The Mathematics of Game Shows

Frank Thorne

September 13, 2016

1 Introduction

We will begin by watching a few game show clips and seeing a little bit of the math behind them.

1.1 Example: The Price Is Right, Contestants' Row

We begin with the following clip from The Price Is Right:

`https://www.youtube.com/watch?v=TmKP1a03E2g`

Game Description (Contestants' Row - The Price Is Right): *Four contestants are shown an item up for bid. In order, each guesses its price (in whole dollars). You can't use a guess that a previous contestant used. The winner is the contestant who bids the closest to the actual price without going over.*

In this clip, the contestants are shown some scuba equipment, and they bid 750, 875, 500, and 900 in that order. The actual price is \$994, and the fourth contestant wins. What can we say about the contestants' strategy?

- As a first step, it is useful to precisely describe the results of the bidding: the first contestant wins if the price is in $[750, 874]$[1]; the second, if the price is in $[875, 899]$; the third, in $[500, 749]$; the fourth, in $[900, \infty)$. If the price is less than \$500, then all the bids are cleared and the contestants start over.

  We can see who did well before we learn how much the scuba gear costs. Clearly, the fourth contestant did well. If the gear is worth anything more than \$900 (which is plausible), then she wins. The third contestant also did well: he is left with a large range of winning prices – 250 of them to be precise. The second contestant didn't fare well at all: although his bid was close to the actual price, he is left with a very small winning range. This is not his fault: it is a big disadvantage to go early.

[1] Recall that $[a, b]$ is mathematical notation for all the numbers between a and b.

- The next question to ask is: could any of the contestants have done better?

  We begin with the fourth contestant. Here the answer is *yes*, and her strategy is **dominated** by a bid of \$876, which would win in the price range $[876, \infty)$. In other words: *a bid of \$876 would win every time a bid of \$900 would, but not vice versa.* Therefore it is better to instead bid \$876 if she believes the scuba gear is more than \$900.

  Taking this analysis further, we see that there are exactly four bids that make sense: 876, 751, 501, or 1. Note that each of these bids, except for the one-dollar bid, screws over one of her competitors, and this is not an accident: Contestant's Row is a **zero-sum game** – if someone else wins, you lose. If you win, everyone else loses.

- The analysis gets much more subtle if we look at the *third* contestant's options. **Assume that the fourth contestant will play optimally.** (Of course this assumption is very often not true in practice.

  Suppose, for example, that the third contestant believes that the scuba gear costs around \$1000. The previous bids were \$750 and \$875. Should he follow the same reasoning and bid \$876? Maybe, but this exposes him to a devastating bid of \$877.

There is much more to say here, but we go on to a different example.

1.2 Deal or No Deal

Here is a clip of the game show **Deal or No Deal**:

`https://www.youtube.com/watch?v=I3BzYiCSTo8`

The action starts around 4:00.

Game Description (Deal or No Deal): *There are 26 briefcases, each of which contains a variable amount of money from \$0.01 to \$1,000,000, totalling \$3,418,416.01, and averaging \$131477.53. The highest prizes are \$500,000, \$750,000, and \$1,000,000.*

The contestant chooses one briefcase and keeps it. Then, one at a time, the contestant chooses other briefcases to open, and sees how much money is in each (and therefore establishes that these are not the prizes in his/her own briefcase). Periodically, the 'bank' offers to buy the contestant out, and give him/her a fixed amount of money to quit playing. The contestant either accepts or says 'no deal' and continues playing.

The **expected value** of the game is the average amount of money you expect to win. (We'll have much more to say about this.) So, at the beginning, the expected value of the game is \$131477.53, presuming the contestant rejects all the deals. In theory, that means that the contestant should be equally happy to play the game or to receive \$131477.53. (Of course, this might not be true in practice.)

Now let's look at the game after he chooses six briefcases. The twenty remaining contain a total of \$2936366, or an average of \$146818. The expected value has gone up, because the contestant eliminated mostly small prizes and none of the three biggest. If he wants to maximize his expected value (and I repeat that this won't necessarily be the case), then all he has to know is that

$$146818 > 51000$$

and so he keeps playing.

The show keeps going like this. After five more cases are eliminated, he again gets lucky and is left with fifteen cases containing a total of $2808416, so an average of $187227. The bank's offer is $125,000 which he refuses. And it keeps going.

1.3 Jeopardy – Final Jeopardy

Here we see the Final Jeopardy round of the popular show Jeopardy:

`https://www.youtube.com/watch?v=DAsWPOuF4Fk`

Game Description (Jeopardy, Final Round): *Three contestants start with a variable amoung of money (which they earned in the previous two rounds). They are shown a category, and are asked how much they wish to wager on the final round. The contestants make their wagers privately and independently.*

After they make their wagers, the contestants are asked a trivia question. Anyone answering correctly gains the amount of their wager; anyone answering incorrectly loses it.

Perhaps here an English class would be more useful than a math class! This game is difficult to analyze; unlike our two previous examples, the players play *simultaneously* rather than *sequentially*.

In this clip, the contestants start off with $9,400, $23,000, and $11,200 respectively. It transpires that nobody knew who said that *the funeral baked meats did coldly furnish forth the marriage tables*. (Richard II? Really? When in doubt, guess Hamlet.) The contestants big respectively $1801, $215, and $7601.

We will save a thorough analysis for later, but we will make one note now: the second contestant can obviously win. If his bid is less than $600, he will end up with more than $22,400.

2 Probability

2.1 Sample Spaces and Events

At the foundation of any discussion of game show strategies is a discussion of *probability*. You have already seen this informally, and we will work with this notion somewhat more formally.

Definition 2.1 1. *A* **sample space** *is the set of all possible outcomes of a some process.*

2. An **event** *is any subset of the sample space.*

Example 2.2 *You roll a die. The sample space consists of all numbers between one and six. Using formal mathematical notation, we can write*

$$S = \{1, 2, 3, 4, 5, 6\}.$$

We can use the notation $\{\dots\}$ *to describe a set and we simply list the elements in it.*

Let E *be the event that you roll an even number. Then we can write*

$$E = \{2, 4, 6\}.$$

Alternatively, we can write

$$E = \{x \in S \ : \ x \text{ is even}\}.$$

Both of these are correct.

Example 2.3 *You choose at random a card from a poker deck. The sample space is the set of all 52 cards in the deck. We could write it*

$$\begin{aligned} S = \{ & A\clubsuit, K\clubsuit, Q\clubsuit, J\clubsuit, 10\clubsuit, 9\clubsuit, 8\clubsuit, 7\clubsuit, 6\clubsuit, 5\clubsuit, 4\clubsuit, 3\clubsuit, 2\clubsuit, \\ & A\diamondsuit, K\diamondsuit, Q\diamondsuit, J\diamondsuit, 10\diamondsuit, 9\diamondsuit, 8\diamondsuit, 7\diamondsuit, 6\diamondsuit, 5\diamondsuit, 4\diamondsuit, 3\diamondsuit, 2\diamondsuit, \\ & A\heartsuit, K\heartsuit, Q\heartsuit, J\heartsuit, 10\heartsuit, 9\heartsuit, 8\heartsuit, 7\heartsuit, 6\heartsuit, 5\heartsuit, 4\heartsuit, 3\heartsuit, 2\heartsuit, \\ & A\spadesuit, K\spadesuit, Q\spadesuit, J\spadesuit, 10\spadesuit, 9\spadesuit, 8\spadesuit, 7\spadesuit, 6\spadesuit, 5\spadesuit, 4\spadesuit, 3\spadesuit, 2\spadesuit\} \end{aligned}$$

but writing all of that out is annoying. An English description is probably better.

Example 2.4 *You choose two cards at random from a poker deck. Then the sample space is the set of all pairs of cards in the deck. For example,* $A\spadesuit A\heartsuit$ *and* $7\clubsuit 2\diamondsuit$ *are elements of this sample space,*

This is definitely too long to write out every element, so here an English description is probably better. (There are exactly 1,326 elements in this sample space.) Some events are easier to describe – for example, the event that you get a pair of aces can be written

$$E = \{A\spadesuit A\heartsuit, A\spadesuit A\diamondsuit, A\spadesuit A\clubsuit, A\heartsuit A\diamondsuit, A\heartsuit A\clubsuit, A\clubsuit A\diamondsuit\}$$

and has six elements. If you are playing Texas Hold'em, your odds of being dealt a pair of aces is exactly $\frac{6}{1326} = \frac{1}{221}$*, or a little under half a percent.*

Our next example is taken from the following TPIR clip:

`https://www.youtube.com/watch?v=TR7Smevj1AQ`

Game Description (Squeeze Play (The Price Is Right)): *You are shown a prize, and a five- or six-digit number. The price of the prize is this number with one of the digits removed, other than the first or the last.*

The contestant is asked to remove one digit. If the remaining number is the price, the contestant wins the prize.

In this clip the contestant is shown the number 114032. Can we describe the game in terms of a sample space?

It is imporrtant to recognize that **this question is not precisely defined. Your answer will depend on your interpretation of the question!** This is probably very much *not* what you are used to from a math class.

Here's one possible interpretation. Either the contestant wins or loses, so we can describe the sample space as

$$S = \{\text{you win}, \text{you lose}\}.$$

Logically there is nothing wrong with this. But it doesn't tell us very much about the structure of the game, does it?

Here is an answer I like better. We write

$$S = \{14032, 11032, 11432, 11402\},$$

where we've written 14032 as shorthand for 'the price of the prize is 14032'.

Another correct answer is

$$S = \{2, 3, 4, 5\},$$

where here 2 is shorthand for 'the price of the prize has the second digit removed.'

Still another correct answer is

$$S = \{1, 4, 0, 3\},$$

where here 1 is shorthand for 'the price of the prize has the 1 removed.'

All of these answers make sense, and all of them require an accompanying explanation to understand what they mean.

The contestant chooses to have the 0 removed. So the event that the contestant wins can be described as $E = \{11432\}$, $E = \{4\}$, or $E = \{0\}$, depending on which way you wrote the sample space. (Don't mix and match! Once you choose how to write your sample space, you need to describe your events in the same way.) If all the possibilities are equally likely, the contestant has a one in four chance of winning.

The contest guesses correctly and is on his way to Patagonia!

