MATH, PROBABILITY AND STATISTICAL RIDDLES

LUDOVICO BESSI

Math, Probability and Statistical riddles

Ludovico Bessi

6th December 2022

Contents

Introduction

About the author

Hey! I am Ludovico. I am working as a Machine learning engineer at Google in the anti abuse space.

However, my background is not in software: I obtained a MSc degree in Applied mathematics with a focus on probability and statistical optimization.

Once upon a time, I wanted to work as a quantitative researcher in a hedge fund, so I practiced a lot of probability riddles commonly asked in those interviews. Turns out, I like riddles a lot, so that was a fun journey.

In the end, I did not go for a quant career, but from time to time I still enjoy solving some riddles related to probability.

I decided to write this little e-book containing the most interesting questions I have solved during my preparation, for mainly two reasons:

1. I like writing math symbols. (This is the main reason)
2. Knowing how to reason about probability is useful.

Why this book?

There are almost infinitely many resources for this kind of riddles. Naming a few in random order that I really liked:

- Quant job interview, questions and answers by Joshi et al.
- Heard on the street: Quantitative questions from Wall street interviews by Crack.
- A practical guide to quantitative finance interviews by Zhou.
- 150 most frequently asked questions on Quant interviews by Stefanica et al.
- Probability and stochastic calculus quant interviews by Stefanica et al.
- Stack overflow probability section.
- Thousand of scattered PDFs from different universities.

However, I always found that the problems are either:

1. Easy and boring
2. Difficult and boring
3. Challenging and fun

This book only contains type 3 problems: just the right amount of difficulty to make them interesting, but leaving you the feeling that you can solve them without giving up.

Learning how to solve this type of questions could be useful if you need to prepare for quant/data science interviews. But the main goal is to just have some fun thinking about probability!

Content of the book

I have divided the book in broadly 3 parts:

- Brainteaser problems (7): everything from brainteasers to funny mathematical expressions.
- Probability problems (38): Playing with dices and coins all the way to almost impossible infinite tile problems (Looking at you, problem number 45!)
- Statistics problems (15): Detecting bias, calculating π in strange ways and looping noodles.

In total, there are 60 problems. Bonus point: I have also drawn pretty digital pictures to make this book a little less dry and a bit more fun!

Every problem is labelled with the following symbols: $\int$, $\iint$, $\iiint$ in order of increasing difficulty. Yes, I like integrals very much!

Notice that there is absolutely no theory in this book. If you are reading a solution that contains terms you don't understand, you will just look them up and (re)learn on the spot what you need: maximum efficiency!

Enjoy!

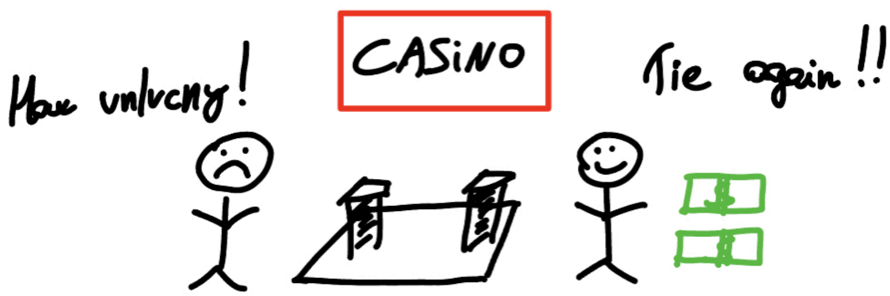
Math questions

This part will contain: brainteasers, strange expressions and difficult integrals. These problems are still asked in some banks, while hedge funds are now focusing more on probability and statistical reasoning. Still, there are some classics here that you really can't miss!

1. Pesky card game $\int$

You have a standard deck of 52 cards. Turn two cards at each time. For each pair: if they are both black they go to the dealer pile, if they are both red they go to your pile and if they are different they are discarded. You win if there are more red cards than black cards, you lose if it's a tie or more black cards. How much do you pay to play the game?

Solution



You should never play this game! No matter the distribution, It is always going to be a tie.

2. Playing with fire! $\iint$

You have two ropes, each of which takes 1 hour to burn. How do you use two ropes to measure 45 minutes?

Solution

A classic. You light the first rope on both end and the second rope on one end. When the first rope burns out, you light out the second end of the second rope. This will make the total time it takes to burn the second rope to be exactly 45 minutes.

3. Counting 0s $\iint$

How many trailing zeros are there in 100!?

Solution

The number of 0s in a number is the number of 5s matched with the 2s. Take 100 as an example: $100 = 10 * 10 = 5 * 2 * *5 * 2$. In the case of 100!, there are $100/5 + 100/25 = 24$ 5s. The number of 2s is: $100/2 + 100/4 = 75$. We take the minimum of the two and that is our answer: 100! has 24 leading zeros.

4. Guessing colors $\int$

A bag has 20 blue balls and 14 red balls. Each time you randomly take two balls out withouth replacement.

If both balls have same color, add a blue ball. If they have a different color, add a red ball.

What will be the color of the last ball left in the bag?

Solution

There are three possible cases:

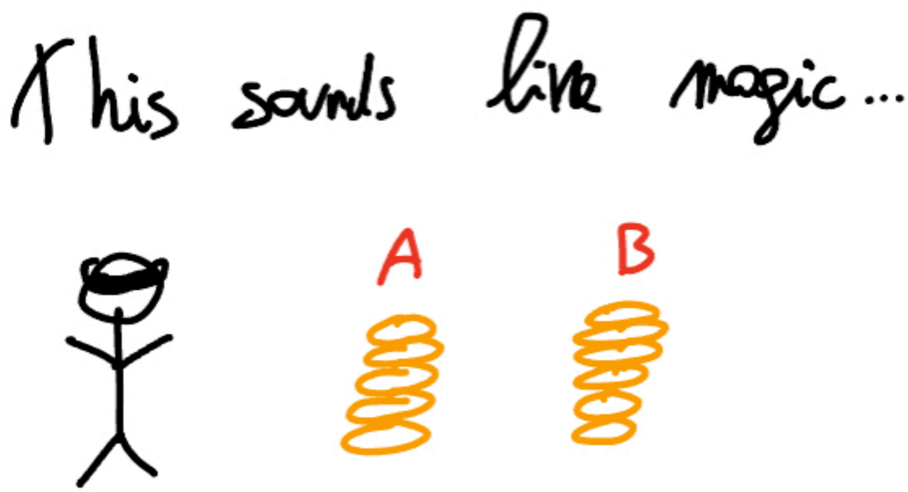
- (B,R) -> (B-1,R)
- (B,R) -> (B+1,R-2)
- (B,R) -> (B-1,R)

Based on this possibilities, you can see that the last ball must be blue because it is not possible to have an odd number of red balls.

5. Tidy up coins, but blindfolded! $\iint$

You have N=1000 coins on the floor. 980 are tails and the remaining 20 are heads. You are blindfolded. Can you separate the coins in two groups such that they have an equal number of heads?

Solution



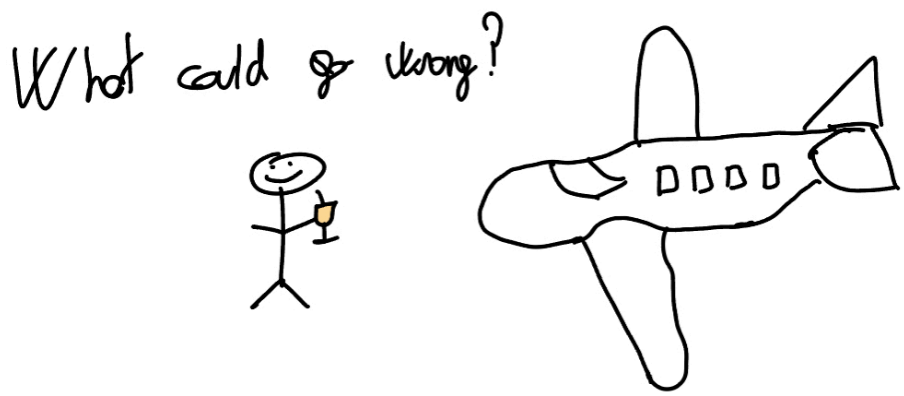
You can take n coins and put them aside. At this point, group A has n coins while group B has $1000 - n$ coins. Suppose group A has m heads. Then group B has $20 - m$ heads. You can now flip all coins in group A: now you have m tails and $n - m$ heads. Setting $n - m = 20 - m$, we have $n = 20$.

