Probability

CHAPTER 5

One of the most crucial skills a data scientist needs to have is the ability to think probabilistically. Although probability is a broad field and ranges from theoretical concepts such as measure theory to more practical applications involving various probability distributions, a strong foundation in the core concepts of probability is essential.

In interviews, probability's foundational concepts are heavily tested, particularly conditional probability and basic applications involving PDFs of various probability distributions. In the finance industry, interview questions on probability, including expected values and betting decisions, are especially common. More in-depth problems that build off of these foundational probability topics are common in statistics interview problems, which we cover in the next chapter. For now, we'll start with the basics of probability.

Basics

Conditional Probability

We are often interested in knowing the probability of an event A given that an event B has occurred. For example, what is the probability of a patient having a particular disease, given that the patient tested positive for the disease? This is known as the conditional probability of A given B and is often found in the following form based on **Bayes' rule:**

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Under Bayes' rule, $P(A)$ is known as the prior, $P(B|A)$ as the likelihood, and $P(A|B)$ as the posterior.

If this conditional probability is presented simply as $P(A)$—that is, if $P(A|B) = P(A)$—then A and B are independent, since knowing about B tells us nothing about the probability of A having also occurred. Similarly, it is possible for A and B to be conditionally independent given the occurrence of another event C: $P(A \cap B|C) = P(A|C)P(B|C)$.

The statement above says that, given that C has occurred, knowing that B has also occurred tells us nothing about the probability of A having occurred.

If other information is available and you are asked to calculate a probability, you should always consider using Bayes' rule. It is an incredibly common interview topic, so understanding its underlying concepts and real-life applications involving it will be extremely helpful. For example, in medical testing for rare diseases, Bayes' rule is especially important, since it is may be misleading to simply diagnose someone as having a disease—even if the test for the disease is considered "very accurate"—without knowing the test's base rate for accuracy.

Bayes' rule also plays a crucial part in machine learning, where, frequently, the goal is to identify the best conditional distribution for a variable given the data that is available. In an interview, hints will often be given that you need to consider Bayes' rule. One such strong hint is an interviewer's wording in directions to find the probability of some event having occurred "given that" another event has already occurred.

Law of Total Probability

Assume we have several disjoint events within B having occurred; we can then break down the probability of an event A having also occurred thanks to the law of total probability, which is stated as follows: $P(A) = P(A|B_1)P(B_1) + \dots + P(A|B_n)P(B_n)$.

The equation above provides a handy way to think about partitioning events. If we want to model the probability of an event A happening, it can be decomposed into the weighted sum of conditional probabilities based on each possible scenario having occurred. When asked to assess a probability involving a "tree of outcomes" upon which the probability depends, be sure to remember this concept. One common example is the probability that a customer makes a purchase, conditional on which customer segment that customer falls within.

Counting

The concept of counting typically shows up in one form or another in most interviews. Some questions may directly ask about counting (e.g., "How many ways can five people sit around a lunch table?"), while others may ask a similar question, but as a probability (e.g., "What is the likelihood that I draw four cards of the same suit?").

Two forms of counting elements are generally relevant. If the order of selection of the n items being counted k at a time matters, then the method for counting possible permutations is employed:

$$n * (n-1) * \dots * (n-k+1) = \frac{n!}{(n-k)!}$$

In contrast, if order of selection does not matter, then the technique to count possible number of combinations is relevant:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Knowing these concepts is necessary in order to assess various probabilities that involve counting procedures. Therefore, remember to determine when selection does versus does not matter.

For some real-life applications of both, consider making up passwords (where order of characters matters) versus choosing restaurants nearby on a map (where order does not matter, only the options). Lastly, both permutations and combinations are frequently encountered in combinatorial and graph theory-related questions.

Random Variables

Random variables are a core topic within probability, and interviewers generally verify that you understand the principles underlying them and have a basic ability to manipulate them. While it is not necessary to memorize all mechanics associated with them or specific use cases, knowing the concepts and their applications is highly recommended.

A random variable is a quantity with an associated probability distribution. It can be either discrete (i.e., have a countable range) or continuous (have an uncountable range). The probability distribution associated with a discrete random variable is a probability mass function (PMF), and that associated with a continuous random variable is a probability density function (PDF). Both can be represented by the following function of x: $f_X(x)$

In the discrete case, X can take on particular values with a particular probability, whereas, in the continuous case, the probability of a particular value of x is not measurable; instead, a "probability mass" per unit per length around x can be measured (imagine the small interval of x and $x + \delta$).

Probabilities of both discrete and continuous random variables must be non-negative and must sum (in the discrete case) or integrate (in the continuous case) to 1:

$$\text{Discrete: } \sum_{x \in X} f_X(x) = 1, \text{ Continuous: } \int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

The cumulative distribution function (CDF) is often used in practice rather than a variable's PMF or PDF and is defined as follows in both cases: $F_X(x) = p(X \le x)$

For a discrete random variable, the CDF is given by a sum: $F_X(x) = \sum_{k \le x} p(k)$; whereas, for a continuous random variable, the CDF is given by an integral:

$$F_X(x) = \int_{-\infty}^{x} p(y)\,dy$$

Thus, the CDF, which is non-negative and monotonically increasing, can be obtained by taking the sums of PMFs for discrete random variables, and the integral of PDFs for continuous random variables.