Notation 2.5 *If S is any set (for example a sample space or an event), write $N(S)$ for the number of elements in it. In this course we will always assume this number is finite.*

Probability Rule: All Outcomes are Equally Likely. Suppose S is a sample space in which all outcomes are equally likely, and E is an event in S. Then the **probability of E, denoted** $P(E)$, is

$$P(E) = \frac{N(E)}{N(S)}.$$

Example 2.6 *You roll a die, so $S = \{1, 2, 3, 4, 5, 6\}$.*

1. *Let E be the event that you roll a 4, i.e., $E = \{4\}$. Then $P(E) = \frac{1}{6}$.*

2. *Let E be the event that you roll an odd number, i.e., $E = \{1, 3, 5\}$. Then $P(E) = \frac{3}{6} = \frac{1}{2}$.*

Example 2.7 *You draw one card from a deck, with S as before.*

1. *Let E be the event that you draw a spade. Then $N(E) = 13$ and $P(E) = \frac{13}{52} = \frac{1}{4}$.*

2. *Let E be the event that you draw an ace. Then $N(E) = 4$ and $P(E) = \frac{4}{52} = \frac{1}{13}$.*

3. *Let E be the event that you draw an ace or a spade. What is $N(E)$? There are thirteen spades in the deck, and there are three aces which are not spades. Don't double count the ace of spades!*

   So $N(E) = 16$ and $P(E) = \frac{16}{52} = \frac{4}{13}$.

Example 2.8 *In a game of Texas Hold'em, you are dealt two cards at random in first position. You decide to raise with a pair of sixes or higher, ace-king, or ace-queen, and to fold otherwise.*

The sample space has 1326 elements in it. The event of two-card hands which you are willing to raise has 86 elements in it. (If you like, write them all out. Later we will discuss how this number can be computed more efficiently!)

Since all two card hands are equally likely, the probability that you raise is $\frac{86}{1326}$, or around one in fifteen.

Now, here is an important example: You roll two dice and sum the totals. What is the probability that you roll a 7?

The result can be anywhere from 2 to 12, so we have

$$S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$$

and $E = \{7\}$. ☠☠☠: Therefore, $P(E) = \frac{N(E)}{N(S)} = \frac{1}{11}$.

Here is another solution. We can roll anything from 1 to 6 on the first die, and the same for the second die, so we have

$$\begin{aligned} S = \{&11, 12, 13, 14, 15, 16, 21, 22, 23, 24, 25, 26, 31, 32, 33, 34, 35, 36, \\ &41, 42, 43, 44, 45, 46, 51, 52, 53, 54, 55, 56, 61, 62, 63, 64, 65, 66\}. \end{aligned}$$

We list all the possibilities that add to 7:

$$E = \{16, 25, 34, 43, 52, 61\}$$

And so $P(E) = \frac{6}{36} = \frac{1}{6}$.

☠☠☠ We solved this problem two different ways and got two different answers. **The point is that not every outcome in a sample space will be equally likely**. We know that a die (if it is equally weighted) is equally likely to come up 1, 2, 3, 4, 5, or 6. So we can see that, according to our second interpretation, all the possibilities are still equally likely because all combinations are explicitly listed. But there is no reason why all the sums should be equally likely.

Note that it is often true that all outcomes are *approximately* equally likely, and we *model* this scenario by assuming that they are. *If our assumptions are close to the truth, so is our answer.*

For example, consider the trip to Patagonia. If we assume that all outcomes are equally likely, the contestant's guess has a 1 in 4 chance of winning. But the contestant correctly guessed that over \$14,000 was implausibly expensive, and around \$11,000 was more reasonable.

Another example comes from the TPIR game **Rat Race**:

`https://www.youtube.com/watch?v=Kp8rhV5PUMw`

Game Description (Rat Race (The Price Is Right)): *The game is played for three prizes: a small prize, a medium prize, and a car.*

There is a track with five wind-up rats (pink, yellow, blue, orange, and green). The contestant attempts to price three small items, and chooses one rat for each successful attempt. The rats then race. If he picked the third place rat, she wins the small prize; if she picked the second place rat, she wins the medium prize; if he picked the first place rat, she wins the car.

(Note that it is possible to win two or even all three prizes.)

Note that except for knowing the prices of the small items, there is no strategy. The rats are (we presume) equally likely to finish in any order.

In this example, the contestant correctly prices two of the items and picks the pink and orange rats.

Problem 1. *Compute the probability that she wins the car.*

Here's the painful solution: describe all possible orderings in which the rats could finish. We can describe the sample space as

$$S = \{POB, POR, POG, PBR, PBG, PRG, \dots, \dots\}$$

where the letters indicate the ordering of the first three rats to finish. Any such ordering is equally likely. The sample space has sixty elements, and twenty-four of them start with P or G. So the probability is $\frac{24}{60} = \frac{2}{5}$.

Do you see the easier solution? To answer the problem we were asked, we only care about the **first** rat. So let's ignore the second and third finishers, and write the sample space as

$$S = \{P, O, B, R, G\}.$$

The event that she wins is

$$E = \{P, G\},$$

and so $P(E) = \frac{N(E)}{N(S)} = \frac{2}{5}$.

Here's a possible solution that was suggested in class. It doesn't work, and it's very instructive to think about why it doesn't work. As the sample space, take all combinations of one rat and which order it finishes in:

$$\begin{aligned} S = \{&\text{Pink rat finishes first,} \\ &\text{Pink rat finishes second,} \\ &\text{Pink rat finishes third,} \\ &\text{Pink rat finishes fourth,} \\ &\text{Pink rat finishes fifth,} \\ &\text{Yellow rat finishes first,} \\ &\text{etc.}\} \end{aligned}$$

This sample space indeed lists a lot of different things that could happen. But how would you describe the event that the contestant wins? If the pink or orange rat finishes first, certainly she wins. But what if the yellow rat finishes third? Then maybe she wins, maybe she loses. There are several problems with this sample space:

- The events are not mutually exclusive. It can happen that **both** the pink rat finishes second, **and** the yellow rat finishes first. A sample space should be described so that **exactly one of the outcomes will occur**.

  Of course, a meteor could strike the television studio, and Drew, the contestant, the audience, and all five rats could explode in a giant fireball. But we're building *mathematical models* here, and so we can afford to ignore remote possibilities like this.

- In addition, you can't describe the event 'the contestant wins' as a subset of the sample space. What if the pink rat finishes fifth? The contestant also has the orange rat. It is ambigious whether this possibility should be part of the event or not.

Altogether, (from the learner's perspective) a very good wrong answer! Once you are very experienced, you will be able to skip straight to the correct answer. When you are just learning the material, your first idea will often be incorrect. Your willingness to critically examine your ideas, and to revise or reject them when needed, will lead you to the truth.

Problem 2. *Compute the probability that she wins both the car and the meal delivery.*

Here we care about the first *two* rats. We write

$$S = \{PO, PB, PR, PG, OP, OB, OR, OG, BP, BO, BR, BG, RP, RO, RB, RG, GP, GO, GB, GR\}.$$

The sample space has twenty elements in it. ($20 = 5 \times 4$: there are 5 possibilities for the first place finisher, and (once we know who wins) 4 for the second. More on this later.) The event that she wins is

$$\{PO, OP\}$$

and $P(E) = \frac{N(E)}{N(S)} = \frac{2}{20} = \frac{1}{10}$.

Problem 3. *Compute the probability that she wins all three prizes.*

Zero. Duh. She only won two rats! Sorry.

2.2 The Addition and Multiplication Rules

The Addition Rule (1). Suppose E and F are two *disjoint* events in the *same sample space* – i.e., they don't overlap. Then

$$P(E \text{ or } F) = P(E) + P(F).$$

Example 2.9 *You roll a die. Compute the probability that you roll either a 1, or a four or higher.*

Let $E = \{1\}$ be the event that you roll a 1, and $E = \{4, 5, 6\}$ be the event that you roll a 4 or higher. Then

$$P(E \text{ or } F) = P(E) + P(F) = \frac{1}{6} + \frac{3}{6} = \frac{4}{6} = \frac{2}{3}.$$

Example 2.10 *You draw a poker card at random. What is the probability you draw either a heart, or a black card which is a ten or higher?*

Let E be the event that you draw a heart. As before, $P(E) = \frac{13}{52}$.
Let F be the event that you draw a black card ten or higher, i.e.,

$$F = \{A\clubsuit, K\clubsuit, Q\clubsuit, J\clubsuit, 10\clubsuit, A\spadesuit, K\spadesuit, Q\spadesuit, J\spadesuit, 10\spadesuit\}.$$

Then $P(F) = \frac{10}{52}$.
So we have

$$P(E \text{ or } F) = \frac{13}{52} + \frac{10}{52} = \frac{23}{52}.$$

Example 2.11 *You draw a poker card at random. What is the probability you draw either a heart, or a red card which is a ten or higher?*

This doesn't have the same answer, because hearts are red. If we want to apply the addition rule, we have to do so carefully.
Let E be again the event that you draw a heart, with $P(E) = \frac{13}{52}$.
Now let F be the event that you draw a diamond which is ten or higher:

$$F = \{A\diamondsuit, K\diamondsuit, Q\diamondsuit, J\diamondsuit, 10\diamondsuit\}.$$

Now together E and F cover all the hearts and all the red cards at least ten, and there is no overlap. So we can use the addition rule.

$$P(E \text{ or } F) = P(E) + P(F) = \frac{13}{52} + \frac{5}{52} = \frac{18}{52}.$$

We can also use the addition rule with more than two events, as long as they don't overlap.

Example 2.12 *Consider the Rat Race contestant from earlier. What is the probability that she wins any two of the prizes?*

Solution 1. *We will give a solution using the addition rule. (Later, we will give another solution using the Multiplication Rule.)*

Recall that her chances of winning the car and the meal delivery were $\frac{1}{10}$. Let us call this event CM instead of E.

Now what are her chances of winning the car and the guitar? (Call this event CG.) Again $\frac{1}{10}$. If you like, you can work this question out in the same way. But it is best to observe that there is a natural symmetry in the problem. The rats are all alike and any ordering is equally likely. They don't know which prizes are in which lanes. So the probability has to be the same.

Finally, what is $P(MG)$, the probability that she wins the meal service and the guitar? Again $\frac{1}{10}$ for the same reason.

Finally, observe these events are all disjoint, because she can't possibly win more than two. So the probability is three times $\frac{1}{10}$, or $\frac{3}{10}$.

Here is a contrasting situation. Suppose the contestant had picked all three small prices correctly, and got to choose three of the rats. In this case, the probability she wins both the car and the meal service is $\frac{3}{10}$, rather than $\frac{1}{10}$. (You can either work out the details yourself, or else take my word for it.)

But this time the probability that she wins two prizes is *not* $\frac{3}{10} + \frac{3}{10} + \frac{3}{10}$, because now the events CM, CG, and MG are not disjoint: it is possible for her to win all three prizes, and if she does, then all of CM, CG, and MG occur!