6. Drunk on a plane $\int$

A line of 100 airline passengers are waiting to board a plane. The first passenger in line is drunk and picks a random seat. After that, a passenger either goes to his seat, or picks one randomly.

You are the last boarding passenger. What's the probability that you end up in your seat?

Solution



Easy solution: Let's think about what happens to the second to last passenger: either they get their seat or they don't. So, for you you the probability is $\frac{1}{2}$.

Another solution:

Let $f(n)$ be the probability of the last passenger getting his seat if we begin with n passengers. With $\frac{1}{n}$ probability, the drunk passenger picks the right seat. Again with $\frac{1}{n}$ probability, the passenger picks the last seat. In any other case, with equal probability $\frac{1}{n}$ another seat is picked, and we continue the dilemma:

$$f(n) = \frac{1}{n} \times 1 + \frac{1}{n} \times 0 + \frac{n-2}{n} \times f(n-1)$$

Knowing that $f(2) = \frac{1}{2}$, then we see that $f(n) = \frac{1}{2} \forall n$.

7. Integral calculation $\int$

What is the value of $\int_{-\infty}^{+\infty} e^{-x^2} dx$?

Hint: this is still a probability based book, do you recognize the integrand function as a famous density function?

Solution

There are two ways to solve this this.

Solution A (more calculations, less elegant)

Solution A uses ideas from Calculus 2, the main trick is to let $I = \int_{-\infty}^{+\infty} e^{-x^2} dx$.

Let's try and calculate I^2.

$$I^2 = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-x^2} e^{-y^2} dxdy =$$

Using polar coordinates: $x = \rho \cos\theta$ and $y = \rho \sin\theta$, you can write:

$$= \int_{0}^{2\pi} \int_{0}^{+\infty} \rho e^{-\rho^2} d\rho d\theta =$$

$$= 2\pi * (-\frac{1}{2} \int_{0}^{+\infty} -2\rho e^{-\rho^2}) =$$

$$= -\pi[e^{-\rho^2}]_0^{+\infty} = -\pi * [0 - 1] = \pi$$

This means that $I = \int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$.

Solution B (less calculations, more elegant)

Remembering that the density function of a gaussian random variable with mean μ and variance σ^2 is: $f(x) = \sqrt{1}\sigma\sqrt{2\pi} * e^{\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$, you can fix $\mu = 0$ and $\sigma = \frac{1}{\sqrt{2}}$. Also, integrating a density function on its support gives you 1 by definition. All in all, you can define the following integral:

$$I^* = \int_{-\infty}^{+\infty} \frac{1}{\frac{1}{\sqrt{2}}\sqrt{2\pi}} e^{-x^2} dx = 1$$

Which brings you to the conclusion that $I = \sqrt{\pi}$.

Probability questions

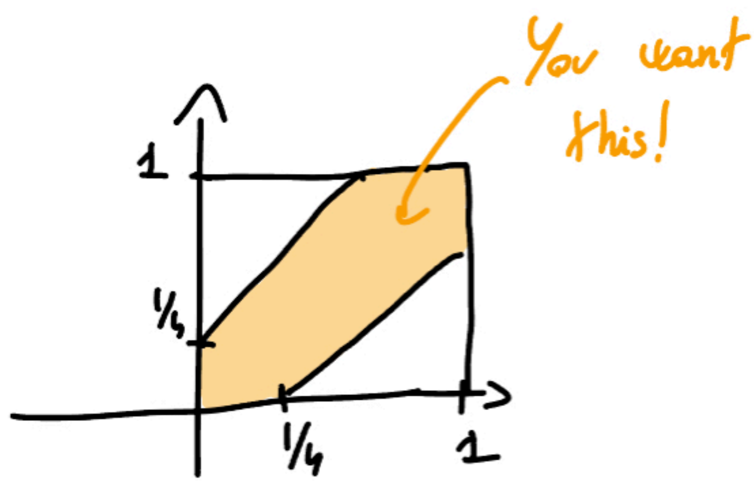
8. Meet your date at a random time to spice things up! $\iint$

You and your date agree to meet between 8pm and 9pm. Each one arrives at a random time and stays there for 15 minutes, then leaves. What is the probabilit that you go on a date tonight?

Solution Let x be your arrival time, and y your date's arrival time. Essentialy, you are trying to answer the following question:

$$\mathbb{P}(|x-y| \leq \frac{15}{60})$$

Let's draw and calculate the area of interest!



With some algebra, you can see that the solution is: $\frac{7}{16}$.

9. Rolling 5s $\iint$

You are rolling a fair six sided dice. What is the expected number of rolls until you roll two consecutive 5s?

Solution

Let X be the event of getting two consecutive 5. Let Y be the event of getting a 5. Then you can write:

$$\mathbb{E}[X] = \frac{1}{6}(1 + \mathbb{E}[X|Y]) + \frac{5}{6}(1 + \mathbb{E}[X])$$

You also know that: $\mathbb{E}[X|Y] = \frac{1}{6}(1+0) + \frac{5}{6}(1 + \mathbb{E}[X])$. Doing some calculations, you get $\mathbb{E}[X] = 42$.

10. Unfair coin $\int$

You have two coins. One is fair and the other one has two tails. You pick a coin at random: what is the probability that the coin you picked is the unfair one given that you got 5 tails in a row?

Solution

This is a straigthforward application of Bayes rule. Let A be the event that the coin is unfair, then you can write:

$$\mathbb{P}(A|TTTTT) = \frac{\mathbb{P}(A \cap TTTTT)}{\mathbb{P}(TTTTT)} = \frac{\frac{1}{2} \times 1}{\frac{1}{2} + \frac{1}{2} \times \frac{1}{2^5}} = \frac{32}{33}$$

11. Expected value of minimum $\int$

Let $X \sim Uniform(0, 1)$ and $Y \sim Uniform(0, 1)$. Find $\mathbb{E}[\min(X, Y)]$

Solution

Let $Z = min(X, Y)$. $F_Z(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(min(X, Y) \leq z)$.

$F_Z(z) = 1 - \mathbb{P}(min(X, Y) > z)$

The minimum of two independent random variables is greater than a value if and only if they are both greater than that value: $F_Z(z) = 1 - \mathbb{P}(X > z) \times \mathbb{P}(Y > z)$. For a uniform random variable, you have $\mathbb{P}(X > z) = 1 - z$.

Taking the derivative, you have: $f_Z(z) = 2(1-z)f$, so you have that: $F_Z(z) = 1 - (1-z)^2$.

Finally, you can calculate the expected value using the definition:

$$\mathbb{E}[Z] = \int_0^1 z \times 2(1-z)dz = \frac{1}{3}$$

12. Make a triangle! $\int\int$

You have a stick of length 1. You cut it uniformly at a point X and Y. What is the probability that the three parts form a triangle?

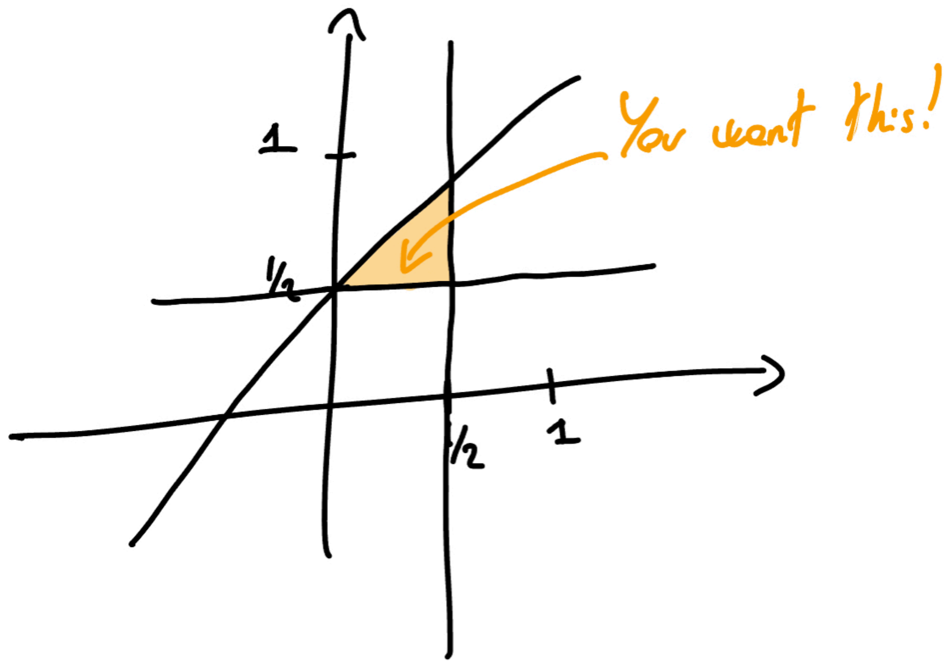
There are two ways to solve this. **Solution A (intuition based)** Three segments form a triangle if every segment is less than the sum of the other two. You can start by noting that if the two points are on the same side of the middle point $M = 0.5$ there is no chance you are going to get a triangle. This happens with probability $\frac{1}{2}$. Say now that the two points are on different side, and $X < Y$ without losing generality. In this case, Y cannot be too far off, otherwise the middle segment will be too big. More precisely, if $Y > X + M$, then the point cannot be too far off. This again happens with probability $\frac{1}{2}$. So, the final answer is $\frac{1}{2} * \frac{1}{2} = \frac{1}{4}$

