
Chapter 4

Chance Makes Magic

Now chance enters the picture, whose mathematical background is provided by the theory of probability, which provides a number of results that are surprisingly counterintuitive. They are paradoxes of a sort, and this mismatch between the audience's expectations and the actual probabilities can be used to create some effective magic tricks.

There is, however, a new point of view that distinguishes the tricks presented here from those described in previous chapters: there is no 100% guarantee that they will actually work. Most of the time, you will experience success, but it will also happen that the magician will have to come up with a convincing reason why not everything went as planned and the trick has to be attempted a second time.

Many readers may hesitate to give chance a chance in their magic shows. But that would be a pity, and they would miss out on an interesting aspect of “the math behind the magic.”

In this chapter, we present two examples of how chance can be used in magic:

- The first trick we present goes back to the physicist Martin Kruskal. It rests on the fact that what are known as random walks show a surprising tendency to join together.

This phenomenon can be explained with elementary laws of probability theory.

- In the second trick, we consider a wager in what seems to be a fair situation, though in fact, the magician has considerably better odds of success. The basis is the surprising behavior of the notion of “better” in this connection.

It would be well at this point to consider three fundamental questions in connection with our tricks: How do mathematicians work with probabilities? How are probabilities calculated? What are “intransitive” orderings? The interested reader will find information on these questions in the following section and at the beginning of Section 4.3. Those who would prefer to go right to the tricks may skip to Section 4.2 and then return later to the foundational questions discussed here.

4.1. What Is Probability?

Chance plays an important role in our lives. Will we win the lottery? Will our train be an hour late? Will we meet a stranger some enchanted evening across a crowded room? Here we shall look only at situations in which we can calculate something concrete. That is the case, anyway, in the first of the examples below. In many other situations, it is necessary to invoke statistical methods to determine the relevant probabilities.

What, then, is “probability”? Let us consider a situation in which chance plays a role and in which some “experiment” is made that in principle could be repeated arbitrarily often. Here are some examples to illustrate the point:

- We throw two dice simultaneously and are interested in the sum of the two numbers that come up.



Will the sum be an even number?

- We fill out a lottery ticket. How many numbers will we guess correctly?
- A thumbtack is tossed into the air, and it lands somewhere on the table below. Will its point be pointing upward or downward?



Will the point be pointing upward or downward?

In such situations, there are certain results that may or may not occur:

- The sum of the dice is an even number. Or the sum is greater than 10.
- On our lottery ticket, we chose exactly three numbers correctly.
- The tack landed with the point pointing downward.

Each of these events will occur or not when we perform the relevant experiment. By the probability of such a result we mean a number p between zero and one with the following property: If one were to repeat the “experiment” a very large number of times, then the portion of the cases in which the given result occurs will be very close to p .

We do not wish to minimize the philosophical and epistemological problems inherent in this definition: Why is it justifiable to associate such a p with a particular result (one can never determine the number exactly)? What exactly do “a large number of times” and “very close” mean in the definition? Possible answers to these questions have been discussed extensively by philosophers and mathematicians without their having arrived at a fully satisfactory solution. The problem does not, however, engage mathematical practitioners particularly intensively. For them, the assumption that a probability can be associated with an event is simply a part of the mathematical model of reality. It works phenomenally well, and the lottery commission, the casino owner, the insurance company, and many others value the reliability of the theoretical predictions. And at some point, it is no longer a high priority to want to know what probability “actually” is.

Similar situations arise in other areas of intellectual pursuit. Physicists, for example, manage quite well with “point masses,” even though everyone knows that “in reality,” there are no such things. Models that rely on idealized assumptions can be found in many places in the natural sciences, and their success has justified this approach over many centuries.

Here are some examples of probabilities:

- The probability that the result of tossing two fair dice is an even number is 0.5. That is, if you were to repeat the experiment very often, the result would be an even number in very close to half of the cases. And the probability that the result will be a number greater than 10 (that is, 11 or 12) is $1/12$. (The justification for this will be given momentarily.)
- The probability of choosing three of six numbers correctly out of a list of 49 numbers, as in some lotteries, for example,

is about 1 in 57. And here are the probabilities of choosing four, five, and six correct numbers:

- probability of four correct: about 1 in 1000;
- probability of five correct: about 1 in 54 000;
- probability of six correct: about 1 in 14 million.