It turns out that in this case the probability that she wins *at least* two is $\frac{7}{10}$, and the probability that she wins *exactly* two is $\frac{3}{5}$.

The Multiplication Rule. The multiplication rule computes the probability that two events E and F **both** occur. Here they are events in **diferent** sample spaces.

The formula is the following:

$$P(E \text{ and } F) = P(E) \times P(F).$$

It is not always valid, but it is valid in either of the following circumstances:

- The events E and F are *independent*.
- The probability given for F assumes that the event E occurs (or vice versa).

Example 2.13 *You flip a coin twice. What is the probability that you flip heads both times?*

We can use the multiplication rule for this. The probability that you flip heads if you flip a coin once is $\frac{1}{2}$. Since coin flips are independent (flipping heads the first time doesn't make *it more or less likely that you will flip heads the second time) we multiply the probabilities to get* $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$.

Alternatively, we can give a direct solution. Let

$$S = \{HH, HT, TH, TT\}$$

and

$$E = \{HH\}.$$

Since all outcomes are equally likely,

$$P(E) = \frac{N(E)}{N(S)} = \frac{1}{4}.$$

We can also use the multiplication rule for more than two events.

Example 2.14 *You flip a coin twenty times. What is the probability that you flip heads every time?*

If we use the multiplication rule, we see at once that the probability is

$$\frac{1}{2} \times \frac{1}{2} \times \cdots \times \frac{1}{2} = \frac{1}{2^{20}} = \frac{1}{1048576}.$$

This example will illustrate the second use of the Multiplication Rule.

Example 2.15 *Consider the Rat Race example again (as it happened in the video). What is the probability that the contestant wins both the car and the meal service?*

Solution. The probability that she wins the car is $\frac{2}{5}$, as it was before. So we need to now compute the probability that she wins the meal service, *given that she won the car.*

This time the sample space consists of *four* rats: we leave out whichever one won the car. The event is that her remaining one rat wins the meal service, and so the probability of this event is $\frac{1}{4}$.

By the multiplication rule, the total probability is

$$\frac{2}{5} \times \frac{1}{4} = \frac{1}{10}.$$

Example 2.16 *Suppose a Rat Race contestant prices all three prizes correctly and has the opportunity to race three rats. What is the probability she wins all three prizes?*

Solution. The probability she wins the car is $\frac{3}{5}$, as before: the sample space consists of the five rats, and the event that she wins consists of the three rats she chooses. (Her probability is $\frac{3}{5}$ no matter which rats she chooses, under our assumption that they finish in a random order.)

Now assume that she wins the first prize. Assuming this, the probability that she wins the meals is $\frac{2}{4} = \frac{1}{2}$ The sample space consists of the four rats *other than the first place finisher,* and the event that she wins the meals consists of the two rats *other than the first place finishers.*

Now assume that she wins the first and second prizes. The probability she wins the guitar is $\frac{1}{3}$: the sample space consists of the three rats *other than the first two finishers,* and the event that she wins the meals consists of the single rat *other than the first two finishers.*

There is some subtlety going on here! To illustrate this, consider the following:

Example 2.17 *Suppose a Rat Race contestant prices all three prizes correctly and has the opportunity to race three rats. What is the probability she wins the meal service?*

Solution. There are five rats in the sample space, she chooses three of them, and each of them is equally likely to finish second. So her probability is $\frac{3}{5}$ (same as her probability of winning the car).

But didn't we just compute that her odds of winning the car are $\frac{1}{2}$? What we're seeing is something we'll investigate much more later. This probability $\frac{1}{2}$ is a **conditional** probability: it assumes that one of the rats finished first, and illustrates what is hopefully intuitive: if she wins first place with one of her three rats, she is less likely to also win second place.

In particular, this reasoning illustrates the following misapplication of the multiplication rule. Suppose we compute again the probability that she wins all three prizes with three rats. She has a $\frac{3}{5}$ probability of winning first, a $\frac{3}{5}$ probability of winning second, and a $\frac{3}{5}$ probability of winning third. By the multiplication rule, the probability that all of these events occur is

$$\frac{3}{5} \times \frac{3}{5} \times \frac{3}{5} = \frac{27}{125}.$$

What is wrong with this reasoning is that these events are *not independent.*

Michael Larson. Here is a bit of game show history. The following clip comes from the game show Press Your Luck on May 19, 1984.

`https://www.youtube.com/watch?v=UzggoA4lLwk`

Here Michael Larsen smashed the all-time record by winning $110,237. The truly fascinating clip starts at around 17:00, where Larson continues to press his luck, to the host's increasing disbelief. On 28 consecutive spins, Larson avoided all the whammies and each time hit a space that afforded him an extra spin. There are eighteen squares on the board, and on average there are approximately five spaces worth money and an extra spin.

Example 2.18 *Assume for simplicity that each time there are exactly five spaces (out of eighteen) that Larson wants to hit, and that the outcome is random and that each square is equally likely to occur.*

If Larson spins twenty-eight times, compute the probability that he hits a good spot every time.

Solution. *This is a straightforward application of the multiplcation rule. The answer is* $\left(\frac{5}{18}\right)^{28}$*, or approximately one in*

$$3,771,117,128,139,603.$$

Either Larson got very, *very*, **very**, **VERY** lucky....... or else the pattern is not random and he figured it out.

Card Sharks. Here is another game show from the eighties that leads to interesting probability computations.

Game Description (Card Sharks): *Each of two contestants receives a lineup of five cards. The first is shown to each contestant, and a marker is placed on the first card. The objective of each* **round** *is to reach the last card.*

A **turn** *by the contestant consists of the following. She starts with the (face-up) card at the marker, and may replace it with a random card if she chooses. She then guesses whether the next card is higher or lower, which is then revealed.*

If is the last card and her guess is correct, she wins the round. Otherwise, she may keep guessing cards for as long as she likes untill one of three things happens: (1) she guesses the last card correctly, and wins; (2) she guesses any card incorrectly, in which case the cards she has guessed are all discarded and replaced with new cards (face down); (3) she chooses to end the turn by moving her marker forward to the last card guessed correctly.

The **round** *begins with a trivia question (I don't describe the rules for that here), and the winner gets to take a turn. If this turn ends with a freeze, the contestants go to another trivia question; if it ends with a loss, the other contestant takes a turn.*

There is also a *bonus round* which we won't discuss here. (We could though; analyzing this would make an interesting term project.)

Here[2] is a typical clip:

`https://www.youtube.com/watch?v=bUv0CRU6t5o`

Here is our objective: *Assuming that the trivia questions are a 50-50 tossup, determine the optimal strategy in all situations.* This problem is somewhat difficult (and our mental

[2] *Summary of the clip*: (**Please note.** The trivia questions are off-color and arguably sexist. This is unfortunately common on this show.) The contestants are Royce and Cynthia. Cynthia wins the first trivia question. Her initial card is a king. She keeps it and guesses lower; the second card is a two. She guesses higher; the third card is a nine. She freezes on position three.

Royce wins the next trivia question. His initial card is an eight; he changes it and gets a four. He guesses higher; the second card is a six. He guesses higher; the third card is a nine. He freezes on position three.

Royce wins the next trivia question. He starts on position three and chooses to replace the nine, and gets a three. He guesses higher; the fourth card is a five. He guesses higher; the fifth card is a king and Royce wins the round.

heuristics for it are fairly spot on). But at least in principle, it is possible to give a complete solution to this problem.

We won't try to achieve this all at once. Instead, we'll ask a number of probability questions to get started:

Example 2.19 *Consider Cynthia's first turn, where she guesses 'lower'. Compute the probability that she is correct.*

Answer. *The sample space consists of the* 51 *cards other than the king of clubs. Of these, only seven are not lower: the four aces, and the three remaining kings. So* $51 - 7 = 44$ *cards are lower, and her chances are* $\frac{44}{51}$.

We also compute the probabilities at the next two rounds. She guesses the third card will be higher than a 2. There are 50 cards remaining, and 47 of them are higher than a 2, so her odds are $\frac{47}{50}$.

The next card was a 9. Of the 49 remaining cards, 27 are lower than a 9 and 19 are higher. (And the three remaining nines are neither higher nor lower – so she would lose no matter what she picked). If she chose to play, her odds of winning the next card would be $\frac{27}{49}$, or slightly better than 50-50. She quite reasonably chooses to freeze and lock in her position.

Now we skip ahead to Royce's second round (when both Royce and Cynthia have frozen on the third of five cards).

Here are several questions we can ask:

- Given that Royce has replaced his nine with a three, compute the probability that he can win the round (assuming he doesn't freeze).
- Before Royce sees the three, compute the probability that he can win the round.
- Given that Royce's card is a five, compute the probability that he wins if he doesn't choose to freeze.
- If Royce chooses to freeze, answers the next trivia question correctly, and gets to go again, compute the probability that he wins on his next attempt.

  (Note that this is not the total probability he wins: he could lose on his next attempt, but then answer another trivia question correctly and get yet another try.)
- If Royce chooses to freeze and Cynthia answers the next trivia question correctly, what is the probability that she wins the next round (if she doesn't freeze)?

These questions get us closer to the question we're *really* interested in: should Royce freeze on the five or not? As is often the case, the question we are interested in is quite difficult and we build up to being able to answer it.

We tackle the first question.

Example 2.20 *Given that Royce has replaced his nine with a three, compute the probability that he can win the round. Assume that he doesn't choose to freeze, and that his higher/lower guess is always optimal.*

Note that there are 48 cards left in the deck: a three, a four, a six, and a nine are all missing.

It is easy to compute the probability that Royce's *first* guess is correct: out of 48 remaining cards, 41 are higher, so the probability is $\frac{41}{48}$. Now, **assuming that Royce's first guess is correct**, what is the probability that his second guess is correct?

Well ...... we don't know. It depends on what the first card **is**. Later, we will see some clever tricks for carrying out this sort of computation more easily. But for now, we outline a 'brute force' computation:

- Royce's first guess will be correct if the first card is a four, five, six, seven, eight, nine, ten, jack, queen, king, or ace.
- Based on Royce's first guess, we can determine what Royce should guess for the second card and the probability that this guess will be correct.

  Let's do an example of this. Suppose the first card is a four; the probability of this occurring is $\frac{3}{48}$. (This reduces to $\frac{1}{16}$, but the pattern will be clearer if we do not reduce our fractions to lowest terms.)

