin the barrel first or take the shot directly. Unfortunately, another bullet will be loaded into the gun if you decide to spin. Do you spin the barrel?

Solution. If we spin, we have a $\frac{3}{6}$ chance of dying as we have randomized the location of the next shot relative to the chambers. If we don't there are five chambers left, two of which have bullets. However, these bullets are staggered by one chamber so the number of configurations of the two bullets is limited. Let's number the chambers in order from 1 through 6. Let's say that the bullets are in chambers 1 and 3. Given that your recruiter survived, he must have fired from chambers 2, 4, 5 or 6, which occur equally likely. If he fired chamber 2 or 6, then the next shot will have a bullet and if he fired 4 or 5, then the next shot will be safe. Then you have 50% chance of dying if we shoot immediately, which is equal to the chance of dying if we decided to spin. Then it does not matter whether we spin or not since we have an equal chance of surviving the next round regardless of how we act.

Bonus. You have x red cards and y blue cards. I flip them over one at a time. The probability of flipping a particular color is proportional to the amount of those colored cards left. You start with 1 and every flip you can bet some proportion of your money on red or blue. If you win the bet, you gain twice your bet, but if you lose the bet, you gain nothing. What is the strategy that maximizes expectancy and minimizes variance?

Solution. I'll be typing up the solution later!

Week One Review

This week, I covered mostly probability games, Markov chains, and some combinatorial brain teasers. I realized that I still have a lot intuition to build on how analytically to approach each problem, which I'll keep in mind in future weeks. I hope to be posting some

more challenging problems as well, since the ones I'm currently working on are not overly difficult. Next week I hope to include some topics on options and market making and reviewing some statistics and algorithms to add some variety to the problems I've been doing. I'm also doing readings on game theory so hopefully I'll be able to include some more nuanced questions on winning strategies and other related topics.

Thanks for tuning in and below are a couple of resources that I used this week that might be helpful:

- *A Practical Guide to Quantitative Finance Interviews*
- [Quant/FM/Data Science Compilation](#)
- [Linear Algebra, Probability, and Financial Math](#)
- [General Problems & Brain Teasers](#)
- *Heard on the Street: Quantitative Questions from Wall Street Job Interviews*

Probability

Quantitative Trading

Interview Prep

7 Comments

 Login ▼

G

Join the discussion...

LOG IN WITH

OR SIGN UP WITH DISQUS 

Name

 3

Share

Best

Newest

Oldest

O

Oscar Martinez


2 years ago

Why is the tennis sets not equal??

If we have 4 possibilities for the first 2 sets, (PP,QQ,PQ,QP) with the first 2 resulting in a 2 set match and the second 2 resulting in a 3 set match.

1 0 Reply 

S

Sujay Somani Oscar Martinez

a year ago

Only true if $p=q=0.5$ try any other value of p and $pp+qq>2pq$ 0 0 Reply 

A

Aniruddha Deb

3 years ago

I think there's a small calculation error in the +/- problem: I'm getting $2/3$ as the answer, and 0.4375 rather than 0.4365 in P^3 . Could you confirm?

1 0 Reply 

H

Huaiyu Li

2 years ago

For Problem 3 of June 25 (noodles) there is a quicker solution. There are n out of $n(2n-1)$ ways to form a loop with the first pair of ends. If connecting the first pair of ends does not make a loop, then we are left with $n-1$ noodles.

Let A_n denote the expected number of loops. $A_n = (1+A_{n-1}) / (2n-1) + (2n-2) / (2n-1) * A_{n-1}$, and $A_n = 1 + 1/3 + \dots + 1/(2n-1)$.

Thanks for the awesome list!

0 0 Reply 

**Nguyễn Hưng**

3 years ago

can you make it in pdf format? Thanks

0

0

Reply

**Mugiwara No Luffy**

3 years ago

Thank you so much for this!

0

0

Reply

**Rohan Raj**

4 years ago

This was helpful, thanks!

0

0

Reply

[Subscribe](#)[Privacy](#)[Do Not Sell My Data](#)