Jane Street

Probability & Markets

Contents

Introduction

The goal of this document is to present an introduction to probability and markets in what we hope is a fairly intuitive and accessible way. Having an intuitive understanding of these things is an important part of our day-to-day jobs, and we feel they help us in thinking about many problems outside of work as well. However, these concepts are not particularly natural, and some familiarity is required before intuition develops. It is worth the effort!

Randomness

There is a lot of uncertainty in the financial markets (stock prices go up and down), and it's often very hard to predict what will happen before it does. We want to be able to reason about and deal with things we don't know. Probability provides a good framework for making decisions when uncertainty is involved.

Lots of things in the world are random in different ways.

– The classical example of a random event is **a die roll.** The roll of a die seems random in a sort of objective way. Without getting into a debate about physics, we can hopefully all agree on the likelihood of different outcomes of the roll; the possibilities are pretty straightforward (a standard die has 6 sides, each of which is equally likely to show up) and nothing besides the act of rolling the die affects the outcome.

– Some **one-off future events,** like the weather tomorrow, are a bit more complicated. Perhaps with perfect information about winds, tides, etc., we can predict future weather, but there would be a lot of factors at play and no one has that information. As a result, it's possible for us to disagree somewhat about the likelihood of rain tomorrow. Maybe I have different information than you, or maybe I've had different experiences in the past that make me draw different conclusions based on the same information. Still, for us each personally, we can say "I think there's an x% chance of rain tomorrow."

– There are also **knowable unknowns,** like the weather at the top of Mt. Everest yesterday at noon. Even though the event in question is in some sense not random (it's already happened!), it's not known to us. Subjectively, then, it is still random to us. We can, and should, still reason about it with all the tools and language of probability.

Although these categories feel different, they are actually very similar. For example, if one were to have precise enough measurements, it's possible one could keep track of the order of cards in a deck as they are shuffled in front of you or infer the temperature tomorrow accurately. Perhaps almost all things are in some sense knowable unknowns.

There are also different types of randomness:

- **binary:** "will it rain tomorrow?" has a % chance of being yes (instead of no)
- **discrete:** a die roll will result in one of a finite list of possible outcomes; this includes binary scenarios
- **continuous:** tomorrow's temperature can be basically anything (although likely within some range)

Counting

In some situations, there are simple methods to figure out the probabilities of different outcomes in a given situation. When there is a **finite number of outcomes, each of which is equally likely,** we can just count up the possibilities:

Notation: we write P[outcome] to indicate the probability of an outcome occurring. P is a function that takes as input the outcome (or set of outcomes) and outputs the probability (from 0 to 1) of that outcome.

P[criteria is fulfilled] = (number of outcomes that fulfill criteria) / (total number of possible outcomes)

So let's take a standard six-sided die (denoted “d6”), which randomly generates a number between 1 and 6, each with an equal chance.

Q: *What's the probability that we get a 1?*
There's only one acceptable outcome (rolling a 1), out of 6 possible outcomes:

P[roll is 1] = 1 / 6

Q: *What's the probability that we roll a prime number?*
Of the possible outcomes, only 2, 3, and 5 are prime:

P[roll is prime] = 3 / 6 = 1 / 2

Here's a trickier one.
Q: *What's the probability that when we roll 2 d6s and add up the results, we get 7?*

A common mistake is to say that there's one acceptable outcome (the sum is 7), out of 11 (the sum could be any integer between 2 and 12), so it's 1/11. It's important to notice that **not all sums are equally likely** — intuitively, it seems like it should be much harder to get a sum of 2, because that requires exactly two 1s, than a sum of 7, which could be 1 + 6, 2 + 5, 3 + 4, etc.

We also remember that the dice are distinct; to illustrate this, imagine a red die and a blue die. There are 6 possible outcomes for the red die, and 6 for the blue die, making a total of 6 x 6 = 36 possible outcomes. This table shows all the sums that could occur:

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Notice that you can roll a red 1 and a blue 6, but you can also flip the colors and roll a red 6 and a blue 1! They are distinct outcomes, and we'll need to include both in your count. This gives us a total of 6 ways to make a sum of 7: [1, 6], [2, 5], [3, 4], [4, 3], [5, 2], and [6, 1]. This is also just 2 x (possible combinations disregarding color).