Knowing the basics of PDFs and CDFs is very useful for deriving properties of random variables, so understanding them is important. Whenever asked about evaluating a random variable, it is essential to identify both the appropriate PDF and CDF at hand.

Joint, Marginal, and Conditional Probability Distributions

Random variables are often analyzed with respect to other random variables, giving rise to *joint PMFs* for discrete random variables and *joint PDFs* for continuous random variables. In the continuous case, for the random variables X and Y varying over a two-dimensional space, the integration of the joint PDF yields the following:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx\,dy = 1$$

This is useful, since it allows for the calculation of probabilities of events involving X and Y.

From a joint PDF, a marginal PDF can be derived. Here, we derive the marginal PDF for X by integrating out the Y term:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$$

Similarly, we can find a joint CDF where $F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$ is equivalent to the following:

$$F_{X,Y}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(u,v)\,dv\,du$$

It is also possible to condition PDFs and CDFs on other variables. For example, for random variables X and Y, which are assumed to be jointly distributed, we have the following conditional probability:

$$f_X(x) = \int_{-\infty}^{\infty} f_Y(y) f_{X|Y}(x \mid y)\,dy$$

where X is conditioned on Y. This is an extension of Bayes' rule and works in both the discrete and continuous case, although in the former, summation replaces integration.

Generally, these topics are asked only in very technical rounds, although a basic understanding helps with respect to general derivations of properties. When asked about more than one random variable, make it a point to think in terms of joint distributions.

Probability Distributions

There are many probability distributions, and interviewers generally do not test whether you have memorized specific properties on each (although it is helpful to know the basics), but, rather, to see if you can properly apply them to specific situations. For example, a basic use case would be to assess the probability that a certain event occurs when using a particular distribution, in which case you would directly utilize the distribution's PDF. Below are some overviews of the distributions most commonly included in interviews.

Discrete Probability Distributions

The *binomial distribution* gives the probability of k number of successes in n independent trials, where each trial has probability p of success. Its PMF is

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

and its mean and variance are: $\mu = np$, $\sigma^2 = np(1-p)$.

The most common applications for a binomial distribution are coin flips (the number of heads in n flips), user signups, and any situation involving counting some number of successful events where the outcome of each event is binary.

The *Poisson distribution* gives the probability of the number of events occurring within a particular fixed interval where the known, constant rate of each event's occurrence is λ. The Poisson distribution's PMF is

$$P(X=k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

and its mean and variance are: $\mu = \lambda$, $\sigma^2 = \lambda$.

The most common applications for a Poisson distribution are in assessing counts over a continuous interval, such as the number of visits to a website in a certain period of time or the number of defects in a square foot of fabric. Thus, instead of coin flips with probability p of a head as a use case of the binomial distribution, applications on the Poisson will involve a process X occurring at a rate λ.

Continuous Probability Distributions

The *uniform distribution* assumes a constant probability of an X falling between values on the interval a to b. Its PDF is

$$f(x) = \frac{1}{b-a}$$

and its mean and variance are:

$$\mu = \frac{a+b}{2},\ \sigma^2 = \frac{(b-a)^2}{12}$$

The most common applications for a uniform distribution are in sampling (random number generation, for example) and hypothesis testing cases.

The *exponential distribution* gives the probability of the interval length between events of a Poisson process having a set rate parameter of λ. Its PDF is $f(x) = \lambda e^{-\lambda x}$ and its mean and variance are:

$$\mu = \frac{1}{\lambda},\ \sigma^2 = \frac{1}{\lambda^2}$$

The most common applications for an exponential distribution are in wait times, such as the time until a customer makes a purchase or the time until a default in credit occurs. One of the distribution's most useful properties, and one that makes for natural questions, is the property of memorylessness the distribution.

The *normal distribution* distributes probability according to the well-known bell curve over a range of X's. Given a particular mean and variance, its PDF is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp -\left(\frac{(x-\mu)^2}{2\sigma^2}\right)$$

and its mean and variance are given by: $\mu = \mu$, $\sigma^2 = \sigma^2$

Many applications involve the normal distribution, largely due to (a) its natural fit to many real-life occurrences, and (b) the Central Limit Theorem (CLT). Therefore, it is very important to remember the normal distribution's PDF.

Markov Chains

A Markov chain is a process in which there is a finite set of states, and the probability of being in a particular state is only dependent on the previous state. Stated another way, the Markov property is such that, given the current state, the past and future states it will occupy are conditionally independent.

The probability of transitioning from state i to state j at any given time is given by a transition matrix, denoted by P:

$$\begin{pmatrix} P_{11} & \cdots & P_{1n} \\ \cdots & & \cdots \\ P_{m1} & \cdots & P_{mn} \end{pmatrix}$$

Various characterizations are used to describe states. A *recurrent state* is one whereby, if entering that state, one will always transition back into that state eventually. In contrast, a transient state is one in which, if entered, there is a positive probability that upon leaving, one will never enter that state again.

A stationary distribution for a Markov chain satisfies the following characteristic: $\pi = \pi P$, where P is a transition matrix, and remains fixed following any transitions using P. Thus, P contains the long-run proportions of the time that a process will spend in any particular state over time.

