**Chapter 1 General Principles**

Let us begin this book by exploring five general principles that will be extremely helpful in your interview process. From my experience **on** both sides of the interview table, these general guidelines will better prepare you for job interviews and will likely make you a successful candidate.

1. **Build a broad knowledge base**

The length and the style of quant interviews differ from firm to firm. Landing a quant job may mean enduring hours of bombardment with brain teaser, calculus, linear algebra, probability theory, statistics, derivative pricing, or programming problems. To be a successful candidate, you need to have broad knowledge in mathematics, finance and programming.

Will all these topics be relevant for your future quant job? Probably not. Each specific quant position often requires only limited knowledge in these domains. General problem solving skills may make more difference than specific knowledge. Then why are quantitative interviews so comprehensive? There arc at least two reasons for this:

The first reason is that interviewers often have diverse backgrounds. Each interviewer has his or her own favorite topics that are often related to his or her own educational background or work experience. As a result, the topics you will be tested on are likely to be very broad. The second reason is more fundamental. Your problem solving skills—a crucial requirement for any quant job -isoften positively correlated to the breadth of your knowledge. A basic understanding of a broad range of topics often helps you better analyze problems, explore alternative approaches, and conic up with efficient solutions. Besides, your responsibility may not be restricted to your own projects. You will be expected to contribute as a member of a bigger team. Having broad knowledge will help you contribute to the team's success as well.

The key here is "basic understanding." Interviewers do not expect you to be an expert on a specific subject—unless it happens to be your PhD thesis. The knowledge used in interviews, although broad, covers mainly essential concepts. This is exactly the reason why most of the books I refer to in the following chapters have the word -introduction" or "first". in the title. If I am allowed to give only one suggestion to a candidate, it will be **know the basics yen' well.**

1. **Practice your interview skills**

The interview process starts long before you step into an interview room. In a sense, the
  
success or thilure of your interview is often determined before the first question is asked.
  
Your solutions to interview problems may fail to reflect your true intelligence and

General Principles

**Chapter 2 Brain Teasers**

knowledge if you are unprepared. Although a complete review of quant interview problems is impossible and unnecessary, practice does improve your interview skills. Furthermore, many of the behavioral, technical and resume-related questions can be anticipated. So prepare yourself for potential questions long before you enter an interview room.

1. **Listen carefully**

You should be an active listener in interviews so that you understand the problems well before you attempt to answer them. If any aspect of a problem is not clear to you politely ask for clarification. If the problem is more than a couple of sentences, jot down the key words to help you remember all the information. For complex problems, interviewers often give away some clues when they explain the problem. Even the assumptions they give inay include some information as to how to approach the problem. So listen carefully and make sure you get the necessary information.

1. **Speak your mind**

When you analyze a problem and explore different ways to solve it, never do it silently. Clearly demonstrate your analysis and write down the important steps involved if necessary. This conveys your intelligence to the interviewer and shows that you are methodical and thorough. In case that you go astray, the interaction will also give your interviewer the opportunity to correct the course and provide you with some hints.

Speaking your mind does not mean explaining every tiny detail. If some conclusions are obvious to you, simply state the conclusion without the trivial details. More often than not, the interviewer uses a problem to test a specific concept/approach. You should focus on demonstrating your understanding of the key concept/approach instead of dwelling

on less relevant details.

1. **Make reasonable assumptions**

lit real job settings, you are unlikely to have all the necessary information or data you'd prefer to have bc.'fore you build a model and make a decision. In interviews, interviewers may not give you all the necessary assumptions either. So it is up to you to make reasonable assumptions. The keyword here is reasonable. Explain your assumptions to the interviewer so that you will get immediate feedback. '1'0 solve quantitative problems, it is crucial that you can quickly make reasonable assumptions and design appropriate frameworks to solve problems based on the assumptions.

e ;,•e now ready to review basic concepts inquantitative finance subject areas and have :lin solving real-world interview problems!

In this chapter, we cover problems that only require common sense, logic, reasoning, and basic—no more than high school level—math knowledge to solve. In a sense, they are real brain teasers as opposed to mathematical problems in disguise. Although these brain teasers do not require specific math knowledge, they are no less difficult than other quantitative interview problems. Some of these problems test your analytical and general problem-solving skills; some require you to think out of the box; while others ask you to solve the problems using fundamental math techniques in a creative way. In this chapter, we review sonic interview problems to explain the general themes of brain teasers that you are likely to encounter in quantitative interviews.

**2.1 *Problem Simplification***

If the original problem is so complex that you cannot come up with an immediate solution, try to identify a simplified version of the problem and start with it. Usually you can start with the simplest sub-problem and gradually increase the complexity. You do not need to have a defined plan at the beginning. Just try to solve the simplest cases and analyze your reasoning. More often than not, you will find a pattern that will guide you

through the whole problem.

**Screwy pirates**

Five pirates looted a chest full of 100 gold coins. Being a bunch of democratic pirates, they agree cm the following method to divide the loot:

The most senior pirate will propose a distribution of the coins. All pirates, *including the mosr senior pirate,* will then vote. If at least 50% of the pirates (3 pirates in this case) accept the proposal, the gold is divided as proposed. If not, the most senior pirate will be fed to shark and the process starts over with the next most senior pirate... The process is repeated until a plan is approved. You can assume that all pirates arc perfectly rational: they want to stay alive first and to get as much gold as possible second. Finally, being blood-thirsty pirates. they want to have fewer pirates on the boat if given a choice

between otherwise equal outcomes.

How will the gold coins be divided in the end?

*Solution..* If you have not studied game theory or dynamic programming, this strategy problem may appear to be daunting. If the problem with 5 pirates seems complex, we can always *,vtari with a simplified version of the problem* by reducing the number of pirates. Since the solution to 1-pirate case is trivial, let's start with 2 pirates. The senior

pirate (labeled as 2) can claim all the gold since he will always get 50% of the votes from himself and pirate 1 is left with nothing.

Brain Teasers A Practical Guide To Quantitative Finance interviews

I..et's add a more senior pirate, 3. He knows that if his plan is voted down, pirate 1 will 1.,,t nothing. But if he offers private 1 nothing. pirate 1 will be happy to kill him. So pirate 3 will offer private 1 one coin and keep the remaining 99 coins, in which strategy the plan will have 2 votes from pirate I and 3.

If pirate 4 is added, he knows that if his plan is voted down, pirate 2 will get nothing. So pirate 2 will settle for one coin if pirate 4 offers one, So pirate 4 should offer pirate 2 one coin and keep the remaining 99 coins and his plan will be approved with 50% of the votes from pirate 2 and 4.

Now we finally come to the 5-pirate case. He knows that if his plan is voted down, both pirate 3 and pirate 1 will get nothing. So he only needs to offer pirate 1 and pirate 3 one coin each to get their votes and keep the remaining 98 coins. If he divides the coins this way. he will. have three out of the five votes: from pirates 1 and 3 as well as himself.

Once we start with a simplified version and add complexity to it, the answer becomes obvious. Actually after the case *n 5,* a clear pattern has emerged and we do not need to stop at 5 pirates. For any 2n +1 pirate case *(n* should be less than 99 though), the most senior pirate will offer pirates 1. 3, and 2n -1 each one coin and keep the rest for

himsel f.

**Tiger and sheep**

One hundred tigers and one sheep are put on a magic island that only has

grass. Tigers

can eat grass, but they would rather eat sheep. Assume: A. Each time only one tiger can eat one sheep, and that tiger itself will become a sheep after it eats the sheep. *B.* All tigers are smart and perfectly rational and they want to survive. So will the sheep be eaten?

*Solution:* 100 is a large [number. so](http://number.so) again let's *start with a simplified version* of *the problem.* it-there is only 1 tiger *(n I* ), surely it will eat the sheep since it does not need to worry about being eaten. How about 2 tigers? Since both tigers are perfectly rational, either tiger probably would do some thinking as to what will happen if it eats the sheep.

ther tiger is probably thinking: if I eat the sheep, I will become a sheerand then I will *be* eaten by the other tiger. So to guarantee the highest likelihood of survival, neither tiger will eat the sheep.

If there are 3 tigers, the sheep will be eaten since each tiger will realize that once it
  
changes to a sheep, there will be 2 tigers left and it will not be eaten. So the first tiger
  
that thinks this through will eat the sheep. If there are 4 tigers, each tiger will understand

that if it eats the sheep, it will turn to a sheep. Since there are 3 other tigers, it will be eaten. So to guarantee the highest likelihood of survival, no tiger will eat the sheep.

Following the same logic, we can naturally show that if the number of tigers is even, the
  
sheep will not be eaten. If the number is odd, the sheep will be eaten. For the case

**4** 5

*n IN,* the sheep will not be eaten.

***2.2 Logic Reasoning***

**River crossing**

Four people, *A, B. C* and *D* need to get across a river. The only way to cross the river is by an old bridge, which holds at most 2 people at a time. Being dark, they can't cross the bridge without a torch, of which they only have one. So each pair can only walk at the speed of the slower person. They need to get all of them across to the other side as quickly as possible, *A* is the slowest and takes 10 minutes to cross; *B* takes 5 minutes; *C*

takes 2 minutes; and *D* takes 1 minute.

What is the minimum time to get all of them across to the other side?'

*Solution:* The key point is to realize that the mil ute person and this should not happen *in* hdooe. EOMg-o back-. So C and *D* should go across and .11 go across (10 min); send *C* back (2rin);

10-minute person should go with the

5-me Erg crossing, kithei.wise one of them first (2 min); then send *D* back (l mini; A *C* and *D* go across again (2 miri).

can send *C* back first and then D back in ell.

ce-14.41r) ,

It takes 17 minutes in total. Alternatively, we the second round, which takes 1? minutes as w

**Birthday problem**

and your colleagues knom, that your bos:- A's birthday is one of the, following

You

10

(-6



dates: 4 ;1‹.-.

Mar 4, NI:or 5, Mar 8 a

Jun 4,3u0-7 • I

Sep 1. Sv 5 f

Dec 1. Dpe- Dec 8

*A* told you only the month 'of his birthday. and told your colleague C only the day. After
  
that, you first said: "TdOn't know A's birthday: *C* doesn't know it either.' After hearing

Hint: The **key is** to realize that .4 and *B* should het across the bridge together

what you said, *C* replied: "I didn't know A's birthday, but now I know it." You smiled and said: "Now I know it, too." After looking at the 10 dates and hearing your COM meats, your administrative assistant wrote down A's birthday without asking any questions. So what did the assistant write?

Brain Teasers A Practical Guide To Quantitative Finance Interviews

*So/w/o/2:* Don't let the "he said, she said" part confuses you. Just interpret the logic behind each individual's comments and try your best to derive useful information from these comments.

*Let D* be the day of the month of A's birthday, we have *D* E 0,2,4,5,7,81, If the

birthday is on a unique day, *C* will know the A's birthday immediately. Among possible Ds, 2 and 7 are unique days. Considering that you are sure that *C* does not know A's birthday, you must infer that the day the *C* was told of is not 2 or 7. Conclusion: the month is not June or December. (If the month had been June, the day C was told of may have been 2; if the month had been December, the day *C* was told of may have been 7.)

Now C knows that the month must be either March or September. He immediately figures out A's birthday, which means the day must be unique in the March and September list. It means A's birthday cannot be Mar 5, or Sep 5. Conclusion: the birthday must be Mar 4, Mar 8 or Sep 1.

Among these three possibilities left, Mar 4 and Mar 8 have the same month. So if the
  
month you have is March, you still cannot figure out A's birthday. Since you can figure
  
out A's birthday, A's birthday must be Sep 1. Hence, the assistant must have written Sep

1.

**Card game**

A casino offers a card game using a normal deck of 52 cards. The rule is that you turn, over two cards each time. For each pair, if both are black, they go to the dealer's pile; If both are red, they go to your pile., if one black and one red, they are discarded. The process is repeated until you two go through all 52 cards. If you have more cards in your pile. you win $100; otherwise (including ties) you get nothing. The casino allows you to negotiate the price you want to pay for the game\_ How much would you be wilting to

pay to play this gaine?2

*Sohition;* This surely is an insidious casino. No matter how the cards are arranged, you
  
and the dealer will always have the same number of cards in your piles. Why? Because
  
each pair of discarded cards have one black card and one red card, so equal number of

hint: Fry to approach the problem using symmetty. Each discarded pair has one black and one red card. What does that tell you as to the number of black and red cards in the rest two piles? red and black cards are discarded. As a result, the number of red cards left for you and the number of black cards left for the dealer are always the same. The dealer always wins! So we should not pay anything to play the game.

7

**Burning ropes**

You have two ropes, each of which takes 1 hour to burn. But either rope has different densities at different points, so there's no guarantee of consistency in the time it takes different sections within the rope to bum. How do you use these two ropes to measure 45 minutes?

*Solution:* This is a classic brain teaser question. For a rope that takes x minutes to burn, if you light both ends of the rope simultaneously, it takes *x12* minutes to burn. So we should light both ends of the first rope and light one end of the second rope. 30 minutes later, the first rope will get completely burned, while that second rope now becomes a 30-min rope. At that moment, we can light the second rope at the other end (with the first end still burning), and when it is burned out, the total time is exactly 45 minutes.

**Defective ball**

You have 12 identical balls. One of the balls is heavier OR lighter than the rest (you don't know which). Using just a balance that can only show you which side of the tray is heavier, how can you determine which ball is the defective one with 3 measurements?3

*Solution:* This weighing problem is another classic brain teaser and is still being asked by many interviewers. The total number of balls often ranges from 8 to more than 100. Here we use *n* to show the fundamental approach. The key is to separate the original 1roup (as well as any intermediate subgroups) into three sets instead of two. The reason is that the •Comparii6T-iof the first two groups always gives inkirnat on about the

third group.

Considering that the solution is wordy to explain, I draw a tree diagram in Figure 2.1 to show the approach in detail. Label the halls 1 through 12 and separate them to three groups with 4 balls each. Weigh balls 1, 2, 3. 4 against balls 5, 6, 7.. 8. Then we go on to explore two possible scenarios: two groups balance, as expressed using an -" sign, or 1,

3

lint: First do it for 9 identical balls and use only 2 measurements. knowing that one is heavier than the

rest.

2, 3, 4 are lighter than 5, 6, 7, 8, as expressed using an "<" sign. There is no need to explain the scenario that 1, 2, 3, 4 are heavier than 5, 6, 7, 8. (Why?)

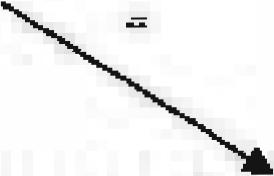
Brain Teasers A Practical Guide To Quantitative Finance Interviews

If the two groups balance, this immediately tells us that the defective ball is in 9, 10, 11 and 12, and it is either lighter *(L)* or heavier *(H)* than other balls. Then we take 9, 10 and 11 from group 3 and compare balls 9, 10 with 8, 11. Here we have already figured out that 8 is a normal ball. If 9, 10 are lighter, it must mean either 9 or 10 is *L* or 11 is *H.* In which case, we just compare 9 with 10. If 9 is lighter, 9 is the defective one and it is *L;* if 9 and 10 balance, then 11 must he defective and *H;* If 9 is heavier, 10 is the defective one and it is *L.* If 9, 10 and 8, 11 balance, 12 is the defective one. If 9, 10 is heavier, than either 9 or 10 is *if,* or 11 is *L.*

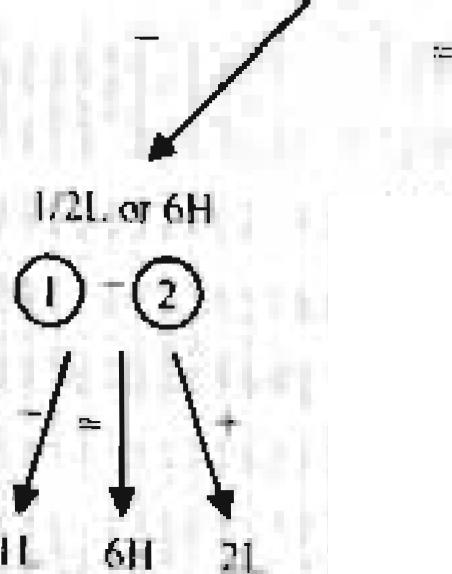
You can easily follow the tree in Figure 2.1 for further analysis and it is clear from the tree that all possible scenarios can he resolved in 3 measurements.

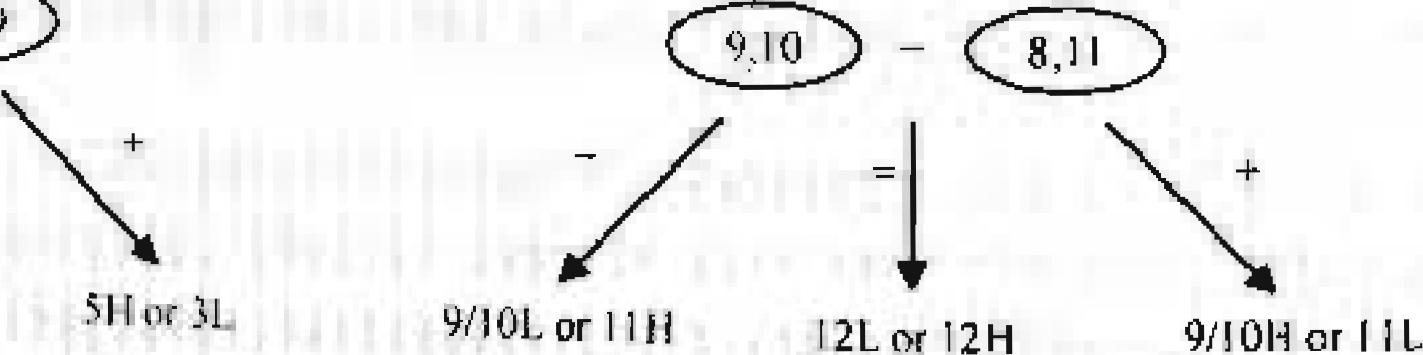


**1/2,314** L or 5/6/7/8 I r



9/10/11/12 1\_ or 1-1

**CLCI 0 -0**



51 f c31.

91101. or 1111 I 2L or 12H

9/101-1(5r I ; L

4L cr 1.8.H

\\4;

811 41. 711

**511 gl** 111-1 N. 1211 121. 1011 111. 911

**Figure 2.1 Tree diagram to identify the defective bail in 12 balls**

In general if you have the information as to whether the defective ball is heavier or

Here k where the symmetry idea conies in. Nothing makes the I , 3, 4 Or 5, 6, 7, 8 labels special. If I, *2.*

3, 4 art: heavier than 5, 6, 7, 8, ler siost exchange the Iabeis of these two grow s. Again we have the case of ", 3. 4 heino, lighter than 5, 6. 7.8. p

8

lighter, you can identify the defective ball among up to 3" balls using no more than *n*measurements since each weighing reduces the problem size by 2/3, If you have no
  
information as to whether the defective ball is heavier or lighter, you can identify the

defective ball among up to (3" -3)/ 2 balls usini no more than *n* measurements.





**Trailing zeros**

How many trailing zeros are there in 100! (factorial of 100)?

*Sohttion:* This is an easy problem. We know that each pair of 2 and 5 will give a trailing zero. If we perform prime number decomposition on all the numbers in 100!, it is obvious that the frequency of 2 will far outnumber of the frequency of 5. So the frequency of 5 determines the number of trailing zeros. Among numbers 1, 2,- • -,99, and

100, 20 numbers are divisible by 5 (5, 10, - .•, 100 ). Among these 20 numbers, 4 are divisible by 52 ( 25, 50, 75, 100). So the total frequency of 5 is 24 and there are 24 trailing zeros.

**Horse race**

There are 25 horses, each of which runs at a constant speed that is different from the other horses'. Since the track only has 5 lanes, each race can have at most 5 horses. If you need to find the 3 fastest horses, what is the minimum number of races needed to identify them?

*Solution:* This problem tests your basic analytical skills. To find the 3 fastest horses, surely all horses need to he tested. So a natural first step is to divide the horses to 5 groups (with horses 1-5, 6-10, 11-15, 16-20, 21-25 in each group). After 5 races, we will have the order within each group, let's assume the order follows the order of numbers (e.g., 6 is the fastest and 10 is the slowest in the 6-10 group)'. That means 1. 6, 11, 16 and 21 arc the fastest within each group.

Surely the last two horses within each group are eliminated. What else can we infer? We know that within each group, if the fastest horse ranks 5th or 4th among 25 horses, then all horses in that group cannot he in top 3; if it ranks the 3rd, no other horse in that group can be in the top 3; if it ranks the 2nd, then one other horse in that group may be in top 3; if it ranks the first, then two other horses in that group may be in top 3.

5 Such an assumption does not affect the generality of the solution. If the order is not as described, just change the labels of the horses.

9

So let's race horses 1, 6, 11, 16 and 21. Again without loss of generality, let's assume the order is 1, 6, 11, 16 and 21. Then we immediately know that horses 4-5, 8-10, 12-15, 16-20 and 21-25 are eliminated. Since I is fastest among all the horses, 1 is in. We need to determine which two among horses 2, 3, 6, 7 and 11 are in top 3, which only takes one extra race.

Brain Teasers

A Practical Guide To Quantitative Finance Interviews

So all together we need 7 races (jri 3 rounds) to identify the 3 fastest horses. **Infinite sequence**

If x"x"x"x"x••• := 2 , where x what is x?

*So/ut/on:* This problem appears to be difficult. but a simple analysis will give an elegant solution. What do we have from the original equation?

lim *x* a x ' x"x A x - • • = 2 \_\_> lim x " x *"x* "x"x- • • =2. In other

words, as ;7 —› co,

*1.--), . frl--ir:r\_*

*FZ—I ferifi.*

adding or minus one x A should yield the same result.

x x x "x - • • .,..,c,\(xAxAxAx...)=xA)= 2 *N.7*

***2.3 Thinking Out of the Box***

**Box packing**

Can you pack 53 bricks of dimensions l x 1 x 4 into a 6x 6 x 6 box?

*Sol*u*tion:* This is a nice problem extended from a popular chess board problem. In that problem, you have a 8x 8 chess board with two small squares at the opposite. diagonal corners removed. You have many bricks with dimension lx 2, Can you pack 31 bricks into the remaining 62 squares? (An alternative question is Whether you can cover all 62 squares using bricks without any bricks overlapping with each other or sticking out of the board, which requires a similar analysis.)

A real chess board figure surely helps the •isualization. As shown in Figure 2.2, when a chess board is filled with alternative black and white squares, both squares at the opposite diagonal corners have the same color. If you put a lx 2 brick on the board, it will always cover one black square and one white square. Let's say it's the two black corner squares were removed, then the rest of the board can fit at most 30 bricks sincq

we only have. 30 black squares left (arid each brick requires one black square), so to ,

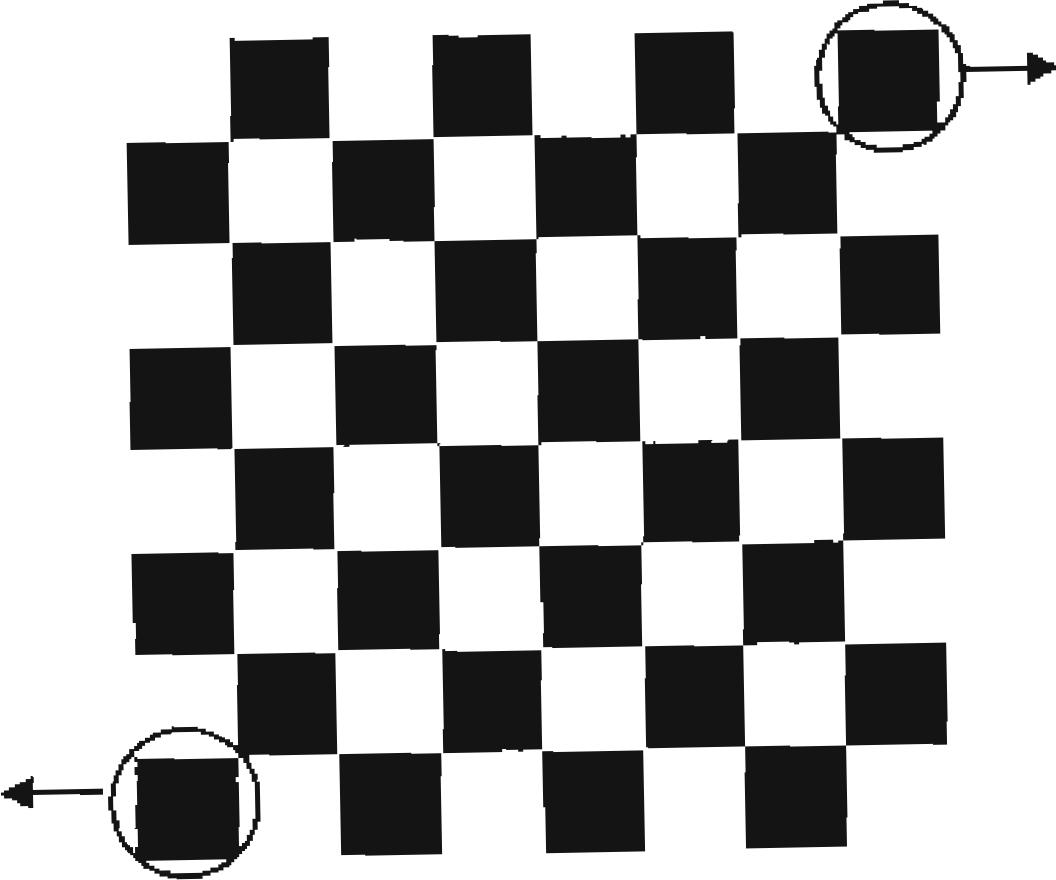
rink 31 bricks is out of the question. To cover all 62 squares without overla ppi ng or

overreaching, we must have exactly 31 bricks. Yet we have proved that 31 bricks cannot

fit in the 62 squares left, so you cannot find a way to fill in all 62 squares without overlapping or overreaching.

10 11

Removed



Removed

Figure 2.2 Chess board with alternative black and white squares

Just as any good trading strateuy, if more and more people get to know it and replicate it, the effectiveness of such a strategy will disappear. As the chess board problem becomes popular, many intervievvees simply commit it to memory (after all, it's easy to remember the answer). So some ingenious interviewer came up with the newer version to test your thinking process, or at least your ability to extend your knowledge to new problems.

If we look at the total volume in this 3D problem, 53 bricks have a volume of 212, which is smaller then the box's volume 216. Yet we can show it is impossible to pack all the bricks into the box using a similar approach as the chess board problem. Let's imagine that the 6x 6 x 6 box is actually comprised of small 2 x 2 x 2 cubes. There should be 27 small cubes. Similar to the chess board (but in 3D), imagine that we have black cubes and white cubes alternates--it does take a little 3D visualization. So we have either 14 black cubes & 13 white cubes or 13 black cubes & 14 white cubes. For any 1x1x 4 brick that we Rack into the box, half (1x1 x 2) of it\_must be in a black 2 x 2 x 2 cube and the other half must be in ti white 2 x 2 x 2 cube. The problem is that each 2 x 2 x 2 cube can only Ix, used by 4 of the 1 xlx 4 bricks. So for the color with 13 cubes, be it black or white, we can only use them for 52 l x 1 x 4 tubes. There is no way to place the 53th brick. So we cannot pack 53 bricks of dimensions 1 x 1 x 4 into a 6x 6 x 6 box.

**Calendar cubes**

You just had tvvo dice custom-made. Instead of numbers 1 -- 6, you place single-digit numbers on the faces of each dice so that every morning you can arrange the dice in a way as to make the two front faces show the current day of the month. You must use both dice (in other words, days I — 9 must be shown as 01 — 09), but you can switch the

order of the dice if you want. What numbers do you have to put on the six faces of each of the two dice to achieve that?

Brain Teasers A Practical Guide To Quantitative Finance Interviews

*Solution:* The days of a month include 11 and 22, so both dice must have 1 and 2. To express single-digit days, we need to have at least a 0 in one dice. Let's put a 0 in dice one first. Considering that we need to express all single digit days and dice two cannot have all the digits from 1 - 9, it's necessary to have a 0 in dice two as well in order to express all single-digit days.