P[sum is 7] = 6 / 36 = 1 / 6

Q: *What's the probability of getting a sum of 10 when you roll 2 d6s? Hint: there's no way to flip [red 5, blue 5]!*

Sometimes events aren't equally likely, for example the probability of the sum of 2 dice rolls being 2 is different than the probability of it being 7. It's occasionally easier to count and think about things when they are equally likely, and a trick you can use is to imagine there are more outcomes which you can treat as such. For example, suppose something you care about has 30% chance to be 0, 20% chance to be 1 and 50% chance to be 2 -- you can imagine there are 10 equally likely events, 3 of which result in 0, 2 of which result in 1 and 5 of which result in 2.

Probabilities

We've already used probabilities, but before going further we'd like to be more precise about this term. This helps in turning problems where it's a little hard to simply intuit the answer into mathematical formulas that we can solve.

In general, we have a set of distinct outcomes that may happen. **Distinct outcomes are mutually exclusive,** meaning that if outcome x happens then outcome y doesn't happen. Each outcome is assigned a probability >= 0 that it will happen.

So,
1. P[X] ≥ 0 for all outcomes X
2. The sum of P[X] across all X's equals 1

We use a probability of 1 of something happening to indicate that the thing will surely happen, so the second condition means that our set of X's includes all possible outcomes.

These definitions extend naturally to sets of outcomes, although there's no guarantee of distinctness here. For example, a die roll could be both even and prime. The probability of a set of outcomes happening is just the sum of the probabilities of each of the outcomes in the set. That is,
$P[A] = \Sigma_{x \text{ in } A} P[x]$

and the properties of Venn Diagrams and sets can now be used, so for example
P[A or B] = P[A] + P[B] - P[A and B]
P[not A] = 1 - P[A]

Independence

One issue that arises is that the set of possible outcomes can start to get large and unwieldy. If I'm thinking about the set of all possible outcomes for a set of 100 coin flips along with the weather tomorrow, then all heads with warm temperature is distinct from all heads with cool temperature is distinct from all heads except the last coin and the temperature is warm, etc. It quickly becomes impractical to list all the outcomes. Very often, and very naturally, we want to break things up into smaller chunks that we can feel are unrelated to each other. In the coin+weather example, we probably don't think either outcome is in any way affected by the result of the other (e.g. coins are not more likely to land heads up on rainy days), and moreover each coin flip is probably unaffected by any other coin flip.

The word used to describe this situation is **independence.** Independence arises when the probabilities of two outcomes are unrelated. A nice property of independent outcomes is that the probability of them both occurring is the same as the product of their probabilities of individually occurring.

For example, the two dice roll example from earlier has 36 outcomes; 1 on first roll and 1 on second roll through to 6 on first roll and 6 on second roll. For each of these outcomes, say x on the first die and y on the second, we have

P[x on first, y on second] = 1/36 = 1/6 * 1/6 = P[x on first die] * P[y on second die]

This extends to sets of outcomes naturally, so P[even on first, odd on second] = 1/2*1/2 = 1/4. When things are independent, it normally helps a lot to treat them as such.

Q: *What's the probability that you draw 2 face cards from a deck of cards, if you put the first card back in before your second draw ("with replacement")?*

Q: *What's the probability that you draw 2 clubs from a deck of cards, if you don't put the first card back in ("without replacement")?*

Q: *What's the probability that the product of 2 rolls of a d6 is odd? 3 rolls?*

Random Variables

Often the set of outcomes you are considering can naturally have numbers attached to them. For example, the number of pips on a die or the price of a security in the future, or how much money you might make or lose as the result of some bet. If you associate a number with each possible outcome, you can then think about the numerical value of the actual outcome, which is unknown, as a **random variable.** The set of numerical outcomes and the probabilities of each is called the **distribution** of the random variable.