Usual questions asked on this topic involve setting up various problems as Markov chains and answering basic properties concerning Markov chain behavior. For example, you might be asked to model the states of users (new, active, or churned) for a product using a transition matrix and then be asked questions about the chain's long-term behavior. It is generally a good idea to think of Markov chains when multiple states are to be modeled (with transitions between them) or when questioned concerning the long-term behavior of some system.

Probability Interview Questions

Easy

5.1. Google: Two teams play a series of games (best of 7 — whoever wins 4 games first) in which each team has a 50% chance of winning any given round (no draws allowed). What is the probability that the series goes to 7 games?

5.2. JP Morgan: Say you roll a die three times. What is the probability of getting two sixes in a row?

5.3. Uber: You roll three dice, one after another. What is the probability that you obtain three numbers in a strictly increasing order?

5.4. Zenefits: Assume you have a deck of 100 cards with values ranging from 1 to 100, and that you draw two cards at random without replacement. What is the probability that the number of one card is precisely double that of the other?

5.5. JP Morgan: Imagine you are in a 3D space. From (0,0,0) to (3,3,3), how many paths are there if you can move only up, right, and forward?

5.6. Amazon: One in a thousand people have a particular disease, and the test for the disease is 98% correct in testing for the disease. On the other hand, the test has a 1% error rate if the person being tested does not have the disease. If someone tests positive, what are the odds they have the disease?

5.7. Facebook: Assume two coins, one fair (having one side heads and one side tails) and the other unfair (having both sides tails). You pick one at random, flip it five times, and observe that it comes up as tails all five times. What is the probability that you are flipping the unfair coin?

5.8. Goldman Sachs: Players A and B are playing a game where they take turns flipping a biased coin, with p probability of landing on heads (and winning). Player A starts the game, and then the players pass the coin back and forth until one person flips heads and wins. What is the probability that A wins?

5.9. Microsoft: Three friends in Seattle each told you it is rainy, and each person has a 1/3 probability of lying. What is the probability that Seattle is rainy, assuming that the likelihood of rain on any given day is 0.25?

5.10. Bloomberg: You draw a circle and choose two chords at random. What is the probability that those chords will intersect?

5.11. Morgan Stanley: You and your friend are playing a game. The two of you will continue to toss a coin until the sequence HH or TH shows up. If HH shows up first, you win. If TH shows up first, your friend wins. What is the probability of you winning?

5.12. JP Morgan: Say you are playing a game where you roll a 6-sided die up to two times and can choose to stop following the first roll if you wish. You will receive a dollar amount equal to the final amount rolled. How much are you willing to pay to play this game?

5.13. Facebook: Facebook has a content team that labels pieces of content on the platform as either spam or not spam. 90% of them are diligent raters and will mark 20% of the content as spam and 80% as non-spam. The remaining 10% are not diligent raters and will mark 0% of the content as spam and 100% as non-spam. Assume the pieces of content are labeled independently of one another, for every rater. Given that a rater has labeled four pieces of content as good, what is the probability that this rater is a diligent rater?

5.14. D.E. Shaw: A couple has two children. You discover that one of their children is a boy. What is the probability that the second child is also a boy?

5.15. JP Morgan: A desk has eight drawers. There is a probability of 1/2 that someone placed a letter in one of the desk's eight drawers and a probability of 1/2 that this person did not place a letter in any of the desk's eight drawers. You open the first 7 drawers and find that they are all empty. What is the probability that the 8th drawer has a letter in it?

5.16. Optiver: Two players are playing in a tennis match, and are at deuce (that is, they will play back and forth until one person has scored two more points than the other). The first player has a 60% chance of winning every point, and the second player has a 40% chance of winning every point. What is the probability that the first player wins the match?

5.17. Facebook: Say you have a deck of 50 cards made up of cards in 5 different colors, with 10 cards of each color, numbered 1 through 10. What is the probability that two cards you pick at random do not have the same color and are also not the same number?

5.18. SIG: Suppose you have ten fair dice. If you randomly throw these dice simultaneously, what is the probability that the sum of all the top faces is divisible by 6?

Medium

5.19. Morgan Stanley: A and B play the following game: a number k from 1-6 is chosen, and A and B will toss a die until the first person throws a die showing side k, after which that person is awarded \$100 and the game is over. How much is A willing to pay to play first in this game?

5.20. Airbnb: You are given an unfair coin having an unknown bias towards heads or tails. How can you generate fair odds using this coin?

5.21. SIG: Suppose you are given a white cube that is broken into $3 \times 3 \times 3 = 27$ pieces. However, before the cube was broken, all 6 of its faces were painted green. You randomly pick a small cube and see that 5 faces are white. What is the probability that the bottom face is also white?

5.22. Goldman Sachs: Assume you take a stick of length 1 and you break it uniformly at random into three parts. What is the probability that the three pieces can be used to form a triangle?

5.23. Lyft: What is the probability that, in a random sequence of H's and T's, HHT shows up before HTT?

5.24. Uber: A fair coin is tossed twice, and you are asked to decide whether it is more likely that two heads showed up given that either (a) at least one toss was heads, or (b) the second toss was a head. Does your answer change if you are told that the coin is unfair?

5.25. Facebook: Three ants are sitting at the corners of an equilateral triangle. Each ant randomly picks a direction and begins moving along an edge of the triangle. What is the probability that none of the ants meet? What would your answer be if there are, instead, k ants sitting on all k corners of an equilateral polygon?