So far we have assigned the following numbers:.

If we can assign all the rest of digits 3, 4, 5, 6, 7, 8, and 9 to the rest of the faces, the
  
problem is solved. But there are 7 digits left. What can we do? Here's where you need to
  
think out of the box. We can use a 6 as a 9 since they will never be needed at the same

0

Dice one

2

Dice tvipo 1

time! So, simply put 3, 4, and *5* ore one dice and 6, 7, and 8 on the other dice, and the Resat riumbers on the two dice are:

**Door to offer**

Dice one
  
Dice two

0

You are facing two doors. One leads to your job offer and the other leads to exit. In front
  
of either door is a guard. One guard always tells lies and the other always tells the truth.
  
You can only ask one guard one yesIno question. Assuming you do want to get the job

offer. what question will you ask?

*Sohaion:* This is another classic brain teaser (maybe a little out-of-date in my opinion).
  
One popular answer is to ask one guard: -Would the other guard say that you are
  
guarding the door to the offer'?" If he answers yes, choose the other door; *if* he answers

no, choose the door this guard is standing in front of

There are two possible scenarios:

1. Truth teller guards the door to offer: Liar guards the door to exit.

ll 2. Truth teller guards the door to exit; Liar- guards the door to offer.

we ask a guard a direct question such as "Are Y ou guarding the door to the offer?" For scenario 1, both guards will answer yes; for scenario 2, both guards will answer no. So a

direct question does not help us solve the problem. The key is to involve both guards in the questions as the popular answer does. For scenario 1, if we happen to choose the truth teller, he will answer no since the liar will say no; if we happen to choose the liar guard, he will answer yes since the truth teller will say no. For scenario 2, if we happen to choose the truth teller, he will answer yes since the liar will say yes; if we. happen to choose the liar guard, he will answer no since the truth teller with say yes. So for both scenarios, if the answer is no, we choose that door; if the answer is yes, we choose the

other door. **Message delivery**

13

You need to communicate with your colleague in Greenwich via a messenger service. Your documents are sent in a padlock box. Unfortunately the messenger service is not secure, so anything inside an unlocked box will he lost (including any locks you place inside the box) during the delivery. The high-security padlocks you and your colleague each use have only one key which the person placing the lock owns. How can you securely send a document to your colleague?6

*Soht,tion:* if you have a document to deliver, clearly you cannot deliver it in an unlocked box. So the first step is to deliver it to Greenwich in a locked box, Since you are the person who has the key to that lock, your colleague cannot open the box to get the document. Somehow you need to remove the lock before he can get the document, which means the box should be sent back to you before your colleague can get the

document.

So what can he do before he sends back the box? He can place a second lock on the box, which he has the key to! Once the box is back to you, you remove your own lock and send the box back to your colleague. He opens his own lock and gets the document.

**Last ball**

A hag has 20 blue balls and 14 red balls. Each time you randomly take two balls out. (Assume each ball in the bag has equal probability of being taken). You do not put these two balls hack. Instead, if both balls have the same color, you add a blue ball to the bag; if they have different colors, you add a red ball to the bag. Assume that you have an unlimited supply of blue and red balls. if you keep on repeating this process, what will be the color of the last ball left in the bag?7 What if the bag has 20 blue balls and 13 red

balls instead?

6 Flint: You can have more than one lock on the box.

7

Flint: Consider the changes in the number of red and blue balls after each step.

*Solution:* Once you understand the hint, this problem should be an easy one. Let *(B,R)*

Brain Teasers A Practical Guide To Quantitative Finance Interviews

represent the number of blue balls and red balls in the bag. We can take a look what will happen after two balls are taken out.

Both balls are blue: *(B R) —> (B —1, R)* Both balls are red: *(13, R) -+ (B +1, 1? —* 2) One red and one blue: ( *B R) —> (B —1, R)*

Notice that *R* either stays the same or decreases by 2, so the number of red balls will never become odd if we begin with 14 red balls. We also know that the total number of balls decreases by one each time until only one ball is left. Combining the information we have, the last ball must be a blue one. Similarly, when we start with odd number of red balls, the final ball must be a red one.

**Light switches**

There is a light bulb inside a room and four switches outside, All switches are currently at off state and only one switch controls the light bulb, You *may* turn any number of switches on or off any number of times you want. How many times do you need to go into the room to figure out which switch controls the light bulb?

*solution:* You may have seen the classical version of this problem with 3 light bulbs inside the room and 3 switches outside. Although this problem is slightly modified, the approach is exact the same. Whether the light is on and off is binary, which only allows us to distinguish two switches. If we have another binary factor, there are 2x 2 4 possible combinations of scenarios, so we can distinguish 4 switches. Besides light, a light bulb also emits heat and becomes hot after the bulb has been lit for some time, So we can use the on/off and cold/hot combination to decide which one of the four switches

controls the light.

Turn on switches 1 and 2; move on to solve some other puzzles or do whatever you like
  
for *a* while; turn off switch 2 and fin on switch 3; get into the room quickly, touch the

bulb and observe \vhether the light is on or off.

The light bulb is on and hot switch t controls the light;

Thc light bulb is off and hot switch 2 controls the light;

[he light bulb is on and cold switch 3 controls the light;

The light bulb is off and cold switch 4 controls the light,

**Quant salary**

Eight quants from different banks are getting together for drinks. They are all interested in knowing the average salary of the group. Nevertheless, being cautious and humble individuals, everyone prefers not to disclose his or her own salary to the group. Can you come up with a strategy for the quants to calculate the average salary without knowing other people's salaries?

14 l5

*Solution:* This is a light-hearted problem and has more than one answer. One approach is for the first quant to choose a random number, adds it to his/her salary and gives it to the second quaint. The second quant will add his/her own salary to the result and give it to the third quant; ...; the eighth quant will add his/her own salary to the result and give it back to the first quant. Then the first quant will deduct the "random" number from the total and divide the "real" total by 8 to yield the average salary.

You may be wondering whether this strategy has any use except being a good brain teaser to test interviewees. It does have applications in practice. For example, a third party data provider collect fund holding position data (securities owned by a fund and the number of shares) from all participating firms and then distribute the information back to participants. Surely most participants do not want others to figure out what they are holding. If each position in the fund has the same fund ID every day, its easy to reverse-engineer the fund from the holdings and to replicate the strategy. So different random numbers (or more exactly pseudo-random numbers since the provider knows what number is added to the fund ID of each position and complicated algorithm is involved to make the mapping one to one) are added to the fund ID of each position in the funds before distribution. As a result, the positions in the same fund appear to have different fund IDs. That prevents participants from re-constructing other funds. Using this approach, the participants can share market information and remain anonymous at

the same time.

***2.4 Application of Symmetry***

**Coin piles**

Suppose that you are blind-folded in a room and are told that there are 1000 coins on the floor. 980 of the coins have tails up and the other 20 coins have heads up. Can you separate the coins into two piles so to guarantee both piles have equal number of heads? Assume that you cannot tell a coin's side by touching it, but you are allowed to turn over

any number of coins.

*Solution:* Let's say that we separate the 1000 coins into two piles with ii coins in one pile
  
and 1000— coins in the other. If there arc *in* coins in the first pile with heads up. there

must be 20—m coins in the second pile with heads up. We also know that there are *n— err* coins in the first pile with tails up. We clearly cannot guarantee that *rn = 10* by simply adjusting *n.*

Brain Teasers A Practical Guide To Quantitative Finance Interviews

What other options do we have? We can turn over coins if we want to. Since we have no way of knowing what a coin's side is, it won't guarantee anything if we selectively flip coins. However, if we flip all the coins in the first pile, all heads become tails and all tails become heads. As a result, it will have *n—m* heads and m tails (symmetry). So, to start, we need to make the number of tails in the original first pile equal to the number of heads in the second pile; in other words, to make /7 = 20 *— n. n-20* makes the equation hold. If we take 20 coins at random and turn them all over, the number of heads among these turned-over 20 coins should be the same as the number of heads among the other 980 coins.

**Mislabeled bags**

You are given three bags of fruits. One has apples in it; one has oranges in it; and one has a mix of apples and oranges in it. Each bag has a label on it (apple,

Unfortunately, your manager tells you that orange or mix).

ALL bags are mislabeled. Develop a strategy

to identify the bags by taking out minimum number of fruits? You can take any number of fruits from any bags.8

*ohllion:* The key here is to use the fact that ALL bags are mislabeled. For example, a bag labeled with apple must contain either oranges only or a mix of oranges and apples. Let's look at the labels: orange, apple, mix (orange + apple). Have you realized that the orange label and the apple label are symmetric? if not, let me explain it in detail: If you pick a fruit from the bag with the orange label and it's an apple (orange --> apple), then the bag is either all apples or a mix. If you pick a fruit from the bag with the apple label and it's an orange (apple orange), then the bag is either an orange bag or a mix. Symmetric labels are not exciting and are unlikely to be the correct approach. So let's try the bag with the mix label and get one fruit from it. If the fruit we get is an orange, then we know that bag is actually orange (It cannot be a mix of oranges and apples since we know the bag's label is wrong). Since the bag with the apple label cannot be apple only. it must he the znix bag. And the bag with the orange label must be the apple hag. Similarly, for the case that apples are in the bag with the mix label, we can figure out all the bags using one single pick.

**Wise men**

A sultan has captured 50 wise men. He has a glass currently standing bottom down. Every minute he calls one of the wise men who can choose either to turn it over (set it upside down or bottom down) or to do nothing. The wise men Fill be called randomly, possibly for an infinite number of times. When someone called to the sultan correctly states that all wise men have already been called to the sultan at least once, everyone goes free. But if his statement is wrong, the sultan puts everyone to death. The wise men are allowed to communicate only once before they get imprisoned into separate rooms (one per room). Design a strategy that lets the wise men go free.

*Solution:* For the strategy to work, one wise man, let's call him the spokesman, will state that every one has been called. What does that tell us? 1. All the other 49 wise men are equivalent (symmetric). 2. The spokesman is different from the other 49 men. So naturally those 49 equivalent wise men should act in the same way and the spokesman

should act differently.

Here is one of such strategies: Every one of the 49 (equivalent) wise men should flip the glass upside down the first time that he sees the glass bottom down. He does nothing if the glass is already upside down or he has flipped the glass once. The spokesman should flip the glass bottom down each time he sees the glass upside down and he should do nothing if the glass is already bottom down, After he does the 49th flip, which means all the other 49 wise men have been called, he can declare that all the wise men have been

called.

**2.5 *Series Summation***

Here is a famous story about the legendary mathematician/physicist Gauss: When he was a child, his teacher gave the children a boring assignment to add the numbers from to 100. To the amazement of the teacher, Gauss turned in his answer in less than a

minute. Here is his approach: En*= 1 +* + • • + 99 + 1001

Oil

In =100± 99+•-•+ 2 + 1

n -1

16 I7

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  | I ■  2n --- 101 + 101 + • • • + 101 + 101---- 101 x 100 | 00 100x101  = |
| The problem struck Inc a word game when I first saw it. But it does test a candidate's attention to  details besides his or her logic reasoning,. skills. | | |  |  |

*N(N +1)*

Brain Teasers A Practical Guide To Quantitative Finance Interviews

This approach can be generalized to any integer AT: En=

2

n=1

The summation formula for consecutive squares may not be as intuitive:

ti

|  |  |  |  |
| --- | --- | --- | --- |
| N(N +1)(2N +1) | *=* | *AO* | *N2 N*  *+—+—* |
| 6 | 3 | 2 6 |

*n- =*

But if we correctly guess that En' = aAT3 +MO + *eN +d* and apply the initial

*13.1*

conditions

"\_-0 0=d

N -1 = *1=a+b+c+d*

*N* 2 *5 - 8a 4b+ 2c +d*

N=3 = 14 = 27a+9b+3c+d

we will have the solution that a = 1/3, b T 1/2, c = 1/6, d = 0. We can then easily show that the same equation applies to all N by induction.

**Clock pieces**

A clock (numbered 1 - 12 clockwise) fell off the wall and broke into three pieces. You find that the sums of the numbers on each piece are equal. What are the numbers on each piece? (No strange-shaped piece is allowed.)

x

*Si)iiizion:* Using the summation equation, 12 12 13 n *= —=* 78. So the numbers on each

2

piece must sum up to 26. Some interviewees mistakenly assume that the numbers on each piece have to be continuous because no strange-shaped piece is allowed. It's easy to see that 5. 6, 7 and 8 add up to 26. Then the interviewees' thinking gets stuck because they cannot find more consecutive numbers that add up to 26.

Such an assumption is not correct since 12 and 1 are continuous on a clock. Once that ‘‘rong assumption is removed, it becomes clear that 12 +1=13 and 11+ 2 =13. So the second piece is 11, 12. 1 and 2; the third piece is 3, 4, 9 and 10.

**Missing integers**

Suppose we have 98 distinct integers from 1 to 100. What is a good way to find out the two missing integers (within II 100])?

*Solution:* Denote the missing integers as x and y, and the existing ones are *z,,-* z9.8 • Applying the summation equations, we have

18 19

100 98 100x101

x-vy=

u., ,=,

21003 1002 100 "

100

2 2 2

*Ln = x + y +* 2

*„A 1=1*

=-> X + y + +

3 2 6

Using these two equations, we can easily solve x and y. If you implement this strategy using a computer program, it is apparent that the algorithm has a complexity of *0(n)* for two missing integers in 1 to *n.*

**Counterfeit coins I**

There are 10 bags with 100 identical coins in each bag. in all bags but one, each coin weighs 10 grams. However, all the coins in the counterfeit bag weigh either 9 or 11 grams. Can you find the counterfeit bag in only one weighing, using a digital scale that tells the exact weight?'

*Solution:* Yes, we can identify the counterfeit bag using one measurement. Take 1 coin out of the first bag, 2 out of the second bag, 3 out the third bag, •-•, and 10 coins out of io

the tenth bag. All together, there are En = 55 coins. if there were no counterfeit coins, they should weigh 550 grams. Let's assume the i-th bag is the counterfeit bag, there will be *i* counterfeit coins, so the final weight will be 550 ±i. Since *i* is distinct for each bag, we can identify the counterfeit coin bag as well as whether the counterfeit coins are lighter or heavier than the real coins using 550 ± *1.*

This is not the only answer: we can choose other numbers of coins from each bag as long as they are all different numbers.

**Glass balls**

You are holding two glass balls in a 100-story building. If a ball is thrown out of the
  
window, it will not break if the floor number is less than X, and it will always break if

9

Hint: In order to find the counterfeit coin bag in one weighing, the number of coins from each bag must be different If we use the same number of coins from two bags, symmetry will prevent you from

distinguish these two bags if one is the counterfeit coin bag.

the floor number is equal to or greater than X. You would like to determine X. What is the strategy that will minimize the number of drops for the worst case scenario? l°

Brain Teasers

A Practical Guide To Quantitative Finance Interviews

*Solution.* Suppose that we have a strategy with a maximum of *AT* throws. For the first throw of bail one, we can try the N-th floor. If the ball breaks, we can start to try the second ball from the first floor and increase the floor number by one until the second ball breaks. At most, there are *N* –1 floors to test. So a maximum of N throws are enough to cover all possibilities. If the first ball thrown out of N-th floor does not break, we have N –1 throws left. This time we can only increase the floor number by N –1 for the first ball since the second ball can only cover *N–* 2 floors if the first ball breaks. If the first ball thrown out of (2N-1)th floor does not break, we have *N* –2 throws left. So we can only increase the floor number by N – 2 for the first ball since the second ball can only cover N –3 floors if the first ball breaks...

Using such logic, we can see that the number of floors that these two balls can cover
  
with a maximum of N throws is A7+ (N – 1)+-- • +1 = N(N +1)/2 . In order to cover 100

stories, we need to have N(Ar +1)/2 ?...100, Taking the smallest integer, we have N =14.

Basically, we start the first ball on the 14th floor, if the ball breaks, we can use the
  
second ball to try floors 1,2,• • •,13 with a maximum throws of 14 (when the 13th or the

14th floor is X). If the first ball does riot break, we will try the first ball on the 14+(14 –1) = 27th floor. If it breaks, we can use the second ball to cover floors 15, 16, 26 with a total maximum throws of 14 as well...

***2.6 The Pigeon Hole Principle***

here is the basic version of the Pigeon Hole Principle: if you have fewer pigeon holes thanpigeons and you put every pigeon in a pigeon hole, then at least one pigeon hole has more than one pigeon. Basically it says that if you have *12* holes andmore than n +1 pigeons, at least 2 pigeons have to share one of the holes. The generalized version is that if you have *n* holes and at least *inn +1* pigeons, at least *m* +1 pigeons have to share one of the holes. These simple and intuitive ideas are surprisingly useful in many problems. Here vie will use some examples to show their applications.

flint: Assume we desie,n a strategy with A' maximum throws. If the first ball is thrown once, the second ball an cover ,V I floors; if the first ball is thrown twice, the second ball can cover ?V - 2 floors...

**Matching socks**

Your drawer contains 2 red socks, 20 yellow socks and 31 blue socks. Being a busy and absent-minded lVIIT student, you just randomly grab a number of socks out of the draw and try to find a matching pair. Assume each sock has equal probability of being selected, what is the minimum number of socks you need to grab in order to guarantee a

20

2t

pair of socks of the same color?

*Solution:* This question is just a variation of the even simpler version of two-color-socks problem, in which case you only, need 3. When you have 3 colors (3 pigeon holes), by the Pigeon Hole Principle, you will need to have 3 + 1 = 4 socks (4 pigeons) to guarantee that at least two socks have the same color (2 pigeons share a hole).

**Handshakes**

You are invited to a welcome party with 25 fellow team members. Each of the fellow members shakes hands with you to welcome you. Since a number of people in the room haven't met each other, there's a lot of random handshaking among others as well. If you don't know the total number of handshakes, can you say with certainty that there are at least two people present who shook hands with exactly the same number of people?

*Solution:* There are 26 people at the party and each shakes hands with from 1—since.
  
everyone shakes hands with you---to 25 people. In other words, there are 26 pigeons and
  
25 holes. As a result, at least two people must have shaken hands with exactly the same

number of people.

**Have we met before?**

Show me that, if there are 6 people at a party, then either at least 3 people met each other before the party, or at least 3 people were strangers before the party,

*Sohaion:* This question appears to be a complex one and interviewees often get puzzled
  
by what the interviewer exactly wants. But once you start to analyze possible scenarios,

the answer becomes obvious.

Let's say that you are the 6th person at the party. Then by generalized Pigeon Hole Principle (Do we even need that for such an intuitive conclusion?), among the remaining 5 people, we conclude that either at least 3 people met you or at least 3 people did not meet you. Now let's explore these two mutually exclusive and collectively exhaustive

scenarios:

Case 1: Suppose that at least 3 people have met you before.

**If two people in this group met each other, you and the pair (3 people) met each other. If no pair among these people met each other, then these people 3 people) did not meet each other. In either sub-case, the conclusion holds.**

**Brain Teasers** A Practical Guide To Quantitative Finance Interviews

**Case 2: Suppose at least 3 people have not met you before.**

**If two people in this group did not meet each other, you and the pair (3 people) did not meet each other. If all pairs among these people knew each other, then these people ( 3 people) met each other. Again, in either sub-case, the conclusion holds.**

**Ants on a square**

**There are 51 ants on a square with side length of 1. If you have a glass with a radius of 1/7, can you put your glass at a position on the square to guarantee that the glass encompasses at least 3 ants?"**

***Solution:* To guarantee that the glass encompasses at least 3 ants, we can separate the square into 25 smaller areas. Applying the generalized Pigeon Hole Principle, we can show that at least one of the areas must have at least 3 ants. So we only need to make sure that the glass is large enough to cover any of the 25 smaller areas. Simply separate the area into 5** x5 **smaller squares with side length of 115 each will do since a circle with radius of 117 can cover a squareI2 with side length 1/5.**

**Counterfeit coins H**

**There are 5 bags with 100 coins in each bag. A coin can weigh 9 grams, 10 grams or 11 grams. Each bag contains coins of equal weight, but we do not know what type of coins a bag contains. You have a digital scale (the kind that tells the exact weight). How many times do you need to use the scale to determine which type of coin each bag contains?**

***Solution:* If the answer for 5 bags is not obvious, let's start with the simplest version of the problem-1 bag. We only need to take one coin to weigh it. Now we can move on to 2 bags. How many coins do we need to take from bag 2 in order to determine the coin types of bag 1 and bag 2? Considering that there are three possible types for bag 1, we will need three coins from bag 2; two coins won't do. For notation simplicity, let s change the number/weight for three types to -1, 0 and I (by removing the mean 10). If**

**" Hint: Separate the square into** 25 **smaller areas; then at least one area has 3 ants in it.**

**12 A circle with radius *r can* cover a square with side length up to** *15\_* **r and 1.414.**

**13 Hint: Start with a simpler problem. What if you have two bags of coins instead of 5, how many**

**you need from each bag to find the type of coins in either bag? What is the minimum di**ff**erence in coin
  
numbers? Then how about three bags? coins do**

**22**

**we only use 2 coins from bag 2, the final sum for 1 coin from bag 1 and 2 coins from bag 2 ranges from -3 to 3 (7 pigeon holes). At the same time we have 9 (3 x 3 ) possible combinations for the weights of coins in bag 1 and bag 2 (9 pigeons). So at least two combinations will yield the same final sum (9>7, so at least two pigeons need to share one hole), and** we **can not distinguish them. If we use 3 coins from bag 2, then the sum ranges from -4 to 4, which is possible to cover all 9 combinations. The following table exactly shows that all possible combinations yield different** sums:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sum** | **I coin, bag 1** | | |  |
| **3 Coins, Bag 2** | **C2**  **N** | **-1** |  | **1** |
| **-1**  0  **1** | **4** -1  **2** | -3  0  **3** | -2  1  **4** |

***C I and C2 represent the weights of coins from bag 1 and respecrrvety.***

**Then how about 3 bags? We are going to have 33 = 27 possible combinations. Surely an indicator ranging from –13 to 13 will cover it and we will need 9 coins from bag 3. The possible combinations are shown in the following table:**

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **ES IA Ill** | **C2 = -1** | | |  |  | **0241** |  |  | **C2=1** |  |
| 9 **Coins. Rag 3** | **C1**  **C)** | **-1** | **0** | **1** | **-1** | **0** | **1** | **-1**  -7  **2**  **11** | **0**  -(  3  **12** | **1**  -5 4 **11** |
| •  -I  0  I | **-13**  -4  5 | **-12  -3**  **6** | **-11**  **-2**  **7** | **-10  -1**  **8** | -9  **0**  **9** | **...-----.**  -8  **1**  **10** |
|  |  |  |

***C I , C1, and C3 represent the weights of coins from ag i, ana rep c si v iy.***

**Following this logic, it is easy to see that we will need 27 coins from bag 4** and 81 **coins from bag 5. So the answer is to take 1, 3, 9, 27 and 81 coins from bags 1, 2, 3, 4, and 5, respectively, to determine which type of coins each bag contains using a single weighing.**

***2.7 Modular Arithmetic***

**The modulo operation—denoted as** x%y **or** x mod y---finds the remainder of division **of
  
number x by another number y. For simpicility, we only consider the case where y is a
  
positive integer. For example, 5%3 = 2. An intuitive property of modulo operation is**

**23**

**that if x, %y =** x, %y, **then (x, x2)%y = 0. From this property we can also show that x%y, (x +1)%y, • • • , and (x + y —1)%y are all different numbers.**

Brain Teasers **A Practical Guide To Quantitative Finance Interviews**

**Prisoner problem**

**One hundred prisoners are given the chance to be set free tomorrow. They are all told that each will be given a red or blue hat to wear. Each prisoner can see everyone else's hat but not his own. The hat colors are assigned randomly acid once the hats are placed on top of each prisoner's head they cannot communicate with one another in any form, or else they are immediately executed. The prisoners will be called out in random order and the prisoner called out will guess the color of his hat. Each prisoner declares the color of his hat so that everyone else can hear it. If a prisoner guesses correctly the color of his hat, he is set free immediately; otherwise he is executed.**

**They are given the night to come up with a strategy among themselves to save as many prisoners as possible. What is the best strategy they can adopt and how many prisoners can they guarantee to save?I4**

***Solution:* At least 99 prisoners can be saved.**

**The key lies in the first prisoner who can see everyone else's hat. He declares his hat to be red if the number of red hats he sees is odd. Otherwise he declares his hat to be blue. He will have a 1/2 chance of having guessed correctly. Everyone else is able to deduce his own hat color combining the knowledge whether the number of red hats is odd among 99 prisoners (excluding the first) and the color of the other 98 prisoners (excluding the first and himself). For example, if the number of red hats is odd among the other 99 prisoners. A prisoner wearing a red hat will see even number of red hats in**

**the other 98 prisoners (excluding the first and himself) and deduce that he is wearing a red hat.**

**The two-color case is easy, isn't it? What if there are 3 possible hat colors: red, blue, and**

**white? What is the best strategy they can adopt and how many prisoners can they guarantee to save?15**

***Solution:* The answer is still that at least 99 prisoners will be saved. The difference is
  
that the first prisoner now only has 1/3 chance of survival. Let's use the following
  
scoring system: red---0, green--1, and blue=2. The first prisoner counts the total score for**

**the rest of 99 prisoners and calculates s%3. If the remainder is 0, he announces red; if the remainder is 1, green; 2, blue. He has 1/3 chance of living, but all the rest of the prisoners can determine his own score (color) from the remainder. Let's consider a prisoner i among 99 prisoners (excluding the first prisoner). He can calculate the total**

**score (x) of all other 98 prisoners. Since (x + O)%3, (x +1)%3, and (x + 2)%3 are all**

**different, so from the remainder that the first prisoner gives (for the 99 prisoners including *i),* he can determine his own score (color). For example, if prisoner *i* sees that there are 32 red, 29 green and 37 blue in those 98 prisoners (excluding the first and himself). The total score of those 98 prisoners is 103. If the first prisoner announces that the remainder is 2 (green), then prisoner *i* knows his own color is green (1) since only 104%3 = 2 among 103, 104 and 105.**

**Theoretically, a similar strategy can be extended to any number of colors. Surely that requires all prisoners to have exceptional memory and calculation capability.**

**Division by 9**

**Given an arbitrary integer,** come up with **a rule to decide whether it** is divisible by 9 **and prove it.**

***Solution:* Hopefully you still remember the rules from your high school math class. Add up all the digits of the integer. If the sum is divisible by 9, then the integer is divisible by 9; otherwise the integer is not divisible by 9. But how do we prove it?**

**Let's express the original integer *as a -= a„10" +* + • + *a,10'* + *ao.* Basically we**

**state that if *an* + + • • • + *a, + at, =*** 9x ( x is **a integer), then the *a* is divisible by 9 as**

**well. The proof is straightforward:**

**For** any *a* = **10n +a„-110"-' + • • • + *a11*** +a0, let *b* = ***— (an + an\_., + - ••+ +*a0). We have *b = a n(10" —1) + a n A(10" — 1) + • • • + a1(10' —1) = a —*** 9 x , which **is divisible by 9 since all (10k *—1), k .1,• •* • ,n are divisible by 9. Because both *b* and 9x are divisible by 9 *a = b +* 9x must be divisible by 9 as well.**

**(Similarly you can also show that *a = (-1)"a„ + (—Di" an\_, + -+ (-1)' at +a***0 ***=11x* is the necessary and sufficient condition for *a* to be divisible by 11.)**

|  |  |
| --- | --- |
| **14 Hint: The first prisoner can see the number of red odd number of counts and the other has even number**  **15 Hint: That a number is odd simply means x%2 = I x6/03 instead.** | **and blue hats of all other 99 prisoners. One color has of counts.**  **. Here we have 3 colors, so you may want to consider** |

**24 25**

**Chameleon colors**

Brain Teasers A Practical Guide To Quantitative Finance Interviews

**A remote island has three types of chameleons with the following population: 13 red chameleons, 15 green chameleons and 17 blue chameleons. Each time two chameleons with different colors meet, they would change their color to the third color. For example, if a green chameleon meets a red chameleon, they both change their color to blue. Is it ever possible for all chameleons to become the same color? Why or why not?''**