There are an infinite number of probability distributions, and quite a large number of them have special names because they're so commonly used. The one you'll hear about the most is probably the **uniform** distribution. If we say a random variable is uniformly distributed, that means all outcomes are equally likely. For example, the roll of a d6 is uniformly distributed, but the sum of 2 d6s is not.

Often, random variables can share the same distribution. For example, if a red and a blue die are rolled, the number of pips on the blue die, B, and the number of pips on the red die, R, are two random variables that have the same distribution, but the actual random variables themselves are not equal. In this case we say B and R are **independent** because being told the outcome of one never gives you any information about the other.

Many new random variables can be constructed from existing ones. For example, B+R is a random variable whose value is the sum of the pips on the red and blue dice, and B*R is a random variable whose value is the product of the pips on the blue and the red dice.

Q: *Are B+R and B*R independent random variables?*

In fact, anything you can do with regular variables can also be done with random variables, giving you new random variables:

– If you have the temperature in Fahrenheit as random variable F, but you want to think about it in Celsius now, you can make C = (F - 32) * (5/9).
– If your commute to work involves driving to the train station (taking D minutes) and then riding the train (taking T minutes), then your total commute could be D+T.

Expected Value

The average of a random variable is the most important single piece of information about a distribution. We call this the **expected value,** and if the random variable is X, we denote the average by E X or, if brackets are needed for clarity, then E[X].

For example, the average number of pips on a roll of a standard 6-sided die is 1*1/6 + 2*1/6 + ... = 3.5. In general, the average can be found in this way, adding up each outcome times its probability. This fact can be written like this:
$E\ X = \Sigma_x\ x * P[X=x]$

where P[X=x] means "the probability that the random variable X ends up being x", e.g. "the probability that the die roll comes up 3". (It's customary to capitalize random variables to distinguish them from plain old variables).

Note, you could have also obtained 3.5 by just thinking to yourself "what is the midpoint of the outcomes?" That's because the distribution of both pips is symmetric around 3.5.

Q: *What is the expected value if the die is reweighted so that 6 occurs half the time?*
Q: *What would the expected value be if the 6 is replaced by a 12?*

Expected values have lots of nice properties:

– For 2 random variables X and Y, E[X+Y] = E X + E Y
This is true even if X and Y are not independent! This can simplify calculations immensely.

– If X and Y are independent, then E[X*Y] = E X * E Y
– In particular, if x isn't random (for example, it's 2 always), then E[x*Y] = x * E Y

– There are some clever ways to assign numbers to outcomes that help with calculations a lot. One nice example is associating outcome with 0 if an event you care about doesn't end up happening and 1 if it does happen. We call a function like this an **indicator function,** and if the event is A, we denote it 1_A. It turns out that E 1_A = P[A] which is really nice. Indicator functions come up often in betting where you get paid out some dollar amount if your sports team wins and lose some dollar amount if they lose. You can write this as a linear function of the indicator function in the following way. If we let A be the event that you win the bet, and you get paid x if you win and y if you lose (y is probably negative), then your payout is y*1_{not A} + x*1_A which means your expected payout is y*P[not A] + x*P[A] which also happens to equal y + (x-y)*P[A]

One reason that expected value is so very useful is if that if you are considering a series of bets and you can think of them independently, then, under almost all situations, your total payout will be not too far from the total of the average payouts. To be precise: the average payout you actually receive will tend to get closer to the average of the expected values as you increase the number of bets. This is called the **"Law of large numbers",** https://en.wikipedia.org/wiki/Law_of_large_numbers

Q: *What's an example of two random variables X and Y such that E[X*Y] doesn't equal E[X]*E[Y]?*

Confidence Intervals

Sometimes we want to know more about a distribution than simply its expected value. For example, if we're making a trade, we might want to know the probability we lose a lot of money, or more generally the range of normal outcomes.

There are a few ways of measuring the range of outcomes of a random variable:

– Look at the literal **minimum and maximum values** of the random variable, e.g. 1 and 6 for a die, regardless of how it is weighted.

– Make a random variable measuring the **deviation from the average** and measure the average of the new random variable. That is, measure the average deviation from the middle. A couple of common measures for the distance between x and y are |x-y| and (x-y)^2 .