Solution B (calculation based)

You can write the three conditions such that three segments form a triangle:

- $x + (y - x) > 1 - y \rightarrow y > \frac{1}{2}$
- $x + (1 - y) > y - x \rightarrow y < \frac{1}{2} + x$
- $(y - x) + (1 - y) > x \rightarrow x < \frac{1}{2}$

Now, let's plot these conditions and calculate the area where they are all true!



You see that the answer is again $\frac{1}{4}$.

13. Coin gaming $\int$

You are a given a coin, what is the expected number of flips needed to get two heads in a row?

Solution

As a slight more general solution, consider the possibility that the coin might not be fair. Say $\mathbb{P}(H) = p$ and $\mathbb{P}(T) = q$. Let A be the number of flips needed to get two heads in a row. Then, you can continue:

$$\mathbb{E}[A] = p \times (1 + \mathbb{E}[A|H]) + q \times (1 + \mathbb{E}[A|T])$$

You know that $\mathbb{E}[A|T] = \mathbb{E}[A]$.

Moreover:

$\mathbb{E}[A|H] = p \times (1 + \mathbb{E}[A|HH]) + q \times (1 + \mathbb{E}[A|HT])$ But: $\mathbb{E}[A|HH] = 0$ and $\mathbb{E}[A|HT] = \mathbb{E}[A]$. Putting everything together and solving for $\mathbb{E}[A]$, you get: $\mathbb{E}[A] = 6$.

14. Fair odds from unfair coin $\int$

You are given an unfair coin. Create a fair game.

Solution

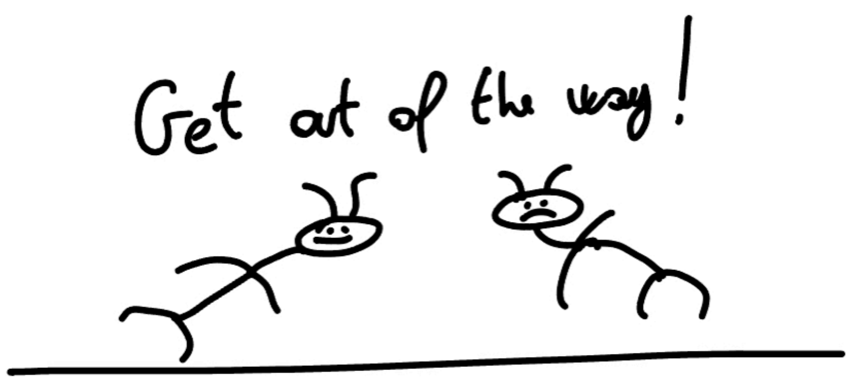
You can flip the coin twice. Discard HH and TT. If you consider consider only HT and TH sequences, then you have a fair game on your hands because they have equal probability of happening.

15. Ants headbutting $\int$

Three ants are sitting at corners of an equilateral triangle. Each picks a random direction. What is the probability that none of them collide?

Bonus question: What if we are talking about a k-sided polygon?

Solution



Without overthinking this, ants do not collide if they all choose to go right or they all choose to go left. This happen with probability:

$$\mathbb{P}(\text{no collision}) = (\frac{1}{2})^3 + (\frac{1}{2})^3$$

With the same reasoning, in a k-sided polygon:

$$\mathbb{P}(\text{no collision}) = (\frac{1}{2})^k + (\frac{1}{2})^k$$

16. Drawing aces! $\int$

How many cards would you expect to draw from a standard deck of 52 cards before seeing an ace?

Solution There are 4 aces in a deck, you can represent that with letter **A**. Then, there are other cards, you can represent them with **C**. Then, any deck can be represented in the following way: CACACACAC. Note: a sequence of cards C can also be empty because you could get an ace in your first draw. The expected value of cards needed to be drawn is all the one contained in C plus the ace, On average, a sequence **C** contains $\frac{48}{5}$. So, final answer is: $\frac{48}{5} + 1$

17. Let's play a dice game $\int$

Alice and Bob are playing a game with a dice. The first that rolls a 6 wins 100€. How much should Alice play to go first?

Solution The probability that Alice wins if she goes first is the probability of getting 6 on the first try (lucky her!) or Bob not winning after he throws, in mathematics:

$$\mathbb{P}(\text{Alice wins on first toss}) = \frac{1}{6} + \frac{5}{6}(1 - \mathbb{P}(\text{Bob wins on first toss}))$$

The probability of Bob winning after Alice fails is exactly the same as the probability of Alice winning on the first toss, so: $\mathbb{P}(\text{Alice wins on first toss}) = \mathbb{P}(\text{Bob wins on first toss})$.

So: $\mathbb{P}(\text{Alice wins on first toss}) = \frac{1}{6}+\frac{5}{6}(1-\mathbb{P}(\text{Alice wins on first toss}))$, meaning $\mathbb{P}(\text{Alice wins on first toss}) = \frac{6}{11}$. So, the expected value of going first versus going second is: $\frac{6}{11} * 100 - \frac{5}{11} * 100 =\sim 9$.

18. Let's play with some coins! $\int\int$

Let's play a coin game! We toss coins until either **HH** or **TH** show up. If we see **HH**, you win 40€. If we see **TH**, I win 20€. Are you happy with this arrengement?

Solution You should start by observing that whenever **T** is drawn, you'd have no chance of winning: either we keep getting **T**'s and nothing happens, or we get a **H** and I win. The only way for you to win is by getting **HH** immediately, which happens with probability $\frac{1}{4}$. The difference in expected values in this game is then:

$$\frac{1}{4} * 40 - \frac{3}{4} * 20 = -5€$$

You should not play this game!

19. Rolling dices $\int\int$

You are rolling a dice three times in a row. What is the probability that you obtain 3 numbers in strictly increasing order?

Solution

To be increasing, first they need to be different. This happens with probability $1 \times \frac{5}{6} \times \frac{4}{9}$. Then, there is only one configuration out of 3! that makes them increasing. All in all, the probability requested is: $1 \times \frac{5}{6} \times \frac{4}{9} \times \frac{1}{3!}$.

20. First toss with information $\int$

A fair coin is tossed n times. Given that there were k heads in the n tossess, what is the probability that the first toss was a head?

Solution A (intuition)

Imagine k was equal to 1. Then you would be quick to say that the probability is just $\frac{1}{n}$. Nothing changes when you generalize to k, so the answer is $\frac{k}{n}$.

Solution B (Bayes rule) Another solution is based on Bayes rule. Let A be the event that there were k heads in n tosses. Let B the event that the first toss was head. You are looking for:

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \times \mathbb{P}(B)}{\mathbb{P}(A)}$$

You know that $\mathbb{P}(B) = \frac{1}{2}$. For $\mathbb{P}(A)$, you need to find all the possible ways to choose k heads out of n heads over all the possible to choose n tosses, which means: $\mathbb{P}(A) = \frac{\binom{n}{k}}{2^n}$. In the same way, you can calculate $\mathbb{P}(A|B)$, just remembering that you know have one coin less: $\mathbb{P}(A) = \frac{\binom{n-1}{k-1}}{2^{(n-1)}}$

21. Recursive coins $\int\int$

A biased coin has probability p of landing on head and probability q on landing on tail. What is the probability that the total number of heads after n tosses is even?

Hint: You need to find a recursive relationship for this!