  Then Royce should clearly guess that the second will be higher. There are 47 remaining cards, of which 38 are higher than a four. So *assuming that the first card is a four*, the probability that Royce wins is $\frac{38}{47}$. Therefore, the probability that *the first card is a four* **and** *Royce wins* is $\frac{3}{48} \times \frac{38}{47}$.
- We will therefore use **both** the addition and the multiplication rules by **dividing into cases**: For each possible first card n (that doesn't lose Royce the round immediately), we compute the probability that the first card **is** n **and that** Royce wins the round. This is the multiplication rule.

  Since all of these possibilities are mutually exclusive, but one of them has to occur if Royce is to win, we see that the probability that Royce wins is the total of the probabilities we computed in the first step. This is the addition rule!

Let's roll up our sleeves and do it. The proof won't be pretty, but it is not as scary as it looks.

- With probability $\frac{3}{48}$ the first card will be a four. Then Royce should guess higher, and with probability $\frac{38}{47}$ the next card will be higher.
- With probability $\frac{4}{48}$ the first card will be a five. Then Royce should guess higher, and with probability $\frac{34}{47}$ the next card will be higher.
- With probability $\frac{3}{48}$ the first card will be a six. Then Royce should guess higher, and with probability $\frac{31}{47}$ the next card will be higher.

- With probability $\frac{4}{48}$ the first card will be a seven. Then Royce should guess higher, and with probability $\frac{27}{47}$ the next card will be higher.
- With probability $\frac{4}{48}$ the first card will be an eight. Then Royce should guess higher, and with probability $\frac{23}{47}$ the next card will be higher.
- With probability $\frac{3}{48}$ the first card will be a nine. Then Royce should guess lower, and with probability $\frac{24}{47}$ the next card will be lower.
- With probability $\frac{4}{48}$ the first card will be a ten. Then Royce should guess lower, and with probability $\frac{27}{47}$ the next card will be lower.
- With probability $\frac{4}{48}$ the first card will be a jack. Then Royce should guess lower, and with probability $\frac{31}{47}$ the next card will be lower.
- With probability $\frac{4}{48}$ the first card will be a queen. Then Royce should guess lower, and with probability $\frac{35}{47}$ the next card will be lower.
- With probability $\frac{4}{48}$ the first card will be a king. Then Royce should guess lower, and with probability $\frac{39}{47}$ the next card will be lower.
- With probability $\frac{4}{48}$ the first card will be an ace. Then Royce should guess lower, and with probability $\frac{43}{47}$ the next card will be lower.

(Note that all of the cases look more or less the same. Often, this is an indication that you can look for shortcuts – but we won't do so here.)

The total probability that Royce wins is therefore

$$\frac{3}{48}\cdot\frac{38}{47}+\frac{4}{48}\cdot\frac{34}{47}+\frac{3}{48}\cdot\frac{31}{47}+\frac{4}{48}\cdot\frac{27}{47}+\frac{4}{48}\cdot\frac{23}{47}+\frac{3}{48}\cdot\frac{24}{47}+\frac{4}{48}\cdot\frac{27}{47}+\frac{4}{48}\cdot\frac{31}{47}+\frac{4}{48}\cdot\frac{35}{47}+\frac{4}{48}\cdot\frac{39}{47}+\frac{4}{48}\cdot\frac{43}{47}.$$

This is equal to $\frac{1315}{2256}$, which is already in lowest terms. Yeah, I know. You were hoping it would be nice and simple, and that in retrospect you could have solved the problem in your head. You couldn't have. Neither could I. Sometimes math is like that.

This is roughly 58.2%, which is not bad at all.

2.3 Permutations and factorials

This video[3] illustrates a playing of the Price Is Right game **Ten Chances**:

[3] *Summary of the clip*: She plays Ten Chances for a pasta maker, a lawnmower, and a car. The digits in the pasta maker are 069, and she guesses the correct price of 90 on her second chance. The digits in the mower are 0689, and she guesses the correct price of 980 on her third chance. (Her third chance overall; she took only once to win the mower.) The digits in the car are 01568, and she guesses the correct price of 16,580 on her first try (and wins).

Barker then hides beyond the prop ... and, uh, (**trigger warning**) the contestant violates his personal space.

https://www.youtube.com/watch?v=iY_gmGcDKXE

Game Description (Ten Chances (The Price Is Right)): *The contestant is shown a small prize, a medium prize, and a large prize. She has ten chances to win as many prizes as she can.*

The price of small prize has two numbers in it, and the contestant is shown three different numbers. She then guesses the price of the first prize. She takes as many chances as she needs to.

Once she wins the small prize, she attempts to win the medium prize. The price of the medium prize has three numbers in it, and the contestant is shown four.

Finally, if she wins the medium prize, she attempts to win the car. Its price has five numbers in it, and the contestant is shown these five.

Example 2.21 *The price of the pasta maker contains two digits from* $\{0, 6, 9\}$. *Suppose that each possibility is equally likely to be the price of the pasta maker.*

If the contestant has one chance, what are her odds of winning?

Solution 1. We can give a straightforward solution by simply enumerating the sample space of all possibilities. It is

$$\{06, 09, 60, 69, 90, 96\}.$$

The contestant's choice describes an event with one of these possibilities in it. Since we hypothesized that each was equally likely to occur, her odds of winning are $\frac{1}{6}$.

Solution 2. We use the multiplication rule. There are three different possibilities for the first digit, and exactly one of them is correct. The probability that she gets the first digit correct is therefore $\frac{1}{3}$.

Now, **assume she got the first digit correct.** (If she didn't, she might have used up the correct second digit already, and be doomed to botch that one also!) Then there are two remaining digits, and the probability that she picks the correct one is $\frac{1}{2}$.

Thus the probability of getting both correct is $\frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$.

Notice, incidentally, that our assumption that the possibilities are equally likely is not realistic. Surely the pasta maker's price is not 06 dollars? Especially since you'd write it 6 and not 06? (Indeed, if you have watched the show a lot, you know that when there is a zero the price always ends with it. Knowing this fact is a *big* advantage.)

Now, she is going to use up at most six of her chances on the pasta maker, so she gets to move on to the mower. Here the price contains three digits from $\{0, 6, 8, 9\}$. This problem can be solved in the same way. The relevant sample space is

$$\{068, 069, 086, 089, 096, 098, 608, 609, 680, 689, 690, 698, 806, 809, 860, 869, 890, 896, 906, 908, 960, 968, 980, 986$$

which has 24 elements in it, so her probability of winning is $\frac{1}{24}$. The analogue of solution 2 gives $\frac{1}{4} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{24}$.

Finally, the price of the car has the digits $\{0, 1, 5, 6, 8\}$ and this time she uses all of them. The sample space is too long to effectively write out. So we work out the analogue of Solution 2: Her odds of guessing the first digit are $\frac{1}{5}$. If she does so, her odds of guessing the second digit is $\frac{1}{4}$ (since she has used one up). If both these digits are correct, her odds of guessing the third digit is $\frac{1}{3}$. If these three are correct, her odds of guessing the fourth digit are $\frac{1}{2}$. Finally, **if** the first four guesses are correct then the last digit is automatically correct by process of elimination. So the probability she wins is

$$\frac{1}{5} \times \frac{1}{4} \times \frac{1}{3} \times \frac{1}{2} \times 1 = \frac{1}{120}.$$

Here the number 120 is equal to 5!, or 5 **factorial**. In math, an exclamation point is read 'factorial' and it means the product of all the numbers up to that point. We have

$$\begin{aligned}
1! &= 1 && = 1 \\
2! &= 1 \times 2 && = 2 \\
3! &= 1 \times 2 \times 3 && = 6 \\
4! &= 1 \times 2 \times 3 \times 4 && = 24 \\
5! &= 1 \times 2 \times 3 \times 4 \times 5 && = 120 \\
6! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 && = 720 \\
7! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 && = 5040 \\
8! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 && = 40320 \\
9! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 && = 362880 \\
10! &= 1 \times 2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8 \times 9 \times 10 && = 3628800,
\end{aligned}$$

and so on. We also write $0! = 1$. Why 1 and not zero? 0! means 'don't multiply anything', and we think of 1 as the starting point for multiplication. (It is the *multiplicative identity*, satisfying $1 \times x = x$ for all x.) So when we compute 0! it means we didn't leave the starting point.

These numbers occur **very** commonly in the sorts of questions we have been considering, for reasons we will shortly see.

Example 2.22 *The lucky contestant wins the first two prizes in only three chances, and has seven chances left over. If each possibility for the price of the car is equally likely, then what is the probability that she wins it?*

The answer is seven divided by $N(S)$, the number of elements in the sample space. So if we could just compute $N(S)$, we'd be done.

Here there is a trick! She guesses 16580, and we know that the probability that this is correct is $\frac{1}{N(S)}$: one divided by the number of total possible guesses. But we already computed the probability: it's $\frac{1}{120}$. Therefore, we know that $N(S)$ is 120, without actually writing it all out!

The mathematical discipline of **combinatorics** is the art of *counting without counting.* We just solved our first combinatorics problem: we figured out that there were 120 ways to rearrange the numbers $0, 1, 5, 6, 8$ without actually listing them. We now formalize this principle.

Definition 2.23 *Let T be a* **string**. *For example,* 01568 *and* 22045 *are strings of numbers, ABC and xyz are strings of letters, and* $\otimes - \oplus\clubsuit$☠ *is a string of symbols. Order matters:* 01568 *is not the same string as* 05186.

A **permutation** *of T is any reordering of T.*

So, for example, if T is the string 1224, then 2124, 4122, 1224, and 2142 are all permutations of T. Note we *do* consider T itself to be a permutation of T, for the same reason that we consider 0 a number. It is called the **trivial permutation**.

We have the following:

Proposition 2.24 *Let T be a string with n distinct symbols. Then there are exactly $n!$ distinct permutations of T.*

In math, a *proposition* (or a *theorem*) is a statement of something true. We have stated lots of true facts in these notes; here the title 'Proposition' indicates that this one is particularly important and worth your attention.

Please read the statement carefully. In particular, the conclusion is only guaranteed to hold when the hypotheses also hold. If the hypotheses don't hold, then the conclusion may or may not be true. For example, if T is the string 122, then the set of all permutations of it is

$$\{122, 212, 221\}$$

which has 3 elements, and $3 \neq 3! = 6$.

Note also that this solves our earlier Ten Chances question. The contestant's guesses are all permutations of the string 01568, of which there are $5! = 120$. The sample space S consists of all 120 permutations. The contestant can make seven guesses, so let E be the set of these 7 permutations. Since we have assumed that each possible guess is equally likely to be correct, her odds (probability) of winning are $\frac{7}{120}$.

We will now offer a **proof** of the proposition. Please don't be too scared by the word 'proof': it just means a convincing explanation of why it is true. This course will not focus on *writing* proofs, but it is good to gain practice reading them.