*Solution: It* **is not possible for all chameleons to become the same color,. There are several approaches to proving this conclusion. Here we discuss two of them.**

**Approach 1. Since the numbers 13, 15 and 17 are "large" numbers, we can simplify the
  
problem to 0, 2 and 4 for three colors, (To see this, you need to realize that if
  
combination (m +1, *n* +1,** *p +1)* can be **converted to the same color, combination**

***(m,n,p)* can be converted to the same color as well.) Can a combination (0,2,4) be converted to a combination (0,0,6)7 The answer is NO, as shown in Figure 2.3:**

|  |  |
| --- | --- |
| **(0, 2, 4)4** |  |

(0, 1, 5)

Figure 2.3 chameleon color combination transitions from (0, 2, **4)**

**Actually combination (1,2,3) is equivalent to combination (0,1,2), which can only be converted to another (0,1, 2) but will never reach (0, 0,3).**

**Approach 2. A different, and more fundamental approach, is to realize that in order** for **all the chameleons to become the same color, at certain intermediate stage, two colors must have the same number. To see this just imagine the stage before a final stage. It must has the combination (1,1,x). For chameleons of two different colors to have the**

**same number, their module of 3 must be the same as well. We start with 15 3x,**

**13 = 3y +1,** and **17 = 3z + 2 chameleon,** when **two chameleons of different colors meet, we will have three possible scenarios:**

(3x + **2,3y,3z +1) = (3x',3y1+1,3e+ 2), one y meets one *z***

(3x, 3y **+1,3z+ 2) (3(x-1)+ 2,3(y +1),3z +1).(3x',3y1+1,3z\*+ 2), onexmeetsonez
  
(3(x —1) + 2,3y,3(z + 1)+1) = (3x',3yi+1,3z '4- 2), one x meetsone y**

**So the pattern is preserved and we will never get two colors to have the same module of 3. In other words, we cannot make two colors have the same number. As a result, the chameleons cannot become the same color. Essentially, the relative change of any pair of colors after two chameleons meet is either 0 or 3. In order for all the chameleons to become one color, at least one pair's difference must be a multiple of 3.**

***2.8 Math Induction***

**Induction is one of the most powerful and commonly-used proof techniques in mathematics, especially discrete mathematics. Many problems that involve integers can be solved using induction. The general steps for proof by induction are the following:**

* **State that the proof** uses induction and **define an appropriate predicate** *P(n)*
* **Prove the base case P(1) , or any other smallest number** *n* **for the predicate to be true.**
* **Prove that** *P(n)* implies *P(n +1)* for every integer n. Alternatively, in a strong induction argument, you prove that P(1), P(2), • • • , and *P(n)* **together** imply *P(n +1) .*

**In most cases, the real difficulty lies not in the induction step, but to formulate the
  
problem as an induction problem and come up with the appropriate predicate** *P(n).* **The**

**simplified version of the problem can often help you identify** *P(n).* **Coin split problem**

**You** split **1000** coins into two piles and count **the number of coins in each pile. If there
  
are x coins in** pile one and **y coins in pile two, you multiple x by y to get xy. Then you**

**split both piles further, repeat the same counting and multiplication process, and add the new multiplication results to the original. For example, you split x to xi and x2, y to** yi

**and y2, then the sum is .xy + xix2 yiy2• The same process is repeated until you only**

**have piles of 1 stone each. What is the final sum? (The final 1's are not included in the sum.) Prove that you always get the same answer no matter how the piles are divided.**

**16** Hint: consider the numbers in module of 3.

26 27

***Solution:* Let *n* be the number of the coins and *f (n)* be the final sum. It is unlikely that a solution will jump to our mind since the number *n =* 1000 is a large number. If you aren't sure how to approach the problem, it never hurts to begin with the simplest cases and try to find a pattern. For this problem, the base case has *n =* 2. Clearly the only split is 1+1 and the final sum is 1. When *n =* 3, the first split is 2+1 and we have xy = 2 and the 2-coin pile will further give an extra multiplication result 1, so the final sum is 3. This analysis also gives the hint that when *n* coins are split into x and n — x coins, the total sum will be *f (n) = x(n*** *x) +* ***f (x) + f (n x).* 4 coins can be split into 2 +2 or 3+1. For either case we can apply x(n — x) + *f (x) + f (n*** *x)* **and yields the same final sum 6.**

**Brain Teasers** A Practical Guide To Quantitative Finance Interviews

**Claim: For n coins, independent of intermediate splits, the final sum is *n(n —1)* 17**

**2**

**So how do we prove it? The answer should be clear to you: have proved the claim for the base cases *n* --- 2,3,4. Assum *n =* 2,• • -, *N —1* coins, we need to prove that it holds for *n = N* apply the equation *f((n) = x(n x) + f (x)+ f (n x).* If *N* coins *N x* coins, we have**

**by strong induction. We**

**e the claim is true for**

**coins as well. Again we
  
are split into x coins and**

***f (N) = x(N x) + f (x)+ f (N -x)***

***= x(N x)+N (N -) +(N x)(N - x -1)*** = ***N (N*** *-1)*

**2 2 2**

**So indeed it holds for *n = N* as well and*\_ (n)* 2 is true for any *17* 2 , Applying**

**the conclusion to *n =1000,* we have *f (n) =1000* x 999 / 2 .**

**Chocolate bar problem**

**A chocolate bar has 6 rows and 8 columns (48 small lx 1 squares). You break it into**

**individual squares by making a number of breaks. Each time, bk one rectangle two smaller rectangles. For example, in the first step yo** an **the 6 x 8 chocolate**

**u c b rea angle into**

**bar into a 6 x 3 one and a 6x 5 one. What is the total number of breaks needed in order to break the chocolate bar into 48 small squares?**

**17 *1(2) =*** 1 , ***f(3} f*** (2) --, **2 and *f (4) — f***(3) = 3 shouldgive you enough hint to **realize the pattern is *f (a) = I +*** 2 + •• • + ***(n — rs( `I}***

**2**

**28**

***Solution:* Let *m* be the number of the rows of the chocolate bar and *n* be the number of columns. Since there is nothing special for the case m = 6 and *n =* 8, we should find a general solution for all *m* and n. Let's begin with the base case where *m* =1 and *n =* I . The number of breaks needed is clearly 0. For *m >1* and *n =1,* the number of breaks is *m* —1; similarly for *m =1* and *n >* 1, the number of breaks is *n* —1. So for any *In* and *n,* if we break the chocolate into m rows first, which takes m —1 breaks, and then break each row into *n* small pieces, which takes *m(n —1)* breaks, the total number of breaks is *(m —1) + m(n —1) = mn —1.* If we breaks it into *n* columns first and then break each column into *m* small pieces, the total number of breaks is also *mn* —1. But is the total number of breaks always *mn* —1 for other sequences of breaks? Of course it is. We can prove it using strong induction.**

**We have shown the number of breaks is *mn* —1 for base cases *n=1* and**

***in =1, n 1.* To prove it for a general *m* x *n* case, let's assume the statement is true for cases where *rows < m, columns 5\_* n and *rows m, columns < n.* If the first break is along a row and it is broken into two smaller pieces *m* x *n,* and *m* x *(n n,),* then the**

**total number of breaks is 1+ *(in x n***, ***—1) + (m x (n ni) —1) = mn —1.* Here we use the
  
results for *rows < In, columns < n.* Similarly, if it is broken into two pieces m, x *n* and
  
*— mi)x n,* the total number of breaks is 1+ (m, x *n —1) + ((m — mi)x n —1) = mn —1.* So**

**the total number of breaks is always *mn* —1 in order to break the chocolate bar into *m x n* small pieces. For the case m = 6 and *n =* 8, the number of breaks is 47.**

**Although induction is the standard approach used to solve this problem, there is actually a simpler solution if you've noticed an important fact: the number of pieces always increases by 1 with each break since it always breaks one piece into two. In the beginning, we have a single piece. In the end, we will have *mn* pieces. So the number of breaks must be *Inn —1.***

**Race track**

**Suppose that you are on a one-way circular race track. There are *N* gas cans randomly placed on different locations of the track and the total sum of the gas in these cans is enough for your car to run exactly one circle. Assume that your car has no gas in the gas tank initially, but you can put your car at any location on the track and you can pick up the gas cans along the way to fill in your gas tank. Can you always choose a starting position on the track so that your car can complete the entire circle?"**

**18 Hint: Start** with N = 1, 2 and solve the problem using induction.

29

Brain Teasers

A Practical Guide To Quantitative Finance interviews

***Solution: If* you get stuck as to how to solve the problem, again start with the simplest
  
cases *(N =*1, 2) and consider using an induction approach. Without loss of generality,**

**let's assume that the circle has circumference of I. For *N =* 1, the problem is trivial. Just
  
start at where the gas can is. For *N =* 2, The problem is still simple. Let's use a figure to**

**visualize the approach. As shown in Figure 2.4A, the amount of gas in can 1 and can 2,
  
expressed as the distance the car can travel, are x, and x, respectively, so x**, **+ x2 =1.**

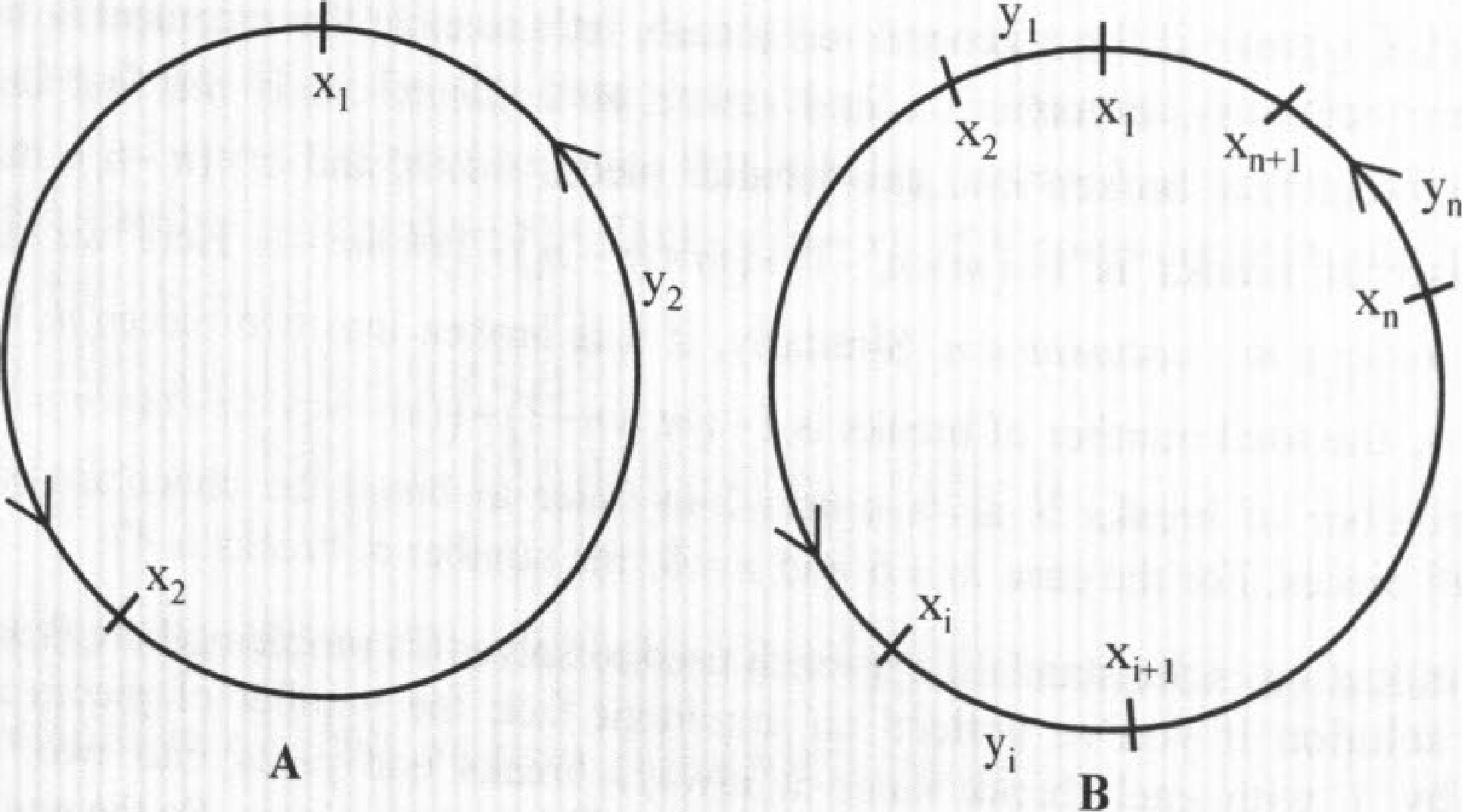
**The corresponding segments are y, and y„ so y, + y2 = 1 . Since x, + x2 =1 and
  
+ *y2* .1, we must have x, yi or x, y, ( < y, and x, < y2 cannot both be true). If
  
x1 y, , we can start at gas can 1, which has enough gas to reach gas can 2, and get more**

**gas from gas can 2 to finish the whole circle. Otherwise, we will just start at gas can 2 and pick up gas can I along the way to finish the whole circle.**

**which the statement holds. So the statement also holds for *N = n +1.* Hence we can always choose a starting position on the track to complete the entire circle for any *N.***

**There is also an alternative approach to this problem that provides a solution to the starting point. Let's imagine that you have another car with enough gas to finish the circle. You put that car at the position of a randomly chosen gas can and drive the car for a full circle. Whenever you reach a gas can (including at the initial position), you measure the amount of gas in your gas tank before you add the gas from the can to your gas tank. After you finish the circle, read through your measurement records and find the lowest measurement. The gas can position corresponding to the lowest measurement should be your starting position if the car has no gas initially. (It may take some thinking to fully understand this argument. I'd recommend that you again draw a figure and give this argument some careful thoughts if you don't find the reasoning obvious.)**

**Figure 2.4 Gas can locations on the cycle and segments betweengas cans The argument for** *N* **2** also **gives us the hint for the inductio**



**A**

B

**show that if the statement holds for *N* n step. Now we want to**

**= */I,* then the same statement also holds for**

***N= n* + I. As shown in Figure 2.4B, we have x, +x2 + • • • + *--* and**

***y1+ y2 + •..+* =1 for *N* -=n+1. So there must exist at least one *1,1-1.n+1,* that**

**has x, y,. That means whenever the car reaches x, it can reach xj+1 with more gas
  
(For *i n +1 ,* it goes to *=*1 instead). In other words, we can actually "combine" *x,* and**

**x,,, to one gas can at the position of x with an amount of gas x + x,+1 (and eliminate**

**,**

**the gas can** *1+1).* **But such combination reduces the *N n +1* problem to** *N* ***n,*** *for* **30**

***2.9 Proof by Contradiction***

**In** a **proof by contradiction or indirect proof, you show that if a proposition were false, then some logical contradiction or absurdity would follow. Thus, the proposition must be true.**

**Irrational number**

**Can you prove that is an irrational number? A rational number is a number that can**

**be expressed as a ratio of two integers; otherwise it is irrational.**

*Solution:* **This is a classical example of proof by contradiction. If Nh is not an irrational number, it can be expressed as a ratio of two integers *HI* and *n.* If *m* and *n* have any common factor, we can remove it by dividing both *m* and *n* by the common factor. So in the end, we will have a pair of m and *n* that have no common factors. (It is called**

**irreducible fraction.) Since *m 1 n= -5,*** we **have *m2 = 2n2 .* So *m2* must be an even
  
number and** *in* **must be an even number as well. Let's express *in* as 2x, where x is an**

**integer, since *in* is even. Then m2 = 4x2 and we also have *n2 =*** 2x2 , **which means *n*must be even as well. But that both *m* and** *n* **are *even* contradicts the earlier statement**

**that *m* and *n* have no common factors. So NE must be an irrational number. Rainbow hats**

**Seven prisoners are given the chance to be set free tomorrow. An executioner will put a
  
hat on *each* prisoner's head. Each hat can be one of the seven colors of the rainbow and
  
the hat colors are assigned completely at the executioner's discretion. Every prisoner can**

**31**

Brain Teasers

**see the hat colors of the other six prisoners, but not his own. They cannot communicate with others in any form, or else they are immediately executed. Then each prisoner writes down his guess of his own hat color. If at least one prisoner correctly guesses the color of his hat, they all will be set free immediately; otherwise they will be executed.**

**They are given the night to come up with a strategy. Is there a strategy that they can guarantee that they will be set free?I9**

*Solution:* **This problem is often perceived to be more difficult than the prisoner problem in the modular arithmetic section. In the previous prisoner problem, the prisoners can hear others' guesses. So one prisoner's declaration gives all the necessary information other prisoners need. In this problem, prisoners won't know what others' guesses are. To solve the problem, it does require an aha moment. The key to the aha moment is given**

7

**by the hint. Once you realize that if we code the colors to 0-6, Ex, %7 must be**

**(**

**•=1**

**among 0, 1, 2, 3, 4, 5 or 6 as well. Then each prisoner i—let's label them as 0-6 as
  
well—should give a guess** *g,* **so that the sum of** *g,* **and the rest of 6 prisoners' hat color
  
codes will give a remainder of *i* when divided by 7, where** *g, is* **a unique number**

4

**between 0 and 6. For example, prisoner 0's guess should make** *go +***Ex, 047 - 0.**

**(**

kit) **,**

**This way, we can guarantee at least one of** *g, =* **x, for i = 0,1,2,3,4,5, 6 .**

(

**We can easily prove this conclusion by contradiction. If** *g, # x „* 7 **then x, %7\*i**

**,.,,**

**(**

**(since** *g, +Ex k %7 # i* **and** *g,* **and x, are both between 0 and 6). But if***g, .x,* **for all**

*k\*?*

*7*

*i =* **0,1,2,3,4,5, and 6, then (Ex,)%7 # 0,1,2,3,4,5,6, which is clearly impossible. So i.1**

**at least one of** *g,* **must equal to x, . As a result, using this strategy, they are guaranteed to be set free.**

19 Hint: Let's assign the 7 colors of rainbow with code 0-6 and *x ,* be the color code of prisoner *I.* Then **E**x, V07 must be 0, 1, 2, 3, 4, 5 or 6. How many guesses can 7 prisoners make?

**Chapter 3 Calculus and Linear Algebra**

**Calculus and linear algebra lay the foundation for many advanced math topics used in quantitative finance. So be prepared to answer some calculus or linear algebra problems—many of them may be incorporated into more complex problems—in quantitative interviews. Since most of the tested calculus and linear algebra knowledge is easy to grasp, the marginal benefit far outweighs the time you spend brushing up your knowledge on key subjects. If your memory of calculus or linear algebra is a little rusty, spend some time reviewing your college textbooks!**

32

**Needless to say, it is extremely difficult to condense any calculus/linear algebra books into one chapter. Neither is it my intention to do so. This chapter focuses only on some of the core concepts of calculus/linear algebra that are frequently occurring in quantitative interviews. And unless necessary, it does so without covering the proof, details or even caveats of these concepts. If you are not familiar with any of the concepts, please refer to your favorite calculus/linear algebra books for details.**

***3.1 Limits and Derivatives***

**Basics of derivatives**

**Let's begin with some basic definitions and equations used in limits and derivatives. Although the notations may be different, you can find these materials in any calculus textbook.**

*(v f x +* **Ax) —** *f(x)*

**Derivative: Let y =** *(x) ,* **then** *f 1(x) =* **-2-111 = lim = itrn**

**dx** Ax—ro Ax-40 **Ax**

**The product rule: If** *u = u(x)* **and v = v(x) and their respective derivatives exist,**

*cl(uv) dv du*

*=u—+v—, (uv)"=u'v+uv'*

*dx cA: dx*

*dv)I , (u), u'v —*v*, — =-*

|  |  |
| --- | --- |
| **The quotient rule:** | *v --u*  *d u) ( du dx v dx* |

*cx v v2*

*dy dy du*

**The chain rule: If y =** *f(u(x))* **and *ti = u(x) ,* then — =** *—*

*cbc du dx*

*dy"* dy

**The generalized power rule: — =** *fly —* **for** *V n #* **0**

**dx dr**

**Some useful equations:**

*ln(ab) =* Ina *+* In *b e' = + \*)"*

Calculus and Linear Algebra A Practical Guide To Quantitative Finance interviews

*ar* = *in.*

ii• slMinx —

x-o)

*lim(1 x)\* =1 + Icc* for any ***k***

lim(In x *1 xr)* = 0 for any *r > 0* lint*e-x =* 0 for any *r*

1C-M

x-3x

*d*

|  |  |  |
| --- | --- | --- |
| *d „ du*  *e e"*  *dx dx* | *—da" = (au* In *a) du dx dr* | *d ,* i ***1 du u'***  —n *u* = =  dx *u dx* ***u*** |

*sin x = cos x—dcos x =* —sin x c—L tan x sec' x

*dx dx dx*

What is the derivative of y = ?1

*Solution:* This is a good problem to test your knowledge of basic derivative formulas—specifically, the chain rule and the product rule,

Let *u =* In y =1n (In xi") =In x x In(lnx). Applying the chain rule and the product rule, we have

*du* d(ln y) \_ I *dy \_ d On x) . . . d* (1n(ln *x))* InOn x) + In x 5

—d —x . dr — —y dr — *dx* x in(in x) + in *x* x

*dx* x xlnx

*d* (1n(ln x))

To derive -----, we again use the chain rule by setting v =1n x : *dx*

d(ln(In x)) = d(ln v) dv \_ I I **I**

dr *dv dx v x* x In x•

*.* 1 *dy* ln(In *x)* In x *dy y* In *x''*

ydr

*.. — = — —* x xlnx *dx - + — =* In(In x) +1) = Ort(Inx)-1-1).

*x x*

**Maximum and minimum**

Derivative *f 1(x)* is essentially the slope of the tangen

the instantaneous rate of change (veloci t line to the curve y = *(.0* and

ty) of y with respect to *x.* At point x = c, If

Hint: To calculate the derivative of functions with the format *Y, it is* common to take natural

logs on both sides and then take the derivative, since d(ln y)/ *dx* y *f(x11 yxdyldx.*

34 *f '(c) >* 0, *f (x)* is an increasing function at *c;* if *f '(c) <* 0, *f"(x)* is a decreasing function at *c.*

**Local maximum or minimum:** suppose that *f (x)* is differentiable at *c* and is defined on an open interval containing *c.* If *f (c)* is either a local maximum value or a local minimum value of *f* (x), then *f '(c) =* 0

**Second Derivative test:** Suppose the secondary derivative of *f (x), f "(x), is* continuous near *c.* If *f i(c) —* 0 and *f "(c) >* 0, then *f(x)* has a local minimum at *c;* if *f '(c) =* 0 and *f "(c) <* 0 , then *f (x)* has a local maximum at *c*.

Without calculating the numerical results, can you tell me which number is larger, *e* or

*?*2

*Solution:* Let's take natural logs of *e'* and *fie.* On the left side we have *nine,* on the

Ine ln

right side we have *e* In *r.* If *e"> >* ge<=>/rxirl e>exIng<4.— > —.

e *z*

In x .

Is it true? That depends on whether *f (x) =* is an increasing or decreasing function

*1 lxxx—* ln x 1— ln x

from *e* to 7r. Taking the derivative of *1(4* we have *f 1(x) =* =

X2 X2 ' which is less than 0 when x > *e* (In x > 1). In fact, *f* (x) has global maximum when x *e* for all x > 0. So ln *e* > In R-

and *eir > ze*

e IT

Alternative approach: If you are familiar with the Taylor's **series, which we** will discuss

1

x x2 X3

in Section 3.4, you can apply Taylor's series to *ex : =E— =1 + + — + — + •* **So**

*n=0 n!* l! 2. **3!**

*ex* > 1+ x, Vx > 0 . Let x= *I e —1,* then *el" e > I e <=> ele >z <7> eff > rre* **L'Hospital's rule**

Suppose that functions *f (x)* and *g(x)* are differentiable at x *a* and that lirn *g '(a) #* 0.

*x—ra*

Further suppose that lien *f (a) =* 0 and lien *g(a) =* 0 or that litn *f (a) -+* -±co and

*x—ra \*a x—ra*

2 Hint: Again consider taking natural logs on both sides: In *a >* In *b a > h* since In x is a

monotonously increasing function.