5.26. Robinhood: A biased coin, with probability p of landing on heads, is tossed n times. Write a recurrence relation for the probability that the total number of heads after n tosses is even.

5.27. Citadel: Alice and Bob are playing a game together. They play a series of rounds until one of them wins two more rounds than the other. Alice wins a round with probability p. What is the probability that Bob wins the overall series?

5.28. Google: Say you have three draws of a uniformly distributed random variable between (0, 2). What is the probability that the median of the three is greater than 1.5?

Hard

5.29. D.E. Shaw: Say you have 150 friends, and 3 of them have phone numbers that have the last four digits with some permutation of the digits 0, 1, 4, and 9. Is this just a chance occurrence? Why or why not?

5.30. Spotify: A fair die is rolled n times. What is the probability that the largest number rolled is r, for each r in 1,...,6?

5.31. Goldman Sachs: Say you have a jar initially containing a single amoeba in it. Once every minute, the amoeba has a 1 in 4 chance of doing one of four things: (1) dying out, (2) doing nothing, (3) splitting into two amoebas, or (4) splitting into three amoebas. What is the probability that the jar will eventually contain no living amoeba?

5.32. Lyft: A fair coin is tossed n times. Given that there were k heads in the n tosses, what is the probability that the first toss was heads?

5.33. Quora: You have N i.i.d. draws of numbers following a normal distribution with parameters μ and σ. What is the probability that k of those draws are larger than some value Y?

5.34. Akuna Capital: You pick three random points on a unit circle and form a triangle from them. What is the probability that the triangle includes the center of the unit circle?

5.35. Citadel: You have r red balls and w white balls in a bag. You continue to draw balls from the bag until the bag only contains balls of one color. What is the probability that you run out of white balls first?

Probability Interview Solutions

Solution #5.1

For the series to go to 7 games, each team must have won exactly three times for the first 6 games, an occurrence having probability

$$\frac{\binom{6}{3}}{2^6} = \frac{20}{64} = \frac{5}{16}$$

where the numerator is the number of ways of splitting up 3 games won by either side, and the denominator is the total number of possible outcomes of 6 games.

Solution #5.2

Note that there are only two ways for 6s to be consecutive: either the pair happens on rolls 1 and 2 or 2 and 3, or else all three are 6s. In the first case, the probability is given by

$$2*\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)^2 = \frac{5}{108}$$

and, for all three, the probability is

$$\left(\frac{1}{6}\right)^3 = \frac{1}{216}$$

The desired probability is given by:

$$2*\frac{5}{216}+\frac{1}{216}=\frac{11}{216}$$

Solution #5.3

First, note that the three rolls must all yield different numbers; otherwise, no strictly increasing order is possible. The probability that the three numbers will be different is given by the following reasoning. The first number can be any value from 1 through 6, the second number has a 5/6 chance of not being the same number as the first, and the third number has a 4/6 chance of not being the prior two numbers. Thus,

$$1*\frac{5}{6}*\frac{4}{6}=\frac{5}{9}$$

Conditioned on there being three different numbers, there is exactly one particular sequence that will be in a strictly increasing order, and this sequence occurs with probability 1/3! = 1/6.

Therefore, the desired probability is given by: $\frac{5}{9}*\frac{1}{6}=\frac{5}{54}$

Solution #5.4

Note that there are a total of $\binom{100}{2} = 4950$

ways to choose two cards at random from the 100. There are exactly 50 pairs that satisfy the condition: (1, 2),..., (50, 100), and since order does not matter, there are exactly 50 * 1 = 100 ways to draw such a pair. Therefore, the desired probability is:

$$\frac{50}{4950} \approx 0.01$$

Solution #5.5

Note that getting to (3, 3, 3) requires 9 moves. Using these 9 moves, it must be the case that there are exactly three moves in each of the three directions (up, right, and forward). There are therefore 9! ways to order the 9 moves in any given direction. We must divide by 3! for each direction to avoid overcounting, since each up move is indistinguishable. Therefore, the number of paths is:

$$\frac{9!}{3!3!3!}=1680$$

Solution #5.6

Let A denote the event that someone has the disease, and B denote the event that this person tests positive for the disease. Then we want: $P(A|B)$

By applying Bayes' theorem, we obtain: $P(A|B)=\frac{P(B|A)P(A)}{P(B)}$

From the problem description, we know that $P(B|A) = 0.98$, $P(A) = 0.001$

Let A' denote the event that someone does not have the disease. Then, we know that $P(B|A') = 0.01$. For the denominator, we have:

$$P(B) = P(B|A)P(A) + P(B|A')P(A') = 0.98(0.001) + 0.01(0.999)$$

Therefore, after combining terms, we have the following:

$$P(A|B)=\frac{0.98*0.001}{0.98(0.001)+0.01(0.999)}=8.93\%$$

Solution #5.7

We can use Bayes' theorem here. Let U denote the case where we are flipping the unfair coin and F denote the case where we are flipping a fair coin. Since the coin is chosen randomly, we know that $P(U) = P(F) = 0.5$. Let $5T$ denote the event of flipping 5 tails in a row. Then, we are interested in solving for $P(U|5T)$, i.e., the probability that we are flipping the unfair coin, given that we obtained 5 tails in a row.