Proof: Suppose T is a string with n distinct symbols, and we want to construct a permutation of T. symbol We first choose the first symbol. Since T has n distinct symbols, we have n choices for the first symbol.

No matter what we choose for the first symbol, there are $n - 1$ choices for the second symbol (all but the one we picked already), so that there are $n \times (n-1)$ choices for the first two.

Similarly, there are $n - 2$ choices for the third symbol, and so on. This continues until the last (the nth) symbol, for which there is exactly one choice. $\square$

In math we often end proofs with a little square. If you like, you can end proofs with the phrase **QED**, which is an abbreviation for 'quod erat demonstrandum' – Latin for 'that which was to be shown'. In practice, saying or writing 'QED' serves the same purpose as a football player spiking the ball after he has scored a touchdown.

If you are especially observant, you will notice that the proof is very similar to our explanation of the multiplication rule for probability. There is a good reason for this: the same principle underlies both, and counting and probability are two sides of the same coin.

We now return to our Ten Chances contestant. Recall that she has seven chances to win the car.

Example 2.25 *Suppose that the contestant has watched The Price Is Right a lot and so knows that the last digit is the zero. Compute the probability that she wins the car, given seven chances.*

Solution. Here her possible guesses consist of permutations of the string 1568, followed by a zero. There are $4! = 24$ of them, so her winning probability is $\frac{7}{24}$.

Her winning probability went up by a factor of exactly 5 – corresponding to the fact that $\frac{1}{5}$ of the permutations of 01568 have the zero in the last digit. Equivalently, a random permutation of 01568 has probability $\frac{1}{5}$ of having the zero ias the last digit.

Now, a smart contestant can do better. Suppose, for example, that she guessed 85610. Mathematically it looks like a good guess .... but she is playing for a Chevy Cavalier. I mean, really. We can rule out the 8 as the first digit, as well as the 6 and the 5.

Example 2.26 *Suppose that the contestant knows that the the last digit is the zero and the first digit is the one. Compute the probability that she wins the car, given seven chances.*

Solution. Her guesses now consist of permutations of the string 568, with a 1 in front and followed by a zero. There are $3! = 6$ of them. Assuming that the assumptions are correct and that she doesn't screw up, she is a sure bet to win the car.

Mathematically, her probability of winning is 1 (which is the same as 100%). Please *don't* answer that her probability is $\frac{7}{6}$. This doesn't make much sense!

Note that it is only true of Ten Chances that car prices always end in zero – not of The Price Is Right in general. Here is a contestant who is very excited until she realizes the odds she is against:

`https://www.youtube.com/watch?v=AAIU6knD7BA`

2.4 Exercises

Most of these should be relatively straightforward, but there are a couple of quite difficult exercises mixed in here for good measure.

1. Card questions. In each question, you choose at random a card from an ordinary deck. What is the probability you –

   (a) Draw a spade?

   (b) Draw an ace?

   (c) Draw a face card? (a jack, queen, king, or an ace)

   (d) Draw a spade or a card below five?

2. Dice questions:

   (a) You roll two dice and sum the total. What is the probability you roll exactly a five? At least a ten?

   Solution. The sample space consists of 36 possibilities, 11 through 66. The first event can be described as $\{14, 23, 32, 41\}$ and has probability $\frac{4}{36} = \frac{1}{9}$. The second can be described as $\{46, 55, 64, 56, 65, 66\}$ and has probability $\frac{6}{36} = \frac{1}{6}$.

   (b) You roll three dice and sum the total. What is the probability you roll at least a 14? (This question is kind of annoying if you do it by brute force. Can you be systematic?)

   Solution. There are several useful shortcuts. Here is a different way than presented in lecture. The sample space consists of $6 \times 6 \times 6 = 216$ elements, 111 through 666. The event of rolling at least a 14 can be described as

   $$\{266(3), 356(6), 366(3), 446(3), 455(3), 456(6), 466(3), 555(1), 556(3), 566(3), 666(1)\}.$$

   The number in parentheses counts the number of permutations of that dice roll, all of which count. For example, 266, 626, and 662 are the permutations of 266. There are 35 possibilities total, so the probability is $\frac{35}{216}$.

   (c) The dice game of *craps* is (in its most basic form) played as follows.

   You roll two dice. If you roll a 7 or 11 on your first roll, you win immediately, and if you roll a 2, 3, or 12 immediately, you lose immediately. Otherwise, your total is called “the point” and you continue to roll again until you roll either the point (again) or a seven. If you roll the point, you win; if you roll a seven, you lose.

   In a game of craps, compute the probability that you win on your first roll and the probability that you lose on your second roll.

   Solution. The probability of winning on your first roll is the probability of rolling a 7 or 11: $\frac{6}{36} + \frac{2}{36} = \frac{8}{36} = \frac{2}{9}$.
   For the second question, I intended to ask the probability that you lose on your *first* roll. Oops. Let’s answer the question as asked. There are multiple possible interpretations, and here is one. Let us compute the probability that you lose on the second round, presuming that the game goes on to a second round. This is the probability of rolling a 6 or $\frac{1}{6}$.

(d) In a game of craps, compute the probability that the game goes to a second round and you win on the second round.

Solution. This can happen in one of six possible ways: you roll a 4 twice in a row, a 5 twice in a row, or similarly with a 6, 8, 9, or 10.

The probability of rolling a 4 is $\frac{3}{36}$, so the probability of rolling a 4 twice in a row is $\left(\frac{3}{36}\right)^2$. Similarly with the other dice rolls; the total probability is

$$\left(\frac{3}{36}\right)^2 + \left(\frac{4}{36}\right)^2 + \left(\frac{5}{36}\right)^2 + \left(\frac{5}{36}\right)^2 + \left(\frac{4}{36}\right)^2 + \left(\frac{3}{36}\right)^2 = \frac{9+16+25+25+16+9}{1296} = \frac{100}{1296} = \frac{25}{324}. \tag{1}$$

(e) In a game of craps, compute the probability that the game goes to a second round and you lose on the second round.

Solution. Multiply the probability that the game goes onto a second round (easily checked to be $\frac{2}{3}$) by the probability $\frac{1}{6}$ computed earlier, so $\frac{1}{9}$.

(f) In a game of craps, compute the probability that you win.

Solution. With probability $\frac{2}{9}$ you win on your first round. We will now compute the probability that you win later, with the point equal to n, for n equal to 4, 5, 6, 8, 9, or 10. We will then add these six results. Write the probability of rolling n on one roll of two dice as $\frac{a}{36}$, so that a is 3, 4, or 5 depending on n.

- As we computed before, the probability of winning on the second round (with point n) is $\left(\frac{a}{36}\right)^2$.
- On each round after the first, there is a probability $\frac{30-a}{36}$ of rolling something other than 7 or the point. This is the probability that the game goes on to another round.
- So, the probability of winning on the third round is the probability of: rolling the point on the first round, going another turn in the second round, rolling the point on the third round. This is $\left(\frac{a}{36}\right)^2 \cdot \left(\frac{30-a}{36}\right)$.
- Similarly, the probability of winning with point n on the fourth round is $\left(\frac{a}{36}\right)^2 \cdot \left(\frac{30-a}{36}\right)^2$, and so on. The total of all these probabilities is

$$\left(\frac{a}{36}\right)^2 \sum_{k=0}^{\infty} \left(\frac{30-a}{36}\right)^k.$$

- For $|r| < 1$, we have the infinite sum formula $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$. Plugging this in, the above expression is

$$\left(\frac{a}{36}\right)^2 \cdot \frac{36}{6+a} = \frac{a^2}{36(6+a)}.$$

So we add this up for $a = 3$ (twice, for $n = 4$ or 5), $a = 4$ (twice), and $a = 5$ (twice). We get

$$2 \cdot \left(\frac{9}{36 \cdot 9} + \frac{16}{36 \cdot 10} + \frac{25}{36 \cdot 11} \right) = \frac{134}{495}.$$

Adding the to the first round probability of $\frac{2}{9}$ we get

$$\frac{2}{9} + \frac{134}{495} = \frac{244}{495}.$$

This is a little less than a half. As expected, the house wins.

3. Consider the game Press Your Luck described above. Assume (despite rather convincing evidence to the contrary) that the show is random, and that you are equally likely to stop on any square on the board.

   (a) On each spin, estimate the probability that you hit a Whammy. Justify your answer.

   (Note: This is mostly not a math question. You have to watch the video clip for awhile to answer it.)

   (b) On each spin, estimate the probability that you *do not* hit a Whammy.

   (c) If you spin three times in a row, what is the probability you don't hit a whammy? Five? Ten? Twenty-eight? (If your answer is a power of a fraction, please also use a calculator or a computer to give a decimal approximation.)

4. Consider the game Rat Race described above.

   (a) Suppose that the contestant only prices one item correctly, and so gets to pick one rat. What is the probability that she wins the car? That she wins *something*? That she wins nothing?

   (b) What if the customer prices all three items correctly? What is the probability that she wins the car? Something? Nothing? All three items?

   (c) Consider now the first part of the game, where the contestant is pricing each item. *Assume* that she has a 50-50 chance of pricing each item correctly. What is the probability she prices no items correctly? Exactly one? Exactly two? All three?

   Comment on whether you think this assumption is realistic.

   Solution. Foobar.

   (d) Suppose now that she has a 50-50 chance of pricing each item correctly, and she plays the game to the end. What is the probability she wins the car?

3 Expectation

3.1 Definitions and examples

We come now to the concept of **expected value**. We will give a few simple examples and then give a formal definition.

Example 3.1 *You play a simple dice game. You roll one die; if it comes up a six, you win 10 dollars; otherwise you win nothing. On average, how much do you expect to win?*

Solution. Ten dollars times the probability of winning, i.e.,

$$10 \times \frac{1}{6} = 1.66\ldots$$

So, for example, if you play this game a hundred times, on average you can expect to win 100 dollars.

Example 3.2 *You play a variant of the dice game above. You roll one die; if it comes up a six, you still win 10 dollars. But this time, if it doesn't come up a six, you lose two dollars. On average, how much do you expect to win?*

Solution. We take into account both possibilities. We multiply the events that you win 10 dollars or lose 2 dollars and multiply them by their probabilities. The answer is

$$10 \times \frac{1}{6} + (-2) \times \frac{5}{6} = 0.$$

On average you expect to break even.