– Describe a range of numbers where seeing a number outside this range is very rare. This is called a **confidence interval,** which is just any continuous range of numbers that contains some predefined percentage of the possible outcomes. Suppose we have a random number between 0 and 100, picked uniformly. [0, 90], [1, 91], etc. are all 90% confidence intervals. Sometimes, we might be interested in a particular one, like the smallest interval, or one that contains the largest possible outcome.

Often times, people tend to be **overconfident.** That is, if they gave a bunch of 95% confidence intervals, less than 95% of them would contain the actual outcome. It's important to try to correct for this, but also don't go overboard!

If you want to practice, you can try quizzing yourself on easily verified facts. Write down some confidence intervals and see how often you're right.

Q: *What is a 95% confidence interval for the number of taxis in New York City? Check your interval after you have constructed it!* https://en.wikipedia.org/wiki/Taxicabs_of_New_York_City

Conditional Probability

Our perceived probabilities of events change as we learn more information about the circumstances.

What's my guess for the weather in London right now? Let's say 60 degrees Fahrenheit. This is my **prior,** or belief about the situation beforehand. If you then tell me that it's currently snowing there, then I'd think it's probably actually a lot colder than 60. As I receive more information (like the fact that it's snowing in London), I **update** my beliefs accordingly.

When we think about probabilities in this way, it's called **conditional probability** (probability of something conditional on something else being true).

Q: *What is the expected value of a d6, if I tell you the outcome is at least 4?*

We know from earlier that the expected value of a d6 is 3.5, but this changes the total possible outcomes. Intuitively, the expected value should be higher now, since we're eliminating the lower rolls, bringing up the average. Now, only 4, 5, and 6 are possible outcomes, and since the relative likelihoods of rolling a 4 vs 5 vs 6 haven't changed (they're still all equally likely), they each occur with 1/3 chance. This would mean that the expected value is 5.

There's also an explicit formula for calculating conditional probabilities, where P[X|Y] is the probability of X being true given that Y is true:

P[A|B] = P[A and B] / P[B]

Q: *Draw a Venn diagram that illustrates that equation.*

We could apply this here:
P[roll=1|roll≥4] = P[roll=1 and roll≥4] / P[roll≥4]
P[roll=1 and roll≥4] is zero, so P[roll=1|roll≥4] = 0
On the other hand:
P[roll=5|roll≥4] = P[roll=5 and roll≥4] / P[roll≥4]
P[roll=5 and roll≥4] is 1/6, because out of the six outcomes of the die roll, only one satisfies the requirements, when a five is rolled. P[roll≥4] is 1/2. This gives us:
P[roll=5|roll≥4] = (1/6)/(1/2) = 1/3

Q: *If I flip 3 coins, what's the probability that I get exactly 2 heads given that I get at least 1 tail?*

P[2 heads|≥1 tail] = P[2 heads and ≥1 tail] / P[≥1 tail]

If we label the coins A, B, and C, we see there are eight possible outcomes of the three coin flips, which we will label 1 through 8:

	A	B	C
1	T	T	T
2	T	T	H
3	T	H	T
4	H	T	T
5	H	H	T
6	H	T	H
7	T	H	H
8	H	H	H

P[2 heads and ≥1 tail]: Outcomes 5, 6, and 7 satisfy both criteria, so this probability is 3/8.
P[≥1 tail]: All outcomes except 8 satisfy this, so this is 7/8.
Overall then, P[2 heads|≥1 tail] = (3/8)/(7/8) = 3/7

Q: *What's the expected value of a d6, given that the outcome is prime?*

Q: *What's the probability the roll of a d6 is no more than 3, given that the outcome is prime?*

The concept of conditional probability is useful for things beyond math problems. It's a constant theme in thinking about trading (see adverse selection later).

Making Markets

Jane Street is a trading firm, so at the end of the day we need to connect all of this knowledge to trading in the markets.