1	2	3	4	5	6	X
8	9	10	11	12	13	14
15	16	17	18	19	20	X
22	23	24	X	26	27	28
29	30	31	X	33	34	35
36	37	38	39	X	X	42
43	44	45	46	47	48	49

The probability in a lottery game with 49 numbers of choosing six numbers correctly is about 1 in 14 million.

It is beyond the scope of this book to explain the mathematics behind these numbers. Such mathematical techniques are taught in college courses on elementary probability and combinatorics.

- The probability of a thumbtack landing with the point pointing upward can be calculated approximately using statistical methods, but not by a direct calculation as in the previous examples. It depends on the dimensions and construction of the tack and even on the material of which the tabletop is constructed.

The human mind has generally a very poor intuition regarding very small probabilities, for example those of winning the lottery. My favorite illustration of this phenomenon is the following: Imagine walking down Broadway in Manhattan, New York, and noting a random

pedestrian. Then dial the Manhattan area code 212 followed by seven random digits. The probability that you have dialed the person you noted is about one in ten million, which is significantly more likely than winning the lottery.

Determining Probabilities

There are, in principle, two ways of determining probabilities. The first is to calculate them. There are two ideas behind this method.

Idea 1: If there is no reason to assume that any of the possible outcomes are more likely to occur than any of the others, then they should all have the same probability. And since all the probabilities must sum to 1, it follows that if there are n equally likely outcomes, then each of them occurs with probability $1/n$. Thus, for example, if a fair die is thrown, the equally likely outcomes 1, 2, 3, 4, 5, 6 each have probability $1/6$, and each lottery selection of six numbers out of 49 has probability $1/13\,983\,816 \approx 7 \cdot 10^{-8}$, since there are 13 983 816 ways of choosing six numbers from the numbers 1, 2, \dots , 49.

Idea 2: Count. In most situations, one is not interested in the probability that a particular event occurs, but rather in whether the result falls among a set of favorable outcomes. For example, you might be interested in the probability that a dice roll will beat your opponent's 10. To calculate this probability, you have only to calculate the number of ways of rolling an 11 or 12 and divide by the number of possible rolls of two dice. Since there are two ways to roll an 11 (5 on the first die and 6 on the second, and vice versa) and one way to roll a 12 (six and six), and 36 ways the dice can fall (six possibilities for the first die times six for the second), the probability of rolling greater than a 10 is $3/36 = 1/12$. Or what is the probability of choosing exactly three numbers correctly in the lottery? To determine this number, you would have to count how many lottery bets of six numbers contain exactly three of six given numbers and then divide by the total number of lottery bets, which is 13 983 816.

In general, we have the following formula: probability = the number of favorable outcomes divided by the number of possible outcomes.

This formula can be used whenever the result of interest is a collection of individual results all of which have the same probability. And by the way, there are 246 820 ways to get exactly three correct numbers out of six in the 49-number lottery, so the probability of getting exactly three numbers correct is

$$\frac{246820}{13983816} \approx 0.0176504.$$

You can also use statistical methods. By definition, the probability of a particular event is the fraction of the experiments in a long series of trials in which that event occurs. For example, if you don't know whether a coin is fair and would like to know the probability of its coming up heads, you could perform the experiment of tossing the coin one thousand times and note how many times the coin comes up heads. If, for instance, heads occurs 612 times, you will estimate the probability of heads to be $612/1000 = 0.612$. The reliability of such an estimate is something that is investigated in the field of mathematical statistics.

4.2. We Are All Friends: The Kruskal Magic Trick

The following trick can be performed with cards or dice. We describe here the version with cards.