Solution

Let $\mathbb{P}(n)$ be the probability of having an even number of heads after n tosses. Then:

$$\mathbb{P}(n) = p \times (1 - \mathbb{P}(n-1)) + q \times \mathbb{P}(n-1)$$

22. Launch the new flavour of pizza? $\int$

You blind fold 500 italians and you make them taste a new chicken-pineapple-tofu pizza. Lucky you, they ave no taste buds! They either like it or not randomly. You also know that 50 of them will for sure not like the pizza, because they do not like eating blinded. You sample 5 people, if they all like your new pizza, you will make it available in all your restaurants.

What is the probability that all the 5 users like it?

Solution

You need to avoid picking those that will always say no, plus getting the random yes from all. This means:

$$\frac{450}{500} * \frac{1}{2} \times \frac{449}{499} * \frac{1}{2} \times \frac{448}{498} * \frac{1}{2} \times \frac{447}{497} * \frac{1}{2} \times \frac{446}{496} * \frac{1}{2}$$

23. Different dice rolls $\int$

You roll a dice three times. What is the probability that the sum of the outcomes is 12, given that the three rolls are different?

Solution Let A be the event that the sum is 12 and let B be the event that the rolls are different. The required probability is:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

You know that $\mathbb{P}(B) = 1 \times \frac{5}{6} \times \frac{4}{6}$. Moreover, there are only 3 different permutations that sum to 12 and all have different values, taking into account permutations you have: $\mathbb{P}(A \cap B) = \frac{3*3!}{6*5*4}$. Now you have everything you need to find the final value!

24. Recursive balls picking $\iint$

Two players are picking without replacement from a black box N red balls and M blue balls. The first one that picks a red ball wins. What is the probability that the the starting players wins?

Hint: Write down a recursive formula!

Solution

This is just the probability of winning instantly plus the probability of the second player not winning on his first draw after first player fails:

$$\mathbb{P}(n, m) = \frac{n}{n+m} + \frac{m}{n+m} \times (1 - \mathbb{P}(n, m-1))$$

25. Intuition challenger $\iint$

A fair coin is tossed twice. Is it more likely that two heads showed up given that at least one toss is head or given that the second toss was head?

Solution

Let A be the event that the first toss is head. Let B be the event that the second toss is head.

In the first case, you want to calculate:

$$\mathbb{P}(A \cap B | A \cup B) = \frac{\mathbb{P}(A \cap B \cap (A \cup B))}{\mathbb{P}(A \cup B)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)} = \frac{1/4}{3/4} = \frac{1}{3}$$

In the second case, you want to calculate:

$$\mathbb{P}(A \cap B|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{1/4}{1/2} = \frac{1}{2}$$

26. Dice re-roll, or not? $\int$

You roll a 6-sided dice up to two times, meaning that you can stop at the first throw. You always win the value in dollars of the latest throw. How much are you willing to pay to play this game?

Solution If you roll a number that is 1,2 or 3, you reroll, otherwise you don't. In case you don't reroll, the average value you are getting is 5. In case we reroll, the average value you get is 3.5. In total, the expected value of the game **G** is then:

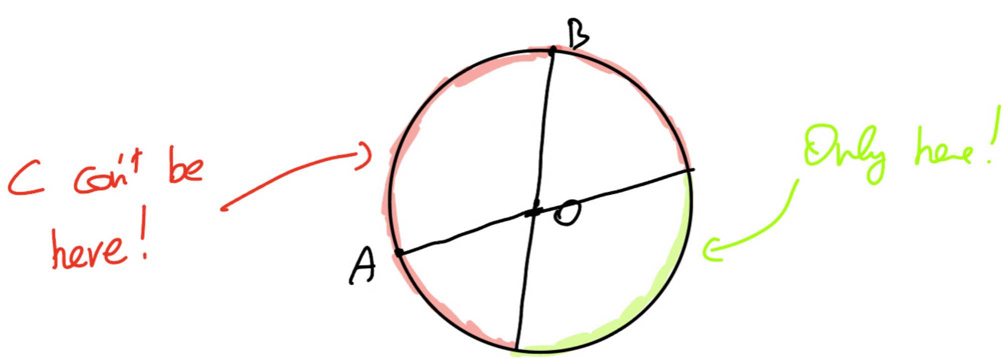
$$\mathbb{E}[G] = \frac{1}{2} * 5 + \frac{1}{2} * 3.5$$

27. Random triangle $\iint$

You pick three points on the unit circle and form a triangle with them. What is the probability that the triangle includes the center of the unit circle?

Solution

Take two points A and B at random on a triangle. Draw the segments passing through those points and the origin. Only 1/4 section of the circumference can have the point C such that the origin is included.



28. Chords intersection $\int$

Say you draw a circle and draw two random chords. What is the probability that they intersect?

Solution

Solution Fix the first chord by setting (randomly) two points AA and BB. Let xx the probability that point C falls between A and B. Then, the solution is: $2\int_0^1 x(1-x)dx = \frac{1}{3}2$ Inside the integral, we have: C between A and B and D not between A and B. Times two because we could also use D instead of C.

29. Points on same semicircle $\int$

You have N points. What's the probability that they all lie in the same semicircle?

Solution

Pick randomly one point, call it A. The probability that all the other points are in the same semicircle is: $\frac{1}{2^{N-1}}$ The first point can be picked randomly between all NN points, so the final answer is: $N * \frac{1}{2^{N-1}}N$

30. Yet another coin toss game

Two players play a game. As soon as the sequence HT happens, the player who flips T wins. What is the probability that the first player wins?

Solution

Let's condition on the first toss: $\mathbb{P}(A) = \frac{1}{2} * \mathbb{P}(A|H) + \frac{1}{2}\mathbb{P}(A|T)$

You know that $\mathbb{P}(A|T) = \mathbb{P}(B) = 1 - \mathbb{P}(A)P(A \mid T) = P(B) = 1 - P(A)$ Because if you toss a T, it's like you become player B.

The case with head can be treated conditioning on the toss of B:

$P(A|H) = \frac{1}{2}\mathbb{P}(A|HH) + \frac{1}{2}\mathbb{P}(A|HT)P(A \mid H)$

You know that by definition $\mathbb{P}(A|HT) = 0$ because this just means that B won. If instead B rolls an H, we have: $\mathbb{P}(A|HH) = \mathbb{P}(B|H) = 1 - \mathbb{P}(A|H)P(A \mid HH) = P(B \mid H) = 1 - P(A \mid H)$

And the final result is: $\mathbb{P}(A) = \frac{4}{9}P(A) =$

31. Scary game (1) $\int$

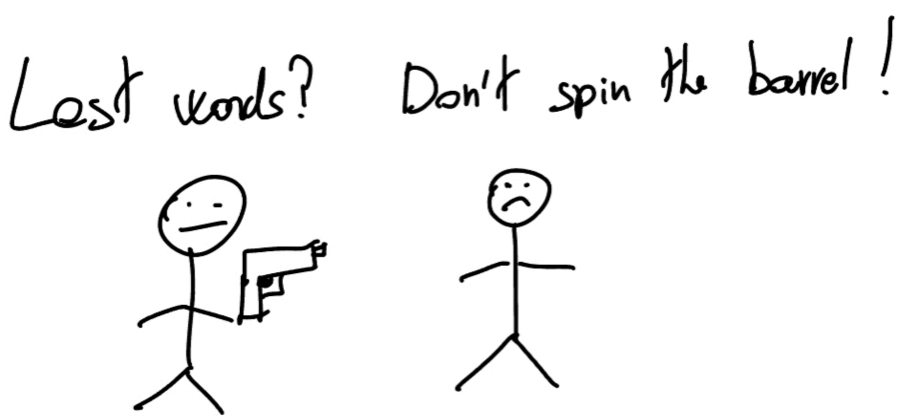
A bullet is put into a 6 chamber revolver. The barrel is randomly spun. Two players take turn pulling the trigger. You can choose to go first or go second. What do you choose?

Solution The player going first dies if bullet is in position 1,3 or 5. The player going second dies if bullet is in position 2,4 and 6. Exact same probabilities!

32. Scary game (2) $\int$

Two bullets are out into a 6 chamber revolver. Your opponent played first and he is alive after the first trigger. You are given the option to spin the barrel. Do you do it?

Solution



If you spin the barrel, the probability of losing is $\frac{2}{6}$, if you don't spin the barrel, then It is $\frac{2}{5}$.

33. Scary game (3) $\int$

Two bullets are put in **two** consecutive positions in a 6 chamber revolver. Your opponent played first and he is alive after the first trigger. You are given the option to spin the barrel. Do you do it?