35

**Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews**

|  |  |  |
| --- | --- | --- |
| *lim g (a) ±00,* | **'**  **then lim =lirn  f (x). L'Hospital's rule converts the limit from**  ***g(x) g ',(x)*** | *dF (x) =f (x), F (a) - y a F(x) .ya+ ff (t)dt cbc* |

**an indeterminate form to a determinate form.**

**What is the limit of e /x2 as x -› co and what is the limit of x2 In x as x 0+ ?**

*Solution:l .* **is a typical example of L'Hospital's**

**rule since lim e 00 and**

\_r Sao

**x**

**lim X2 = 00. Applying L'Hospital's rule, we have**

**lim = lim --= lim r(x) = lim —ex .**

**"a** *g(x) x g'(x) x->.c.* **2x**

**The result still has the property that lim** *f (x) = lime = Go* **and lim g(x) lim 2x =Go, so we can apply the L' Hospital's rule again: N->z Jr--)acp**

*x-m g(x) ,--.-', x' i--.(i g'(x)* **-,--oi, 2x x->-** *d (2x) I dx -,--* **2**

*um =.- = um --., = am = lim — = um- — = am— = ac .*

*f(x) „ ex i: f(x) ex* **12** *d(ex )1 dx ,. e*

**At first look, L'Hospital's rule does not appear to be applicable to lim x2 In x since it's not in the format of limiLc). However, we can rewrite the original limit as lim**

**in x**

***g(x)***

**and it becomes obvious that Jim cc and Urn In x = -co. So we can now apply**

**L'Hospital's rule: -°.**

**lim x2 Inx = lim lnx lirn d(lnx)/dx 1/x , x2**

**--------**

**lira iirn = 0**

**x-49' x`,01.** *d(x-2) cbc - .0.* **-2/ x3 x-'1:14 -2**

**Basics of integration 3.2 *Integration***

**antiderivative of** *1(x)*

**Again, let's begin with some basic definitions and equations used in integration.**

**If we can find a function** *F(x)* **with derivative *f(x),* then we call *F(x)*** an

*If 1(x) = P(x) f(x) =f FP(x)dx =[F(x F(b)- F (a)*

**36**

**k+1**

**The generalized power rule in reverse: *Sukdu****+ c (k* **1), where** *c* **is any**

*k +* **1**

**constant.**

**Integration by substitution:**

**if** *(g (x)) • g'(x)dx = f f(u)du* **with** *u = g(x), du = gi(x)dx* **Substitution in definite integrals:** *(g(x)) • gi(x)dx = (h1* ***(Q)*** *f (u)du*

***g***

**Integration by parts: *Judy = uv - ivdu***

1. **What is the integral of ln(x)?**

*Solution:* **This is an example of integration by parts. Let** *u =* **In x and v = x , we have** *d(uv) = vdu udv = (x* **X 11 *x)dx* + In***xdx*,

**fin *xcix* = xln x - fdr = xlnx x + c, where *c* is any constant.**

1. **What is the integral of sec(x) from x = 0 to x = / 6?**

*Solution:* **Clearly this problem is directly related to differentiation/integration of trigonometric functions. Although there are derivative functions for all basic**

*d*

**trigonometric functions, we only need to remember two of them: dx sinx - cos x,** *dcos x =* **-sin *x.* The rest can be derived using the product rule or the quotient rule. For** *dx*

**example,**

*d* **sec x d(1 / cos x)sin x**

*dr dx cosy x*

**= = sec x tan *x,***

*d* **tan x d(sin *x* / cos x)cos' x + x**

**- Secs X.**

**cost x**

*dx dx*

**d(sec x + tan x) - sec x(sec x + tan x) .** *• •*

**dx**

**37**

**Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews**

38 **39**

|  |  |
| --- | --- |
| **Since the (sec x + tan x) term occurs in the derivative, we also have *d* In I sec x + tan x I sec x(sec x + tan x) — sec x**  ***dx* (sec x + tan x)**  **Jsec x = In sec x + tan x I *+c***  **and**  *ri* **6**  **sec x = ln(sec(ar / 6) + tan(/r / 6)) — In(sec(0) + tan(0))** |  |

**Applications of integration**

***A.* Suppose that two cylinders each with radius 1 intersect at right angles and their centers also intersect. What is the volume of the intersection?**

**Solution: This problem is an application of integration to volume calculation. For these
  
applied problems, the most difficult part is to correctly formulate the integration. The**

**general integration function to calculate 3D volume is *V =*** *f* ***A(z)dz* where *A(z)* is the**

**cross-sectional area of the solid cut by a plane perpendicular to the z-axis at coordinate Z. The key here is to find the right expression for cross-sectional area *A* as a function of *z.***

**Figure 3.1 gives us a clue. If you cut the intersection by a horizontal plane, the cut will**

**be a square with side-length \f(202 --(24 . Taking advantage of symmetry ,** we can
  
**calculate the total volume as**

**2x 1[00' *—(2z)*2*clz .8x[r2 z z3*310 = 16/3r3**

**An alternative approach requires even better 3D imagination. Let's imagine a sphere that is inscribed inside both cylinders, so it is inscribed inside the intersection as well. The sphere should have a radius of r /2. At each cut perpendicular to the z-axis, the circle from the sphere is inscribed in the square from** the intersection as **well. So**

***=* Since it's true for all *z* values, we have**

**-0**

**73 —41/**

**s„-0,,n.. *0* 4*— intetsection VIllterSeclic►* =16/3r' =16/3.**

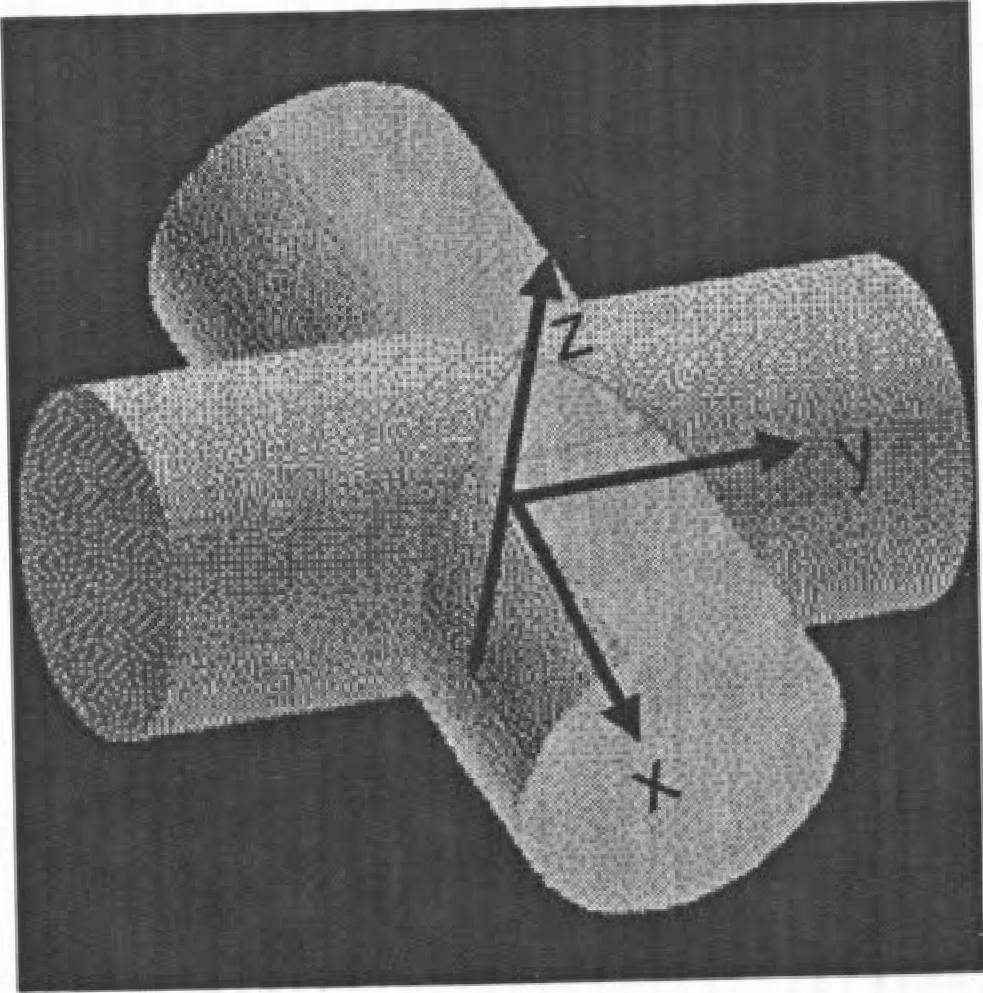


Figure 3.1 Interaction of two cylinders

***B.* The snow began to fall some time before noon at a constant rate. The city of Cambridge sent out a snow plow at noon to clear Massachusetts Avenue from MIT to Harvard. The plow removed snow at a constant volume per minute. At 1 pm, it had moved 2 miles and at 2 pm, 3 miles. When did the snow begin to fall?**

***Solution:* Let's denote noon as time 0 and assume snow began to fall *T* hours before noon. The speed at which the plow moves is inversely related to the vertical cross-sectional area of the snow: v = *c, / A(t),* where v is the speed of the plow,** *c,* **is a constant representing the volume of snow that the plow can remove every** hour **and *A(t)* is the cross-sectional area of the snow. If *t* is defined as the time after noon, we also have *AO) = c2 (t + T) ,* where c2 is the rate of cross-sectional area increase per hour (since the**

**ci** *c*

snow falls at a constant rate). **So v =— *c2(t +T)* —** where ***c =* Taking the**

***I +T c*2**

integration, **we have**

**From these two equations,** we get

--cln(14-2)=2,

***=c1n12+71=3***

***T )***

***c dt = cln(l+T)— clnT***

***T + t***

***1.2T +t = c*** ln(2+T)—cln ***T***

*(i+T), (2\_12* ***T T***

Overall, this question, although fairly straightforward, tests analytical skills, integration knowledge and algebra knowledge.

**Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews**

**Expected value using integration**

Integration is used extensively to calculate the unconditional or conditional expected value of continuous random variables. In Chapter 4, we will demonstrate its value in probability and statistics. Here we just use one example to show its application:

If *X* is a standard normal random variable, *X - N(0,* **1),** what is *E[X X* > IA?

*Solution:* Since *X* - N(0, 1), the probability density function of x is *f (x) \_r*and we have *E[X I X >* 0] = *xf (x)dx = f*x*-217, dx*

Because *d(--1 f* 2x2)= -x and *je dy= eu c,* where *c* is an arbitrary constant, it is obvious that we can use **integration by substitution** by letting *u = -11* 2x2 . Replace *e-112''* with *eu* and *xdx* with *-du,* we have

rx *2'2 - e" du - eu — (0 l)*

where *Le] 0* is

.42z L 0

determined by x = Ou = 0 and x = co *u* = .

*E[X X* > 0] = *11-sETT*

**Changing Cartesian integrals into polar integrals:** The variables in two-dimension plane can be mapped into polar coordinates: x = r cos *0,* y = *r* sin *O.* Tthe integration in a continuous polar region *R* is converted to

*llf (x, y)dxdy = if .f(r* cos *0,r* sin *0)r dr dO.*

Calculate f *e-x'udx*

*Solution:* Hopefully you happen to remember that the probability density function (pdf) of the standard normal distribution is *f()c)=-Cx' 12* By definition, we have

*27t-*

f (x)dx 77 e ca.

1 \_i2/2 \_

2A- 2

**x2/2dx.1 fe-x'j2dx.**

If you've forgotten the pdf of the standard normal distribution or if you are specifically

1

asked to prove E;\_i=e-x2/2*dx =1,* you will need to use polar integrals to solve the

*.427r*

problem:

•ic

*rdrd*9

-x 12CIX 2 .04-y2)/2

**"-Y**

f e-Y212dY = Le **cix**

r *r*-0-2 cos26+r2 sin2 0)/2 = f fir *e-r2'2rdra. - f e-r2I2d(-r2* 2) 12 *dO*

***3.3 Partial Derivatives and Multiple Integrals***

|  |  |
| --- | --- |
| *= -[e* | *—.P1121*0 *x.* I pil2R. 271.  v **Jo** |

|  |  |
| --- | --- |
| Since | r/2  *e'* 2*cbc = Le*—y212*dy , we* have E. *e*-x212ch NETT f e--x2/ *2*  **2** 7. |

**Partial derivative: w.** *f(x,y) f*—(xo, yo ) = litre (x° AY- *Y'})- f (x°' Y°) =*

i3x

*f a cif*

a2 *f af af*

*axay a, ay ay ax*

**order partial derivatives: — = —(—),**

**8x2 &x ax**

***3.4 Important Calculus Methods***

**Taylor's series**

One-dimensional Taylor's series expands function *f (x)* as the sum of a series using the derivatives at a point x xo :

**The general chain rule:** Suppose that w= *f (xi , x2,. • ,x,)* and that each of variables , x,, is a function of the variables *t„ t, • t, T. If* **all** these functions have

continuous first-order partial derivatives, then — • • - +'

47 *x aw* (3x *aw* ax,,, *for*

*at, 5x, a ax 2 a ax„,*

each i, 1

*f(x) = f (x0)* f Vcc,)(x *xo)* +—(x- x0)2 +...+ - (x - xo)" +...

**40 41**

2! *n!*

**Calculus and Linear Algebra**

**If x0 =0, *f(x)= f(0) + f 1(0)x + f "(o)* x2 + • ^** *f (0) e +.••*

**2! *n!***

**Taylor's series are often used to represent functions in power series terms. For example,**

**Taylor's series for three conunon transcendental functions, *,* sin x and cosx , at**

**x0 = 0 are**

**2 3**

**, X *X* X**

**=1-F--F—+—+•••**

**►mi) 1! 2! 3!**

**of the nth-degree Taylor**

**3 X57**

**•**

**n=0 (2n +1)! —x 3! ! 7!• •**

**(-- Ir X2" x2 X4 x6 COS X ----- 1 -- +— - •**

***(2n)!* 2! 4! 6!**

**The Taylor's series can also be expressed as the sum**

**sin x (-1)nX2n+1**

**polynomial 7;(x) = *(x0) + f '(x0)(x — x) +f "(x) (x — x0)2 + • • + f(")(xo)***

***(x —* xor and**

***n!***

**a remainder *R„(x): f (x) = Tn(x)+ Rn(x).***

**For some x between *x0* and *x***. ***R„ (x) =(5c-) x xo*** *I* ***f 0+1)***

***(n +1)! .* Let *M* be the maximum of**

**MX I X 0 I**

**x**

`**1)0)1 for all between x0 and x, we get constraint IR„(x)I <**

**(n+1)!**

***A.* What is ?**

***Solution:* The solution to this problem uses Euler' *e?* s formula, *e'9 = cos + I* sin t9, which**

**can be proven using Taylor's series. Let's look at the proof. Applying Taylor's series to *e'9,* cos and sin *0 ,* we have**

**est) 19**

**=1++ 0)2 00)3 (i9)4**

***0* 02 0304 *05***

**41 ----**

**1! 2! 3! —1+1—/—+ —+/—+•-•**

**1! 2! 3! 4! 5!**

***03 05B'***

***07***

**3! 5! 7!**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***cos 0***  sin | ***02 04 06***  ***i--+---+***   1. **4! 6!**   H\_ **0' *0'***   1. **5! 7!** | **•** | **isinu *n*** | **1!** |

**A Practical Guide To** Quantitative Finance Interviews

**Combining these three series, it is apparent that *e'0 = cos 0 + i* sin *0.***

**43**

**When *0 = 7r,* the equation becomes *ebr = cos it + i* sin ir = —1. When *0 = it I* 2, the equation becomes *eiff /2 =COS(T C12)+ i* sin *(ir I 2) = i.3* So In *i =* In (eig/2)=*11 - 1*2.**

**Hence, ln = *i* ln = *i(i7r I* 2) = *I* 2 = *CITI2 .***

***B.* Prove (1+ x)" 1+ *nx* for all x > —1 and for all integers 77.. 2 .**

***Solution:* Let *f (x)= (1+ x)" It* is clear that 1+ *nx* is the first two terms in the Taylor's series of *f(x)* with x0 = 0. So we can consider solving this problem using Taylor's series.**

**For x0 = 0 we have (1 + x)" =1 for Vn 2. The first and secondary derivatives of *f (x)* are *f '(x) =* n(1 + x)n-I and *f "(x) = n(n —1)(1+* x)"-2. Applying Taylor's series, we have**

*f (x) = f (;) + f 1(. x)(x — x0) + (x x )2 = f (0) + f '(0)x +*  **"(i) X2**

**2! ° 2! ,**

**---- 1+ *nx + n(n* —1)(1+ i)n-2x2**

**where if x<0 and *x.?\_5(-0* if x>0.**

**Since x > —1 and n 2, we have *n >* 0, (n-1) > 0, (1 +7)'r-2 >0, X2 0.**

**Hence, *n(n —* 1)(1+ x)" x' > 0 and *f (x) = (1 + x)" > 1 + nx***

**If Taylor's series does not jump to your mind, the condition that *n* is an integer may give you the hint that you can try the induction method. We can rephrase the problem as: for every integer *n* 2 , prove (1+ x)" 1 + *nx* for x > —1 .**

**The base case: show (1 + x)" 1+ *nx x > —1* when n = 2, which can be easily proven since (1+x)2 \_1+2x+x2 1+2x, Vx > —1.**

**The induction step: show that if (1 + x)" 1+ nx,Vx > —1 when *n = k,* the same statement holds for *n = k +1: (1 + x)" ?\_1 + (k + 1)x, V x > —1.* This step is straightforward as well.**

**Clearly they satisfy equation *(eL" =*** *e* **= —1.**

(I +x)" = (1+ x)k (I +

Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews

“l+loc)(1+x)=1+(k+Dx+kx2, Vx> **—1**

...1+(k+1)x

So the statement holds for all integers 2 when x > —1.

**Newton's method**

Newton's method, also known as the Newton-Raphson method or the Newton-Fourier method, is an iterative process for solving the equation *f (x) =* 0. It begins with an initial value x0 and applies the iterative step ;Jr, *=* xn *f ('r*„*)n)* to solve *f (x) =* 0 if xi, x2,... *f '(x„)*

Convergence of Newton's method is not guaranteed, especially when the starting point is far away from the correct solution. For Newton's method to **converge, it is** often necessary that the initial point is sufficiently close to the root; *\_f(x)*

converge.4

....\_. - - - must be

differentiable around a the root. When it does converge, the convergence **rate** is quadratic,
  
 (xn, —xf)2

,

which means —-----=-- < 5 <1, where *x* i

(; —xf)2 is the solution to *1 (x) =* 0.

*A.* Solve x2 = 37 to the third digit.

*Solution:* Let *f(x) =* x2 —37, the original problem is equivalent to solving *f(x)* = xo = 6 is a natural initial guess. Applying Newton's method, we have

*f (xo) x2 —* 37 z 36-37

= xo — xo *o*

— o 6.083.

*f 'Oro)* 2; 2 X 6

(6.0832 = 37.00289, which is very close to 37.)

**If you do not** remember Newton's method,you can directly apply Taylor's function *Nr.; (x) =* with *.f l(x) = ix ” :*

*1(37) f(36)+* '(36X37 —36) = 6+1/12 - 6.083 series **for**

**4** The iteration equation comes from the first-order Tay*)* lor's series:

*f f (x.) + t(x.)(x,,, x). 0 x ,,*

*ftx*

*f '(x.)*

**44**

**Alternatively, we can use algebra since it** is obvious that the solution should be slightly higher than 6. We have (6 + **y)2** 37 **y2 + y** *—*

*I 2 —1=* **0. If we ignore** the y2 term, which is small, then y = 0.083 and x = 6+ y = 6.083 .

*B.* Could you explain some root-finding algorithms to solve *f (x) =* **0 ? Assume *.f*** *(x)* is a differentiable function.

*Solution:* Besides Newton's method, the bisection method **and the secant method are two** alternative methods for root-finding. 5

**Bisection method** is an intuitive root-finding algorithm. It starts with two initial values
  
*ao* and bo such that *f (00 <* 0 and *f (bo) >* 0. Since *f (x)* is differentiable, there must be
  
an x **between** *ao* **and bo that makes** *f (x) =* 0. At each step, we check the sign of
  
*f ((an + bn) 1* **2). If** *f ((an + b„)12)* < 0, we set *b„#, =* bn **and** *a„+, = (an + bn) 1* 2; If
  
*f ((an + b„) 12)* > 0, we set *an+,* = *a„* **and** *k+,* = *(an + b, ) 1* 2; If *f ((an + b,,)/* 2) = 0, or its
  
absolute value is within **allowable error, the iteration stops and** *x = (a„ + b„) 12.* The
  
, — x ,

bisection method converges linearly, xn+ *<1,* **which means it** is slower than

x,, *—xf*

Newton's method. But once you **find an** *ao 1 bo* pair, convergence is guaranteed.

**Secant method** starts with two initial values xo, x, and applies the iterative step

*xn* —n-,

*f(;).* It replaces the *f '(x„)* in Newton's method with a

x„, = *x n*

*f (x n) — f (xn\_i)*

linear **approximation** *f (x„ )— f (x„\_,)* **Compared with Newton's method,** it does not **•** xn— x„, require the calculation of **derivative** *f '(x„,) ,* **which makes it valuable** *if f '(x)* is difficult to calculate. **Its convergence rate is (l + -J)/ 2, which makes it faster than the bisection method but slower than Newton's method.\_\_Similar to Newton's method, convergence is not guaranteed if initial values are not close to the root.**

**Lagrange multipliers**

**The method of Lagrange multipliers** is a common technique used to find local maximums/minimums of a multivariate function with one or more constraints. 6

**5** Newton's method is also used in optimization—including multi-dimensional optimization problems—to find local minimums or maximums.

45

**Separable differential equations**

**Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews**

**Let** *f (xi,* x2, ;) be a **function of** *n* **variables** x = (x1, x2, -•-, x„) **with gradient**

**vector *Vf* (x) = , • •,**f-÷-,). **The necessary condition for maximizing or
  
minimizing** *f* **(x) subject to a set of** *k* constraints

gi(x1,x2,•••,;)=0, g2(x,,x2,•••,x„)= 0, •.•, gk(x,,x2,--,x„)= 0

is that *V:f(* x ) + Ayg, (x) + A2Vg2(x)+ • + *V g k (x) =* **0, where , • • • , *2k* are called the** Lagrange multipliers.

A separable differential equation has the form *—dy = g(x)h(y).* **Since it** is separable, we

*dx*

can express the original equation as *—dy =* ***g(x)dx.* Integrating both sides, we have the** *h(y)*

solution j---dy *= fg(x)dx*

*h(y)*

1. **Solve ordinary differential equation y'+** 6xy = 0, y(0). **1**

What is **the distance from the origin to the plane** 2x + 3y + 4z =12 **?**

*Solution:* The distance *(D)* from the origin to a plane is **the minimum distance** between the origin and points on the plane. Mathematically, the **problem can be expressed as**

**min** *D2 = f(x,y,z) = + y2 + z2*

S*i. g(x, y, z) =* 2x +3y + 4z-12 = 0

Applying the Lagrange multipliers, we have

Tzef+1=2x+2,1=0

+A-1.2y+3A=0

2x+4A, = 0

2x+3y-F4z-12 =0

In general, for a plane with equation *ax + by + cz = d,* the distance to the

*D -*

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | *=-24/29*  x=24/29  y = 36/29 *D N[04)2* )2 *(g)2*  *z =*48129 |  |
|  |
|  |  |  |

origin is

*Solution:* **Let** *g(x)* = -6x and *h(y) = y,* **we have d***y = -6xdx.* Integrate both sides of
  
the equation: = .1-6xdx In y = -3x2 *+c* y = *e-3x2',* where *c* is a constant.

Plugging in the **initial condition y(0) = 1, we** have *c* **0 and y =** *.*

1. Solve ordinary differential equation y = *x - y* 7

x + y

*Solution:* **Unlike the last example, this equation is not separable in its current form. But we can use a change of variable to turn it into a separable differential equation. Let z = x+ y ,** then the **original differential equation is converted to**

***d(z* -x) *x-(z*** *-x) dz* 2x

-1= -1 *zdz = 2xdx fzdz =* f2xdr + *c*

*dx dx z*

*(x + y)2 = z2 = 2x2+c* y2+ 2xy-x2 = ***c***

**First-order linear differential equations**

***3.5 Ordinary Differential Equations***

seen in interviews.

In this section, we cover four typical differential equation patterns that are commonly

A **first-order differential linear** equation has the form —dY + *P(x)y = Q(x).* The standard

*dx*

**approach to solving a first-order differential equation is to identify** a suitable function
  
***1(x),* called an integrating factor, such that** *I(x)(y'+ P(x)y) = 1(x) y'+ 1(x)P(x)y*

**which** 7 **Hint: Introduce variable *z* = x y.**

**6 The method of Lagrange multipliers is a special case of Karush-Kubn-Tucker (KKT) conditions,**

**47**

**reveals the necessary conditions for the solutions to constrained nonlinear optimization problems.**

**46**

Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews

*= (1(x)y)" ;* **Then we have** *(I (x)y)' = I (x)Q(x)* **and we can integrate both sides to solve 1.** If ri and r2 are real and r1 **r2 , then the general solution is y =** *cie" + c2er2x ;*

*11(x)Q(x)dx*

**for y:** *1(x)y = 1(x)Q(x)dx y =* 2. If *r,* **and r2 are real** and *r,* = *r2 = r,* then the general solution is y *cierx + c2xerx ;*

*1(x)*

3. **If ri and r2 are complex numbers** *a±i13,* then the general solution is

**The integrating factor, *1(x),* must satisfy *di(x)****=* ***1(x)P(x),* which means** *1(x)* **is a y =** *e" (c1 cos fix +* c, sin *fix) .*

dx

**It is easy to verify that the general solutions indeed satisfy the homogeneous linear solutions by taking the first and secondary derivatives of the general solutions.**

separable differential **equation with** general solution *1(x) = efr(x)dx* **8**

Solve ordinary different equation y'+y —.— 1

2-, y(1) = I, where x > 0.

x x

***Solution:* This is a typical** example of fi**rst-order linear equations with** *P(x) = 1 —* **and**

*x*

*1 fx f*

*Q(x) = .* So /(x) *=e"p ood = e'(i = elnx* = x and we have *1(x)Q(x) = 1.*

*x*

x

*1(x)(y'+ P(x)y) = (xy)' = 1(x)Q(x) =1****I*** *x*

**Taking integration on both sides,** xy = f(1/x)dx *=* **In x +** *c* **y = In x +**

x

**Plugging in y(1) = 1, we** get *c =1* and y = In *x* +1

x

**What is the solution of ordinary differential equation y "+** y'+ y = 0?

*Solution:* **In this specific** case, we *have a=b=c=1* **and** *b2 4ac =* **—3 < 0 ,** so we have complex roots r = —1/ 2 ± I / 2/ ( *a = —11* 2, f3 = .1S / 2), **and the general solution to the differential equation is therefore**

**y = *ec" (et* cos** *fix + c2* **sin *fix)* =**1/2x (c1 COS(s512x)+ *c*2 **sin(,A I 2x)) .***e*

**Nonhomogeneous linear equations**

Unlike a homogenous linear equation *a —d2y***+** *b—dy+ c =* 0, a nonhomogeneous linear

dx2 *dx*

**2y**

**equation** *a* **d** *—+b* dy *--+c=d(x)* **has no closed-form solution. But if we can** fi**nd a**

*dx dx*

*dx*

**Homogeneous linear equationsd2y *dy***

**A homogenous linear equation is a second-order differential equation with the** form particular solution yp(x) for *a* dx2 *+b —dx+ c = d(x),* **then y = *yp* (x) +** *Yg (x),* **where**

*,***x) ---i-** *d2 y f* 4

a*(*x) ***d2y dy***

***d* +** *b(x) dY* ***+ c(x) 0***

***chic y,(x)* is the general** solution of the **homogeneous equation *a --i-dx + b dx+ c —* 0, is a**

*--****-. .***

**d2y** *dy*

**It is easy to show that, if yl and y2 are linearly independent solutions to the homogeneous linear equation, then any y(x) cly,(x) c2 .Y2 (x),**

**arbitrary constants, is a solution to the homogeneous linear equation as well.**

**where ci and C2 are When** *a, b* **and** *c (a # 0* **) are constants instead of functionsof x**, **the homogenous linear equation has closed form solutions: ,**

**Let ri and r2 be the roots of the characteristic equation** *ar2 + br + c* **019**

**8 The constant *c* is not** needed in **this case since it just scales both sides of the equation by a factor. 48**

general solution of the nonhomogeneous equation **a —(11.2 +** *b—dx+ c = d(x) .*

9 *—b ±* ***117--.72c***

A quadratic **equation** *ar*e ***+ br c =* 0 has roots given by quadratic** formula ***r 2a .* You**

**should either commit the formula to memory or be able to derive it using *(r + b I 2af = (b2 — 4ac) I 4a2 .***

**49**

**Although it may be difficult to identify a particular solution *yr(x)* in general, in the special case when *d(x)* is a simple polynomial, the particular solution is often a polynomial of the same degree.**

**Calculus and Linear Algebra**

**A Practical Guide To Quantitative Finance Interviews**

**What is the solution of ODEs y"+ y =1 and y y'+ y = x ?**

***Solution:* In these ODEs, we again have *a = b = c =1* and *b2 —4cie =* —3 < 0, so we have complex solutions *r = —11*2 ± / 2i *(a = —112,* /3 — / 2 ) and the general solution is y = *e-112x (CI COS(15 /* 2x) + *C2 sin( IS /* 2x)).**

**What is a particular solution for y"+ yp+ y = 1? Clearly y =1 is. So the solution to *y"+ y =1* is**

***y = yr,(x)+ yg(x)=* cos(X 2x) + c2 sin(N5 / 2x)) +1.**

**To find a particular solution for y "+ y'+ y = *x,* Let y,, (x) = mx + *n,* then we have**

**y"+ y'+ y = 0 + m + (mx + *n) = x m =1, n = —1.* So the particular solution is x —1 the solution to y"-f- y1+ y = x is**

**y = (x)+ yg(x) *e-112 x (C1* cos(Z/ 2x)+ c2 sina-3-/ 2x)) + (x**

**Inner product/dot product: the inner product (or dot product) of two *R"* vectors x and y is defined as x,y, = y**

i=1

***\In***

.s' '1/(X — AT (X — **Y)**

1=1

**Euclidean norm:** H. Lx, = x p o

x; ox— Yll =

**x y**

**Then angle *0* between *k'* vectors x and y has the property that cos *0 = IlxIIIIYII• x* and y are orthogonal if xry = O. The correlation coefficient of two random variables can be viewed as the cosine of the angle between them in Euclidean space ( *p =* cos *0).***

**There are 3 random variables x, y and *z.* The correlation between x and y is 0.8 and the correlation between x and** *z* **is 0.8. What is the maximum and minimum correlation between y and *z?***

***Solution:* We can consider random variables x, y and** *z* **as vectors. Let *0* be the angle between x and y, then we have cos *0 = p****x,****. =* 0.8. Similarly the angle between x and *z is 0* as well. For y and *z* to have the maximum correlation, the angle between them needs to be the smallest. In this case, the minimum angle is 0 (when vector y and *z* are in the same direction) and the correlation is 1. For the minimum correlation, we want the maximum angle between y and z, which is the case shown in Figure 3.2.**

**cb**

**If you still remember some trigonometry,**

***3.6 Linear Algebra***

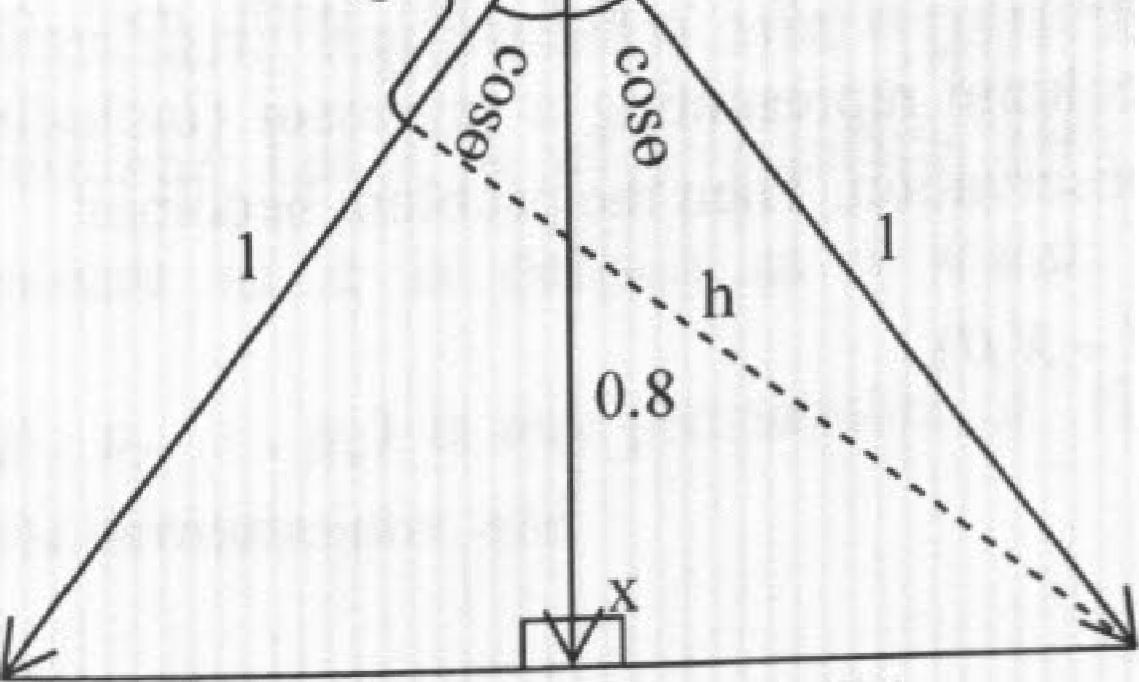
**Linear algebra is extensively used in applied quantitative finance because of its role** in
  
**statistics, optimization, Monte Carlo simulation, signal processing, etc. Not surprisinglY,
  
it is also a comprehensive mathematical field that covers many topics. In this section. we**

**discuss several topics that have significant applications in statistics and numerical methods.**

**Vectors**

**An *nx* 1 (column) vector is a one-dimensional array. It can represent the coordinates of a point in the *R"* (n-dimensional) Euclidean space.**

**all you need is that**

**cos(20) = (cos 0)2 — (sin 0)2**

**50 51**

**= 0.82 0.62 = 0.28**

**Otherwise, you can solve the problem using Pythagoras's Theorem:**

**0.8x 1.2 = 1** *xhh=* **0.96 cos20 =NiF— 0.962 = 0.28**

**0.6 0.6**

**Figure 3.2 Minimum correlation and maximum angle between vectors** y **and** *z*

A Practical Guide To Quantitative Finance Interviews To minimize the function *f(8),* taking the first derivative' 1 of *f (13)* with respect to )8, we have *f(11)=2XT (Y – X 13).* 0 (XIX)ft *= A°* Y, where (X' X) is a *pxp* symmetric matrix and XLY is a *p* x1 column vector.