The trick: A participant shuffles a deck of cards and lays them face up in a long row or a snake pattern (see the figure below). The cards are used for random walks. You begin at one of the first cards and proceed the number of steps as indicated on the card (aces count one step, jacks 2, queens 3, and kings 4). This is done until you would run off the end of the line (for example, if you arrive at a 4, but there are only two cards left in the line). The big surprise: no matter from which of the first few cards you begin, you (almost) always end up on the same card. And if you repeat the experiment but beginning near the end and going backward, you achieve a similar result. Here are some examples based on the arrangement of cards shown in the

figure:

- Start at the beginning on $\diamond J$:
 $\diamond J \heartsuit A \diamond 2 \heartsuit 5 \diamond A \diamond K \clubsuit 4 \clubsuit 2 \heartsuit 2 \spadesuit 5$
- Start near the beginning on $\heartsuit 3$:
 $\heartsuit 3 \spadesuit 3 \spadesuit 2 \diamond 3 \clubsuit Q \clubsuit 4 \clubsuit 2 \heartsuit 2 \spadesuit 5$
- Start at the end on $\clubsuit 3$:
 $\clubsuit 3 \spadesuit 5 \clubsuit J \diamond 4 \clubsuit Q \diamond 3 \clubsuit K \heartsuit A \heartsuit 3$
- Start near the end on $\spadesuit Q$:
 $\spadesuit Q \spadesuit J \heartsuit Q \spadesuit A \diamond 4 \clubsuit Q \diamond 3 \clubsuit K \heartsuit A \heartsuit 3$



If you begin at either of the first two cards, the last card you reach is $\spadesuit 5$, while if you begin with either of the last two cards, you end up with $\heartsuit 3$.

Preparation: The pack of cards used for this trick must have no high-value cards. So remove all cards with values 6, 7, 8, 9, 10. This improves the odds, already quite good, that the trick will work.

Execution: There should be a good number of cards—more is better than less—so use a bridge deck or even a canasta deck. There should be about 25 to 50 cards.

The cards are then shuffled by a participant and laid out face up in a snaking line, as described and pictured above. The rules for the random walk, as explained above, are now told to the audience. And now the fun begins. You, the magician, begin with a practice run. You begin the walk (for example with a game token, checker, or chess piece) on the first card and proceed according to the rules. When you can move no farther, the token remains on the last card it reached.

And now the audience can take turns with their own game tokens. They begin at cards number 2, 3, ... (always at a card near the beginning), and they will end up, with overwhelming probability, at the same card as the magician's. It may be explained simply as a surprising phenomenon, but you could also prepare the ground by casting a magic spell: "I will use my magical powers to bring your token to the same square as mine."

The math behind the magic: To understand the underlying mathematics, you will have to absorb an additional fact from probability theory:

The Product Theorem Rule for Probabilities

Suppose we are interested in outcomes A and B in some experiment involving chance. As an example, let us consider tossing three dice, one red, one white, and one blue, and suppose that outcome A is "the red die shows an even number," while outcome B is "the sum of the white and blue dice is greater than 10." Then A and B each have a probability: for A , it is $1/2$, while for B , it is, as we have shown earlier, $3/36$.

But now we would like to know the probability that both events occur on a single throw of the three dice, that is, that both A and B occur: A : the red die shows an even number, and B : the sum of the white and blue dice is greater than 10. We can determine this "joint" probability using the following rule:

Suppose A and B are independent outcomes in the sense that each has no effect on the other. This is surely the case in our example, since the red die "knows nothing" about what comes up on the other two dice, and vice versa. Then the probability that both A and B occur is the product of the individual probabilities for A and B .

In our example, then, the probability that the red die shows an even number and the white and blue dice sum to a number greater than 10 is equal to $1/2 \cdot 3/36 = 1/24$ (that is, a bit more than four percent).

More generally, if events A_1, \dots, A_n are independent in the sense described above, then the probability that they all occur is the product of the individual probabilities.

As an aside, the precise definition of independent probabilities is somewhat technical, and it takes some getting used to even for college math majors. Here are two illustrative examples:

- If a number of dice are thrown in succession, the results on the individual dice have no influence over each other. For the layperson, the following fact is difficult to accept: even if no six has appeared in the last ten throws, the probability that the next throw will be a six is the same as ever: $1/6$.
- Now let us consider throwing two dice—a red one and a blue one—and let A be the result that the red die comes up six, and B the result that the sum of the two dice is greater than 10. In this case, the two events are clearly not independent. If I know that the red die shows a six, then the odds are greatly improved that the sum of the two dice will be large. And conversely, if I know that the red die shows one of 1, 2, 3, 4, then I don't even need to look at the blue die to know that the sum will not be greater than 10.