We know $P(5T|U) = 1$, since, by definition, the unfair coin always results in tails. Additionally, we know that $P(5T|F) = 1/2\text{^}5 = 1/32$ by definition of a fair coin. By Bayes' theorem, we have:

$$P(U|5T)=\frac{P(5T|U)*P(U)}{P(5T|U)*P(U)+P(5T|F)*P(F)}=\frac{0.5}{0.5+0.5*1/32}=0.97$$

Therefore, the probability we picked the unfair coin is about 97%.

Solution #5.8

Let $P(A)$ be the probability that A wins. Then, we know the following to be true:

1. If A flips heads initially, A wins with probability 1.
2. If A flips tails initially, and then B flips a tail, then it is as if neither flip had occurred, and so A wins with probability $P(A)$.

Combining the two outcomes, we have: $P(A) = p + (1 - p)^2P(A)$, and simplifying this yields $P(A) = p + P(A) - 2pP(A) + p^2P(A)$ so that $p^2P(A) - 2pP(A) + p = 0$

and hence: $P(A)=\frac{1}{2-p}$

Solution #5.9

Let R denote the event that it is raining, and Y be a "yes" response when you ask a friend if it is raining. Then, from Bayes' theorem, we have the following:

$$P(R|YYY)=\frac{P(YYY|R)P(R)}{P(YYY)}$$

where the numerator is given by:

$$P(YYY|R)P(R)=\left(\frac{2}{3}\right)^3\left(\frac{1}{4}\right)=\frac{2}{27}$$

Let R' denote the event of no rain; then the denominator is given by the following:

$$P(YYY) = P(YYY|R)P(R) + P(YYY|R')P(R') = \left(\frac{2}{3}\right)^3\left(\frac{1}{4}\right)+\left(\frac{1}{3}\right)^3\left(\frac{3}{4}\right)$$

which, when simplified, yields: $P(YYY)=\frac{11}{108}$

Combining terms, we obtain the desired probability: $P(R|YYY)=\frac{\frac{2}{27}}{\frac{11}{108}}=\frac{8}{11}$

Solution #5.10

By definition, a chord is a line segment where the two endpoints lie on the circle. Therefore, two arbitrary chords can always be represented by any four points chosen on the circle. If you choose to represent the first chord by two of the four points, then you have:

$$\binom{4}{2}=6$$

choices of choosing the two points to represent chord 1 (and, hence the other two will represent chord 2). However, note that in this counting, we are duplicating the count of each chord twice, since a chord with endpoints p1 and p2 is the same as a chord with endpoints p2 and p1. That is, chord AB is the same as BA, (likewise with CD and DC). Therefore, the proper number of valid chords is:

$$\frac{1}{2}\binom{4}{2}=3$$

Among these three configurations, only one of the chords will intersect; hence, the desired probability is:

$$p=\frac{1}{3}$$

Solution #5.11

Although there is a formal way to apply Markov chains to this problem, there is a simple trick that simplifies the problem greatly. Note that, if T is ever flipped, you cannot then reach HH before your friend reaches TH, since the first heads thereafter will result in them winning. Therefore, the probability of you winning is limited to just flipping an HH initially, which we know is given by the following probability:

$$P(HH)=\frac{1}{2}*\frac{1}{2}=\frac{1}{4}$$

Therefore, you have a 1/4 chance of winning, whereas your friend has a 3/4 chance.

Solution #5.12

The price you would be willing to pay is equal to the expectation of the final amount. Note that, for the first roll, the expectation is

$$\sum_{i=1}^{6}\frac{i}{6}=\frac{21}{6}=3.5$$

Therefore, there are two events on which you need to condition. The first is on getting a 1, 2, or 3 on the first roll, in which case you would roll again (since a new roll would have an expectation of 3.5, and so, overall, you have an expectation of 3.5. The second is on if you roll a 4, 5, or 6 on the first roll, in which case you would keep that roll and end the game, and the overall expectation would be 5, the average of 4, 5, and 6. Therefore, the expected payoff of the overall game is

$$\frac{1}{2}*3.5+\frac{1}{2}*5=4.25$$

Therefore, you would be willing to pay up to $4.25 to play.

Solution #5.13

Let D denote the case where a rater is diligent, and E the case where a rater is non-diligent. Further, let $4N$ denote the case where four pieces of content are labeled as non-spam. We want to solve for $P(D|4N)$, and can use Bayes' theorem as follows to do so:

$$P(D\mid 4N)=\frac{P(4N\mid D)*P(D)}{P(4N\mid D)*P(D)+P(4N\mid E)*P(E)}$$

We are given that $P(D) = 0.9$, $P(E) = 0.1$. Also, we know that $P(4N|D) = 0.8 \cdot 0.8 \cdot 0.8 \cdot 0.8$ due to the independence of each of the 4 labels assigned by a diligent rater. Similarly, we know that $P(4N|E) = 1$, since a non-diligent rater always labels content as non-spam. Substituting into the equation above yields the following:

$$\frac{P(4N\mid D)*P(D)}{P(4N\mid D)*P(D)+P(4N\mid E)*P(E)}=\frac{0.8^4*0.9}{0.8^4*0.9+1^4*0.1}=0.79$$

Therefore, the probability that the rater is diligent is 79%.