Definition 3.3 *Consider a random process whose outcome can be described as a real number. Suppose that the possible outcomes are $a_1, a_2, \ldots a_n$, which occur with respective probabilities $p_1, p_2, \ldots, p_n$. Then the* **expected value** *of this process is*

$$\sum_{k=1}^{n} a_k p_k = a_1 p_1 + a_2 p_2 + \cdots + a_k p_k.$$

If the outcomes represent the amount of money you win (positive) or lose (negative), then the expected value is the amount you should expect to win on average.

Example 3.4 *You roll a die and win a dollar amount equal to your die roll. Compute the expected value of this game.*

Solution. The possible outcomes are that you win 1, 2, 3, 4, 5, or 6 dollars, and each happens with probability $\frac{1}{6}$. Therefore the expected value is

$$1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{21}{6} = 3.5.$$

As another example, consider the Deal or No Deal clip from the introduction.

`https://www.youtube.com/watch?v=I3BzYiCSTo8`

This is quite simple to analyze, and indeed we did so in the introduction.

For example, after the second round, he has eliminated 11 briefcases and 15 remain, which contain a total of $2,808,416, or an average of $187,227. If he keeps playing all the way until the end, the expected value is equal to the average of the remaining briefcases. The bank offers him a flat payment of $125,000 to quit. If he wants to maximize his expected value, he should refuse this, and indeed he does.

We now consider some expected value computations arising from the popular game show **Wheel of Fortune**.

Game Description (Wheel of Fortune, Simplified Version): *The contestants play several rounds where they try to solve word puzzles and win money. (The contestant who has won the most money then gets to play in a bonus round.)*

The puzzle consists of a phrase whose letters are all hidden. In turn, each contestant either **attempts to solve the puzzle** *or* **spins the wheel**. *If the contestant attempts to solve, he states a guess; if is correct, he wins all the money in his bank, and if it is wrong, play passes to the next player.*

The wheel contains lots of spaces with various dollar amounts or the word 'bankrupt'. When the contest spins, the wheel comes to rest on one of these spaces. If 'bankrupt', the contestant loses all his money from this round and play passes to the next contestant. Otherwise, the contestant chooses a letter. If that letter appears in the puzzle (and has not yet been guessed), then each of these letters is revealed and the contestant wins the amount of money on his space for each time it appears. If the letter does not appear, the contestant wins nothing and play passes to the next contestant.

These rules are incomplete: the contestants can 'buy a vowel'; there are non-monetary prizes on the board which work differently (you don't win more than one of them if a letter appears multiple times), other spaces like 'lose a turn', and so forth.

Consider the episode shown in this clip:

`https://www.youtube.com/watch?v=A8bZUXi7zDE`

Robert wins the first round in short order. After guessing only two letters (and buying a vowel) he chooses to solve the puzzle. Was his decision wise?

Let us make some assumptions to simplify the problem and set up an expected value computation:

- Robert wants to maximize the expected value of his winnings this round.

  This is not completely accurate, especially in the final round; the contestants are interested in winning *more than the other two contestants*, because the biggest winner gets to play the bonus round. But it is reasonably close to accurate, especially early in the running.

- Robert definitely knows the solution to the puzzle.

  So, if he chooses to spin again, it's to rack up the amount of prizes and money he wins.

- If Robert loses his turn, then he won't get another chance and will therefore lose everything.

  In fact, there is a chance that each of the other two contestants will guess wrongly or hit the 'bankrupt' or 'lose a turn' spots on the wheel. But this puzzle doesn't look hard: the first word *don't* is fairly obvious; also, the second word looks like *bet*, *get*, or *let* and B, G, and L are all in the puzzle. Robert is wise to assume he won't get another chance.

- We won't worry too much about the 'weird' spots on the board.

  The $\frac{1}{3}$-sized million dollar wedge is not what it looks like: it sits over (what I believe is) a \$500 wedge now, and offers the contestant the opportunity to win \$1,000,000 in the bonus round *if* he goes to the bonus round *and* doesn't hit bankrupt before then *and* solves the bonus puzzle correctly *and* chooses the million dollars randomly as one of five prizes. It's a long shot, although three contestants have indeed won the million.

So we freeze-frame the show and we count what we see. Out of 24 wedges on the wheel, there are:

- 16 ordinary money wedges on the wheel, with dollar amounts totalling \$12,200.
- Two 'bankrupt' wedges, a 'lose a turn' wedge, and an additional two thirds of a bankrupt wedge surrounding the million.
- A one-third size wedge reading 'one million'.
- The cruise wedge. This isn't relevant to the contestant's decision, because he wins the cruise and reveals an ordinary wedge underneath. We can't see what it is, so let's say \$500.
- Two other positive wedges.

Let us now compute the expected value of another spin at the wheel. There are (with the cruise wedge) 17 ordinary wedges worth a total of \$12,700. If the contestant hits 'bankrupt' or 'lose a turn' he loses his winnings so far (\$10,959 including the cruise). Let us guess that the million wedge is worth, on average, \$5,000 to the contestant and that the other two are worth \$2,000 each. His expected value from another spin is

$$\frac{1}{24} \cdot 12700 + \frac{2\frac{2}{3}}{24} \cdot (-10959) + \frac{2}{24} \cdot 2000 + \frac{\frac{1}{3}}{24} \cdot 5000 = -\$452.39.$$

It is clear by a large margin to solve the puzzle and lock in his winnings.

Remark 3.5 *You may be wondering where the* $\frac{1}{24} \cdot 12700$ *came from. Here is one way to see it: the seventeen wedges have an average of* $\frac{12700}{17}$ *dollars each, and there is a* $\frac{17}{24}$ *probability of hitting one of them. So the contribution is*

$$\frac{12700}{17} \times \frac{17}{24} = \frac{12700}{24}.$$

Now let us suppose that there was some consonant appearing in the puzzle twice. In that case Robert would know that he could guess it and get *double* the amount of money he spun. So, in our above computation, we double the 12700. (We should probably increae the 2000 and 5000 a little bit, but not double them. For simplicity's sake we'll leave them alone.) In this case the expected value of spinning again is

$$\frac{1}{24} \cdot 12700 \cdot 2 + \frac{2\frac{2}{3}}{24} \cdot (-10959) + \frac{2}{24} \cdot 2000 + \frac{\frac{1}{3}}{24} \cdot 5000 = -\$76.77,$$

so slightly positive. If Robert has the stomach to risk his winnings so far, he should consider spinning again.

For an example where Robert arguably chooses unwisely, skip ahead to 10:45 on the video (the third puzzle) where he solves the puzzle with only \$1,050 in the bank. In the exercises, you are asked to compute the expected value of another spin. Note that there are now two L's and two R's, so he can earn double the dollar value of whatever he lands on. There is now a \$10,000 square on the wheel, and hitting 'bankrupt' only risks his \$1,050. (His winnings from the first round are safe.)

There is one factor in favor of solving now: an extra prize (a trip to Bermuda) for the winner of the round. If it were me, I would definitely risk it. You do the math, and decide if you agree.

(But see the fourth run, where I would guess he knows the puzzle and is running up the score.)

The game **Punch a Bunch** from The Price Is Right has a similar (but much simpler) mechanic:

Game Description (Punch-a-Bunch (The Price Is Right)): *The contestant is shown a punching board which contains 50 slots with the following dollar amounts: 100 (5), 250 (10), 500 (10), 1000 (10), 2500 (8), 5000 (4), 10,000 (2), 25,000 (1). The contestant can earn up to four punches by pricing small items correclty. For each punch, the contestant punches out one hole in the board.*

The host proceeds through the holes punched one at a time. The host shows the contestant the amount of money he has won, and he has the option of either taking it and ending the game, or discarding and going on to the next hole.

So, if you just get one punch, there is no strategy: you just take whatever you get. In this case the expected value is the total of all the prizes divided by 50, or $\frac{103000}{50} = 2060$.

Here is a typical playing:

`https://www.youtube.com/watch?v=25THBiZNPpo`

The contestant gets three punches, throws away 500 on his first punch, 1000 on his second, and gets 10,000 on his third. Was he right to throw away the 1000?

Clearly yes, as the expected value of one punch is 2,060. Indeed, in this example it is a little bit higher: there is \$101,500 in prizes left in 48 holes, for an average of \$2,114.58. You don't have to do the math exactly: just remember that two of the small prizes are gone, so the average of the remaining ones goes up slightly.

So let's figure out optimal strategy for this game. The last two rounds are easy.

- On your last round, there is no strategy: you take whatever you get.
- On your next-to-last round, throw away anything less than \$2500. You should keep the \$2500 prize if you're trying to maximize your expected value. It's pretty close though; I wouldn't fault anyone who tried for the big prize. (If nothing else, it would make better TV.)
- What about your third-to-last round?

We are going to compute the expected value of the *next-to-last round*. We'll assume that this is also the contestant's *first* round; otherwise, the contestant will have thrown away one or two small prizes and the expected value will be slightly higher. (This is another example of where we simplify our problem by making such an assumption. In this case, the assumption is very nearly accurate.)

- The contestant might win \$25,000 ($\frac{1}{50}$ chance), \$10,000 ($\frac{2}{50}$ chance), \$5,000 ($\frac{4}{50}$ chance), or \$2,500 ($\frac{8}{50}$ chance). As we discussed earlier, the contestant should keep it and end the game.
- The contestant might draw a card less than \$2,500 ($\frac{35}{50}$ chance). As we discussed earlier, the contestant should throw it away. In this case, the contestant expects to win \$2,060 (in fact, slightly more, as previously discussed) on average from the last punch.

So the expected value of the next-to-last round is

$$25000 \cdot \frac{1}{50} + 10000 \cdot \frac{2}{50} + 5000 \cdot \frac{4}{50} + 2500 \cdot \frac{8}{50} + 2060 \cdot \frac{35}{50} = \$3,142.$$

So we see that on the contestant's third-to-last round, he should throw away the $2,500 cards in addition to everything cheaper, and only settle for $5,000 or more. The expected value of the third-to-last round is

$$25000 \cdot \frac{1}{50} + 10000 \cdot \frac{2}{50} + 5000 \cdot \frac{4}{50} + 3142 \cdot \frac{43}{50} = \$4,002.12.$$

Therefore, if the contestant gets four punches, his strategy on the first round should be the same: to keep anything $5,000 or more, and throw everything else away. The expected value of a four-round game is

$$25000 \cdot \frac{1}{50} + 10000 \cdot \frac{2}{50} + 5000 \cdot \frac{4}{50} + 4002 \cdot \frac{43}{50} = \$4,741.72.$$

A contestant can win only up to four punches. But we see that if the contestant got more, he would eventually throw away the $5,000 cards too.