We are going to employ this principle for our random walk. Let us imagine two random walks beginning on the first and second cards. What is the probability that they will end up in the same place? To answer this question, let us combine the following two observations:

- Suppose that each of the two walks lands on one and the same particular card before the end. Then it is clear by the rules of the game that they will land on the same cards from that point on, and in particular, they will end on the same card.

- “The walks meet” is the negation of “they never meet,” and therefore, the probability that they meet is equal to 1 minus the probability that they never meet.

For the trick to succeed, it would be good to have the probability of the walks not meeting as small as possible.

And how large is this probability? It depends on whether the walks fail to meet at the first card, fail to meet at the second card, and so on. Altogether, there are k steps to consider, where k depends on the total number of cards.¹

Thus it is clear that one should lay out as many cards as possible to ensure that the number of steps is sufficiently large.

The analysis for each individual step is as follows. Player 1, beginning at card 1, has visited certain cards that are at most five steps apart (since we have chosen 5 as the maximal step size). If player 2, starting at card number 2, has not yet arrived at one of the squares visited by player 1, then player 2 can see ahead, at a distance at most 4, a square that the other player has occupied. Thus player 2 has a chance at arriving in the next step at a square visited by player 1, and that probability is at least $1/5$. And conversely, the chance of a miss is at most $4/5$. Considering now the product rule given above, we arrive at the following conclusion. The probability that player 2 starting at square 2 never lands on a card visited by player 1 starting at square 1 is at most $(4/5)^k$ if k is the expected number of steps. As k increases, this quantity gets very small very quickly.²

We have shown, then, that the probability that both players end up on the same last card is quite high. It is positively influenced by the number of cards (hence a higher value of k) and the maximal step size. (If h is the maximal step, then $((h-1)/h)^k$ is an upper bound on the probability of not coinciding in k steps, and this value grows smaller as k grows larger.)

¹If the highest card value is 5 and there are 50 cards, then there will be at least 10 steps. So k is equal to at least 10.

²For example, $(4/5)^{20} = 0.0115\dots$. That is, the probability that the two players will not meet within twenty steps is just over one percent. The probability that they won't meet in ten steps is at most about ten percent.

Let me end with a small confession. I have cheated a bit. The conditions under which the product rule applies have not, strictly speaking, been satisfied, since the likelihood of encountering a particular card depends on what cards have already been encountered. If, for example, all the fours have appeared near the beginning, then you know that no more of them will appear. Thus the probabilities of encountering the various cards are not independent. However, the effect on the relevant probabilities is minimal.

Presentation: We have already made one suggestion for the presentation, in the section on execution: magician and participant end up on the same card. But you could also have several participants, beginning, say, on the first, second, and third cards, and you could compel them “by magic” to end on the same card.

A spectacular variant was demonstrated by the mathematician Steve Humble at an event for the popularization of mathematics in Cracow in 2012. He had prepared some gigantic playing cards, which were laid out on the ground. The random walks were then actually executed by several people by walking along the path of cards, and it was quite a surprise when they all found themselves together on the same final spot. It was rather crowded: “We are all friends!”



We all are friends!

Variants: 1. In this variant, the magician lays out the cards from a well-shuffled deck in a long row. In doing so, he pays particular attention to the first card, from which the walk will commence. If the card is a five, he lays out five additional cards, without considering their values. Then he lays out as many cards as the value of the last of the five cards, and so on. At some point, he comes to a card whose value is such that there is an insufficient number of remaining cards to lay out. It is this card that the magician takes note of, and then he lays down the remaining cards. This should all happen casually, as though the cards were simply being laid down one after the other.

Now it is time for a prediction: where will a random walker starting at one of the first cards end up? (The starting position can be determined by a throw of a die.) A participant then carries out the walk, and with high probability, the prediction will have been correct.

2. The magician could also shuffle a pack of cards and assert that she can memorize the order of the cards. There should be a large number of cards, because that will both heighten the impression on the audience and increase the probability of success. The magician fans through the cards face up, presumably to memorize the cards. But what she is really doing is mentally making the random walk, inconspicuously, and without altering the order of the cards. She notes the final card of the walk.

Now a participant enters the game. He throws a die and removes the number of cards from the top of the pack indicated by the number on the die. With the new top card as starting position, he makes the random walk, either by laying out the cards or by laying a few cards at a time on the table. The audience checks that it is done correctly; the magician has turned her back on the procedure. The chances are good that the last card in the walk is the one noted by the magician.