Solution The probability that the next chamber contains a bullet is $\frac{1}{4}$, so chance of survival is $\frac{3}{4}$. If you spin the barrel, change of survival is: $1-\frac{2}{6}=\frac{2}{3}$.

34. Product of uniforms $\int$

Say you have two random variables $X, Y \sim$ Unif(0,1) that are independent. What is the probability that their product is greater then $\frac{1}{2}$?

Solution

You can calculate the area under the curve $y=\frac{1}{2x}$ contained in the square of length 1:

$$1 \times \frac{1}{2} + \int_{\frac{1}{2}}^{1} \frac{1}{2x} dx$$

35. XY distribution $\int$

If $X \sim \mathbb{N}(0,1)$ and $Y=1$ with probability $\frac{1}{2}$, $Y=-1$ with probability $\frac{1}{2}$, what is the distribution of $Z=XY$?

Solution

Just a straightforward calculation:

$$\mathbb{P}(Z=z)=\mathbb{P}(XY=z)=\mathbb{P}(X=z)*\mathbb{P}(Y=1)+\mathbb{P}(X=-z)*\mathbb{P}(Y=-1)=\mathbb{P}(X=z)$$

36. Many coin games $\int\int\int$

You have 4 different coin games:

- **G1**: one fair coin toss. If you flip head, you win 200€. If you flip tail, you lose 100€.

- **G2**: 100 fair coin tosses. If you flip head, you win 2€. If you flip tail, you lose 1€.
- **G3**: You start with a capital of 50€. 100 fair coin tosses. If you flip head, you win 2€. If you flip tail, you lose 1€. If you get to 0, you stop playing.
- **G4**: *100* fair coin tosses, you can bet an amount you choose from initial stake. You either win and double the bet or lose and lose the bet.
- What is the expected value of **G1**?
- What is the expected value of **G2**?
- Is It better to play **G3** or **G2**?
- How would you play **G4** in order to maximize $\mathbb{E}[G4]$ and $\mathbb{E}[\log G4 + 100]$?

Solution You can calculate the expected value of G_1 as:

$$\mathbb{E}[G_1] = \frac{1}{2} * 200 - \frac{1}{2} * 100 = 50$$

Crucially, the expected value of G_2 is exactly the same: $\mathbb{E}[G_2] = \mathbb{E}[G_1]$. G_2 has less variance. G_3 has the additional constraint of potentially stopping and losing, so It's better go with G_2 even though the probability of bankrupcy is quite low.

For the last question, suppose you always bet the maximum amount. On the first flip, you win: $100 + 2 * 100 = 300$. If you win, you bet 300 and win $300 + 300 * 2 = 900$. The total profit is $100 * 3^100 - 100$. This happens with probability $\frac{1}{2^{100}}$. Otherwise, you just lose 100. In this way, you have maxized the expected value:

$$\mathbb{E}[G_4] = \frac{1}{2^100} * (100 * 3^100 - 100) - (1 - \frac{1}{2^100}) * 100 = alot$$

To maximize $\mathbb{E}[\log G_4 + 100]$, let's say you bet x amount. The expected gain is:

$$\frac{1}{2}(\log(1-x)) + \frac{1}{2}(\log(1+2x))$$

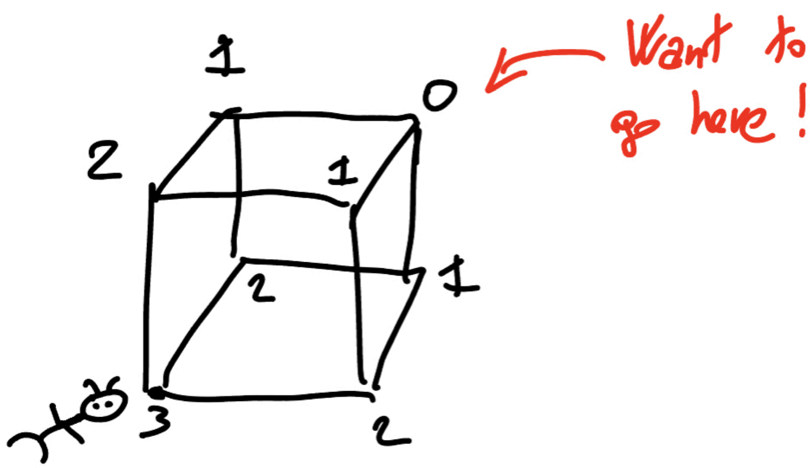
Taking the derivative, you find that $x = \frac{1}{4}$ maximizes the required expected value.

37. Ant on a cube $\iint$

Suppose there is an ant on a vertex of a cube. What is the expected number of steps the ant needs to take before reaching the opposite vertex?

Solution

Pick a random vertex. We are three steps away from the opposite vertex, so we label it with **3**. The vertexes neighbouring this vertex are two steps away from the opposite vertex, so we labe them with **2**. Beautiful diagram to explain this:



Let's now define E_i as the expected number of steps to be made to reach a vertex labeled with i. You are interested in finding E_3. Starting with E_0, clearly $E_0 = 0$ as the ant is sitting on that vertex. After the ant jumps, it for sure lands on a vertex labeled with 2, so you have: $E_3 = 1 + E_2$. One step closer! From a 2-vertex, the ant jumps with probability $\frac{2}{3}$ to a 1-vertex and with probability $\frac{1}{3}$ to a 3-vertex, so you have: $E_2 = 1 + \frac{2}{3}E_1 + \frac{1}{3}E_3$. Similarly for a 1-vertex: $E_1 = 1 + \frac{2}{3}E_2 + \frac{1}{3}E_0 = 1 + \frac{2}{3}E_2$

Putting everything together, you have: $E_1 = 7$, $E_2 = 9$ and $E_3 = 10$.

38. Another dice game! $\int$

Two players each roll two standard dice, first player A, then player B. If player A rolls a sum of 6, they win. If player B rolls a sum of 7, they win. What is the probability that A wins?

Solution

The probability of rolling a sum of 6 with two dices is: $p = \frac{5}{36}$. The probability of of rolling a sum of 7 with two dices is $r = \frac{1}{6}$. The probability of A winning is:

$$\sum_{k=0}^{\infty} p(1-p)^k(1-r)^k = \frac{p}{p+r-pr}$$

The explanation is that you are summing up all probabilities of all events: A rolls 6 immediately; A fails, B fails, and then A rolls and so on. These probabilities are independent probabilities and you can sum them up.

39. Lots of coin throws $\int\int$

Flip 10000 fair coins. You are offered a 1-1 bet that the sum is less (strictly) than 5000. How much would you bet? How much would you bet if someone tells you that the sum of the coins is less than 5100?

Solution There are 2^10000 possible outcomes of the sequence. The probability that there are less than 5000 heads is the sum of all binomial coefficients with n $<$ 5000 divided by the total possible outcomes. This gives you the probability of 0.496. The expected value of a bet B is then:

$$0.496 * B - B * (1 - 0.0496)$$

Onto the second question: we can use Baeysian formula, which gives us a probability of 0.508. This additional information gives us an edge!

40. Betting on an unfair coin $\int\int$

You have 100€. You are playing a game where you betting X dollars on a biased coin flip with a 90% probability of being head. You make twice the amount if it's head and lose all the bet amount if it is tails. How much do you think you should bet on each flip if you are going to play for 100 flips?

Solution If you just wanted to maximize the *expected* winnings, you should bet maximum every time. However, this will happen only with $(0.9)^1 00$ probability.

Let's see a different strategy. Let's try to maximize the worst case winnings. Let $X_0, X_1 ... X_n$ be the current wealth at each stage of the game and let $r_0, r_1, ..., r_1 00$ the proportion of current wealth that you place on each bet, i.e on bet i you bet $r_i X_{i-1}$.

You know that on a given turn, the ratio of our wealth is $\frac{X_i}{X_{i-1}} = 1 \pm r_i$ depending if you win or lose the i turn.

You want to maximize (X_i/X_{i-1}) for each i. You are going to maximize $log(X_i/X_{i-1}) = 0.9 * (1 + r_i) + 0.1 * (1 - r_i)$ as It's easier to work with. With some calculus, you can see you maximize this with $r_i = 0.8$. Which is equivalent to the Kelly criterion in this case.

41. Minimum of random variables $\iint$

Let $X_1, X_2, ..., X_n$ be uniform indepent random variables $\sim U(0, 1)$. What is the expected value of the minimum?