Solution #5.14

This is a tricky problem, because your mind probably jumps to the answer of 1/2 because knowing the gender of one child shouldn't affect the gender of the other. However, the phrase "the second child is also a boy" implies that we want to know the probability that both children are boys given that one is a boy. Let B represent a boy and G represent a girl. We then have the following total sample space representing the possible genders of 2 children: BB, BG, GB, GG.

However, since one child was said to be a boy, then valid sample space is reduced to the following: BB, BG, GB.

Since all of these options are equally likely, the answer is simply 1/3.

Solution #5.15

Let A denote the event that there is a letter in the 8th drawer, and B denote the event that the first 7 drawers are all empty.

The probability of B occurring can be found by conditioning on whether a letter was put in the drawers or not; if so, then each drawer is equally likely to contain a letter, and if not, then none contain the letter. Therefore, we have the following:

$$P(B)=\left(\frac{1}{2}\right)\left(\frac{1}{8}\right)+\left(\frac{1}{2}\right)(1)=\frac{9}{16}$$

For A and B to both occur, we also know that: $P(A\cap B)=\left(\frac{1}{2}\right)\left(\frac{1}{8}\right)=\frac{1}{16}$

Therefore, we have: $P(A\mid B)=\frac{P(A\cap B)}{P(B)}=\frac{1}{9}$

Solution #5.16

We can use a recursive formulation. Let p be the probability that the first player wins. Assume the score is 0-0 (on a relative basis).

If the first player wins a game (with probability 0.6), then two outcomes are possible: with probability 0.6 the first player wins, and with probability 0.4 the score is back to 0-0, with p being the probability of the first player winning overall.

Similarly, if the first player loses a game (with probability 0.4), then with probability 0.6 the score is back to 0-0 (with p being the probability of the first player winning), or, with probability 0.4, the first player loses. Therefore, we have: $p = 0.6^2 + 2(0.6)(0.4)p$

Solving this yields the following for p: $p \approx 0.692$

The key idea to solving this and similar problems is that, after two points, either the game is over, or we're back where we started. We don't need to ever consider the third, fourth, etc., points in an independent way.

Solution #5.17

The first card will always be a unique color and number, so let's consider the second card. Let A be the event that the color of card 2 does not match that of card 1, and let B be the event that the number of card 2 does not match that of card 1. Then, we want to find the following:

$$P(A\cap B)$$

Note that the two events are mutually exclusive: two cards with the same colors cannot have the same numbers, and vice versa. Hence, $P(A\cap B) = P(A)P(B|A)$

For A to occur, there are 40 remaining cards of a color different from that of the first card drawn (and 49 remaining cards altogether). Therefore,

$$P(A)=\frac{40}{49}$$

For B, we know that, of the 40 remaining cards, 36 of them (9 in each color) do not have the same number as that of card 1.

Therefore, $P(B\mid A)=\frac{36}{40}$

Thus, the desired probability is: $P(A\cap B)=\frac{40}{49}*\frac{36}{40}=\frac{36}{49}$

Solution #5.18

Consider the first nine dice. The sum of those nine dice will be either 0, 1, 2, 3, 4, or 5 modulo 6. Regardless of that sum, exactly one value for the tenth die will make the sum of all 10 divisible by 6. For instance, if the sum of the first nine dice is 1 modulo 6, the sum of the first 10 will be divisible by 6 *only* when the tenth die shows a 5. Thus, the probability is 1/6 for any number of dice, and, therefore, the answer is simply 1/6.

Solution #5.19

To assess the amount A is willing to pay, we need to calculate the expected probabilities of winning for each player, assuming A goes first. Let the probability of A winning (if A goes first) be given by $P(A)$, and the probability of B winning (if A goes first but doesn't win on the first roll) be $P(B')$.

Then we can use the following recursive formulation: $P(A) = \frac{1}{6} + \frac{5}{6}(1 - P(B'))$

Since A wins immediately with a 1/6 chance (the first roll is k), or with a 5/6 chance (assuming the first roll is not a k), A wins if B does not win, with B now going first.

However, notice that, if A doesn't roll side k immediately, then $P(B') = P(A)$, since now the game is exactly symmetric with player B going first.

Therefore, the above can be modeled as follows: $P(A) = \frac{1}{6} + \frac{5}{6} - \frac{5}{6}P(A)$

Solving yields $P(A) = 6/11$, and $P(B) = 1 - P(A) = 5/11$. Since the payout is \$100, then A should be willing to pay an amount up to the difference in expected values in going first, which is $100 * (6/11 - 5/11) = 100/11$, or about \$9.09.

Solution #5.20

Let $P(H)$ be the probability of landing on heads, and $P(T)$ be the probability of landing tails for any given flip, where $P(H) + P(T) = 1$. Note that it is impossible to generate fair odds using only one flip. If we use two flips, however, we have four outcomes: HH, HT, TH, and TT. Of these four outcomes, note that two (HT, TH) have equal probabilities since $P(H) * P(T) = P(T) * P(H)$. We can disregard HH and TT and need to complete only two sets of flips, e.g., HHT wouldn't be equivalent to HT.

Therefore, it is possible to generate fair odds by flipping the unfair coin twice and assigning heads to the HT outcome on the unfair coin, and tails to the TH outcome on the unfair coin.

Solution #5.21

The only possible candidates for the cube you selected are the following: either it is the inside center piece (in which case all faces are white) or a middle face (where 5 faces are white, and one face is green).