Calculus and Linear Algebra

**QR decomposition**

QR decomposition: For each non-singular *nx n* matrix A, there is a unique pair of orthogonal matrix Q and upper-triangular matrix *R* with positive diagonal elements such that A - *QR*

QR decomposition is often used to solve linear systems *Ax = b* when A is a non-singular
  
matrix. Since *Q* is an orthogonal matrix, Q-1 = *01* and *QRx = b = Rx = Q b.* Because *R*is an upper-triangular matrix, we can begin with xn (the equation is simply

*(Q1 Ii)„),* and recursively calculate all x1, = *n,n* ,1 .

0'

•

Let A = *(XT)()* and *b = X7* Y, then the problem becomes Afi" = *b,* which can be solved using QR decomposition as we described.

Alternatively, if the programming language has a function for matrix inverse, we can directly calculate /3 as )6 = Ar1'xy.'2

Since we are discussing linear regressions, its worthwhile to point out the assumptions behind the linear least squares regression (a common statistics question at interviews):

1. The relationship between Y and Xis linear: Y X,8

IC the programming language you are using does not have a function for the linear least squares regression, how would you design an algorithm to do so?

*Solution:* The linear least squares regression is probably the most -widely used statistical
  
analysis method. Let's go over a standard approach to solving linear least squares

regressions using matrices. A simple linear regression with *a* observations can be

expressed as

y, + /31 + • + *13p\_Ix „p\_i+ V =1j - • • ,n,* where x,, 1,*V i,* is the

term and x, • , are *p –* 1 exogenous regressors.

y Xig+ *E,* where Y [Y,1„, - - and *E* rel — , 1 are both *n xl* column

vectors: is a *nx p* matrix with each column representing a regressor (including the

intercept) and each row representirm an observation. Then the problem becomes

min *1(11.)* rain I ,2 min( y *A-A) — X 13)*

intercept

The goal of the linear least squares regression is to find a set of =

*n*ryar 1,

that makesEs: the smallest. Let's express the linear regression in matrix format:

***d***

1. *E[0. 0, V i =1,- • -,n.*
2. var(e,) *i* =1, • , *n* (constant variance). and *Eteicji=* 0, *i* (uncorrelated
     
   errors).
3. No perfect multicollinearity: *p(x„r, j) ±-1, j* where *p(x„xj)* is the
     
   correlation of regressors x, and
4. *g* and x-, are independent.

Surely in practice, some of these assumptions are violated and the simple linear least squares regression is no longer the. best linear unbiased estimator (BLUE). Many econometrics books dedicate significant chapters to addressing the effects of assumption violations and corresponding remedies.

**Determinant, eigenvalue and eigenvector**

Determinant: Let *A* be an *nx n* matrix with elements {,41, j}, where i, j = -,n. The

where

determinant of A is defined as a scalar: det(A)– *Ev(p)a. a 0*

*p ="--- (pp P2, " • pp)* is any permutation of (1, 2, )7); the sum is taken over all

possible permutations; and

**A** nonsingular matrix Qis called an orthogonal matrix if • Q is orthogonal if acrd only if the

'—'1411111S (and ro\A'. 01 Q tbrni an orthonormal set of vectors in *W .* The Gram-SchnlidRI orthonoonalliation process (often improved to increase numerical stability) is often used for Qss. deem ition. Please refer to a linear algebra textbook if you are interested in the Gram-Schmidt Proc.'

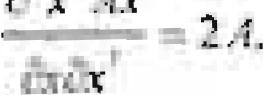
51

equations for vectors/matrices arc

" To do that, you do need a little knowledge about matrix derivatives. Some of the important derivative

a):

*--. (.4' 4- ..1).1c,*



*Cx AX*

*a(Ax + b)' r e) .*

.r7.v

The matrix inverse introduces large numerical error if the matrix is close to sinaular or badly scaled.

53

**Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| For  det | *yI(p)=*  ,  -1,  example,  *db* | if *p* can be coverted if p can be coverted  determinants of  *= ad -bc,* det | to natural to natural 2 x2 and  *(*  *a b*c  d *e f g h* | order by even number of exchanges order by odd number of exchanges  3x 3 matrices can be calculated as  *= aei + bfk +cdh-ceg - afh-bcii .)3* | 1  If matrix *A =*  [21 2  *Solution:* This is a using three related approaches: Approach *A:* Apply the Let *A* be an eigenvalue | , what are the eigenvalues simple example  definition of  and x -- [x,  x2 | and eigenvectors o f *A?*  of eigenvalues and eigenvectors. It can be solved  eigenvalues and eigenvectors directly.  be its corresponding eigenvector. By definition, we |

Determinant properties: det(A') = det(A), det(AB) = det(A)det(B), det(A-') **det(A)**

|  |  |  |  |
| --- | --- | --- | --- |
| have  *Ax ,`*2 | lir x, [2x, + x2  = Ax =  2ILx, Lx 4-2x**2**  **LAX2 .** | 2x, *+ x2 = AXI* 2x2 = Ax2 | 3(x, + x2) = A(x, + x2) |

**54 55**

**1**

**Elgenvalue: Let** *A* be an nxn matrix. A real number A is called an eigenvalue of A if there exists a nonzero vector *x ire R"* such that *Ax = Ax.* Every nonzero vector *x* satisfying this equation is called an **eigenvector of** *A* associated with the eigenvalue A.

Eigenvalues and eigenvectors are crucial concepts in a variety of subjects such as ordinary differential equations, **Mark0v** chains, principal component analysis (PCA). etc. The importance of determinant lies in its relationship to eigenvalues/eigenvectors.

The determinant of matrix *A-21,* where 1 is an nx *n* identity matrix with ones on the main diagonal and zeros elsewhere, is called the **characteristic polynomial of A,** The equation det(A *2.1)-. 0* is called the **characteristic equation of A.** The eigenvalues of .4 are the real roots of the characteristic equation of *A.* Using the Characteristic equation.

**we** can also show

*„* that Ai A. • • • *A* det(A) and EA, - *troce(A)=*

,.1 i.1

is **diagonalizable** if and only if it has linearly independent eigenvectors. 15 Let • • An be the eigenvalues of *A, x,. x,. • x,* be the corresponding cigenvectors. and *X* .[x, x2 • x,j ], then

V

*D A XDX1 24'*

An\_

1 x 'AX I

**IS**

**' ' In practice, determinant is usually not solved by the sum of all permutations because Lomputatimally inefficient 1,1J decomposition and cofactors are often used to calculate deter-inn**

**llTSCead.**

**Determinant can also be applied to matrix inverse and linear equations as well.**

**[fall *n* eigenvalucs** are **real and distinct, then the eigenveetors are independent and A is diagonalizable'**

So either A = 3, in which case xl = x2 (plug A = 3 into equation 2x, + x, = *Axi)* and the *ll* , or x, + x2 - 0, in which case the *ll*

and 2. = l (plug x2 = -x1 into equation

corresponding normalized eigenvector is

|  |  |
| --- | --- |
|  |  |
| normalized eigenvector is |
|  |

2x, + x2 = Ax, ).

Approach *B:* Use equation det(A - *21).* 0.

det(A = 0 (2 - A)(2 - A) -1 - 0. Solving the equation, we have 2 =1 and

the eigenvalues to *Ax = Ax,* we can get the corresponding

eigenvectors.

lyingecAtoprsp.

**11**

Approach *C:* Use equations 2 .4 • • • 1„ = det(A) and EA, = *trace(A),\_* det(A)=2x2-1x1= 3 and *trace(A)=2x2=4.*

*A.,x 2,*

o we have

*=* 3

Sve - A' =1 . Again apply the eigenvalues to *Ax = Ax,* and we

can get the corresponding eigcnvectors.

**Positive semidefiniteldefinite matrix**

Calculus and Linear Algebra A Practical Guide To Quantitative Finance Interviews

When A is a symmetric *nxn* matrix, as in the cases of covariance and correlation matrices, all the eigenvalues of *A* are real numbers. Furthermore, all eigenvectors that belong to distinct eigenvalues ofA are orthogonal.

Each of the following conditions is a necessary and sufficient condition to make a symmetric matrix *A* **positive semideiinite:**

**1,** *xr Ax?.* 0 for any *nxi* vector x .

1. **All** eigenvalues of*A* are nonnegative.
2. All the upper left (or lower right) submatrices A, *K =1, n* have nonnegative
     
   determinants.16

Covariance/correlation matrices must also be positive semidefinite. If there is no perfect linear dependence among random variables, the covariance/correlation matrix must also be positive definite. Each of the following conditions is a necessary and sufficient condition to make a symmetric matrix *A* **positive definite:**

1. Ax > 0 for any nonzero *nx* I vector x
2. All eigertvalues of *A* are positive.
3. All the upper left (or lower right) submatrices *A io K =1, • , n* have positive determinants.

There are 3 random variables x, *y* and *r.* The correlation between x and y is 0.8 and the correlation between x and *z* is 0.8. What is the maximum and minimum correlation between y and *z?*

*Solution:* The problem can he solved using the positive semidefiniteness property of the correlation matrix.

Let the correlation between *y* and *z* be p , then the correlation matrix for x, y and z is

|  |  |  |  |
| --- | --- | --- | --- |
|  | 1 0.8 0.S | | |
| P = | 0.8 | 1 | *p* |
|  | 0.8 | *p* | *1* |

A necessary, but not sufficient, condition for matrix A to be positive sernidi Finite is that A has *no* negative diagonal elements. det(P) =I x det

— 0.8x det .+0.8xdet

([ p

1-0.8 0.81) 0.8 0.8

*1 J)* 1

56 57

*[[p*

1 *p*

= — p2)— 0.8x (0.8— 0.8p)+ 0.8x (0.8p-0.8) = —0.28+ .28p— p2

*(p —1)(p —0.28)* 0 0.28

So the maximum correlation between y and *z* is 1, the minimum is 0.28.

**LU decomposition and Cholesky decomposition**

Let *A* be a nonsingular *n* x *n* matrix. **LU decomposition** expresses *A* as the product of a lower and upper triangular matrix: *A = LU .1* 7

LU decomposition can be use to solve *Ax =1)* and calculate the determinant of *A:*

01

*LUx = b Ux = y, Ly =b;* det(A) = det(L)det(U) n

1-1

*When A* is a symmetric positive definite matrix, **Cholesky decomposition** expresses *A*as *A= R7 R,* where *R* is a unique upper-triangular matrix with positive diagonal entries.

Essentially, it is a LU decomposition with the property *L =UT .*

Cholesky decomposition is useful in Monte Carlo simulation to generate correlated random variables as shown in the following problem:

How do you generate two N(0,1) (standard normal distribution) random variables with correlation *p* if you have a random number generator for standard normal distribution?

*Solution:* **Two N(0,1) random variables xt,** x, with a correlation *p* can be generated from independent N(0,1) random variables z, z2 using the following equations:

= zi

x2 + ,02:2

It is easy to confirm that var(x,)= var(zi) =1, var(x2) = *p2* var(zi) + (1— *p2)* vai(z2) = **1,** -

and eov(x,,.x-,) = cov(z,,pz, +. z.2) = cov(z..1, *pri)= p .*

This approach is a basic example using Cholesky decomposition to generate correlated
  
random numbers. To generate correlated random variables that follow a n-dimensional

17 .

decomposition occurs naturally in Gaussian elimination.

Calculus and Linear Algebra

multivariate normal distribution *X =[XI, X 2,• • • ,X*,,]'*-'* E) with mean

*= [Pi P21— ' ,P„]7.* and covariance matrix I, (a *17x n* positive definite mat ri x)18, we can decompose the covariance matrix 2, into *RT R* and generate n independent ,V(0, random variables *z, z,,* • • •, *z* Let vector *Z = z2,••• ,* then *X* can be generated

as X + R T *Z.* 19

Alternatively, X can also be generated using another important matrix decomposition called **singular value decomposition (SVD):** For any nx *p* matrix X, there exists *a* factorization of the form A' = *UDV'* , where *U* and *V* are *n x p* and *p x p* orthogonal matrices, with columns of *U* spanning the column space of X, and the columns of V spanning the row space; D is a *px p* diagonal matrix called the singular values of X. For a positive definite covariance matrix, we have V= *U* and E *UD(11.* Furthermore, *D* is the diagonal matrix of eigenvalues k and *U* is the matrix or *N* corresponding eig,envectors. Let *D1.2* be a diagonal matrix with diagonal elements Nrri F, then it is clear that *D =(D")2 (D112 )(D'I2 )1* and

*UDItz(UDI''Y* Again, if we generate a vector of *n* independent N(0, **1)** random variables Z I. ?\ , *z ]'*, X ean be generated as X = *p + (U )Z*

**Chapter 4 Probability Theory**

Chances are that you will face at least a couple of probability problems in most quantitative interviews. Probability theory is the foundation of every aspect of quantitative finance. As a result, it has become a popular topic in quantitative interviews.

Although good intuition and logic can help you solve many of the probability problems, having a thorough understanding of basic probability theory will provide you with clear and concise solutions to most of the problems you are likely to encounter. Furthermore, probability theory is extremely valuable in explaining some of the seemingly-counterintuitive results. Armed with a little knowledge, you will find that many of the interview problems are no more than disguised textbook problems.

So we dedicate this chapter to reviewing basic probability theory that is not only broadly tested in interviews but also likely to be helpful for your future career. The knowledge is applied to real interview problems to demonstrate the power of probability theory. Nevertheless, the necessity of knowledge in no way downplays the role of intuition and logic. Quite the contrary, common sense and sound judgment are always crucial for analyzing and solving either interview or real-life problems. As you will see in the following sections, all the techniques we discussed in Chapter 2 still play a vital role in solving many of the probability problems.

Let's have some fun playing the odds.

***4.1 Basic Probability Definitions and Set Operations***

First let's begin with some basic definitions and notations used in probability. These definitions and notations may seem dry without examples—which we will present momentarily—yet they are crucial to our understanding of probability theory. In addition, it will lay a solid ground for us to systematically approach probability

problems.

**Outcome** (o.): the outcome of an experiment or trial.

**Sample space/Probability space (0):** the set of all possible outcomes of an experiment.

As I have emphasized in Chapter 3, this book does not teach probability or any other math topics due to the space limit. —it is not my goal to do so: either. The book gives a summary of the frequently-tested knowledge and shows how it can be applied to a wide range of real interview problems. The knowledge used in this chapter is covered by most introductory probability books. It is always helpful to pick up one or two classic probability books in case you want to refresh your memory on some of the topics. My personal favorites are ***First Course in Probability*** by Sheldon Ross and ***Introduction to Probability*** by



58

The probability density of multivariate normal distribution is fr "P( - (A I- • .

)

det(IS Jar general, if *p,* where A and *h* are constant, then the covariance matricel,

Dim itri P. Bertsekas and John N. Tsitsiklis.

**Coin toss game**

Probability Theory A Practical Guide To Quantitative Finance Interviews

P(w): Probability of an outcome *(P(w)* 0, *Vro* **E** Q, *P(w)* ).

Event: A set of outcomes and a subset of the sample space. *P( A):* Probability of an event *A, P(A)= EP(0).*

*Au B* Union *A* u *B* is the set of outcomes in event *A* or in event *B* (or both).

A *n /3* or *AB :* Intersection *A n B* (or *AB)* is the set of outcomes in both A and /3. A': The complement of A, which is the event "not *A".*

**Mutually Exclusive:** *A r B =*Icl) where (I) is an empty set.

*N*

For any mutually exclusive events E1, *E„• • - EN, P N =* E

**Random variable:** A **function** that maps each outcome (to) in the sample space **(a)** into the set of real numbers.

Let's use the rolling of a six-sided dice to explain these definitions and notations. A roll of a dice has 6 possible outcomes (mapped to a random variable): 1, 2, 3, 4, 5, or 6. So the sample space 0 is {1,2,3,4,5,61 and the probability of each outcome is 116 (assuming a fair dice), We can define an event *A* representing the event that the outcoinc is an odd number *A,* {1, 3. 5), then the complement of *A* is *A' =* {2, 4, 6}. Clearly *P(A) =* PM+ P(3)+ P(5) =1 / 2 Let ***B*** be the event that the outcome is larger than *3:*

*B =* (4, 5, 6). Then the union is A u ***B ==*** {1. 3, *4. 5,* 6) and the intersection is *13 =* {5} One popular random variable called indicator variable (a binary dunini) variable) for event *A* is defined as the following:

*(1, if x e* 11, 3. 51

,

. Basically 1, when A occurs and *=0* if *A'* occurs. The

Lo. x e U, 3,5a

expected value of I. is *E[1] = Pf A)*

Nov. time for some examples.

60 61

Two gamblers are playing a coin toss game. Gambler A has *(n +* 1) fair coins; *B* has *n*

fair coins. What is the probability that A will have more heads than *B* if both flip all their coins?2

*Solution:* We have yet to cover all the powerful tools probability theory offers. What do we have now? Outcomes, events, event probabilities, and surely our reasoning capabilities! The one extra coin makes *A* different from *B.* If we remove a coin from *A, A* and *B* will become symmetric. Not surprisingly, the symmetry will give us a lot of nice properties. So let's remove the last coin of *A* and compare the number of heads in *A's* first *n* coins with B's *n* coins. There are three possible outcomes:

*E,: A's n* coins have more heads than *B's n* coins;

*E2* A's n coins have equal number of heads as *B's n* coins;

*E*3 *A's ii* coins have fewer heads than B's *n* coins.

By symmetry, the probability that *A* has more heads is equal to the probability that ***B*** has more heads. So we have *P(E1)= P(E3).* Let's denote *P(E1) = P(E3)= x* and *P(E2)* = y.

Since E 1, we have 2x + y =1. For event *E1, A* will always have more heads

than *B* no matter what A's (n +1)th coin's side is; for event E3 , *A* will have no more heads than ***B*** no matter what A's *(n +1)th* coin's side is. For event *E2 A's (n +1)th*

coin does make a difference. If it's a head, which happens with probability 0.5, it will
  
make *A* have more heads than *B.* So the (n + 1)th coin increases the probability that A

has more heads than *B* by 0.5y and the total probability that *A* has more heads is x + 0.5y = x + 0.5(1- 2x) = 0.5 when *A* has *(n+* 1) coins.

**Card game**

A casino offers a simple card game. There are 52 cards in a deck with 4 cards for each

jack queen kiR acc

value 2, 3, 4, 5, 6, 7, 8, 9, 10, *J, Q, K, A .* Each time the cards are thoroughly shuffled

(so each card has equal probability of being selected). You pick up a card from the deck and the dealer picks another one without replacement. If you have a larger number, you win; if the numbers are equal or yours is smaller, the house wins- as in all other casinos, the house always has better odds of winning. What is your probability of winning?

Hint: What are the possible results (events) if We compare the number of heads in A's first *17* coins with B's ***n*** coins? By making the number of coins equal, we can take advantage of symmetry. For each event,

what will happen if A's last coin is a head? Or a tail?

Probability Theory

*Solution:* One answer to this problem is to consider all 13 different outcomes of your
  
card. The card can have a value 2, 3, • *„A* and each has 1/13 of probability. With a
  
value of 2, the probability of winning is 0/51; with a value of 3, theprobability of

winning is 4/51 (when the dealer picks a 2); ...; with a value of *A.* theprobability of

p

winning is 48/51 (when the dealer picks a 2, 3, • , or *K).* So your probability of winning is

Although this is a straightforward solution and it elegantly uses the sum of an integer
  
sequence, it is not the most efficient way to solve the problem. If you have got the core

different outcomes:

spirits of the coin tossing problem, you may approach the problem by considering three *E, :* Your card has a number larger than the dealer's;

*E,:* Your card has a number equal to the dealer's•

*El:* Your card has a number lower than the dealer's.

Again by symmetry, *P(E). P(E.,).* So we only need to figure out *p(Eo.* the

probability that two cards have equal value, Let's say you have randomly selected a card. Among the remaining 51 cards, only 3 cards will have the same value as your card. ,So the probability that the two cards have equal value is 3/51. As a result, the probabillt). that you win isi)(E1).--(1— *P(E2))I 2=* (1-3/51)/2 8/17.

Ix .2\_+±+...\_i\_ 41\_

(

x(0+1+•--+12)=-4 xl2x13 8

13 51 51 51 13x5l 13x51 2 17

=

**Drunk passenger**

A line of /00 airline passengers are waiting to board a plane. They each hold a ticket *!()* one of the 100 seats on that flight. For convenience, let's say that the n-th passenger in line has a ticket for the seat number *n.* Being drunk, the first person in line

random seat (equally likely for each seat). Al? of the other passengers are sober, and vkl\_ /, go to their proper seats unless it is already occupied; In that ease, they will random!) choose a free seat. You're person number 100. What is the probability

in' our seat (i.e., scat 4100) '1)3 picks.1

that you end LIP *Solution: Let's* consider seats *#* I and 4100. There are two possible outcomes:

I fyou are ti:‘,,ing to use complicated conditional probability to solve the problem,

If you decide to start with a simpler version or the problem, starting with two passilgers ore

problem m

incrrac.inp die number of passengers to show. a pattem by induction, vou can solve the

;11 have officientl:.-.. Bur ilie problem is much simpler than that Focus on events and symmetry and yoll Vi•

go back and

intuitiyc A Practical Guide To Quantitative Finance Interviews

63

*E!:* Seat 41 is taken before 4100;
  
*E2 :* Seat 4100 is taken before 41.

If any passenger takes seat 4100 before 41 is taken, surely you will not end up in you own seat. But if any passenger takes 41 before 4100 is taken, you will definitely end up in you own seat. By symmetry, either outcome has a probability of 0.5. So the probability that you end up in your seat is 50%.

in case this over-simplified version of reasoning is not clear to you, consider the following detailed explanation: If the drunk passenger takes #1 by chance, then it's clear all the rest of the passengers will have the correct seats, If lie takes 4100, then you will not get your seat. The probabilities that he takes 41 or 4100 are equal. Otherwise assume that he takes the *n-th* seat, where *n* is a number between 2 and 99. Everyone between 2 and *(n-1)* will get his own seal. That means the *n-th* passenger essentially becomes the

new "drunk" guy with designated seat #1.. If he chooses 41, all the rest of the passengers will have the correct seats. If he takes 4100, then you will not get your seat. (The probabilities that he takes 41 or #100 are again equal.) Otherwise he will just make another passenger down the line the new "drunk" guy with designated seat 41 and each new "drunk" guy has equal probability of taking *41* or 4100. Since at all jump points there's an equal probability for the "drunk" guy to choose seat 41 or 100, by symmetry, the probability that you, as the 100/1 passenger, will seat in #100 is 0.5.