First, some jargon:

When you want to effect a trade, you'll need to specify the

1. thing you want to trade (which is often implied by the context, so let's ignore this for now)
2. **direction** - whether you want to buy it or sell it
3. **price** at which you're willing to transact, and
4. **quantity** - how many shares/contracts/etc.; also called **size**

If you want to buy something, like a stock or a contract, you can say or otherwise indicate a **bid** to buy. If you want to sell, it's an **offer** to sell.

If you want to try to buy 10 widgets for a price of $2 each, say "I'm **2 bid for 10.**"

If you want to try to sell 10 widgets for a price of $4 each, say "I have **10 at 4."**

Now, if you're happy to do both of the above, you can say both phrases, but it's much more common to say "I'm **2 at 4, 10 up."** Note that this requires you to be willing to buy or sell for the same quantity; if you have different quantities depending on direction, then you'll have to say the bid and offer separately.

Saying one of the above phrases creates an **order,** which stays active and anyone can trade against it until it either trades or you say "I'm **out."**

Once an order is created, someone else can trade against it, by saying **"sold"** to **hit** (sell) to your bid, or **"take 'em"** to **lift** (buy) your offer. Saying "sold" or "take 'em" trades against the entire quantity available. To trade only 5 (or any **partial** amount) of the available 10 widgets, say "I **buy 5"** or "I **sell 5."**

If someone trades against your order, it's called a **fill,** and your order is **filled.** If your order is fully filled (someone traded against all available quantity), then your order is gone and cannot be traded against again. If your order is partially filled, the remaining quantity is still eligible for trading, unless otherwise specified.

Now, the strategy:

If you know the expected value of something, you should be happy to buy for less or sell for more.

Consider a contract that pays out a dollar for every pip on the outcome of a d6 roll. You'll get $1 if you roll a 1, $2 if you roll a 2, etc.

Where should you buy this contract? Well, there are lots of considerations, like:

– *How much is this contract worth on average?*

We know from above that the expected value of a d6 is 3.5, so if you play this game a lot (for free!), you'll expect to make about $3.5 on average per game. Unless you're doing something elaborate (like hedging another trade to reduce its risk — see later), you'll want to make sure you buy below the expected value.

– *Can you lose a lot of money? How bad is that?*

The worst case scenario is that you buy this for $3.5, and it settles to $1. You'll lose $2.5, which happens 1/6 of the time. That's probably okay, so you should be making trades here solely based on expected value. If the potential loss is actually a large portion of your capital, though, you'll want to modify your trading to reduce the **risk.** It's bad to lose a lot of money, because that means in the future when there are great opportunities to trade, you won't have as much capital to do those trades.

– *How much are you making?*

This is just to say, sure, you can break even by bidding $3.5, but then you won't make anything on average. If you bid $1, probably no one's going to sell to you. You will want to balance your likelihood of a trade and the expected profit.

Here, after considering all these factors, I might make a market "3 at 4, 10 up."

Try making some markets on these contracts:

Q: *The outcome of a 20-sided die roll (d20).*

Q: *The maximum of 3 d6 rolls.*

Q: *The outcome of a d6, with a choice to reroll once if you don't like the first roll for any reason (but you can't go back to the first roll if the second turns out worse). Hint: who gets to decide to re-roll? Does it matter?*

Q: *The temperature in an hour.*

Q.: *1,000,000 times the outcome of a six sided die. How would you feel if I hit your bid and the roll ended up a 1?*

Adverse Selection

When you're trading, you're trading against a human. Someone who is presumably at least as good at thinking about markets as you. If you buy a security, that means someone else thought it was a good idea to sell it, which should understandably make you wary.

It helps to understand why they're doing the trade:

– Are they just like you, trying to make a good expected value decision? This is scary; there's a good chance you did your calculations wrong, or misunderstood the contract terms, or something.

– Do they care about the price? Sometimes people want to invest long term (or your interviewer commits to doing a trade to make your scenario more interesting), and they tend to care less about the exact price at which they're doing a trade.

Adverse selection is the idea that the trades you get to do are worse than they would appear, because someone else is selecting into the trade against you. In general, you should be more careful in cases where the calculations are complicated or the answer is hard to determine (you'd probably be sad to do a trade with a meteorologist about the weather in a week) than something where you're more certain of the value (you're probably fine buying a d6 roll for $3).