The former can be placed in six different ways, and the latter can only be placed in one particular way. Since all cubes are chosen equally randomly, let A be the event that the bottom face of the cube picked is white, and B be the event that the other five faces are white.

Note that there is a 1/27 chance that the piece is the center piece and a 6/27 chance that the piece is the middle piece. Therefore, the probability of B happening is given by the following:

$$P(B) = \frac{1}{27}(1) + \frac{6}{27}\left(\frac{1}{6}\right)$$

Then, using Bayes' rule: $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{27}}{\frac{1}{27} + \frac{6}{27} * \frac{1}{6}} = \frac{1}{2}$

Solution #5.22

Assume that the stick looks like the following, with cut points at X and Y

----------X----|----Y--------

Let M (shown as | above) denote the stick's midpoint at 0.5 of the stick's 1-unit length. Note that, if X and Y fall on the same side of the midpoint, either on its left or its right, then no triangle is possible,

because, in that case, the length of one of the pieces would be greater than 1/2 (and thus we would have two sides having a total length strictly less than that of the longest side, making forming a triangle impossible). The probability that X and Y are on the same side (since the breaks are assumed to be chosen randomly) is simply 1/2.

Now, assume that X and Y fall on different sides of the midpoint. If X is further to the left in its half than Y is in its half, then no triangle is possible in that case, since then the part lying between X and Y would have a length strictly greater than 0.5 (for example, X at 0.2 and Y at 0.75). This has a 1/2 chance of occurring by a simple symmetry argument, but it is conditional on X and Y being on different sides of the midpoint, an outcome which itself has a 1/2 chance of occurring. Therefore, this case occurs with probability 1/4. The two cases represent all cases in which no valid triangle can be formed; thus, it follows that probability of a valid triangle being formed equals $1 - 1/2 - 1/4 = 1/4$.

Solution #5.23

Note that both sequences require a heads first, and any sequence of just tails prior to that is irrelevant to either showing up. Once the first H appears, there are three possibilities. If the next flip is an H, HHT will inevitably appear first, since the next T will complete that sequence. This has probability 1/2.

If the next flip is a T, there are two possibilities. If TT appears, then HTT appeared first. This has probability 1/4. Alternatively, if TH appears, we are back in the initial configuration of having gotten the first H. Thus, we have:

$$p = \frac{1}{2} + \frac{1}{4}p$$

Solving yields $p = \frac{2}{3}$

Solution #5.24

Let A be the event that the first toss is a heads and B be the event that the second toss is a heads. Then, for the first case, we are assessing: $P(A \cap B|A \cup B)$, whereas for the second case we are assessing: $P(A \cap B|B)$

For the first case, we have: $P(A \cap B \mid A \cup B) = \frac{P(A \cap B) \cap P(A \cup B)}{P(A \cup B)} = \frac{P(A \cap B)}{P(A \cup B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$

And, for the second case, we have: $P(A \cap B \mid B) = \frac{P(A \cap B) \cap P(B)}{P(B)} = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$

Therefore, the second case is more likely. For an unfair coin, the outcomes are unchanged, because it will always be true that $P(A \cup B) > P(B)$, so the first case will always be less probable than the second case.

Solution #5.25

Note that the ants are guaranteed to collide unless they each move in the exact same direction. This only happens when all the ants move clockwise or all move counter-clockwise (picture the triangle in 2D). Let P(N) denote the probability of no collision, P(C) denote the case where all ants go clockwise, and P(D) denote the case where all ants go counterclockwise. Since every ant can choose either direction with equal probability, then we have:

$$P(N) = P(C) + P(D) = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4}$$

If we extend this reasoning to k ants, the logic is still the same, so we obtain the following:

$$P(N)=P(C)+P(D)=\left(\frac{1}{2}\right)^k+\left(\frac{1}{2}\right)^k=\frac{1}{2^{k-1}}$$

Solution #5.26

Let A be the event that the total number of heads after n tosses is even, B be the event that the first toss was tails, and B' be the event that the first toss was heads. By the law of total probability, we have the following: $P(A) = P(A|B)P(B) + P(A|B')P(B')$

Then, we can write the recurrence relation as follows: $P_n = (1-p)P_{n-1} + p(1-P_{n-1})$

Solution #5.27

Note that since Alice can win with probability p, Bob, by definition, can win with probability $1-p$. Denote $1-p$ as q for convenience. Let B_i represent the event that Bob wins i matches for $i = 0, 1, 2$. Let $B*$ denote the event that Bob wins the entire series. We can use the law of conditional probability as follows:

$$P(B^*) = P(B^* \mid B_2) * P(B_2) + P(B^* \mid B_1) * P(B_1) + P(B^* \mid B_0) * P(B_0)$$

Since Bob wins each round with probability $1-p$, we have: $P(B_2) = q^2$, $P(B_1) = 2pq$, $P(B_0) = p^2$

Substituting these values into the above expression yields: $P(B^*) = 1 * q^2 + P(B^*) * 2pq + 0 * p^2$

Hence, the desired probability is the following: $P(B^*)=\dfrac{q^2}{1-2qp}$

Solution #5.28

Because the median of three numbers is the middle number, the median is at least 1.5 if at most one of the 3 is strictly less than 1.5 (since the other 2 must be strictly greater than 1.5). Since each is uniformly randomly distributed, then the probability of any one of them being strictly less than 1.5 is given by the following:

$$\frac{1.5}{2}=\frac{3}{4}$$

Therefore, the chance that at most one is strictly less than 1.5 is given by the sum of probabilities for exactly one being strictly less than 1.5 and none being strictly less than 1.5

$$p=\binom{3}{1}\left(\frac{3}{4}\right)\left(\frac{1}{4}\right)^2+\left(\frac{1}{4}\right)^3=\frac{10}{64}$$

Solution #5.29

Let p be the probability that a phone number has the last 4 digits involving only the above given digits (0, 1, 4 and 9).