**N points on a circle**

Given *N* points drawn randomly on the circumference of a circle, what is the probability that they are all within a sernicircle74

*Solution:* Let's start at one point and clockwise label the points as 1, 2, • • • *N .* The

probability that all the remaining *N —I* points from 2 to *N* are in the clockwise semicircle starting at point I (That is, if point 1 is at 12:00, points 2 to *N* are all between 12:00 and 6:00) is 1/2k 1. Similarly the probability that a clockwise semicircle starting at any point *1,* where *i* E {2,•-•, *N}* contains all the other *N* —1 points is also

1/

Claim: the events that all the other N —1 points arc in the clockwise semicircle starting
  
at point *1, i* 2.•-•, N are mutually exclusive. In other words, if we. starting at point *i*

and proceeding clockwise along the circle, sequentially encounters points *i +1, i +*2,..., N, 1, • • *i* —1 in half a circle, then starting at any other point *j,* we cannot encounter all

4 flint: Consider the events that starting from a point *it,* you can reach all the rest of the points on the circle clockwise, *n* N.) in a semicircle. Are these events mutually exclusive?

other points within a clockwise semicircle. Figure 4.1 clearly demonstrates this
  
conclusion. if starting at point *i* and proceeding clockwise along the circle, we

Probability Theory A Practical Guide To Quantitative Finance Interviews

sequentially encounter points *i* +1, *i* 2, *N, 1, i —1* within half a circle, the

clockwise arc between *i* —1 and r must be no less than half a circle. If we start at any other point, in order to reach all other points clockwise, the clockwise arc between

and *i* are always included. So we cannot reach all points within a clockwise semicircle starting from any other points. Hence, all these events are mutually exclusive and we have

*N ti r* N

P UE, *'/P(E,)/INE,*

1x112A`-' =NUN-I

*'1/4,11* I I

-File same argument can be extended to any arcs that have a length less than half a circle. If the ratio of the arc length to the circumference of the circle is x ( x < 112 ), then the probability of all N points fitting into the arc is Nx

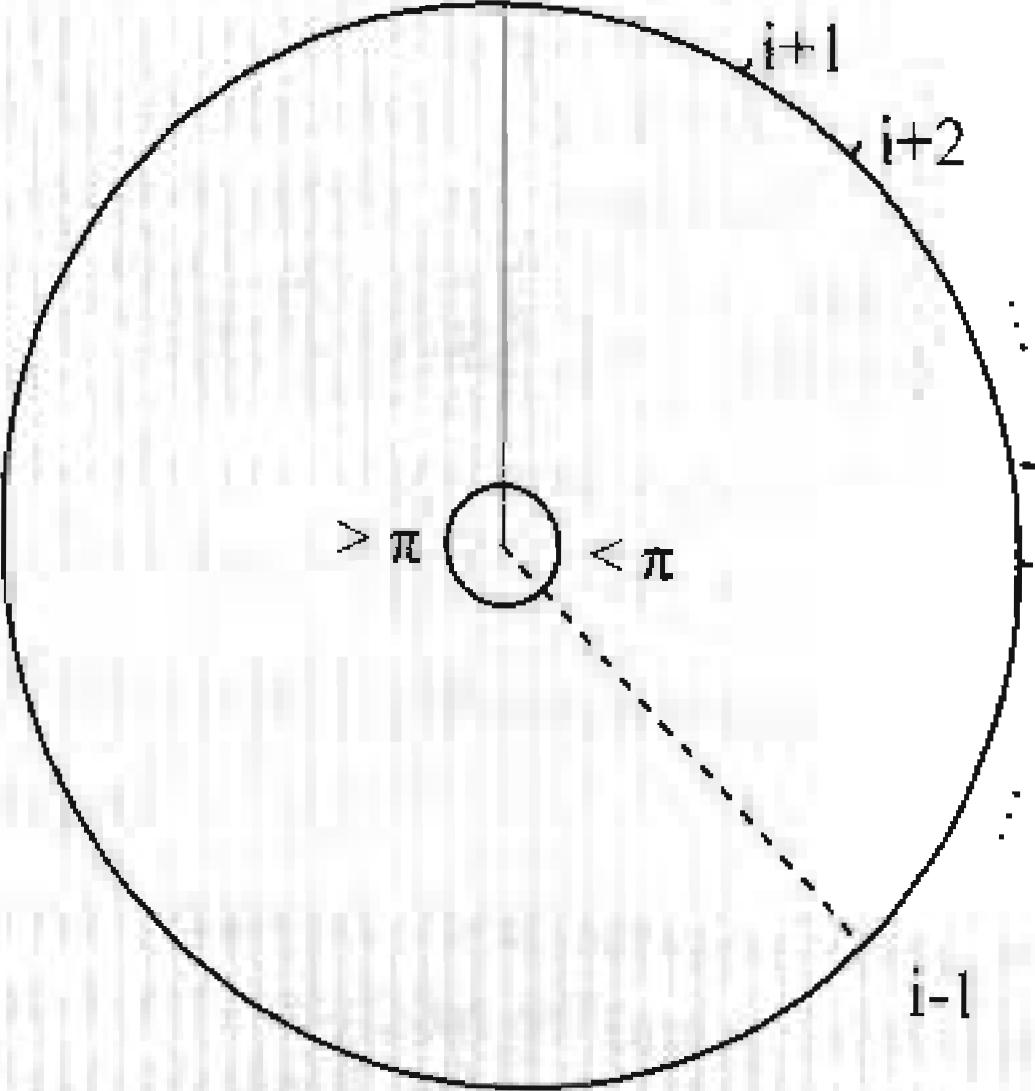


Figure 4.1 ***N*** points fall in a clockwise semicircle starting from i

* possible first entries,
* *n,* possible second entries for each first entry,
* n3 possible third entries for each combination of first and second entries, etc. Then there are a total of *n1 • n?..-nk* possible outcomes.

**Permutation:** A rearrangement of objects into distinct sequence (i.e., order matters).

Property: There are *n!* different permutations of *n* objects. of which

*ni!n2!...nr!* r; are

alike, *n2* are alike, • , *nt.* are alike.

Combination; An unordered collection of objects (i.e., order doesn't matter).

*n!*

Property: There are = different combinations of *n* distinct objects taken

*r (n— r)!r!*

*r* at a time.

*(11.\*

***k*** A-

x y***w***

**pinornial theorem: (x÷** =

*k=0*

**Inclusion-Exclusion Principle: *P(E* kiE2)** *P(E1)-1- P(E,) — P(EIE,)*

*P(Eiu* ***E2L) E3)= P(E3+ P(E2)+*** *P(Ea)— P(EiE2)— P(E,E3)— P(E2E3)+* E-2 E3
  
and more generally,

*P(Eiu* ***E24 )...0 EAT) =ZP(Ei)—EP(Ei,E,$)+-•.+(-1)--1***

*,=1*

*+(-1)N+1 P(EE2\* • • E.,)*

where I *1'(E, E„,- • Et)* has ( terms.

***4.2 Combinatorial Analysis***

Many problems in probability theory can be solved by simply counting the number (3"1: different ways that a Certain event can occur. The mathematic theory of counting ;ii

Often referred to as Combinatorial analysis (or combinatorics). In this section. we "I cover the basics of combinatorial analysis.

**Ilasic principle of counting:** Let *S* be a set of length-k sequences. **If** there are

**Poker hands**

Poker is a card game in which each player gets a hand of 5 cards. There are 52 cards in a
  
deck. Each card has a value and belongs to a suit. There are 13 values,

64 65

jack quail kW!' aix :spade t Lir• 1 v:311 :II din

2, 3, 4, 5. 6, *7.* 8. 9, 10. *.1, Q. K. A,* and four suits, 4 , **V , •**

What are the probabilities of getting hands with four-of-a-kind (four of the five cards with the same value)? Hands with a full house (three cards of one value and two cards of another value)? Hands with two pairs?

Probability Theory A Practical Guide To Quantitative Finance Interviews

*Solution:* The number of different hands of a five-card draw is the number of 5-element subsets of a 52-clement set, so total number of hands = ,52

5 = 2,598.960.

**Hands with a four-of-a-kind:** First we can choose the value of the four cards with the same value, there are 13 choices. The 5th card can be any of the rest 48 cards **(12 choices for values and** 4 choices for suits). So the number of hands with four-of-a kind is 13>c 48 =624 .

**Hands with a Full House: In** sequence we need to choose the value of the triple. 13 choices; the suits of the triple, (41 choices; the value of thepair, 12 choices; and the 3)

of hands with full house is

suits of the pair,

choices, So the

number

(4

2)

134 3 x12x =13x4x12x6,---3,744.

4

2/ i

,3\ 14\

**Hands with Two Pairs:** in sequence we need to **choose the values of the two** pairs.

(4

choices; the suits of the first pair, choices; the suits of the second pair. 1

2i i\ 2) v-? i
  
choices; and the remaining card, 44 (52 —4 x 2, since the last cards can. not have the same value as either pair) choices. **So the number** of hands with two pairs is

rin **,4 '4.**

2 . (2 **\2./**

K X x4=78x6x6x44,--123552.

**To** calculate the probability of each, we only need to divide the number of hands of each **kind by the total** possible number of hands,

**Hopping rabbit**

A rabbit sits at the bottom of a staircase with *n* stairs. The rabbit can hop up only one ortwo stairs at a tilne, How many different ways are there for the rabbit to ascend to **top of the** stairs?'

I-lint: Consider an induction approach. Before the final hop to reach the n-th stir,either, the (n-l)th stair or the (n-2)th stair assuming > ). the rabbit can *be al*

*Solution:* Let's begin with the simplest cases and consider solving the problem for any number of stairs using induction. For ii =1 , there is only one way and *f (1) =* 1. For *n = 2,* we can have one 2-stair hop or two 1-stair hops. So .f(2) = 2. For any *n >* 2, there are always two possibilities for the last hop, either it's a 1-stair hop or a 2-stair bop. In the former case, the rabbit is at *(n-1)* before reaching *n,* and it has *f(n —1)* ways to reach *(n-1).* In the latter case, the rabbit is at *(n-2)* before reaching *n,* and it has *f(n* — 2) ways to reach (n 2). So we have *f (n) = f (n — 2) + f (n* —1). Using this function we can calculate f *(n)* for *n --* 3, 4, • • 6

**Screwy pirates 2**

Having peacefully divided the loot (in chapter 2), the pirate team goes on for more looting and expands the group to II pirates. To protect their hard-won treasure, they gather together to put all the loot in a safe. Still being a democratic bunch, they decide that only a majority — any majority — of them (>6) together can open the safe. So they ask a locksmith to put a certain number of locks on the safe. To access the treasure, every lock needs to be opened. Each lock can have multiple keys; but each key only opens one lock. The locksmith can give more than one key to each pirate.

What is the smallest number of locks needed? And how many keys must each pirate carry?7

*Solution:* This problem is a good example of the application of combinatorial analysis in information sharing **and** cryptography. A general version of the problem was explained in a 1979 paper *"How to Share a Secret"* by Adi Shamir. Let's randomly select 5 pirates from the 11-member group; there must be a lock that none of them has the key to. Yet any of the other 6 pirates must have the key to this lock since any 6 pirates can open all locks. In other words, we must have a -special" lock to which none of the 5 selected pirates has a key and the other 6 pirates all have keys. Such 5-pirate groups are randomly selected. So for each combination of 5 pirates, there must be such a "special" lock. The minimum number of locks needed is ill) =1! = 462 locks. Each lock has 6 keys,

5 ) 5!6?

which are given to a unique 6-member subgroup. So each pirate must have

462x6

= 252 keys. That's surely a lot of locks to put on a **safe and a lot of keys for** each pirate to carry.

y*ou* may have recognized that the sequence is a sequence of Fibonacci numbers.

7 Hint; every subgroup of 6 pirates should have the same key to a unique lock that the other 5 pirates do

not have.

66 67

**Chess tournament**

Probability Theory A Practical Guide To Quantitative Finance lnterviews

A chess tournament has 2" players with skills 1 > 2 > >211. It is organized as a

knockout tournament, so that after each round only the winner proceeds to the next round. Except for the final, opponents in each round are drawn at random. Let's also assume that when two players meet in a game, the player with better skills always wins. What's the probability that players 1 and 2 will meet in the Ima178

*Solution:* There are at least two approaches to solve the problem. The standard approach applies multiplication rule based on conditional probability, while a counting approach is far more efficient. (We will cover conditional probability in detail in the next section.) Let's begin with the conditional probability approach, which is easier to grasp. Since there are 2" players. the tournament will have *n* rounds (including the final). For round 1, players 2,3,• 2" 1 each have probability to be 1's rival, so the probability that

2" —1

1 and 2 do not meet in round 1 is —" 2 — 2 x (2" —1) . Condition on that 1 and 2 do not 2" —1 *?n* \_1

meet in round 1. 2" •1 players proceed to the 2nd round and the conditional probability

that I and 2 will not meet in round 2 is 2:1—\_21 — 2 x (2'2 —1). We can repeat the same

2

2"-' —1

process until the *(n-1)111* round, in which there are 22 *(= r 12'2)* players left and the conditional probability that 1 and 2 will not meet inround *(n —I)* is 2' —2 2A(22 --1)

2' —1 )2

Let he the event that 1 and 2 do not meet in round 1;

*E.* be the event that I and 2 do not meet in rounds 1 and 2;

he the event that 1 and 2 do not meet in round *n —1*

Apply the multiplication rule. we have

PO and 2 meet in the nth game) )x *P(E,* )x • • - x *E* E2 • • E

2x(2' • 1) 2x(2' —1) 2 x (2' -- 1) 2"

--X ----T--------X•••X

*2''* —1 22 — 1 2" — 1

Hint: Consider separating the players to two 2" subgroups. What will happen if player

'ante group? Or not in the same group? I and 2 ill Ole

Now let's move on to the counting approach. Figure 4.2A is the general case of what happens in the final. Player 1 always wins, so he will be in the final. From the figure, it is obvious that 2" players are separated to two 2"-1-player subgroups and each group will have one player reaching the final. As shown in Figure 4.2B, for player 2 to reach the final, he/she must be in a different subgroup from I . Since any of the remaining players in 2,3,-2" are likely to be one of the (2"-' —1) players in the same subgroup as player 1 or one of the 2"-' players in the subgroup different from player 1, the probability that 2 is in a different subgroup from 1 and that 1 and 2 will meet in the final

2'1

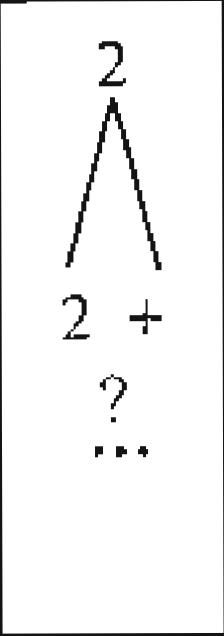
()8 69

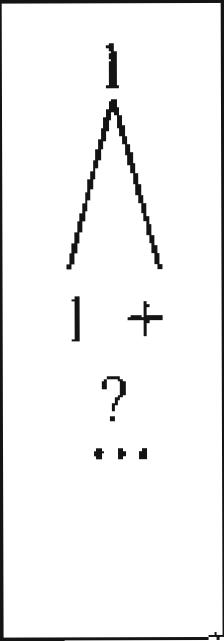
is simply err . Clearly, the counting approach provides not only a simpler solution but

also more .insight to the problem.

General Case 1 & 2 in the Final

1 1





|  |  |  |
| --- | --- | --- |
|  |  |  |
| 1 +  7 |  |
|  |  |  |

nth round

(n-1)th round

nth round

(n-1 )th round

2111 players 2n-1 players

2°.' players 2n-' players
  
B

A

Figure 4.2A The general case of separating 2" players into 2n-1-player subgroups; 4.2B The special case with players 1 and 2 in different groups

**Application letters**

You're sending job applications to 5 firms: Morgan Stanley, Lehman Brothers, UBS, Goldman Sachs, and Merrill Lynch. You have 5 envelopes on the table neatly typed with names and addresses of people at these 5 firms. You even have 5 cover letters personalized to each of these firms. Your 3-year-old tried to be helpful and stuffed each cover letter into each of the envelopes for you. Unfortunately she randomly put letters



*Solution.* This problem is a classic example for the Inclusion-Exclusion Principle. In fact, a more general case is an example in Ross' textbook *First Course in Probability.*

Let's denote by E,. i *=1,•* •-..5 the event that the i-th letter has the correct envelope. Then

*(c*

*P uft,* is the probability that at least one letter has the correct envelope and

1 i

i , ••,

1-- PuE is the probability that all letters have the wrong envelopes,
  
 i\_i ..,,

be calculated using the Inclusion-Exclusion Principle:

P(5 ) 5
  
uE, 1-, *ZP(E,)–ZP(E,jE) + • •- + (-1)6 P(E,* Ey - E5) '1<1

1

It's obvious that *P(Er) =* 5 *–, i =1,• • • ,* So I *P(Ei)* =1

*P(E,,E,,)* is *the* event that both letter and letter *i,* have the

probability that i, has the correct envelope is 1/5; Conditioned on that

envelope, the probability that i, has the correct envelope envelopes left). So *P(E, Er )* 1\_ 5 I (5 – 2)!

5–I

(5)

There are 1 5 !

2!(5 –

members of *P(E ) ,* so we have

J

(5– 2)! 5! 1

*P(E, x ,*

5! 7!(5-2)! 2!

rn:.ft complement is that at least one letter is mailed to the correct firm.

|  |  |
| --- | --- |
| (5 \  So the probability that all 5 letters are mailed to the wrong firms is 1-- *p* uE,  ***r.l j*** | 11  \_30*•* |

**Birthday problem**

How many people do we need in a class to make the probability that two people have the same birthday more than 1/2? (For simplicity, assume 365 days a year.)

*Solution:* The number is surprisingly small: 23. Lets say we have n people in the class. Without any restrictions, we have 365 possibilities for each individual's birthday. The basic principle of counting tells us that there are 365' possible sequences.

We want to find the number of those sequences that have no duplication of birthdays. For the first individual, we can choose any of the 365 days; but for the second, only 364 remaining choices left, ..., for the *rth* individual, there are 365 – r +1 choices. So for *n* people there are 365x 364 x • - -x (365– n +1) possible sequences where no two

365 x 364 x • x (365 **— 1)**

individuals have the same birthday. We need to have <1/2

365'

for the odds to be in our favor. The smallest such *n* is 23.

**100th digit**

What is the 100th digit to the right of the decimal point in the decimal representation of (1 + -5)300o

*Solution:* If you still have not figure out the solution from the hint, here is one more hint: (1 + *-,Er +0* is an integer when *n =* 3000 .

Applying the binomial theorem for (x + y)" , we have

***A [ \***

***(I -I- 4***

***=*** E - 1"-kT2k

k-O *ki*

`.■-■ -■-■■-■=MT

tint: (1 - *)2 ± (1 \_ i•* 6 . What will happen to (1— NEP' as ra becomes large?

5 \

*p uE.* 1 can \ I .1

correct envelope. "rile

iI has the correct

is 1/ 4 (there are only Lai

5!

Similarly we have E *P(E, E, E.* 1 *, P(E E,* and

*t*

*1)(E, ' 1-51*

4!

|  |  |
| --- | --- |
| *nji,,,*  **--2,..11-'41S7,** |  |

70

71

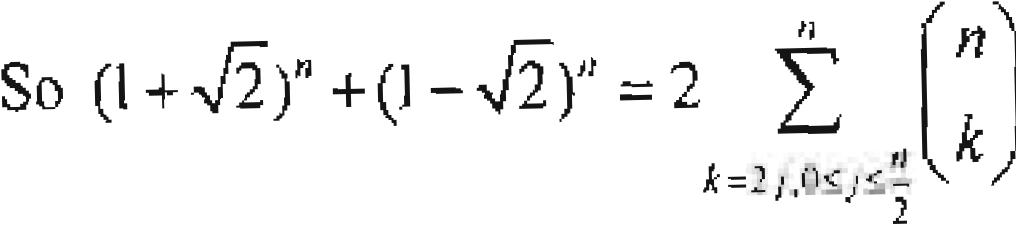
|  |  |  |  |
| --- | --- | --- | --- |
| Probability Theory  4  into envelopes without realizing that the letters are personalized. What is the probability that all 5 cover letters are mailed to the wrong .firms?9 | *P(s L\_\_)E* | 1 1 1 1 19  2! 3! 4! 5! 30 | A Practical Guide To Quantitative Finance Interviews |

Probability Theory A Practical Guide To Quantitative Finance Interviews

|  |  |  |  |
| --- | --- | --- | --- |
| i  k(n)1"-4(-4-21. k |  | A *n -k*  2 E -Nt2  *\,k)*  *f‹..;* | ***Law of total probability:*** for any mutually exclusive events {F,I, i =1,2,• • *-,n ,* whose  *fa* |
|  |
| 2 | union is the entire sample space *(F, nF1=c1),Vi# j; U F 0,),* we have |

r.1

*1"-k* which is always an integer. It is easy to



-\

*P(E) P(EF,) P(EF2)+ P(EF„) = iP(E F:)P(F,)*

1.1

see that 0 <(1-,12) << 10'. So the 100th digit of (1 + NT2,)" must be 9.

**Cubic of integer**

Let x be an integer between 1 and 1012, what is the probability that the cubic of x ends with 11?11

*Solution:* All integers can be expressed as x = c *+ 1 Ob,* where *a* is the last digit of x. Applying the binomial theorem, we have x3 *+10b)3 = a3 +30a2b* 300ab2 +1000/,'.
  
The unit digit of *x3* only depends on *a:.* So a3 has a unit digit of 1. Only *a = 1* satisfies

this requirement and *ti3 =1.* Since a3 =1, the tenth digit only depends on *30a2b*

So we must have that 3b ends in 1, which requires the last digit of *b* to be

Consequently, the last two digits of x should be 71, which has a probability of 1% for integers between 1 and 1012.

***4.3 Conditional Probability and Bayes' formula***

Many financial transactions an responses to probability adjustments based on new—and most likely incomplete—information. Conditional probability surely is one of the most popular test subjects in quantitative interviews. So in this section, we focus on basic conditional probability definitions and theorems.

**Conditional probability *P(A I*** *B) :* If *P(B) >* 0, then *P(Al B) =* P(AB)

of *B outcm P(B)*

that are also .4 outcomes. is the fraction

*Rd t: P(EE, P(EI)P(E2 E[)P(Ei F* ')'

*P(E I FONFI)+ P(E I F2)P(F2) + r,i)P(F„)*

**Independent events:** *P(EF) = P(E)P(F) P(Er)= P(E)P(Fc ) .*

1 lint: last two digits of x3 only depend on the last two digits of r.

72 73

Independence is a symmetric relation: X is independent of Y <:=> Y is independent of X.
  
*P(E LFi)P(F I) if*

**Sayer Formula:** *P(FJ 1 E) =* rr

*n,* are mutually

*P(E 1 Fi.)P(F,)*

exclusive events whose union is the entire sample space.

As the following examples will demonstrate, not all conditional probability problems have intuitive solutions. Many demand logical analysis instead.

**Boys and girls**

Part A. A company is holding a dinner for working mothers with at least one son. Ms. Jackson, mother thser :4 c r

hi

invited. p - ted. What is the rrobability that both

children are *So/ution:* The sample space of two children is given by c2= *{(b•b),(h,g),(g,b),(g, g)},* (e *(g,b)* means the older child is a girl and the younger child a boy), and each outcome has the same probability. Since Ms. Jackson is invited, she has at least one son. eLheitld B children eeatrhebeovyesn e have

one of the children is a boy and A be the event that both

*P(A (Th B)*

*P(B) P({(b,b),(b,g),(g,b)})* 314 3

Part *B.* your new colleague, Ms. Parker is known to have two children. If you see her walking with one of her children and that child is a boy, what is the probability that both children are boys?

*Solution:* the other child is equally likely to be a boy or a girl (independent of the boy you've seen), so the probability that both children are boys is 1/2.

Probability Theory A Practical Guide To Quantitative Finance Interviews

Notice the subtle difference between part *A* and part *B.* In part *A.* the problem essentially asks given there is at least one boy in two children, what is the conditional probability that both children are boys. Part *B* asks that given one child is a boy, what is the conditional probability that the other child is also a boy. For both parts. we need to assume that each child is equal likely to be a boy or a girl.

**All-girl world?**

In a primitive society, every couple prefers to have a baby girl. There is a 50% chance that each child they have is a girl, and the genders of their children are mutually independent. If each couple insists on having more children until they get a girl and once they have a girl they will stop having more children, what will eventually happen to the fraction of girls in this society?

*Solution:* h was surprising that many interviewees—include many who studied probability—have the misconception that there will be more girls. Do not let the word "prefer" **and** a wrong intuition misguide you. The fraction of baby girls are driven by, nature, or at least the *X* and Y chromosomes, not by the couples' preference. You only need to look at the key information: 50% and independence, Every new-born child has equal probability of being a boy or a girl regardless of the gender of any other children. So the fraction of girls born is always 50% and the fractions of girls in the society will stay stable at 50%.

**Unfair coin**

You are given 1000 coins. Among them, **I** coin has heads on both sides. The other 999 coins are fair coins. You randomly choose a coin and toss it 10 times. Each time..,;he coin turns up heads. What is the probability that the coin you choose is the unfair one,

*Solution:* This is a classic conditional probability question that uses Bayes' theorem, Let

chosen.

*A* be the event that the chosen coin is the unfair one, then Ac is the event that the

coin is a lair one. Let *13* be the event that all ten tosses turn up heads. Apply F3aYes

*P(B A)P(A) P(B A)P(A)*

theorem vse have P(.-11B) =

***P(B) P(B1*** *A)P(.4)+ P(B: A'* )P(Ac )

The priors are P(A) -=1/1000 and *P(A`)* 999/1000. If the coin is unfair, it alwaYs
  
turns up heads. so *MI A)* **I. If** the coin is fair each time it has 1/2 probability turning

up heads. So *PO I = (1 12)r° =1*11024. Plug in all the available information and we

have the answer:

74 75

*P(B I A)P(A) 111000x1*

0

*P(A I B)* .

*.4)P(A)+ P(B I Ar)P(Ac) 111000x1+999/1000x111024*

**Fair probability from an unfair coin**

If you have an unfair coin, which may bias toward either heads or tails at an unknown probability, can you generate even odds using this coin?

*Solution:* Unlike fair coins, we clearly can not generate even odds with one toss using an unfair coin. How about using 2 tosses? Let *p1* be the probability the coin will yield head, and Pr be the probability the coin will yield tails ( *PH +* Pr Consider two independent tosses. We have four possible outcomes *TN, HT, TH* and *TT* with probabilities *P(Hil)*

***= A D*** *H.****D***

*H P(HT) = pi, pr, PITH) = pr. pH,* and *P(TT) = p rp7. .*

So we have *P(I-IT)= P(TH).* By assigning *HT* to winning and *TH* to losing, we can generate even odds.12 0 ki o i e,1 r. •

**Dart game**

Jason throws two darts **at** a dartboard, aiming for the center. The second dart lands
  
farther from the center than the first. If Jason throws **a** third dart aiming for the center,
  
what is the probability that the third throw is farther from the center than the first?

Assume Jason's skillfulness is constant.

*Solution:* A standard answer directly applies the conditional probability by enumerating all possible outcomes. If we rank the three darts' results from the best (A) to the worst (C), there are 6 possible outcomes with equal probability:

**12**

should point out that this simple approach is not the most efficient approach since I am disregarding the cases 1-11-i and TT. When the coin has high bias (one side is far more likely than the other side to occur), the method may take many runs to generate one useful result. For more complex algorithm that increasing efficiency, please refer to *Tree Algorithm s for I-Viral...v.1' Tqssing* by Quentin F. Stout and Bette I... Warren. Annals of Probability 11: (I 90 I. pp. 212-222

Probability Theory

A Practical Guide To Quantitative Finance Interviews

don't know anyone else's birthday and all birthdays are distributed randomly throughout the year (assuming 365 days in a year), what position in line gives you the largest chance of getting the free ticket?'

*Sohetion:* If you have solved the problem that no two people have the same birthday in an n-people group, this new problem is just a small extension. Assume that you choose to be the n-th person in line. In order for you to get the free ticket, all of the first n-1 individuals in line must have different birthdays and your birthday needs to be the same as one of those *n -1* individuals.

*p(n) = p(first n -1 people have no same birthday)x p(yours among those n -1 birthdays)* 365 x 364 x • • -(365 - n 2) *xn -1*

365"' 365

It is intuitive to argue that when *n* is small, increasing *n* will increase your chance of getting the free ticket since the increase of *p(yours among those n -1 birthdays)* is more significant than the decrease in *p(firsi n-1 people have no scone birthday).* So when *n* is small, we have *P(n) > P(n -1).* As *n* increases, gradually the negative impact of *p(first n -1 people haven° same birthday)* will catch up and at a certain point we wanidll P(n) >

have + .

(n(+1),;P(n). So we need to find such an *n* that satisfies *P(n) > P(n -1)*

*po \_1) =* 365 . \_364 x 365 - *(n xn*-2

365 365 365 365

**P** 365 364 365-(n-2).n-1

*P(n* , i) —.365 x 365

365 364 365 - *(n* - 2) .365 - *(n 1) n*

*(n-')) n -1 n*

*P(n)> P(n -1)*

x

365 365 365 n -3n -363 < 0

*n* > 0

***—***1 365 *-in -n n n2 — -365*

You should be the 20th person in line.

.

Flint: If you are the to-th person in Line, to get the free ticket, the first *0-I)* people in line must not have the same birthday, and you must have the same birthday as one of them.

77

365 365

*— x •* x65

* ---i 365
    
  -

•

n = 20

Hence,

*P(n) > P(n +1) > x*

365 365 365

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Outcome | 1 | 2 | 3 | 4 | 5 | 6 |
| **1**st throw | A | B | A | C | B | C |
| 2nd throw | B | A | C | A | C | B |
| 3rd throw |  | CC |  | BB | A | A |

The information from the first two throws eliminates outcomes 2, 4 and 6. Conditioned on outcomes **1,** 3, and 5, the outcomes that the 3rd throw is worse than the **1st** throw are outcomes 1 and 3. So there is 2/3 probability that the third throw is farther from the center than the first.