Solution Let M be the the minimum of the random variables. You know by definition that:

$$\mathbb{E}[M] = \int_0^\infty \mathbb{P}(M > t)dt$$

In this case, $\mathbb{P}(M > t)$ means probability of all n variables being greater than t, which happens with probability $(1-t)^n$. Noting that $\mathbb{P}(M > t) = 0$ for $t > 1$, then:

$$\mathbb{E}[M] = \int_0^1 (1-t)^n = \frac{1}{n+1}$$

42. Tossing and multiplying $\iiint$

A fair coin is tossed n times. What is the expected product of the number of heads and tails.

Solution

Let $X_i = 1$ if head, 0 otherwise. Let $n = 100$. Let $H_n = X_1 + X_2 + ... + X_n$ be the number of heads. Let $T_n = n-(X_1+X_2+...+X_n)$ You are interested in finding:

$$\mathbb{E}[H_n T_n] = \mathbb{E}[H_n(n - H_n)] = 100\mathbb{E}[H_n] - \mathbb{E}[H_n^2]$$

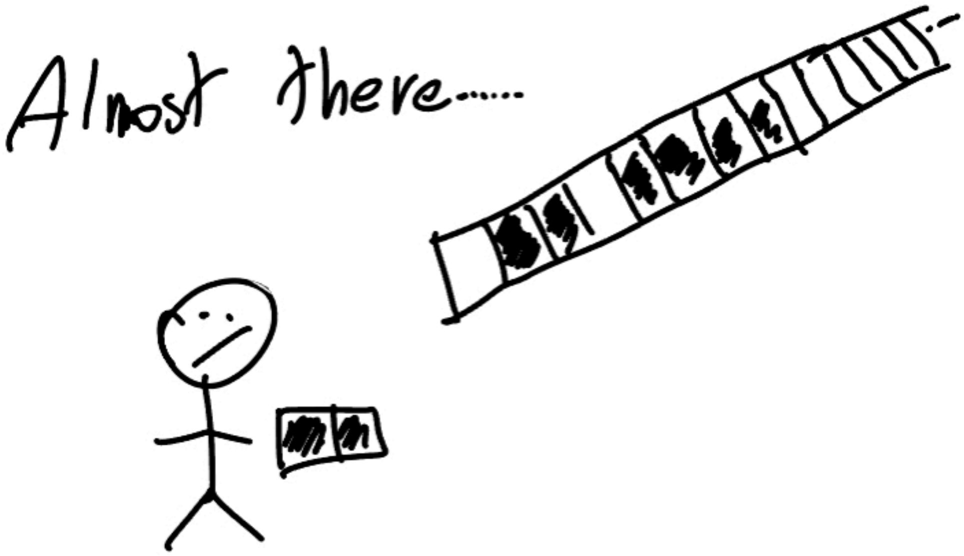
H_n is a binomial distribution, so we have $\mathbb{E}[H_n] = np$ and $\mathbb{E}[H_n^2] = np(1-p)$. All in all you have:

$$\mathbb{E}[H_n T_n] = (n^2 - n)p(1 - p)$$

43. Infinite board $\iiint$

You start with an empty $1xn$ board. Every turn, randomply place a $1x2$ tile. You keep placing tiles until it's not possible anymore. What is the expected fraction of the board that is covered when the process if finished, as $n \to \infty$?

Solution



Let X_n be the random variable denoting the final number of squares covered by dominos on the $1xn$ board, i.e. $\theta_n = \mathbb{E}[X_n]$. We know that $\theta_0 = \theta_1 = 0$ because we cannot fit any tile. We also know that $\theta_2 = 1$ as there is only space for one tile only. I am going to split the solution in three different steps:

- Step 1: Find a recoursive formulation for θ_n.
- Step 2: Set up a differential equation to solve a function that depends on the coefficients.
- Step 3: Solve the differential equation, apply taylor series to find θ_n blocked there.

Let's start with **Step 1**.

Let T_i the event that we place the tile in position i and $i+1$. We have that $\mathbb{P}T_i = \frac{1}{n-1}$ for every i. Then we can condition the expected value depending on where we placed the tile:

$$\mathbb{E}[X_n] = \sum_{i=3}^{n-1} \mathbb{E}[X_n|T_i] * \mathbb{P}(T_i) = \frac{1}{n-1} \sum_{i=0}^{n-1} \mathbb{E}[X_n|T_i]$$

Let's focus on a single term $\mathbb{E}[X_n|T_i]$. At this point, we placed a tile in position i, plus we can still place some tiles on the left ($i-1$ spaces remaining) or right ($n-i-1$ spacecs remaining):

$$\mathbb{E}[X_n|T_i] = 2 + +\theta_{i-1} + \theta_{n-i-1}$$

We can then write the following recursive formulation:

$$\theta_n = 2 + \frac{2(\theta_0 + \theta_1 + ... + \theta_{n-1})}{n-1}$$

We now have a recursive formulation, but we have still some way to go. Onto **Step 2**!

At this point, we need some way to study the sequence θ_n. What better way than using these coefficients in a infinite sum? Let's define $F(x) = \sum_{n=0}^{\infty} \theta_n x^n = \sum_{n=2}^{\infty} \theta_n x^n$. Now, let $G(x) = \frac{F(x)}{x}$. We define $G(0) = 0$ to make this differentiable. It will also come in handy later! The idea is to somehow derive information on θ_n by using $G(x)$. Let's start by taking the derivative:

$$G'(x) = (\sum_{n=2}^{\infty} \theta_n x^{n-1})' = \sum_{n=2}^{\infty} \theta_n (n-1) x^{n-2}$$

Now, we rewrite the recursive formulation found above as: $(n-1)\theta_n = 2(n-1) + 2(\theta_0 + \theta_1 + ... + \theta_{n-2})$ and conveniently use it here:

$$G'(x) = \sum_{n=2}^{\infty} [2(n-1) + 2(\theta_0 + \theta_1 + ... + \theta_{n-2})] x^{n-2} =$$

$$= \sum_{n=2}^{\infty} 2(n-1) x^{n-2} + 2 \sum_{m=0}^{\infty} (\theta_0 + ... + \theta_m) x^m$$

We know from calculus that

$$2 \sum_{n=2}^{\infty} (n-1) x^{n-2} = \frac{2}{(1-x)^2}$$

With some algebra, we can see that the second element can be written as:

$$\sum_{i=0}^{\infty} \frac{2\theta_i}{1-x} = 2\frac{F(x)}{1-x} = 2\frac{xG(x)}{1-x}$$

. All in all, we have the following differential equation:

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

I will let you solve the differential equation. Using G(0) = 0 to find the constant C = -1, then we have that G(x) is:

$$F(x) = xG(x) = \frac{x}{(1-x)^2} - \frac{xe^{-2x}}{(1-x)^2} = \frac{x}{(1-x)^2}(1-e^{-2x})$$

Finally, we are at **Step 3**. Let's use the taylor expansion of F(x) around 0 to write F(x) again as an infinite. This way, the coefficient we find in the Taylor expansion will be equal to θ_n. We have that the Taylor expansion of $\frac{x}{(1-x)^2}$ is $\sum_{n=0}^{\infty} nx^n$ while the Taylor expansion of e^{-2x} is $\sum_{n=0}^{\infty} \frac{(-2)^n}{n!}x^n$. Then, we have that

$$F(x) = \sum_{n=0}^{\infty}(n - \sum_{i=0}^{n} i\frac{(-2)^{n-i}}{(n-i)!})x^n$$

Which means we finally managed to find

$$\theta_n = n - \sum_{i=0}^{n} i\frac{(-2)^{n-i}}{(n-i)!}$$

The study of the limit itself is not super interesting, but for the curious reader we have that: $\lim_{n\to\infty} \theta_n = 1 - e^{-2}$.

44. Sine function and expectations $\int\int$

Let X be a random variable uniformly distributed on $[0, \pi]$. What is the value of: $\mathbb{E}[X|\sin X]$

Solution Conditioning on the value of $\sin X$ means knowing the value of $\sin X$. Knowing the value of $\sin X$ gives you two possibility for X: either $\arcsin X$ or $\pi - \arcsin X$. Given that that the value is normally distributed the two possibilities are equally likely, so the expected value is just the average:

$$\mathbb{E}[X|\sin X] = (\arcsin X + \pi - \arcsin X)/2 = \frac{\pi}{2}$$

45. Minus * Minus = Plus! $\int$

Let X and Y be two independent random Gaussian variables. What is $\mathbb{E}[X|XY]$ equal to?