We know that the total number of possible last 4 digit combinations is $10^4 = 10000$, since there are 10 digits (0-9). There are 4! ways to pick a 4-digit permutation of 0, 1, 4 and 9.

Therefore, we have: $p=\dfrac{4!}{10000}=\dfrac{3}{1250}$

Now, since you have 150 friends, the probability of there being exactly 3 with this combination is given by:

$$\binom{150}{3}p^3(1-p)^{147}\approx 0.00535$$

Therefore, the odds of this happening are very low.

Solution #5.30

Let B be the event that all n rolls have a value less than or equal to r. Then we have:

$$P(B_r)=\frac{r^n}{6^n}$$

since all n rolls must have a value less than or equal to r. Let A be the event that the largest number is r. We have: $B_r = B_{r-1} \cup A_r$ and, since the two events on the right-hand side are disjoint, we have the following: $P(B_r) = P(B_{r-1}) + P(A_r)$

Therefore, the probability of A is given by: $P(A_r)=P(B_r)-P(B_{r-1})=\dfrac{r^n}{6^n}-\dfrac{(r-1)^n}{6^n}$

Solution #5.31

Let p be the probability that the amoeba(s) die out. At any given time step, the probability of dying out eventually must still be p.

For case (1), we know the probability of survival is 0 (for one amoeba).

For case (2), we know the probability of dying out is p.

For case (3), there are now two amoebas, and both have a probability p of dying.

For case (4), each of the three amoebas has a probability p of dying.

Putting all four together, we note that the probability of the population dying out at $t = 0$ minutes must be the same as the probability of the population dying out at $t = 1$ minutes. Therefore, we have:

$$p=\frac{1}{4}(1+p+p^2+p^3)$$

and solving this yields: $p=\sqrt{2}-1$

Solution #5.32

Note that there are $\binom{n-1}{k}$ ways to choose k heads with the first coin being a T, and a total of

$$\binom{n}{k}$$

ways to obtain k heads. So, the probability of having a tails first is given by: $\dfrac{\binom{n-1}{k}}{\binom{n}{k}}=\dfrac{n-k}{n}$

and, therefore, the probability of obtaining a heads first is given by the following:

$$1-\frac{n-k}{n}=\frac{k}{n}$$

Solution #5.33

Let the n draws be denoted as $X_1, X_2, \ldots X_n$

We know that, for any given draw i, we have the following:

$$P(X_i>Y)=1-P(X_i\le Y)=1-P\left(\frac{X_i-u}{\sigma}\le\frac{Y-\mu}{\sigma}\right)=1-\Phi\left(\frac{Y-\mu}{\sigma}\right)$$

Additionally, the probability that k of those draws are greater than Y follows a binomial distribution with the value above being the p parameter:

$$p = 1 - \Phi\left(\frac{Y - \mu}{\sigma}\right)$$

Then, the desired probability is given by: $\binom{n}{k} p^k (1-p)^{n-k}$

Solution #5.34

Note that, without loss of generality, the first point can be located at (1, 0). Using the polar coordinate system, we have the two other points at angles: θ, ϕ, respectively.

Note that the second point can be placed on either half (top or bottom) without loss of generality. Therefore, assume that it is on the top half. Then, $0 < \theta < \pi$

If the third point is also in the top half, then the resulting triangle will not contain the center of the unit circle. It will also not contain the center if the following is the case (try drawing this out): $\phi \geq \theta + \pi$

Therefore, for any given second point, the probability of the third point making the resulting triangle contain the center of the unit circle is the following:

$$p = \frac{\theta}{2\pi}$$

Therefore, the overall probability is given by the integrating over possible values of θ, where the constant in front is to take the average:

$$\frac{1}{\pi}\int_0^\pi p\,d\theta = \frac{1}{4}$$

Solution #5.35

In order to run out of white balls first, all the white balls must be drawn before the r-th red ball is drawn. We can consider the draws until $w + r - 1$ (we know the last ball must be red), and count how many include w white balls.

The first white ball has $w + r - 1$ options, the second white ball has $w + r - 2$ options, etc., until the drawing of the w-th white ball: $(w + r - 1)(w + r - 2)...(r)$, which can be written as a factorial:

$$\frac{(w+r-1)!}{(r-1)!}$$

Similarly, there are $r!$ ways to arrange the drawing of the remaining r red balls. We know the total number of balls is $r + w$, so there are $(r + w)!$ total arrangements. Therefore, the probability is:

$$\frac{\frac{(w+r-1)!}{(r-1)!}r!}{(r+w)!} = \frac{r}{w+r}$$

A more intuitive way to approach the problem is to consider just the last ball drawn. The probability that the ball is red is simply the chance of it being red when picking randomly, which is the following:

$$\frac{r}{w+r}$$