This approach surely is reasonable. Nevertheless, it is not an efficient approach. When the number of darts is small, we can easily enumerate all outcomes. What if it is a more complex version of the original problem:

Jason throws *n (n >* 5) darts **at** a dartboard, aiming for the center. Each subsequent dart

is faahet from the center than the first.dart. If Jason throws the *(n 4* 1)th dart. whir( k the

. \_

probability that it is also fartho from the center than

This question is equivalent **to** a simple question: what is the probability that the (n-t-Oth throw is not the best among *all (n +1)* throws? Since the 1st throw is the best among the first *n* throws, essentially I am saving the event that the *(n* -i-l)th throw is the best of all *(n + 1)* throws (let's call it An\_,, ) is independent of the event that the 1st throw is the best of the first u throws (let's call it A, ). In fact, An+, is independent of the order of the first *n* throws. Are these two events really independent? The answer is a resounding yes. If it

is not obvious to you that *A„..I* is independent of the order of the first *n* throws, let's look at it another way: the order of the first n throws is independent of An+, . Surely this claim is conspicuous. But independence is symmetric! Since the probability of A,;+1 is

*11(n +1),* the probability that *(n + 1)th* throw is not the best is *n 1(n +* 1).13

For the original version. three darts are thrown independently, they each have a 13e chance of being the best throw. As long as the third dart is not the best throw, it will . worse than the first dart. Therefore the answer is 2/3.

**Birthday line**

**.:\t** a movie theater. a whimsical manager announces that she will give a free ticket to the
  
a
  
first person in line whose birthday is the same as someone who has already bough.t,u

ticket. You are given the opportunity to choose any position in line. Assuming that yc

Fere v can again use symmetry argument: each throw is equally likely to be the best.

76

**Dice order**

Probability Theory A Practical Guide To Quantitative Finance interviews

's 1-lint: To obtain 3 points in strictly increasing order. the 3 points tnust be different. I

78 'or 3 different points

in a sequence, strictly increasing order is one of the possible permutations.

79

Hint: tf there are at least I blue candy and I green candy lett the last red candy must have been

16

removed before the last blue candy and the last green candy in the sequence of 60 candies. What is the

probability that the blue candy is the last one in the 60-candy sequence? Conditioned on that, what is the
  
probability that the last green candy is the last one in the 30-candy sequence ( 10 red, 20 g,r(Tn)? What if

the green can is the last one in the 60-candy sequence?r

We throw 3 dice one by one. What is the probability that we obtain 3 points in strictly increasing order?'

*Sohilion:* To have 3 points in strictly increasing order, first all three points must be different numbers. Conditioned on three different numbers, the probability of strictly increasing order is simply 1/3! = 1 /6 (one specific sequence out of all possible permutations). So we have

P = P(diffcrent numbers in all three throws) x P(increasing order13 different numbers) --(1x(ix!)>+=5/54

**Monty Hall problem**

Monty Hall problem is a probability puzzle based on an old American show *Lei is Make a Deal .* The problem is named after the show's host. Suppose you're on the show now, and you're given the choice of 3 doors. Behind one door is a car; behind the other two, goats. You don't know ahead of time what is behind each of the doors.

You pick one of the doors and announce it. As soon as you pick the door, Monty opens one of the other two doors that he knows has a goat behind it. Then he gives you the option to either keep your original choice or switch to the third door. Should you switch? What is the probability of winning a car if you switch?

*Solid ion:* If you don't switch, whether you win or not is independent of Monty's action of showing you a goat, so your probability of winning is 1/3. What if you switch? Many Nvould argue that since there are only two doors left after Monty shows a door with goat. the probability of winning is 1/2. But is this argument correct?

lf you look at the problem from a different perspective, the answer becomes clear. Using a switching strategy, you win the car if and only if you originally pick a door with a goat. which has a probability of 2/3 (You pick a door with a goat, Monty shows a door With another goat, so the one you switch to must have a car behind it). If you originally, picked the door with the car, which has a probability of 1/3, you will lose by switching. So your probability of winning by switching is actually 2/3.

**Amoeba population**

There is a one amoeba in a pond. After every minute the\_ amoeba may die. stay the same, split into two or split into three with equal probability:ji its offspring, if it has any, will behave the same (and independent of other amoebas). What is the probability the amoeba population will die out?

*Solution:* This is just another standard conditional probability problem once you realize
  
we need to derive the probability conditioned on what happens to the amoeba one
  
minute later. Let *P(E)* be the probability that the amoeba population will die out and

apply the law of total probability conditioned on what happens to the amoeba one minute later:

*P(E) = P(E FOP (F1) P(E I F2)P(F,)+ • P(E I F„)P(F„)*

For the original amoeba, as stated in the question, there are four possible mutually
  
exclusive events each with probability 1/4. Let's denote *F,* as the event the amoeba dies;

F2 as the event that it. stays the same; F3 as the event that it splits into two; *€4* as the event that it splits into three. For event *F, P(E I F) =* 1 since no amoeba is left. P(E1 F2) = *P(E)* since the state is the same as the beginning. For *F1,* there are two

amoebas; either behaves the same as the original one. The total amoeba population will
  
die only if both amoebas die out. Since they are independent, the probability that they

both will die out is *P(E)2* Similarly we have *P(F4)* = *P(E)'.* Plug in all the numbers, the equation becomes *P(E).* 1/ 4x1+1/ 4x *P(E) +1* / 4 x *P(E)2* +1/ 4x *P(E)3.* Solve this equation with the restriction 0 < *P( E) <* I, and we will get *P(E) = \12. —)* 0.414 (The other two roots of the equation are 1 and —1).

**Candies in a jar**

You are taking out candies one by one from a jar that has 10 red candies, 20 blue candies, and 30 green candies in it. What is the probability that there are at least I blue candy and

1 green candy left in the jar when you have taken out all the red candies.Solution: At first look, this problem appears to be a combinatorial one. However, a conditional probability approach gives a much more intuitive answer. Let *T,. Th* and *T,*

Probability Theory



be the rAttrnberthat the last red, blue, and green candies are taken out regectivel. To have at least I blue candy and 1 green candy left when all the rest candies are taken out, we need to have *Tr <7;* and *Tr <Tg.* In other words, we want to derive *P(7; <T, (mT,.C Tg).* There are two mutually exclusive events that satisfy *Tr <* 7j, and *Tr CT:* 7;, and *Tr <7.g <Tb.*

*P(Tr <* 7;, ra'r < *= P(T,. < Ti, <Tg)+ P(T, <7; <TO*

*Tr. < <T,* means that the last candy is green (1-, -= 60) Since each of the 60 candies

are equally likely to be the last candy and among them 30 are green ones, we have

*P(7:* -- 60) —3. Conditioned on T*=* 60, we need *P(7:* < I 7i, 60). Among the 30 60

red and blue candies, each candy is again equally likely to be the last candy and there are

*< < T ) = —30*x-20.

20 blue candies, so *P(T, <T,iT =* 60) = —20 and *P(T,*

30

we have *13(7;T <T* 20 30x**g**

60 40

Hence,

g 60 30

|  |  |  |  |
| --- | --- | --- | --- |
| 20  60 | x | 30  40 | 7  12 |

*P(T,*

3 20

< *()Tr* < *P(T,* < *P(T, <1 )*

*; <1=-'*0 +

60 x 30

**Coin toss game**

Two players, *A* and *B,* alternatively toss a fair coin *(A* tosses the coin first, then *B* tosses the coin, then A, then *B...).* The sequence of heads and tails is recorded. If there is a head followed by a tail *(HT* subsequence), the game ends and the person who tosses the tail wins. What is the probability that A wins the ganne?17

*Solution:* Let P(A) be the probability that *A* wins; then the probability that *B wills is NB),* –1'(4). Let's condition 1)( A) on *A's* first toss, which has 10 probability °In (heads) and 2 probability of *T* (tails).

P(t1) =112P(/11 *11)÷112P(A T)*

*If r* first Fos:. i:; T. il-!en B essentially becomes, the first to toss (An *H* is required for the II I sag iequenee), So we have *P(A1T) PO) –*

*If A's first* toss ends in *11,* lees further condition on B's first toss, *B* has 1/2 probabiiiiti..vv of getting. *i;* in that case .4 loses, For the 1// probability that *B* gets *11,*

Hiat: condition **oni** the result of A's first toss and use symmetry. *essentia*

80

A Practical Guide To Quantitative Finance InteNiews

becomes the first one to toss an *H.* In that case, *A* has (1– *P(A1H))* probability of winning. So *P(Al H* ) =1 / 2 x 0 +1/2(1– *POI H)) H)* =1 / 3

Combining all the available information, we have

P(A)-112x113+1/2(1–F(A)) P(A)=419.

Sanity check: we can see that *PO) < I 12* , which is reasonable since A cannot win in his first toss, yet *B* has 114 probability to win in her first toss.

**Russian roulette series**

Let's play a traditional version of Russian roulette. A single bullet is put into a 6-chamber revolver. The barrel is randomly spun so that each chamber is equally likely to be under the hammer. Two players take turns to pull the trigger—with the gun unfortunately pointing at one's own head. -without further spinning until the gun goes off and the person who gets killed loses. If you5E1e, (Tf-Th'e players, **C;111** choose to go -first or second, how will you choose? And what is your probability of loss?

*Solution:* Marty people have the wrong impression that the first person has higher probability of loss. After all, the first player has a 1/6 chance of getting killed in the first round before the second player starts. Unfortunately, this is one of the few times that intuition is wrong. Once the barrel is spun, the position of the bullet is fixed. If you go first, you lose if and only if the bullet is in chamber 1, 3 and 5. So the probability that you lose is the same as the second player, 112. In that sense, whether to go first or seco does not matter.nd Now, let's change the rule slightly. We will spin the barrel again after every trigger pull. Will you choose to be the first or the second player'? And what is your probability of loss? *So/II/ion:* The difference is that each run now becomes independent. Assume that the first player's probability of losing is *p,* then the second player's probability of losing is 1– *p.* Let's condition the probability on the first person's first trigger pull. He has 1/6 P of losing in this run. Otherwise, he essentially becomes the second player in the game with new (conditional) probability of losing 1– *p.* That happens with

6/11. So you should

probability 516. T *)(*

hat gives us *p =11' 6 4-0 p)X* 5/6 *P*

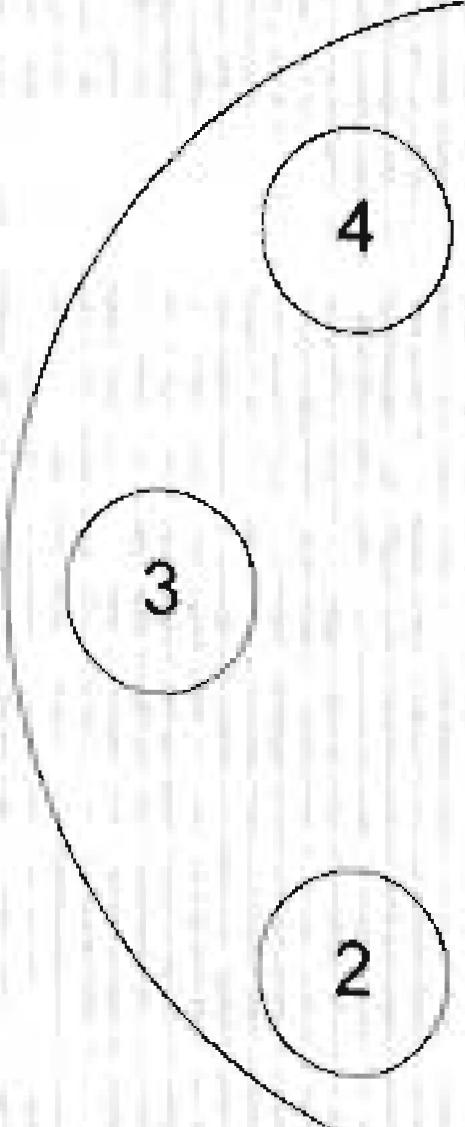
*=* choose to be the second player and have 5/11 probability of losing.

If instead of one bullet, two bullets are randomly put in the chamber. Your opponent played the first and he was alive after the first trigger pull. You are given the option whether to spin the barrel. Should you spin the barrel?

81

*Soitition:* if you spin the barrel, the probability that you will lose in this round is 2/6.1i you don't spin the barrel, there are only 5 chambers left and your probability of losing in this round (conditioned on that your opponent survived) is 2/5. So you should spin the barrel.

Probability Theory A Practical Guide To Quantitative Finance Interviews



What if the two bullets are randomly put in two consecutive positions? If your opponent survived his first round, should you spin the barrel?

*Sol talon:* Now we have to condition our probability on the fact that the positions of the two bullets are consecutive. As shown in Figure 4.3, let's label the empty chambers as 1, 2, 3 and 4; label the ones with bullets 5 and 6. Since your opponent survived the first round, the possible position he encountered is 1, 2, 3 or 4 with equal probability. With 1/4 chance, the next one is a bullet (the position was 4). So if you don't spin, the chance of survival is 3/4. If you spin the barrel, each position has equal probability of being chosen, and your chance of survival is only 2/3. So you should not spin the barrel.

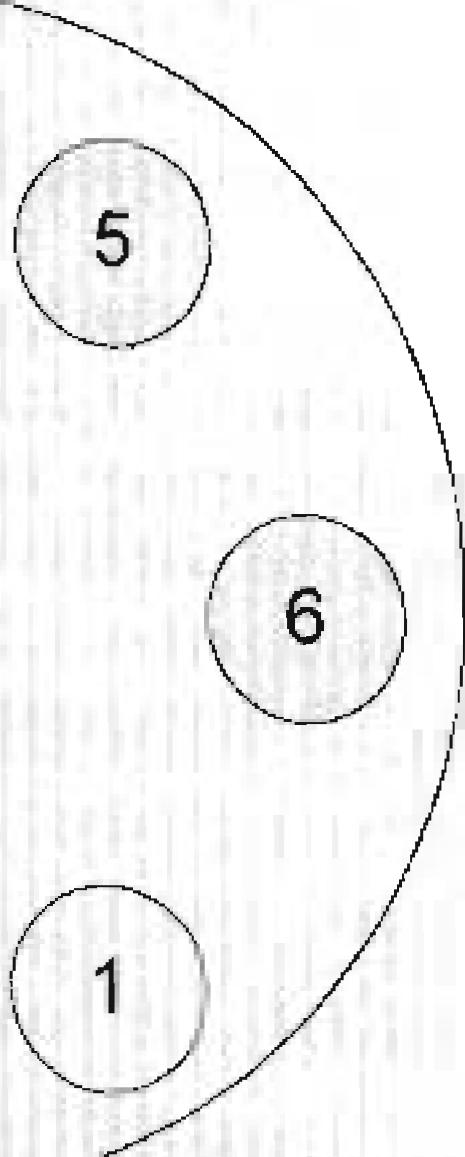


Figure 4.3 Russian roulette with two consecutive bullets.

**Aces**

Fifty-two curds arc randomly distributed to 4p

lavers with each player oetting 13 cards. What is the probability that each of them will have an ace?

*Sobitirm:* The problem can be answered using standard counting methods. To distribute 52 cards to 4 players with 13 cards each has 52! permutations. If each Piga

13!13!13!13!

82

needs to have one ace, we can distribute the aces first, which has 4! ways. Then we

48!

distribute the rest 48 cards to 4 players with 12 cards each, which has permutations. So the probability that each of them will have an Ace is

12!12!12!12!

48! 52! 52 39 26 13

4i x x — x —

12!12!12!12! 13113!13!13! 52 51 50 49

The logic becomes clearer if we use a conditional probability approach. Let's begin with any one of the four aces; it has probability 52/52 =1 of belonging to a pile, The second ace can be any of the remaining 51 cards, among which 39 belong to a pile different from the first ace. So the probability that the second ace is not in the pile of the first ace is 39 /51. Now there are 50 cards left, among which 26 belong to the other two piles. So the conditional probability that the third ace is in one of the other 2 piles given the first two aces are already in different piles is 26/50. Similarly, the conditional probability that the fourth ace is in the pile different from the first three aces given that the first three aces are in different piles is 13/49 . So the probability that each pile has an ace is 1x39 26 13

—x-----x-
  
51 50 49

**Gambler's ruin problem**

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

*Solution:* This is a classic textbook probability problem called the Gambler's Ruin-Problem. Interestingly, it is still widely used in quantitative interviews.

**•**

From any initial state *i* (the dollars the gambler has), 0 *N* let *p* be the probability

that the gambler's fortune will reach N instead of 0\_ The next state is either *i* +1 with probability *p* or *i* –1 with probability *q.* So we have

*PP,,,* + *17,1 P, =1(P, – P, (P, – Pr-2 ) = 5- (PI– Po )*

*p P ) P*

We also have the boundary probabilities *Po =* 0 and p, =1.

So starting from *P, ,* we can successively evaluate *P,* as an expression of P*,*

83

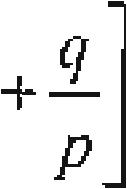
Probability Theory

A Practical Guide To Quantitative Finance Interviews

.P((4,1)j(3,1))xP1+P((4,1)1(3,2))xP3,=-2 x1-1-0x 1 =1

*= P2+ (IP0 =• P2=1 P,*

|  |  |
| --- | --- |
|  | *+9-411p*  *P* |



P,t)
  
P4,2
  
.P41

3 2 2 3

.N(4,2) (3,1))xP3,1+P((4,2) (3,2))xP„ .1x1+ 1 x 1 =1

3 2 3 2 3

P((4,3) (3,1)) x P3) +P(0,3)1 (3,2)) x *P*•

*2 =0 x + 2 X1 =1* 2 3 2 3

The results indicate that *P,, k =* Vk *n —1 ,* and give the hint that the law of

*P,* +•-- +

\ *P* )

*n-1*

total probability can be used in the induction step.

Induction 1

Extending this expression to *P.* we have

step: given that *Po = Vk =1,* 2,• • • , *n —1,* we need to prove

*n —1*

*n.* To show

it, simply apply the law of total

*if (11 p -1*

*—(ei py Pp f p* 1/2

1 *— y I pyv*

*iIN, if p =1/2*

*(*

*(q I py'r*

*ql p#1*

*P*

*1—q1 p*

Px

*[*

*{l—qIP*

pi *if. qlpi*

,\_ 1 *— (q I pYv ' P, =-*

*11 N , if q I p -1*

k =(12 + 1) — 1 *n*

*k* =1, 2,

1

probability;

*k*

PA+1k = *P(miss (n,k))P„k P(scorel(n,*

1—*k 1 k —1 1—*

*n n n n-1 n*

**Basketball scores**

A basketball player is taking 100 free throws. She scores one point if the ball passes, through the hoop and zero point if she misses. She has scored on her first throw aii.d, missed on her second. For each of the following the

the probability of her scoring l the fraction of throws she has made so far. For example, if she has scored 23 points atiel the 40th throw, the probability that she will score in the 41th throw is 23/40. After 100 50 baskets?'8

throws (including the first and the second), what is the probability that she •scores exactly *Soh/lion:* I.et (n, k}. *I 15 k < n ,* be the event that the player scores *k* baskets after 11 throws and *1),,. , P ((n.k)). The sotution is* surprisingly simple if we use an induction approach starting with *n -* 3. The third throw has 1/2 probability of scoring. So we have, P., Lf2 and P. .1/2. For

probahilitv the ease when *n =* 4, let's apply the law of total

-.

' !lint; Again, do not let the nturiber 100 scares you. Start with smallest n, solve the pro Men: trY I° rind a pallernb incrousing ***07;*** and prove the pattern using induction.

84

The equation is also applicable to the and *—1* although in these cases = 0

*rr*

85

and

(1-- *=* 0, respectively. So we *17*

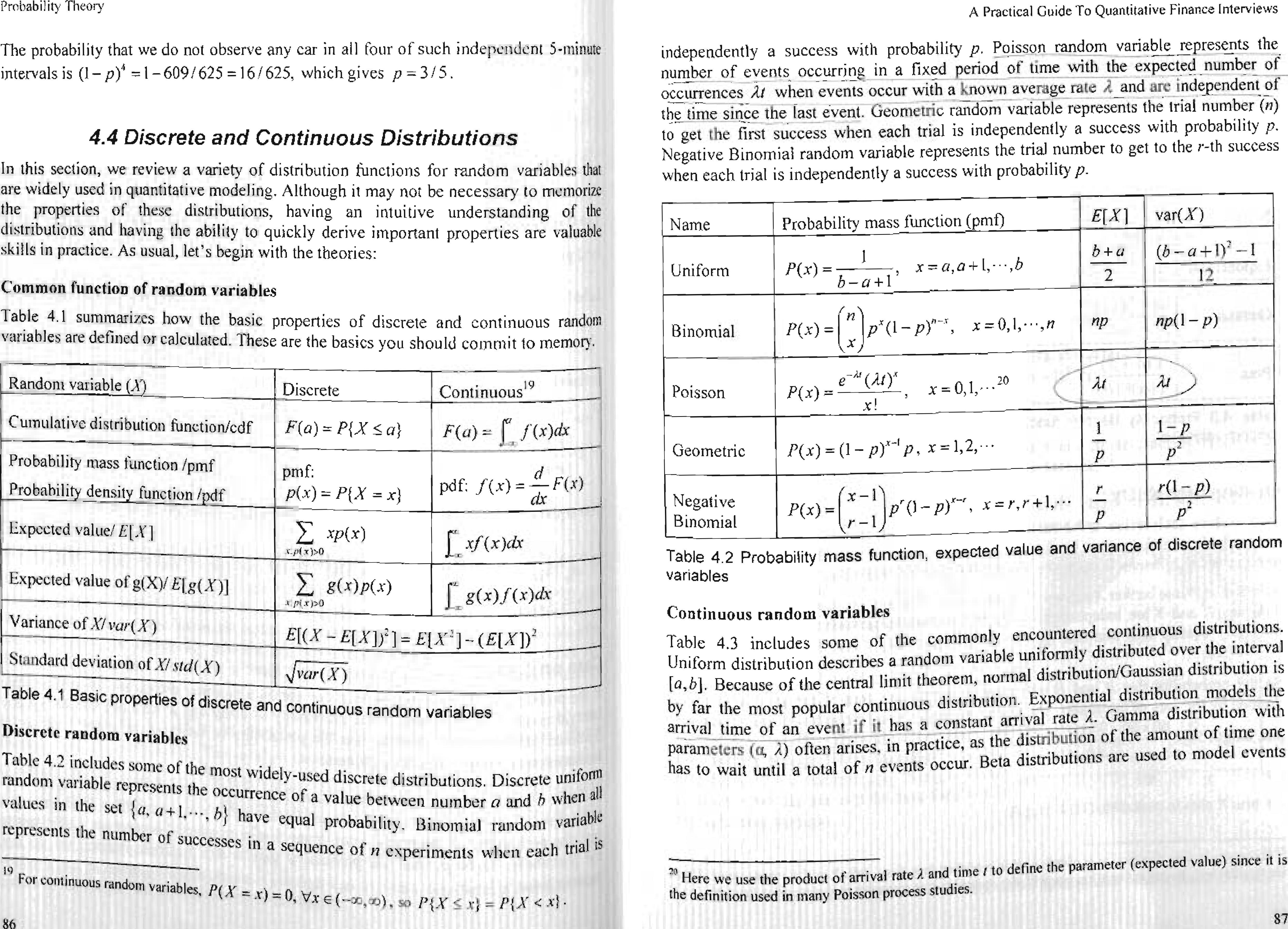
Hence, Piocoo =1 /99 .

|  |  |  |
| --- | --- | --- |
| have Puk | 1 | Vk =1,2,•••,n —1 and *Vn>").* |
| *n —1* |

**Cars on road**

If the probability of observing at least one car on a highway during any 20-minute time interval is 609/625, then what is the probability of observing at least one car during any 5-minute time interval? Assume that the probability of seeing a car at any moment is uniform (constant) for the entire 20 minutes.

*solution:* We can break down the 20-minute interval into a sequence of 4 non-overlapping 5-minute intervals. Because of constant default probability (of observing a car), the probability of observing a car in any 5-minute interval is constant. Lets denote the probability to he *p,* then the probability that in any 5-minute interval we do not observe a car is 1— *p .*



x •

-

*P(x)=*

*17*

)11-X

Pxo -

*P(x)=*

X!

firobabilit), Theory

The probability that we do not observe any car in all four of such independcrit 5-minute intervals is (1— *p)4* **1-609/625** =16/625, which gives *p =* 315 .

***4.4 Discrete and Continuous Distributions***

In this section, we review a variety of distribution functions for random variables that are widely used in quantitative modeling. Although it may not be necessary to memorize the properties of these distributions, having an intuitive understanding of the distributions and having the ability to quickly derive important properties are valuable skills in practice. As usual, let's begin with the theories:

Common function **of random variables**

Table 4.1 summarizes how the basic properties of discrete and continuous random variables are defined or calculated. These are the basics you should commit **to** mernmy.

Table 4.1 Basic properties of discrete and continuous random variables **Discrete random variables**

Table 4.2 includes some of the most widely -used discrete distributions. Discrete unifon.ni random variable represents the occurrence of a value between number *a* and *h*

values in the set *a, a 41.-- .1)}* have equal probability. Binomial random variabl'

represents the number of successes in a sequence of *n* experiments when each trial IS

1,)

For continuous random variables, *P(X x) =* 0, Vx c (-

A Practical Guide To Quantitative Finance Interviews

independently a success with probability *p.* Poisson random variable represents the number of events occurrinp, in a fixed period of time with the expected number of occurrences At when events occur with a known ave; Age - and ..•.• independent of the time since -me last event. Geometric random variable represents the trial number (/I) to get •-ie first success when each trial is independently a success with probability *p.* Negative Binomial random variable represents the trial number to get to the Nth success when each trial is independently a success with probability *p.*

var(X)

*b+ a*

2

*np(1 — p)*

20

Negative *P(x)=*

Binomial *r —1*

Table 4.2 Probability mass function, expected value and variance of discrete random variables

Continuous random variables

Table 4.3 includes some of the commonly encountered continuous distributions. Uniform distribution describes a random variable uniformly distributed over the interval [0,b1. Because of the central limit theorem, normal distribution/Gaussian distribution is

by far the mostpopular continuous distribution. .Exponential distribution models the arrival time of an eve' if it has a constant arrival rate A. Gamma distribution with param ocrs (a, ,) often arises., in practice, as the distribution of the amount of time one has to wait until a total of *n* events occur. Beta distributions arc used to model events

Here we use the product of arrival rate and time to define the parameter (expected value) since it is ,

the definition used in many Poisson process studies.

Name Uniform

Binomial

Poisson

Geometric

Probability mass function (pmf)

*At*

p

r
  
*p*

*(b* — *a* +11' - 1

*)*

*1— p P­r°*

PiX • ri \_ PO' <

fi 87

Discrete

Continuous

*F(a)* f(x)dx

Pdf:*F(x)*

*ax*

Variance of .N7 var( X ) EVA-*XD:*

Standard deviation of X/ *sid(*X)

*1=E X j- (E[X])2*

Random variable (X)

Cumulative distribution functionlcdf

Probability mass function /pint' Probability density function ipdf

Expected value/ E[X1

Expected value of g( El.g( X )

*F(a)= P{X*

*g(x)f(x)dx*

*P(x) =*

*x -a,a+1,„b*

*b-a+1*

*e-2-r (At).< x 0,1,.*

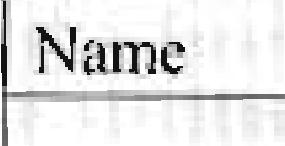
*P (x)*

that are constrained within a defined interval. By adjusting the shape parameters a and it can model different shapes of probability density functions.'I

Probability Theory

A Practical Guide To Quantitative Finance Interviews

Probability density function (pde



I ni form

*E[X]*

*b + a*

2

var( X)

0\_02
  
12

*< x < b*

*a*

*\_V- 11)2*

Normal

271-Ci *e 2*X E 00)

Exponential

*Ae- x > 0*

1 / 2 1, A.