Who Wants To Be a Millionaire? Here is a typical clip from the show:

`https://www.youtube.com/watch?v=sTGxOqp3qB8`

The rules in force for this episode were as follows.

Game Description (Who Wants to be a Millionaire?): *The contestant is provided with a sequence of* 15 *trivia questions, each of which is multiple choice with four possible answers. They are worth an increasing amount of money: 100, 200, 300, 500, and then (in thousands) 1, 2, 4, 6, 16, 32, 64, 125, 250, 500, 1000. (In fact, in this epsiode, the million dollar question was worth $2,060,000.)*

At each stage he is asked a trivia question for the next higher dollar amount. He can choose to answer, or to not answer and to keep his winnings so far. If he answers correctly, he goes to the next level. If he answers incorrectly, the game is over. At the $1,000 and $32,000 level his winnings are protected: he is guaranteed of winning at least that much money. Beyond that, he forfeits any winnings if he ventures an incorrect answer.

He has three 'lifelines', each of which may be used exactly once over the course of the game: '50-50', which eliminates two of the possible answers; 'phone a friend', allowing him to call a friend for help; and 'ask the audience', allowing him to poll the audience for their opinion.

In general we want to ask the following question:

Question. The contestant is at level x, and (after using any applicable lifelines) estimates that he has a probability δ of answering correctly. Should he guess or not?

Let us assume that $x \geq 32000$ (that's the interesting part of the show). Note that if $x = 32000$, he should always guess since he is risking nothing.

Suppose then that $x = 64000$, and for now we'll consider *only the next question.* We will work with δ as a variable, and so our answer will be of the form 'He should guess if he believes his probability of answering correctly is greater than [something].' His winnings will be 32000 if he is incorrect and 125000 if he is right; and these events have probability $1-\delta$ and δ respectively. Therefore, the expected value of guessing is

$$(1-\delta)\cdot 32000+\delta\cdot 125000=32000+\delta\cdot 93000.$$

When is this greater than 64000? We solve the inequality $32000+93000\delta>64000$, which is equivalent to $93000\delta>32000$, or $\delta>\frac{32000}{93000}=\frac{32}{93}$. This is a little bit bigger than $\frac{1}{3}$. So, random guessing would hurt the contestant, but if (for example) he can eliminate two of the answers, it makes sense for him to guess.

At the level $x=125000$, our computations are similar. This time we have to solve the inequality

$$32000+\delta\cdot(250000-32000)>125000,$$

which is equivalent to $\delta>\frac{93}{218}$. This is bigger, which makes sense: proportionally he is risking more – he would go down two levels, rather than just one.

Of course, working with only one question at a time is a little bit misleading. For example, consider the \$125,000 question. Even after phoning a friend (and using the last of his lifelines), he has no idea. If he will only go one more question, it is clearly correct to walk, but *what if the \$250,000 question is something he definitely knows?*

Let us go one step further in our analysis (and you can see how to do still better). Suppose that the contestant estimates that there is a 40% chance that the \$250,000 question is one he will know the answer to. If he does, he will guess it correctly, quit the next turn, and walk away with \$500,000. If he doesn't, he won't venture a guess and will walk away with \$500,000.

In this case, reaching the \$250,000 level is worth

$$0.4\times 500000+0.6\times 250000=350000.$$

So the contestant is risking \$93,000 to win another \$225,000. The expected value of guessing is

$$(1-\delta)\cdot 32000+\delta\cdot 350000=32000+\delta\cdot 318000,$$

and our inequality is

$$32000+318000\delta>125000,$$

which is equivalent to $\delta>\frac{93}{318}$. In this case it still doesn't make sense for him to randomly guess, but if his guess is even slightly better than random it does. (Moreover, if the contestant estimates that there is a small chance that he would know the answer to the \$500,000 question, this would mean that even a random guess was called for.)

3.2 Linearity of expectation

Example 3.6 *You roll two dice and win a dollar amount equal to the sum of your die rolls. Compute the expected value of this game.*

Solution. (**Hard Solution**). The possible outcomes and the probabilities of each are listed in the table below.

2	3	4	5	6	7	8	9	10	11	12
$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

The expected value is therefore

$$2\times\frac{1}{36}+3\times\frac{2}{36}+4\times\frac{3}{36}+5\times\frac{4}{36}+6\times\frac{5}{36}+7\times\frac{6}{36}+8\times\frac{5}{36}+9\times\frac{4}{36}+10\times\frac{3}{36}+11\times\frac{2}{36}+12\times\frac{1}{36}$$
$$= \frac{2+6+12+20+30+42+40+36+30+22+12}{36} = \frac{252}{36} = 7,$$

or exactly 7 dollars.

You should always be suspicious when you do a messy computation and get a simple result.

Solution. (**Easy Solution**). If you roll one die and get the dollar amount showing, we already computed that the expected value of this game is 3.5.

The game discussed now is equivalent to playing this game twice. So the expected value is $3.5 \times 2 = 7$.

Similarly, the expected value of throwing a thousand dice and winning a dollar amount equal to the number of pips showing is (exactly) $3,500.

Here is another problem that illustrates the same principle.

Example 3.7 *Consider once again the game of Rat Race. Suppose that our contestant gets to pick two out of five rats, that first place wins a car (worth $16,000), that second place wins meal service (worth $2,000) and that third place wins a guitar (worth $500).*

The hard solution would be to compute the probability of every possible outcome: the contestant wins the car and the meals, the car and the guitar, the guitar and the meals, the car only, the meals only, the guitar only, and nothing. What a mess!!! Instead, we'll give an easier solution.

Solution. Consider only the first of the contestant's rats. Since this rat will win each of the three prizes for the contestant with probability $\frac{1}{5}$, the expected value of this rat's winnings is

$$16000 \times \frac{1}{5} + 2000 \times \frac{1}{5} + 500 \times \frac{1}{5} = 3700.$$

The second rat is subject to the same rules, so the expected value of its winnings is also \$3700. Therefore, the total expected value is $\$3,700 + \$3,700 = \$7,400$.

Indeed, the expected value of the game is $3,700 per rat won, so this computation gives the answer no matter how many rats she wins.

There is a subtlety going on in this example, which is noteworthy because we **didn't** worry about it. Suppose, for example, that the first rat fails to even move from the starting line. It is a colossal zonk for the contestant, who must pin all of her hopes on her one remaining rat. Does this mean that her expected value plummets to \$3,700? *No!* It now has a one in *four* chance of winning each of the three remaining prizes, so its expected value is now

$$16000 \times \frac{1}{4} + 2000 \times \frac{1}{4} + 500 \times \frac{1}{4} = 4625.$$

Conversely, suppose that this rat races out from the starting block like Usain Bolt, and wins the car! Then the expected value of the remaining rat goes *down*. (It has to: the car is off the table, and the most it can win is \$2,000.) Its expected value is a measly

$$2000 \times \frac{1}{4} + 500 \times \frac{1}{4} = 625.$$

This looks terribly complicated, because **the outcomes of the two rats** are independent. If the first rat does poorly, the second rat is more likely to do well, and vice versa.

The principle of **linearity of expectation** says that our previous computation **is correct, even though the outcomes are not independent.** If the first rat wins the car, the second rat's expected value goes down; if the first rat loses or wins a small prize, the second rat's expected value goes up; and these possibilities average out.

Principle of Linearity of Expectation. Suppose that we have a random process which can be broken up into two or more separate processes. Then, the total expected value is equal to the sum of the expected values of the smaller processes.

Often, games can be broken up in multiple ways. In the exercises you will redo the Rat Race computation a different way: you will consider the expected value of winning just the car, just the meals, and just the guitar – and you will verify that you again get the same answer.

We can now compute the expected value of Rat Race as a whole! Recall that Rat Race begins with the contestant attempting to price three small items correctly, and winning one rat for each item that she gets right.

Example 3.8 *Assume for each small item, the contestant has a 50-50 chance of pricing it correctly. Compute the expected value of playing of Rat Race.*

Solution. Recall from your homework exercises that the probability of winning zero, one, two, or three rats is $\frac{1}{8}$, $\frac{3}{8}$, $\frac{3}{8}$, and $\frac{1}{8}$. Since the expected value of Rat Race is \$3.700 per rat won, the expected value of the race is respectively \$0, \$3,700, \$7,400, and \$11,100. Therefore the expected value of Rat Race is

$$0 \times \frac{1}{8} + 3700 \times \frac{3}{8} + 7400 \times \frac{3}{8} + 11000 \times \frac{3}{8} = 5550.$$

This solution is perfectly correct, but it misses a shortcut. We can use linearity of expectation again!

Solution. Each attempt to win a small item has probability $\frac{1}{2}$ of winning a rat, which contributes \$3,700 to the expected value. Therefore the expected value of each attempt is $3700 \times \frac{1}{2} = 1850$. By linearity of expectation, the expected value of three attempts is

$$1850 + 1850 + 1850 = 5550.$$

3.3 A further expected value example

The St. Petersburg Paradox. You play a game as follows. You start with \$2, and you play the following game. You flip a coin. If it comes up tails, then you win the \$2. If it comes up heads, then your stake is doubled and you get to flip again. You keep flipping the coin, and doubling the stake for every flip of heads, until eventually you flip tails and the game ends.

How much should you be willing to pay to play this game?

To say the same thing another way, your winnings depend on the number of consecutive heads you flip. If none, you win \$2; if one, you win \$4; if two, you win \$8, and so on. More generally, if you flip k consecutive heads before flipping tails, you win 2^{k+1} dollars. Unlike most game shows, you never risk anything and so you will certainly continue flipping until you flip tails.

We first compute the probability of every possible outcome:

- With probability $\frac{1}{2}$, you flip tails on the first flip and win \$2.
- With probability $\frac{1}{4}$, you flip heads on the first flip and tails on the second flip: the probability for each is $\frac{1}{2}$ and you multiply them. If this happens, you win \$4.
- With probability $\frac{1}{8}$, you flip heads on the first two flips and tails on the third flip: the probability for each is $\frac{1}{2}$ so the probability is $\left(\frac{1}{2}\right)^3$. If this happens, you win \$8.
- Now, we'll handle all the remaining cases at once. Let k be the number of consecutive heads you flip before flipping a tail. Then, the probability of this outcome is $\left(\frac{1}{2}\right)^{k+1}$: we've made $k+1$ flips and specified the result for each of them.

  Your winnings will be 2^{k+1} dollars: you start with \$2, and you double your winnings for each of the heads you flipped.