This last card is not revealed to the magician, but she asserts that she can repeat the walk in her thoughts, since she has memorized the cards. She acts as though she were doing so, and then confidently announces what the last card was, and with high probability it will be the card at which the participant's walk ended.

3. If you have a large number of dice available—say about thirty—you can use them for the Kruskal trick. Roll them all at once and arrange them in a row. Then proceed as above: begin with one of the first dice and proceed as many dice beyond as the number shown on the die. In this variant as well, all of your walks beginning at one of the first few dice will end (with very high probability) at the same die. In this variant, moreover, the conditions of the product rule are precisely satisfied.

4. If need be, you could perform the trick with a single die. Throw it a number of times (30, for instance) and note the result on a piece of paper. Then proceed as in the previous variant, of course taking the walk not on dice, but on the numbers written on the paper.

Here is an example, in which the die produced the following numbers:

2 4 3 1 4 6 4 5 2 3 1 4 6 2 3 1 4 3 5 3 2 4 3 1 4 5 2 1 4 2

Every walker beginning at one of the initial squares ends up at the last **5**. For instance, if you begin at **2**, you obtain the following walk: **2 3 6 4 1 4 2 3 5**, while if you begin at **4**, you experience **4 6 4 1 4 2 3 5**.

4.3. I Always Win!

There are many situations in which comparisons are made: Peter is taller than Paul; Thérèse runs faster than Bertha; the cost of living is greater in Boston than in Dubuque. It is usually the case that two such pieces of information of the same type allow a third conclusion:

- If you know that Peter is taller than Paul and that Paul is taller than Federico, then you may conclude that Peter is taller than Federico.
- Thérèse runs faster than Bertha, and Bertha is faster than Patience. Then Thérèse is faster than Patience.

- Boston is more expensive than Dubuque, and Dubuque is more expensive than Plainfield. It follows that Boston is more expensive than Plainfield.

In mathematics, one calls an ordering in which such conclusions can always be drawn *transitive*. The most familiar example is the usual ordering of numbers: from $x < y$ and $y < z$ one may always conclude that $x < z$.

This is so obvious that in many situations in which there is a notion of “better,” we at once assume that the order is transitive. Yet surprisingly, that is not the case. An example that most of you will remember from your childhood is the game rock, paper, scissors, often played by two children in decision-making. At the count of three, each child reveals a hand formed to represent a rock (fist), paper (flat palm), or scissors (first two fingers spread out), and the winner is determined by the following rules:

- Scissors is better than paper (scissors cuts paper).
- Paper is better than rock (paper wraps rock).
- Rock is better than scissors (rock crushes scissors).

This relation “better than” is clearly not transitive; in particular, there is no best choice: whatever you choose, there is something else that is “better.”

You might think that such a phenomenon is confined to the odd children’s game and does not appear in formal mathematics. But that is untrue. In fact, in probability theory there are many orderings that are not transitive.

Here is an example that is well known among mathematicians:

Intransitive Dice

Imagine some ordinary dice, but instead of the usual 1 through 6 on their faces, there could be other numbers. You and a friend each have such a die. The game that you are playing is quite simple: you each throw your die at the same time, and the higher number wins. (The dice have been constructed in such a way that none of the numbers on one die appears on the other. There is therefore always a winner.)

It makes sense to call one die, W_1 , better than the other, W_2 , if the player holding W_1 wins more than fifty percent of the time over a long sequence of plays. But now for a surprising phenomenon: it is possible to number three dice W_1, W_2, W_3 in such a way that

- W_2 is better than W_1 ;
- W_3 is better than W_2 ;
- W_1 is better than W_3 .

The relation “is better than” is therefore not transitive in this case. In particular, you can ask one player which of the dice he would like to use in the game, and then the other player can always choose a better one.

How do we calculate this in a concrete case? Imagine that we have two dice in front of us, W and W' . One of them, W , will be given the numbers a_1, \dots, a_6 , and the other, W' , the numbers a'_1, \dots, a'_6 . What are the odds of winning for each competitor? Each result (a_i, a'_j) is possible, where i and j range through the numbers 1 to 6. That makes 36 cases, and we have to determine whether $a_i > a'_j$ or $a'_j > a_i$ occurs more frequently. In the first case, W will be the better die, while in the second case, it will be W' . (Should it happen that $a_i > a'_j$ occurs at the same frequency as $a'_j > a_i$, then the winning probabilities will be the same.)