*a/2* a/22

Gamma

Beta

*(xr*

*x O, r(o)*

-.y y a - I

*1(a)*

*a afi*

|  |  |  |  |
| --- | --- | --- | --- |
| *l'(a +) x„ 0 x)fl\_, F(a)1(#)* | | | 0 < x 1 |
|  |  |  |  |

*a + 6 (a fi* 1)(a +fit j

Table 4.3 Probability density function, expected value and variance of continuous random variables

**Meeting probability**

Two bankers each arrive at the station at some random time between 5;00 am and 6:00 am (arrival time for either banker is uniformly distributed). They

stay exactly

minutes and then leave. What is the probability they will meet on a given day? *Soizaion:* Assume banker A arrives X minutes after 5:00 am and *B* arrives Y minutes afte!

5:00 ant and Y are independent uniform distribution between 0 and 60. Since bol11,
  
onl: stay exactly five minutes,. as shown in Figure 4.4, *A* and *B* meet if and onlY . X r 5.

So the probability that *A* and *B* will meet is simply the area of the shadowed regi.°11, divided by the area of the square (the rest of the region can be combined to a square %On , .„ 60x0-2x(1/2x55x55) (60+55)x(60-55) 23

siie lengtn

60x 60

60 x 60 144

,nt If Yoti

`for cxamp1e, hem distribution is widely used in modeling loss given default in risk managernt,
  
dre familiar with liayesian statistics, you will also recognize it as a popular conjugate prior function.

88

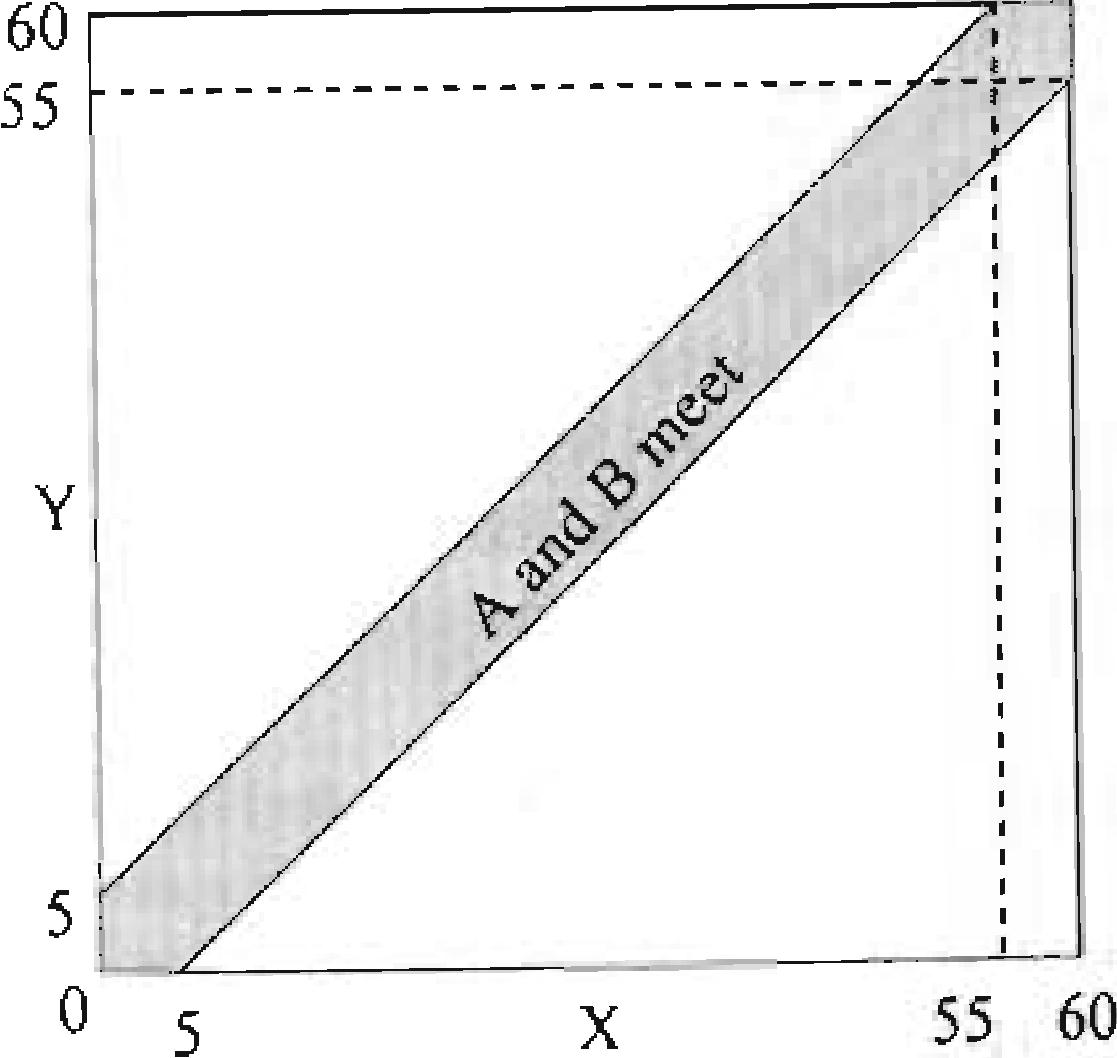


Figure 4.4 Distributions of Banker A's and Banker B's arrival times

**Probability of triangle**

A stick is cut twice randomly (each cut point follows a uniform distribution on the stick),

22

what is the probability that the 3 segments can form a triangle.

*Solution:* Without loss of generality, let's assume that the length of the stick is I . Let's also label the point of the first cut as x and the second cut as y

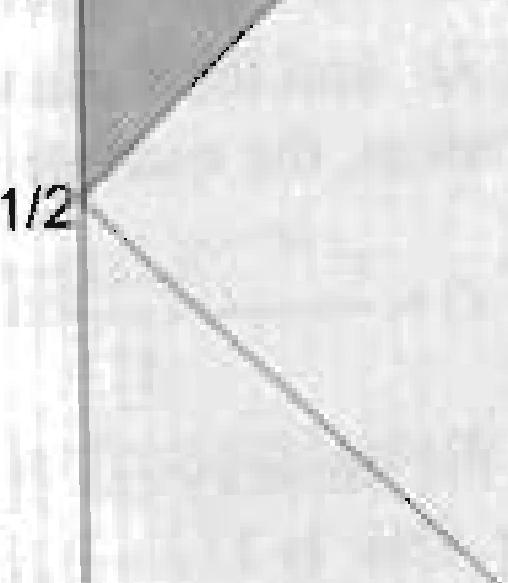
1 —

„Al

I fx < y, then the three segments are x, y-x and

*-1-y.* The conditions to form a triangle are

*x y-x 1-y*



*x* +(y—x)>1—yy>1/2.

x+(l—y)>y—xy<1/24-x

(y x) +(1— y) > x x < 1 / 2

The feasible area is shown in Figure 4.5. The
  
case for x < y is the left gray triangle. Using

1 *i2* X 1

symmetry, we can see that the case for x > y is

the right gray triangle. Figure 4.5 Distribution of cuts X and Y

22 •

I lint: Let the first cutpoint be x, the second one bey, use the figure to show the distribution of x and y.

89

The total shadowed area represents the region where 3 segments can form a triangle, which is 1/4 of the square. So the probability is 1/4.

Probability Theory A Practical Guide To Quantitative Finance Interviews

**Property of Poisson process**

You are waiting for a bus at a bus station. The buses arrive at the station according to a Poisson process with an average arrival time of 10 minutes (A = 0,11 min). If the buses have been running for a\_ long time and you arrive at the bus station at a random toe, what is your expected waiting time? On average, how many minutes ago did the last bus leave?

*Solution:* Considering the importance of jump-diffusion processes in derivative pricing and the role of Poisson processes in studying jump processes, let's elaborate more on exponential random variables and the Poison process. Exponential distribution is widely used to model the time interval between independent events that happen at a constan average rate (arrival rate) *A:* 0)

0 *(t* < 0) . The expected arrival time is 1/

and the variance is 1/ , Using integration, we can calculate the cdf of an exponential

distribution to be *F(t) P(r = – e-ri* and *P(r > t) =* where r is the randoni variable for arrival time, One unique property of exponential distribution is memorylessness: PIT > s+ r r >,51 *1)(1- >1'1.23* That means if we have waited for s time units, the extra waiting time has the same distribution as the waiting time when we start at time 0.

[\\Then](file://///Then) the arrivals of a series of events each independently follow an

exponential

distribution with arrival rate i4, the number of arrivals between time 0 and *i* can e

b

e--`1>tx e.d

modeled as a Poisson process *P(N(t)--,--* x). x = 0, 1, • • 24 The expCct

!

number of arrivals is *At* and the variance is also *At .* Because of the memoryless nature of exponential distribution, the number of arrivals between time *s* and *t* is also a Poisson *(.*

process P( *s)* = x) *1(1s))*

Taking advantage of the memoryless property of exponential distribution, we know,t\_11:,( the expected \Nailing time is 1/ A, =10 min If you look back in time, the menloOL property stills applies. So on average, the last bus arrived 10 mintites ago as well'

*Po- > s tI r> s} (x*

More rigorously. N(t) is defined as a right-continuous function.

This is another example that your intuition may misguide you. You may be wondering that if the last bus on average arrived 10 minutes ago and the next bus on average will arrive 10 minutes later, shouldn't the average arrival time be 20 minutes instead of 10? The explanation to the apparent discrepancy is that when you arrive at a random time, you are more likely to arrive in a long time interval between two bus arrivals than in a short one. For example, if one interval between two bus arrivals is 30 minutes and another is 5 minutes, you are more likely to arrive at a time during that 30-minute interval rather than 5-minute interval, In fact, if you arrive at a random time, the

E[X2]

90 9!

expected residual life (the time for the next bus to arrive) is for a general

2E[]

distribution.25

**Moments of normal distribution**

If X follows standard normal distribution ( X – N(0, 1)), what is E[r] for *n --* I. 2, 3 and 4? L c , . .. \_ i !i-. 1r )‹.- 6-

.

.\_.

*sohllion:* The first to fourth moments of the standard normal distribution are essentially ,

the mean, the variance, the skewness and the kurtosis. So you probably have remembered that the answers are 0, 1, 0 (no skewness), and 3, respectively.

1

. Using simple symmetry we

Standard normal distribution has pdf *f (x) =*,----- e

*V271-*

have

1

*e*

*2 dx =* 0 when *n* is odd. For *n* , integration by parts arc

*27/-*

often used, To solve E[V] for any integer *n,* an approach using moment generating

functions may be a better choice. Moment generating functions are defined as Ni(t) *E[ef•* iEer(x), *if* x *is discrete*

*–0,3 f (x)dx, f x is continuous*

Sequentially taking derivative of MO, we get one frequently-used property of Al(t): *111'(t) —d ElelY* '(0) = E[X],

*"(0 —d ET* Are' 1. *E[X: M 11(0)*

*tit*

25 'Fhe residual life is explained in Chapter 3 *of-Discrete Stochastic Process"* by Robert G. Gallager,

Probability Theory

and .t1"(0) = *an, n ?\_* 1 in general.

We can use this property to solve *E[X ]* for X — N(0, 1). For standard normal

-x` 2/2  *e* /212

distribution *M E[eul. e ch ef*

*ax* e .

-1.--c\* 2 7T

ff 7-4 2 is the pdf of normal distribution X — *N (t , 1)* E

Taking derivatives, we have

*11'(1) to* Al = 0, A/ "(0= *ef' + e.'21' AP(0) = =1,*

*(i) = +2tet t'ef, `.2 = 31(t '2 + t'*

and )1.44(t)----- 3e' *+3t2ef22* + *31' + 1 A44 (0) - 3e° =* 3

***4.5 Expected Value, Variance & Covariance***

I..xpeeted value, variance and covariance are indispensable in estimating returns mid The basic knowledge includes the following:

risks of any investments. Naturally, they are a popular test subject in interviews as v,e1. I f *,1* is finite for all *i = l*.. • *n.* then *E[...)(1* + • • + = E[X, - • • E[x „1. The relationship holds whether the x, 's are independent of each other or not

lf X and Yace independent, then *E[g(X)11(1)1,- E[g(x)]E[h(y*)]

**Covariance: Cove vi'. Y1** *ER X — E[X] ElY1)1* **Ef *E[X]E[Y]***

**Correlation: p(X,** *C* **—** *oq X , Y )*

**Var (** X) *fiar(r)*

f X and Y are independent. Coy( X. r)-,--- 0 and */0( X )* 0.26

**General rules of variance and covariance:**



21y cova.,x-i)

* dent tk- :►erse is not true. pa, Y)---= 0 only means X and Y are uncorrelated they may well b rh-,n e en

A Practical Guide To Quantitative Finance Interviews

**Conditional expectation and variance** lurs.uu f vOL 0:-1-

Sy)

93

•

For discrete distribution: *E[g(X)IY yll=Eg(x)p y) = g(x)p( X* **=xIY=y)**

For continuous distribution: E[g(X) Y = y] g(x) [co, (x j yv\_x

**Law of total expectation:** *E[X y = y[p(1' = y),* for discrete *Y*

*E[E[X 111]= -* Y *E[X iY Ai; (y)(i.y,* for continuous Y

**Connecting noodles**

You have 100 noodles in your soup bowl. Being blindfolded, you are told to take two ends of some noodles (each end on any noodle has the same probability of being chosen) in:rcyloeusr. bowl and connect them, You continue until there are no free ends, The number of loops formed by the noodles this way is stochastic. Calculate the expected number of circles.

*Sohdion..* Again do not be frightened by the large number 100. If you have no clue how
  
to start, let's begin with the simplest case where *n ---- 1 .* Surely you have only one choice
  
(to connect both ends of the noodle), so *E[ (1)] =1 .* How about 2 noodles? Now you

**4 x3**

**have** 4 ends (2 x 2 ) and you can connect any two of them. There are-- 6

2 2

tcs000gnet-ihtbheieneraxtapioleindesty.ediAerriduo.minbcgeirtrcholeefm, 2 combinations will connect both ends of the sirrie noodle

snicslodle. The other 4 choices will yield a single noodle.

and

216x(1+ *E[f (01) +* 4/6 x *ET f (1)].* 113+*E[*

*(1)1= /* 3 +1

**We now** move on to 3 noodles with 6 2

x 15 choices. Among them. 3 choices

will yield 1 circle and 2 noodles; the other **i2** choices will yield 2 noodles so

*E[f* (3)] = 3/15 x(1+ *Lif* (2)])+12 /15 x (2)) 1/5+ *E[f(2)]..-* 115+1 '3 • 1.

See the pattern? For any *n* noodles, we will have *(n)1 --- +* I /3+1 r " • +1 '(2n
  
which can be easily proved by induction. Plug 100 in, we will have the answer.

Actually alter the 2-noodle case, you probably have found the key to this question. If

Probability Theory A Practical Guide To Quantitative Finance Interviews

2 n

you start with *n* noodles, among = *n(217* ***-*** ) possible combinations, we have

, 2

***17*** 1 2n -2

probability to yield 1 circle and *n -1* noodles and probability

*n(2n-1)= 2n -1 2n -1*

to yield *n -1* noodles only, so *E[f* 1 *(17)] = E[f(n -1)]+ .* Working backward, you

-1

can get the final solution as well.

**Optimal hedge ratio**

**You just** bought one share of stock A and want to hedge it by shorting stock *B.* Flew many shares of *B* should you short to minimize the variance of the hedged position? Assume that the variance of stock As return is *crA2;* the variance of B's return is *o--:* their correlation coefficient is *p.*

*Soiution;* Suppose that we short *h* shares of B, the variance of the portfolio return is var().; *- 1v)* - *2phu (71,+/I'a1*

The best hedge ratio should minimize varfrA *-hrB).* Take the first order partial

var

derivative with respect to *h* and set it to zero:

*= -2po- o- +2ha2 = 0 h - -*

*p*

*ah /I 13*

To confirm its the minimum . we can also check the second-order partial derivative:

|  |  |
| --- | --- |
| var ,  > 0. So Indeed when  variance. | *h 4 13-1 ,* the hedge portfolio has the minimum  *0- 1* |

1/2 chance to get Y **E** 14,5, 61, in which case you get expected face value 5 and extra throw(s). The extra throw(s) essentially means you start the game again and have an extra expected value *E[X] .* So we have *E[X I Y* **E** (4,5,6)] = 5 + *E[X].* Apply the law of

total expectation, we have *E[X] = EIE[X !Y]]= +x* 2 +-Ix (5+ *E[X]) E[X]. 7 .27*

**Card game**

What is the expected number of cards that need to be turned over in a regular 52-card deck in order to see the first ace?

*Solution:* There are 4 aces and 48 other cards. Let's label them as card 1,2,-- •,48. Let

1, if card i is turned over before 4 aces

*X,*

0, otherwise

The total number of cards that need to be turned over in order to see the first ace is

48

X E x-„ so we have *E[X]* =1+ 48E *E[X J.* As shown in the following sequence,

each card *i* is equally likely to be in one of the five regions separated by 4 aces: 1 *A* 2 *A.* 3 *A* 4 A 5

So the probability that card *i* appears before all 4 aces is 1/5, and we have *E[X]=1/5.*

48

Therefore, *E[X]= 1 +EE[,V i]=1+* 481 5 =10.6 .

This is just a special case for random ordering of ***177*** ordinary cards and *n* special cards.

The expected position of the first special card is 1+ ZE[X,1= I +

*+1*

*1=I*

**Dice game**

Suppose that ou roll a dice. For each roll. you are paid the face value. if a roll 14i **"2.5.4.` ,.,**

**5**

or 6, you can roll the dice again. Once you get 1. 2 or 3, the ame stos, What Is expected payoff of this game? g p

*I be Solution.. This* is an example of the law of total expectation. Clearly your payoff wil*l different* depending on the outcome of first roll. Let *FIX]* he your expected paY(-)ff

he the outcome of your first throw,. You have I /2 chance to get Y ***E t I,*** 2. , its whic.1

the expected value is the expected face value so El- XlYc 11,2.1} I '7- Y°11 haVe **Sum of random variables**

**94** 95

Assume that *X* • " , and .,1(11 are independent and identically-distributed (I1D)

random variables with uniform distribution between 0 and 1. What is the probability

that *S„* --- + X2 +. + X„ < 1 ?"

27

Yoli will also see that the problem can be solved using Wald's equality in Chapter 5.

Hint: start with the simplest case where n =1, 2. and 3. Tr), to find a general formula and prove it using

induction.

*Solution..* This problem is a rather difficult one. The general principle to start with the
  
simplest cases and try to find a pattern will again help you approach the problem; even
  
though it may not give you the final answer. When *n =1, P(S1* is 1. As shown in

Probability Theory A Practical Guide To Quantitative Finance Interviews

we have *P(S„,1*

|  |  |
| --- | --- |
| 4(:) is the cross-sectional area. | all distinct types is *E[X]=ZE[Xil•* |

Figure 4.6, when *n* --,. 2, the probability that X,2(21 is just the area under FYI + X, l within the square with side length I (a triangle). So *P(S,* 12, When *n=* 3, the probability becomes the tetrahedron ABCD under the plane *X,* + X, + X,

within the cube with side length 1. The volume of tetrahedron ABCD is 116.29 So P(S, 1)=1 /6. Now we can guess that the solution is 1I n!. To prove it, let's again

resort to induction. Assume *P(Si, <1)=11 n!.* We need to prove that

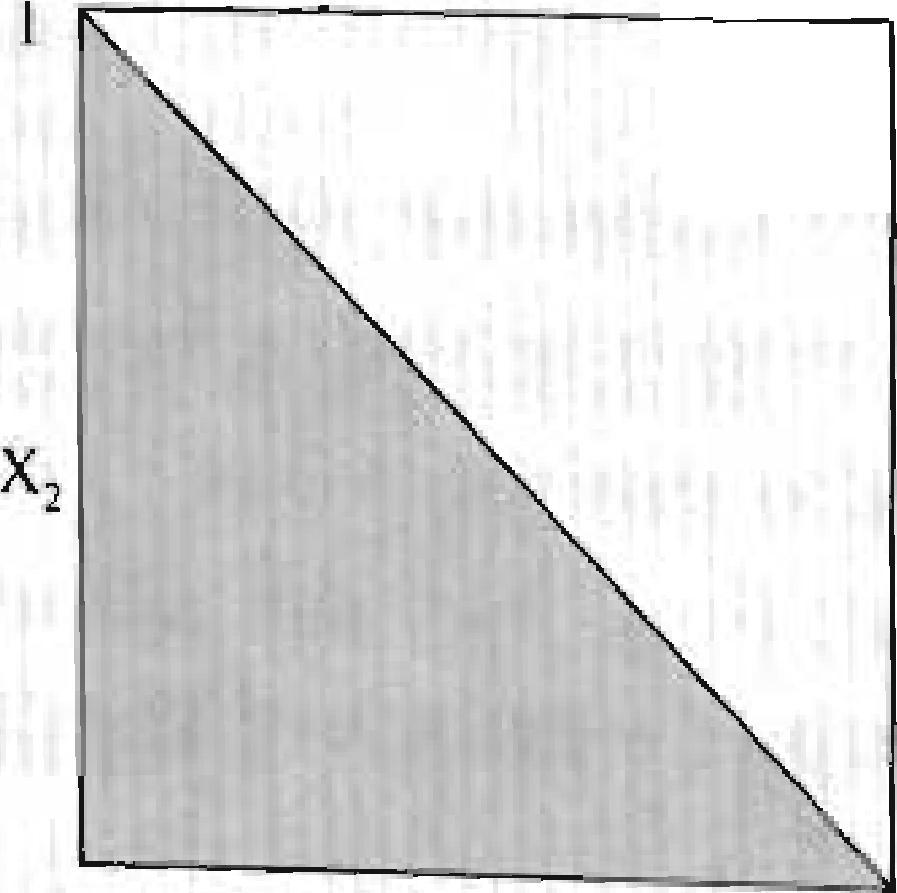
*P(S„., 1) =10+1)1.*

n=2

Figure 4.6 Probability that S,-, s 1 when n = 2 or n= 3.

XI

n ---- 3



BO

X,



tir



Here we can use probability by conditioning. Condition on the value of X we have

P(. 1l *f (X ....1)1)(S. 1 X ,)(.1X* where *f* (X,+1) is the probability densit).

function of , so *.1 1.* But how do we calculate *P(Si,* X)? The Cases

have provided us with some clue. For S < 1— x„,., instead of

we cs,.s,mtially need to shrink every dimension of „ the -dimensional simp1ex3° from 1 to

Ali 1(,)c! - - I / 2 -'(/z = l '6, where

n-Siinplekis th o-dimensional analog of a triangle.

|  |  |
| --- | --- |
| ou can derive it b. | integration: |

96

*"*

*)*

1— *X* (1— *„+,*

So its volume should be instead of 1. Plugging in these results,

*n? n!*

(1—)' 1 [ (1— )rr+I **1** 1 **1**

**X ---**

*oa-11*

*! n! n +1 0 n!* )7+1 (n+1)!

So the general result is true for *n +1* as well and we have *P(S„ n!.*

**Coupon collection**

There are *N* distinct types of coupons in cereal boxes and each type, independent of prior selections, is equally likely to be in a box.

1. If a child wants to collect a complete set of coupons with at least one of each type, how many coupons (boxes) on average are needed to make such a complete set?
2. if the child has collected *n* coupons, what is the expected number of distinct coupon types?3

*Solution:* For part *A,* let *X „ 1=h* 2, • *N ,* be the number of additional coupons needed
  
to obtain the i-th type after *(1-1)* distinct types have been collected. So the total number

of coupons needed is X ='1+ X2 + - • + XA; =

? =,

For any *i i* —1 distinct types of coupons have already been collected. It follows that a
  
new coupon will be of a different type with probability 1—(i —1)1 *N= (N —i+DIN.*

Essentially to obtain the i-th distinct type, the random variable X, follows a geometric distribution with *p = (N —1+1)/ N* and *E[X j= N l(N —i+1).* For example, if we simply have *Xi* = *E[Xi]=* .

*N N* E[X] E[xi] *N \_ Ni \_1 + 1 + \_1).*

*N-1+1 Of N-1 1)*

31 Hint: For P *,* ,art *A* let X be the number of extra coupons collected to get the i-th distinct coupon after

i- I types of distinct coupons have been collected. Then the total expected number of coupons to collect For part B, which is the expected probability (P) that the i-th

coupon type is not in *the n* coupons?

97

Probability Theory

For part *B,* let be the number of distinct types of coupons in the set of *n* coupons. We

introduce indicator random variables *1„* 2, • N , where

*1,* I, if at least one coupon of the i-th type is in the set of *n* coupons = 0, otherwise

|  |  |
| --- | --- |
| So we have 1- *1, +12+...+1* |  |

. A' -1

For each collected coupon, the probability that it is not the i-th coupon type is

Since all ii coupons are independent, the probability that none of the *n* coupons is the i-th

*N -1Y*

coupon type is *P(1 =* 0) ------ and we have *EU]. P(1 =1) --,--1 I*

ti

*E[ 1 N N ( N -1* 32

*N )*

**Joint default probability**

tf there is a 50% probability that bond A will default next year and a 30% probability that bond *B* will default. What is the range of probability that at least one bond defaults and what is the ranue of their correlation?

*Soluition:* The range of probability that at least one bond defaults is easy to find. To have the largest probability. we can assume whenever A defaults, *B* does not default., whenever *B* defaults, A does not default. So the maximum probability that at least olle bond defaults is 50%-i-30% 80%. (The result only applies if *P(A)-1-- PU3).5-1* For the minimum, we can assume whenever A defaults, *B* also defaults. So the minimum

probability that at [cast one bond defaults is 50%.

To calculate the corresponding correlation, let /I and be the indicator 1b the event

that bond A.13 defaults next year and be their correlation. Then we have

/ - vac(/'. ) p„ x (1 - p 0.25. 'ar(1)

**13**

**A** 511)1ilar Litivstion: )011randotn put 18 balk into 10 boxes, what is the expected number of ell\*

boxes?

A Practical Guide To Quantitative Finance Interviews

*P(A or B delbults)---= Ai+ E[18]- E{1,418)*

99

*E[15]-(E[14141/3]-cov(1* ))

= 0.5 + 0.3 - (0,5x 0.3 -

NE.Y1 2pAn

For the maximum probability, we have 0.65 -•NATT21 12pAfl \_ 0.8 p4„ = .

For the minimum probability, we have 0.65 -*0.21 12p,48* = 0.5 *ii9,8=1317 .*

in this problem, do not start with *P(A or B defaults)=* 0.65 -J.7211 and try to set

p. = ±1 to calculate the maximum and minimum probability since the correlation

cannot be ±1. The range of correlation is restricted to [---577, V3/7]

***4.6 Order Statistics***

Let *X* be a random variable with cumulative distribution function *F',. (x).* We can derive the distribution function for the minimum Y„ = min(X,, X2 • ' • , X„ ) and for the maximum *z„ =* max(X,, *X„•* • • , of *n* 11D random variables with cdf *Fx* (x)as

*P()1 ?\_x).--(P(X x))” 1- Fr (x) = (1- F. (x))" fr,(x) njA. .(x)(1-*

*x)=(P(X x))" F4(x)=(Fx (x))" fz f*

*,(x) = of ( F ( O*

**Expected value of max and min**

Let .x,,x2,.„,

*X„* be HD random variables with uniform distribution between 0 and 1. What are the cumulative distribution function, the probability density function and expected value of Z„ = X,,)? What are the cumulative distribution function the probability density function and expected value of Y„ =.min(X,. X,,,• • • , X,,)? *Suloion:* This is a direct test of textbook knowledge. For uniform distribution on [0,11. FX (x) ----- x and *f (x) = 1 .* Applying *F,-(x)* and (x) to Z„ = max(X,, X2,. • • , X„) we have

*P(Z (P(X* e x)Y *p/* (X) (/";. *..*

I'., *qi:v(X)(FX* (X))p = ► *u*IS