Solution This problem requires a bit of a trick. Let $A = -X$ and $B = -Y$. Clearly also A and B are independent. Then, you have that: $\mathbb{E}[X|XY] = \mathbb{E}[X|AB] = \mathbb{E}[-A|AB] = -\mathbb{E}[X|XY]$, which means that $\mathbb{E}[X|XY] = 0$!

Statistics questions

46. Unbiased vs consistent $\int$

Can you showcase an example of: 1. an unbiased but not consistent estimator 2. a biased but not consistent estimator

Solution Let $(x_1, \ldots, x_n) \sim \mathbb{N}(\mu, \sigma^2)$ be random samples from a normal distribution.

Then: 1. Picking x_1 as an estimator makes it unbiased because its expected value is μ by definition, but It is clearly not consistent because the limit of the sequence $x_1, x_1, \ldots, x_1$ does not converge to μ in probability.

2. Estimating now the variance, taking $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$ gives us a biased but consistent estimator.

47. Drawing normally until we find value greater than *y*. $\int$

You have a random variable $X \sim \mathbb{N}(0, 1)$. What is the expected number of draws until you get a value greater or equal than y?

Solution Let A be the event "sample a value greater than y". You can write:

$$\mathbb{P}(X \geq y) = 1 - \mathbb{P}(X < y) = 1 - \Phi(y) = p$$

The draws are independent, so this can be seen as a geometric distribution with probability of success p. By definition, your success draw will happen after $\frac{1}{p}$ draws.

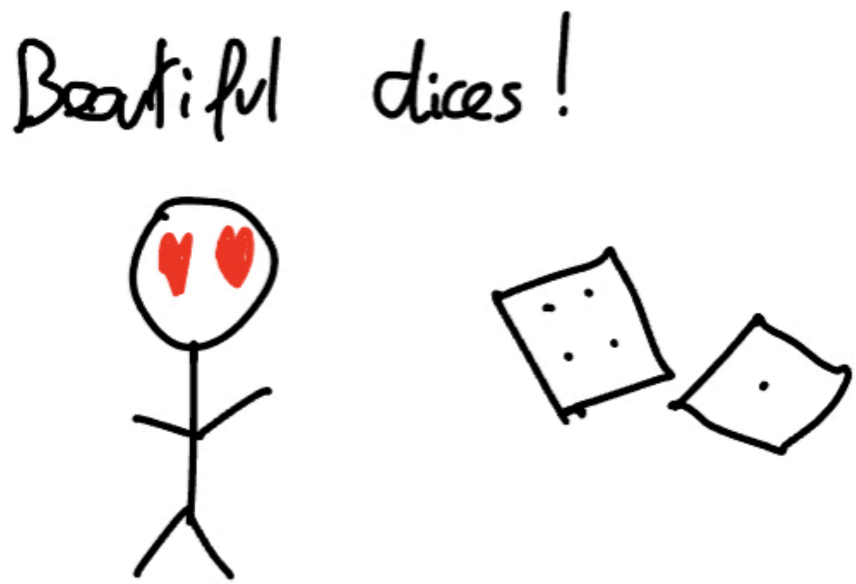
48. Dice lover $\int$

You have two games you need to choose from:

- In game **A**, you roll two dices and you win the product of the rolls.
- In game **B**, you roll one dice and you win its value squared.

Which game do you choose?

Solution



Let's calculate expected values of the two games. Let X be the value of the dice rolled.

- Game **A** has expected value: $\mathbb{E}[X] * \mathbb{E}[X]$.
- Game **B** has expected value: $\mathbb{E}[X^2]$.

Now the trick. You know that the variance of any non constant random variable is always greater than 0 and you also remember that the variance is defined as $\mathbb{E}[X^2] - \mathbb{E}[X]^2$. Putting this two ideas togheter you have that:

$$\mathbb{E}[X^2] \geq \mathbb{E}[X]^2$$

Meaning that you will choose game **B**.

49. Do we live in a simulation? $\iint$

You are given an oracle that acts as a Bernoulli random generator, you see either 0 or 1 as output with a certain probability p that is unknown to you. How could you use this oracle to generate values from a standard normal distribution?

Solution

Let's generate N samples following Bernoulli(p). Thanks to the central limit theorem, you have that:

$$\frac{\sqrt{N}(\overline{X_N} - \mu)}{\sigma} \to \mathbb{N}(0,1)$$

In this setting, you have that $\mu = p$, $\sigma = \sqrt{p(1-p)}$ and $\overline{X_N} = \frac{x_1 + \dots + X_N}{N}$. Then, you can write that $\overline{X_N} \sim \mathbb{N}(p, \frac{1(1-p)}{N})$.

To get a value that is distributed as a standard normal distribution, just take the Z score:

$$Z(\overline{X_N}) = \frac{\overline{X_N} - p}{\sqrt{\frac{p(1-p)}{N}}}$$

50. Correlation is not causation, and certainly not independence! $\iint$

Give an example where X and Y are uncorrelated but not independent.

Solution

Let's go back to some definitions: - Independent: $\mathbb{P}(X,Y) = \mathbb{P}(X) \times \mathbb{P}(Y)$ - Uncorrelated: $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$.

This kind of examples are best built using simple homemade distributions, integer values and with some dependence throw in. For example, let X take values $1, 0, -1$ with equal probabilities. Let $Y = 1$ if $X = 0$ and $Y = 0$ otherwise. Clearly, Y depends on X, so X and Y are not independent. Still, we have that:

$$\mathbb{E}[XY] = \frac{1}{3}(-1)(0) + \frac{1}{3}(0)(1) + \frac{1}{3}(1)(0) = 0$$

51. Coin toss game with advantage (or not?) $\int$

Alice has $n+1$ coins and Bob has n coins. What is the probability that Alice has more heads than Bob, assuming that the coins are all fair?

Solution

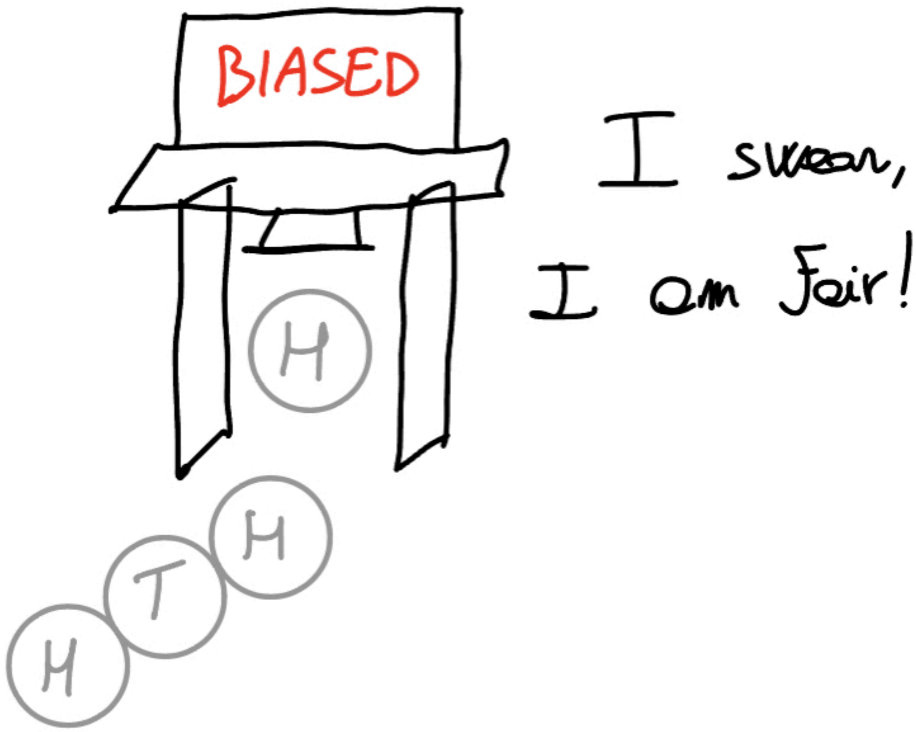
A very intuitive solution: the first n coins of Alice match all the n coins of Bob. This means that now Alice has 1 coin and Bob has 0 coins, then the probability that Alice has **strictly** more heads than Bob is $\frac{1}{2}$.