Here is an example: let W be given the numbers 1, 2, 4, 10, 11, 12, while W' will have 3, 5, 6, 7, 8, 9. Then in 19 cases, a_i is bigger than a'_j , while in 17 cases, a'_j will be bigger than a_i . We conclude that W is better than W' .

After this preparation with two dice, let us move on to three, W_1, W_2, W_3 , where we shall see the paradox of the intransitive dice. We assign numbers to the dice as follows:

- W_1 is given the numbers 7, 8, 9, 10, 11, 12.
- W_2 has the numbers 1, 2, 13, 14, 15, 16.
- W_3 has the numbers 3, 4, 5, 6, 17, 18.

With an analysis like that given above, it is easy to see that indeed, W_2 is better than W_1 , W_3 is better than W_2 , and W_1 is better than W_3 .

We are now ready to see a magic trick that is based on intransitivity.

The trick: A participant and the magician play against each other. The game is played with a well-shuffled deck of cards.³ This is how the game is played:

- The participant chooses a list of three items from among the two colors red and black. For example, he might choose red–black–red.
- Now the magician chooses a list of three colors (it must be different), say red–red–black.
- And now the game begins. The cards are turned up one at a time on the table, and we are interested only in the color: red or black. The player whose pattern appears first wins a point. The turned-up cards are set aside, and the game continues with the remaining cards. (If the entire deck is used up, the cards are reshuffled and the game begins with the entire pack.)
- The first player to achieve five points is declared the winner.

The big surprise: the magician's odds—with a proper choice of sequence—are better than those of the participant, and therefore he will win in the majority of cases.

Preparation: All that is necessary is a sufficient number of playing cards.

³It is also possible to use a coin, a variant that will be described below.

Execution: The trick is based on an intransitive situation: the relation “pattern M_2 has a greater probability of success than M_1 ” is not transitive. And indeed, the magician can always choose a sequence better than the one chosen by the participant.

In order to provide a clear description of how the choice should be made, let us introduce the abbreviation R for red and B for black. And here is the rule:

If the participant chooses X–Y–Z, where each of X, Y, Z represents R or B, then the magician should choose Y^{opp} –X–Y, where Y^{opp} is the opposite of Y; that is, $Y^{opp} = R$ if $Y = B$, and $Y^{opp} = B$ if $Y = R$. For example, if the participant chooses B–R–B, then the magician should choose B–B–R.

Since that is somewhat abstract, here is a complete table of the eight possible participant’s choices with the magician’s responses, together with the magician’s odds of winning. (The interested reader will find the associated calculations in the appendix.)

Participant	RRR	RRB	RBR	RBB	BRR	BRB	BBR	BBB
Magician	BRR	BRR	RRB	RRB	BBR	BBR	RBB	RBB
Odds of Winning	7/8	3/4	2/3	2/3	2/3	2/3	3/4	7/8

The picture below shows an example in which the magician has a 3/4 (i.e., 75%) chance of winning:



If the participant chooses B–B–R, the magician chooses R–B–B.

The math behind the magic: From a naive point of view, there seems to be no reason that the choice of pattern should make a difference, since the cards appear in random order. But a little thought

Bets can be placed on any number of events: Which will appear first, red or black? First the sequence red–red or black–black? But those are too simple, and it is therefore better to consider sequences of length three. The game continues as described above.

Variants: In principle, the deck of cards can be replaced with a coin or a die. If we denote the sides of the coin by H (heads) and T (tails), then you can generate a sequence of H's and T's by a series of coin tosses (e.g., T–H–H–T–T–T–H–T–H). The participant chooses a sequence of length three (say T–T–H), and the magician then chooses a sequence that gives her a better than even chance of winning (here it would be H–T–T). You have simply to use the rule given above, with R replaced by H, and B by T.

You could also play with a single die by considering not the values of the numbers that turn up but only whether the number is even (E) or odd (O). The participant and the magician choose a sequence of E's and O's, where the optimal strategy for the magician is as indicated above.