The idea of conditional probability is relevant here. When you make a market, and someone trades against you, that should prompt you to **update** your beliefs! If you bought, the correct value of the security is probably a bit less than you initially believed.

When making markets, it also helps to think: "Conditional on my order being filled, what do I think the expected value of the security is? Given this updated expected value, would I be happy to have done the trade?"

Groucho Marx's oft-quoted comment "I don't want to belong to any club that will accept me as a member" is an appreciation of adverse selection.

Problem Solving Tools

Recursion

Sometimes, in solving a problem, you'll run into an infinite series, where each term is smaller than the previous in some way. This often happens in games with repetitive play, like this:

Q: *Suppose you flip a coin until you see heads. How many flips do you expect to need?*

One way to set up your equation is to say, there's a 1/2 chance you only need 1 flip, a (1/2 * 1/2 = 1/4) chance you'll need 2 flips, and so on:

$$E[\text{flips}] = 1/2 * 1 + 1/4 * 2 + 1/8 * 3 + 1/16 * 4 + \ldots$$
$$= \Sigma_{i = 1 \text{ to } \infty} \; i/2\hat{\ }i$$

If you know the formula for evaluating that, great! If not, no worries, we can instead formulate this recursively.

Think about what actually happens when you flip. You expend 1 flip, and then you're in a situation where there's a 1/2 chance you're done (no more flips!), and a 1/2 chance you have to start all over again. If you have to reflip, you still expect the same number of flips as you did before this flip, because coin flips are independent of each other. When a process has the same distribution no matter what successes or failures you had earlier, we consider it **memoryless,** and often times it helps to think about it recursively.

This gives us:

$$E[\text{flips}] = 1 + 1/2 * 0 + 1/2 * E[\text{flips}]$$

Solving the above gives us E[flips] = 2. (In fact, in general, if something has a probability p, you'll expect to need 1 / p trials to get your first success.)

Q: *How many flips to you need to see 3 heads in a row? Hint: you can either write one big recursive equation, or figure out how many flips for 2 heads first and use that.*

Edge cases

When you come up with a solution that involves a generalization, it helps to try extreme values as inputs to check for correctness. What happens if you put in 0? 1? A negative number? Infinity? Do you end up with division by zero or some other crazy result?

Thinking about these edge cases can also give you insight into how to solve a problem by simplifying it. Small cases are often helpful -- for example, if you're struggling to calculate the result of a game with 100 players, you can take a step back and consider how 2 players would play, then add a player and see how that changes.

Interview Myths

There are lots of Jane Street interview tips floating around on the internet. Here are some common myths:

– *You have to be really good at mental math*

Numeracy certainly helps in our interviews, not only in getting to an answer, but also in deciding if your answer is reasonable (e.g. check the order of magnitude!). At the same time, we're not going to judge you harshly for not being able to mentally multiply double-digit numbers in 5 seconds.

– *You have to be really good at complicated math*

Although many of our questions are easier if you know a lot of math principles, we try to ask creative problems that you can reason through with minimal outside knowledge. In fact, we often prefer the simple, intuitive answer to the one that involves a bunch of esoteric theorems.

– *You "fail" the interview if you get the wrong answer*

The interview is meant to be a collaboration between you and your interviewer. If you're confused about the question, ask for clarification! Sometimes even when you try your best, you might not end up with the correct solution (let alone the most efficient one), and that's okay. We care about how you think about and approach tough problems, which includes how you take hints and respond to being wrong.

The flip side of this is, of course, that simply getting the right answer isn't enough. Your interviewer will likely want you to explain how you got to your answer and possibly ask follow-up questions.

– *You should do X*

You might've been told that interviewers want you to behave in a certain way -- super confident and assertive, really fast at answering questions, etc. For us, we just want you to be yourself. It's totally fine to say that you're not sure about your answer, and we'd much rather that you take your time and get to a better answer than say the first thing that comes to mind (unless specifically asked!).

All of your interviewers have gone through the same process, so we try to be sympathetic :)