52. Bias detector $\int$

You have an unfair coin: $\mathbb{P}(H) = 0.6$ and $\mathbb{P}(T) = 0.4$. How many coin flips are needed to detect that the coin is indeed unfair?

Similar question you can solve that follows the same idea: A coin is flipped 1000 times, it shows 600 times head. Is it biased?

Solution



There are many ways to go about this, but the main idea is to use the central limit theorem. You know that $n\frac{\overline{X_N}-p}{\sqrt{p(1-p)}} \sim \mathbb{N}(0,1)$. Then, $\overline{X_N} \sim \mathbb{N}(p, \frac{p(1-p)}{n})$

We can then use the Z-score and say: if the Z-score for the empirical X_N is above say 2, than we say that the coin is indeed biased. Meaning, we just need to solve for n the following equation:

$$\frac{\overline{X_N} - p}{\sqrt{\frac{p(1-p)}{n}}} > 2$$

53. Find the density function of... the area of a circle!? $\int$

You pick the radius of a circle from a uniform distribution with values between 0 and 1. What is the probability density of the area of the circle?

Solution In this kind of problems, It is usually better to calculate the CDF of the required distribution. Then, you have the every information about your distribution: if you want the density function you can just take the deriative.

The density function of the radius is just $f_R(r) = r$ for $0 < r < 1$.

Let's now find the CDF:

$$F_A(a) = \mathbb{P}(A \leq a) = \mathbb{P}(\pi r^2 \leq a) = \mathbb{P}(r \leq \sqrt{\frac{a}{\pi}}) = \int_0^{\sqrt{\frac{a}{\pi}}} dr = \sqrt{\frac{a}{\pi}}$$

54. I give you a random number generator, you give me... π?!? $\int\int$

Estimate π. Hint: use a random number generator.

Solution

A great classic. Suppose you have a square of size 1, with bottom left corner on the origin. Draw a circle with radius 1 and center on the origin:

Now, generate random points inside the circle. You have that:

$$\frac{\text{number of points inside the circle}}{\text{total number of points}} = \frac{\text{Area of quarter of circle with radius 1}}{\text{Area of square with size 1}} = \frac{\frac{\pi}{4}}{1}$$

55. Expecting a big value? $\int$

What is the expected value of the max of two dice rolls?

Solution

Let X and Y be the values of the first and second roll respectively. You are interested in finding the expected value of $Z = \max(X, Y)$.

You are dealing with discrete distributions, so you can just write:

$$\mathbb{E}[Z] = \sum_{i=1}^{6} i \times \mathbb{P}(Z = i)$$

All that's left is to find $\mathbb{P}(Z = i)$. We have three cases: 1. The max value is on the first roll 2. The max value is on the second roll 3. The two rolls are equal

And you can express this as follows:

$$\mathbb{P}(Y = i) = \mathbb{P}(X = i, Y < i) + \mathbb{P}(X = i, Y < i) + \mathbb{P}(X = Y = i)$$

Then:

$$\mathbb{P}(Y = i) = \frac{1}{6} * \frac{(i-1)}{6} + \frac{1}{6} * \frac{(i-1)}{6} + 1 * \frac{1}{6}$$

56. Keep on playing! $\iint$

You have a N-sided dice. You keep rolling and sum values as long as the current roll is larger than the previous. What is the expected value of the sum?

Solution

Let's find the expected value of the sum if you first roll a dice with value r. For sure you have a value of r in the sum, because that was the first value. Then, with equal probability you get values higher than $r+1$. Values lower than r *can* be rolled, but they are not summed up and the game ends. Putting this idea into a formula:

$$\mathbb{E}[r] = r + \frac{1}{n}\mathbb{E}[r+1] + \frac{1}{n}\mathbb{E}[r+1] + \cdots + \frac{1}{n}\mathbb{E}[n]$$

You also know that $\mathbb{E}[n] = n$ by definition.

Let's then solve it starting from the back:

$$\mathbb{E}[n-1] = n - 1 + \frac{1}{n}\mathbb{E}[n] = n - 1 + \frac{1}{n} \times n = n$$

$$\mathbb{E}[n-2] = n-2+\frac{1}{n}\mathbb{E}[n] = n-2+\frac{1}{n}\mathbb{E}[n-1]+\frac{1}{n}\mathbb{E}[n] = n-2+\frac{1}{n}\times n+\frac{1}{n}\times n = n$$

It is then clear to see that $\mathbb{E}[r] = n \ \forall r$.

57. Keep on playing part 2! $\iint$

Assume you draw independently from a uniform number generator that generates values between 0 and 1. Keep drawing as long as the sequence is monotonically increasing. What is the expected length of draws?

Solution

Let's define $X_i = 1$ if draw x_i is in the list. Draw x_i is in the list if the sequence of draws $x_1, x_2, \dots, x_i$ is monotonically increasing. There is always one configuration, meaning this happens with probability $\frac{1}{i}$. Then, let's put this to work. If you think about it, with this set up, you now need to find the value of the following expression: $\mathbb{E}[\sum_{i=1}^{\infty} X_i]$. This looks a bit daunting, but luckily you have independece, which means that:

$$\text{Expected number of draws} = \mathbb{E}[\sum_{i=1}^{\infty} X_i] = \sum_{i=1}^{\infty} \mathbb{E}[X_i] = \sum_{i=1}^{\infty} \frac{1}{i!} = \sum_{i=0}^{\infty} \frac{1}{i!} - \frac{1}{0!} = e - 1$$

58. Keep on playing part 3! $\iiint$

Assume you draw independently from a uniform number generator that generates values between 0 and 1. Keep drawing until the sum exceeds 1. How many times have you drawn?

Hint: Solve the general problem: "until the sum exceeds v" and try to write down a recursive formulation that allows you to solve an easy differential equation.

Solution

The question asked is a special case of a more general question: you keep drawing until the sum exceeds value v. Let's solve the general problem!

You are then interested in finding the expected value of the smallest n such that $\sum_{i=1}^{n} U_i > v$. Let's call this event N_v. Now onto the recusive formulation: suppose we sample a value x. Then, we can write:

$$N_v = 1 + N_{v-x}$$

Meaning that when you sample a value x, you add 1 more sample and you ask again the same question by shifting the exceeding value to $v - x$. It is a tricky passage, It's OK to spend some time to mull this over. :)

Now, let's calculate the expected value:

$$E(v) = \mathbb{E}[N_v] = 1 + \mathbb{E}[N_{v-x}]$$

As we are solving the general v value case, I define the expected value as a function of v. Later on, you are going to calculate the value $E(1)$ to answer the original question. Then, you have:

$$E(v) = \mathbb{E}[N_v] = 1+\mathbb{E}[N_{v-x}] = 1+\int_0^v E(v-x)dx = 1-\int_v^0 E(y)dy = 1+\int_0^t E(y)dy$$

Taking the derivative both sides:

$$E'(v) = E(v)$$

Meaning that $E(v) = e^v$. Then, final answer: $E(1) = e$. Congrats for making it so far, time for a short break to let your brain rest! :)

59. Looped noodles $\iint$

Suppose you have 100 noodles. At each step, you randomly pick two ends and connect them. What is the expected number of loops?

Solution

There are two cases: either you pick two ends of the same noodle or you don't. Let's express this in mathematical terms, by defining $\mathbb{E}[n]$ as the expected number of loops when you have n noodles.

$$\mathbb{E}[n] = \frac{n}{\binom{2n}{n}} \times [1 + \mathbb{E}[n-1]] + (1 - \frac{n}{\binom{2n}{n}}) \times [0 + \mathbb{E}[n-1]]$$

Doing some calculation, you have:

$$\mathbb{E}[n] = \frac{1}{2n-1} + \mathbb{E}[n-1]$$

Hardest part is done! Now, you can also recognize that you have $\mathbb{E}[1] = 1$ and you can recursively solve for $n = 100$.

60. Special sequence $\int\int$

Say you toss a coin n times. How many sequences of 5 heads followed by 5 tails do you observe on average?

Solution

You can define $X_i = 1$ if i is the starting point of the special sequence. Then, you are interested in:

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

And you also know that: $\mathbb{E}[X_i] = (\frac{1}{2})^{10} \times (n-10)$ which you can substitue in.

Closing notes

Thank you!

If you reached this far, thank you! I hope you have enjoyed solving these 60 riddles as much as I liked: writing them, classifying the difficulty using integrals and drawing silly images! :)

Ludo

MATH, PROBABILITY AND STATISTICAL RIDDLES

LUDOVICO BESSI