We now compute the expected value of this game. This time there are infinitely many possible outcomes, but we do the computation in the same way. We multiply the probabilities by the expected winnings above, and add:

$$\$2 \cdot \frac{1}{2} + \$4 \cdot \frac{1}{4} + \$8 \cdot \frac{1}{8} + \$16 \cdot \frac{1}{16} + \cdots = \$1 + \$1 + \$1 + \$1 + \cdots = \infty$$

The expected value of the game is infinite, and you should be willing to pay an infinite amount of money to play it. This does not seem to make sense.

By contrast, consider the following version of the game. It has the same rules, only the game has a maximum of 100 flips. If you flip 100 heads, then you don't get to keep playing, and you're forced to settle for 2^{101} dollars, that is, \$2,535,301,200,456,458,802,993,406,410,752.

The expected value of *this* game is a mere

$$\$2 \cdot \frac{1}{2} + \$4 \cdot \frac{1}{4} + \$8 \cdot \frac{1}{8} + \$16 \cdot \frac{1}{16} + \cdots + \$2^{100} \cdot \frac{1}{2^{100}} + \$2^{101} \cdot \frac{1}{2^{100}} = \$1 + \$1 + \$1 + \$1 + \cdots + \$1 + \$2 = \$102.$$

Now think about it. If you won the maximum prize, and it was offered to you in \$100 bills, it would weigh[4] 2.5×10^{25} kilograms, in comparison to the weight of the earth which is only 6×10^{24} kilograms. If you stacked them, you could reach any object which has been observed anywhere in the universe.

Conversely, suppose that it were offered to you in \$100,000,000,000,000 (100 trillion) dollar bills[5] This is much more realistic to imagine; it's roughly equivalent to the Himalayan mountain range being made of such banknotes, all of which belong to you. If you went out to lunch, you'd probably leave a generous tip. You'd have to, because it's not like they can make change for you.

In summary: **This is obviously ridiculous.** You can read more on the Wikipedia article, but the point is that *the real-life meaning of expected values can be distorted by extremely large, and extremely improbable, events.*

3.4 Exercises

1. Watch the Deal or No Deal clip from the introduction. Fast forward through all the talk and choosing briefcases if you like, but pay attention to each time the bank offers him a buyout to quit. Compute, in each case, the expected value of playing the game out until the end. Does the bank ever offer a payout larger than the expected value?

   What would you decide at each stage? Explain.

2. Consider again a game of Rat Race with two rats, played for prizes worth \$16,000 (car), \$2,000 (meals), and \$500 (guitar).

   (a) Compute the expected value of the game, considering only the car and ignoring the other prizes. (This should be easy: she has a 2 in 5 chance of winning the car.)

   (b) Compute the expected value of the game, considering only the meals.

   (c) Compute the expected value of the game, considering only the guitar.

[4] more precisely: have a mass of

[5] Such banknotes were actually printed in Zimbabwe. See, for example, `https://en.wikipedia.org/wiki/Zimbabwean_dollar`.

(d) By linearity of expectation, the expected value of the game is equal to the sum of the three expected values you just computed. Verify that this sum is equal to \$7,400, as we computed before.

The next questions concern the Price is Right game **Let 'em Roll**. Here is a clip:

`https://www.youtube.com/watch?v=g5qF-W9cSpo`

Game Description (Let 'em Roll (Price Is Right)):

The contestant has five dice to roll. Each die has \$500 on one side, \$1,000 on another, \$1,500 on a third, and a car symbol on the other three. The contestant rolls all five dice. If a car symbol is showing on each of them, she wins the car. Otherwise, she wins the total amount of money showing. (Car symbols count nothing, unless she wins the car.)

By default, the contestant gets one roll, and may earn up to two more by correctly pricing small grocery items. After each roll, if she gets another roll, she may either keep all the money showing, or set the dice showing 'car' aside and reroll only the rest.

3. First, consider a game of Let 'em Roll where the contestant only gets one dice roll.

   (a) Compute the probability that she wins the car.

   (b) Compute the expected value of the game, considering the car and ignoring the money. (The announcer says that the car is worth \$16,570.)

   (c) Compute the expected value of the game, considering the money and ignoring the car.

   (d) Compute the total expected value of the game.

4. (a) Now watch the contestant's playing of the game, where after the second round she chooses to give up \$2,500 and reroll. Compute the expected value of doing so. Do you agree with her decision?

   (b) Suppose that after two turns she had rolled no car symbols, and \$1,500 was showing on each of the five dice. Compute the expected value of rerolling, and explain why she should *not* reroll.

   (c) Construct a hypothetical situation where the expected value of rerolling is within \$500 of not rerolling, so that the decision to reroll is nearly a tossup.

5. If the contestant prices the small grocery items correctly and plays optimally, compute the expected value of a game of Let 'em Roll.

   (Warning: if your solution is simple, then it's wrong.)

4 Counting

We now consider a variety of clever counting methods, which will be useful in sophisticated probability computations.

4.1 The Multiplication Rule

Just as there was a multiplication rule for probability, there is a multiplication rule for counting as well. It is as follows.[6]

The multiplication rule. Suppose that an operation consists of k steps, and:

- The first step can be performed in n_1 ways;
- The second step can be performed in n_2 ways (regardless of how the first step was performed);
- and so on. Finally the kth step can be performed in n_k ways (regardless of how the preceding steps were performed).

Then the entire operation can be performed in $n_1 n_2 \dots n_k$ ways.

Example 4.1 *In South Carolina, a license tag can consist of any three letters followed by any three numbers. (Example: TPQ-909) How many different license tags are possible?*

Solution. There are 26 possibilities for the first letter, 26 for the second, and 26 for the third. Similarly there are 10 possibilities for each number. So the total number of possibilities is $26^3 \cdot 10^3 = 17576000$.

Note that big states with more people than South Carolina have started using different license plate schemes, because they ran out of possible tags.

Example 4.2 *How many license tags are possible which don't repeat any letters or numbers?*

Solution. There are still 26 possibilities for the first letter, and now 25 for the second and 24 for the third: we must avoid the letters that were previously used. Similarly there are 10, 9, and 8 possibilities for the three numbers. The total number of possibilities is

$$26 \cdot 25 \cdot 24 \cdot 10 \cdot 9 \cdot 8 = 11232000.$$

These computations may be used to solve probability questions. For example:

Example 4.3 *What is the probability that a random license tag doesn't repeat any letters or numbers?*

[6] We adopt the wording of Epp, *Discrete Mathematics with Applications*, 4th ed., p. 527.

This follows from the previous two computations. The result is

$$\frac{11232000}{17576000} = .639\ldots$$

Example 4.4 *On a game of Ten Chances, Drew Carey feels particularly sadistic and puts all ten digits – zero through nine – to choose from in the price of the car. The price of the car consists of five different digits. How many possibilities are there?*

Solution. There are 10 possibilities for the first digit, 9 for the second, 8 for the third, 7 for the fourth, and 6 for the fifth, for a total of

$$10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = 30240$$

possibilities. Good luck to the poor sucker playing this game.

Example 4.5 *As above, but suppose you know that the first digit is not zero. Now how many possibilities are there?*

Solution. This time there are only 9 possibilities for the first digit. There are still 9 possibilities for the second, no matter what the first digit was, and 8, 7, 6 for the last three in turn. The total is

$$9 \cdot 9 \cdot 8 \cdot 7 \cdot 6 = 27216.$$

Example 4.6 *As above, but suppose you know that the first digit is not zero* **and** *that the last digit is zero. Now how many possibilities are there?*

Solution. There are 9 possibilties for the first digit, 9 for the second, 8 for the third, 7 for the fourth, and ... **either** 0 **or** 1 **for the last depending on whether we've used the zero.** No good! We can't use the multiplication rule this way!

To use the multiplication rule, we pick the numbers in a different order: the first digit first (anything other than the zero, 9 ways), then the last digit (must be the zero, so 1 way), and then the second, third, and fourth digits in turn (8, 7, and 6 ways), for a total of

$$9 \cdot 8 \cdot 7 \cdot 6 = 3024$$

ways.

Alternatively, we could have picked the last digit before the first, and we can pick the second, third, and fourth digits in any order. It is usually best to find one order which works and stick to it.

4.2 Permutations and combinations

Recall that a **permutation** of a **string with n symbols** is any reordering of the string. If the symbols are all distinct, then there are $n!$ possible permutations of it. We justified this earlier, and it is an example of the multiplication rule: there are n ways to choose the first symbol, $n-1$ to choose the second, $n-2$ to choose the third, and so on.

Implicitly, we also discussed what are called r-permutations. If $r \leq n$, then an r-permutation of a string of length n is a reordering of r of the n symbols. For example, 16820, 98561, and 37682 are 5-permutations of the string 1234567890. We discussed these in our Ten Chances examples above, and there are 30240.

Notation. Write $P(n,r)$ for the number of r-permutations of a string with n distinct symbols.

We have the following formula:

$$P(n,r) = \frac{n!}{(n-r)!}$$

Why is this true? It comes from the multiplication rule. There are n possibilities for the first symbol, $n-1$ possibilities for the second, and so on: one less for each subsequent symbol. There are $n-r+1$ possibilities for the rth symbol: we start at n and count down by 1 $r-1$ times. So we see that

$$P(n,r) = n \cdot (n-1) \cdot (n-2) \cdot (n-3) \cdots (n-r+1).$$

Why is this equal to $\frac{n!}{(n-r)!}$? Our expression is the same as $n!$, except that the numbers from $n-r$ down to 1 are all absent. So we've left out a product equalling $(n-r)!$ from the definition of $n!$, and so it equals $\frac{n!}{(n-r)!}$.

Example 4.7 *Compute $P(n,r)$ for all possible n and r with $r \leq n \leq 6$ and notice any patterns.*

Solution. (To be written up here)

Combinations. Combinations are like permutations, only the order doesn't matter. If we start with a string (or a set) with n distinct elements, then an r-**combination** is a string or r of these elements *where order doesn't matter*, or equivalently a subset of r of these elements.

Example 4.8 *Write out all the* 3*-combinations of* 12345*.*

Solution. They are: 123, 124, 125, 134, 135, 145, 234, 235, 245, and 345. There are ten of them.

Here, we could have equivalently written 321, 213, or $\{1,2,3\}$ (for example) in place of 123, because when counting **combinations** it is irrelevant which order the symbols come in.

When counting permutations it **is** relevant, so **please always be careful to pay attention to exactly what you are counting!**

Note that a string with n distinct elements, where order doesn't matter, is the same thing as a set of n distinct elements. We won't worry about distinguishing these too carefully, although in advanced mathematics and in computer programming it is important to be precise.

Example 4.9 *Write out all the* 2*-combinations of* 12345*.*

Solution. They are: 45, 35, 34, 25, 24, 23, 15, 14, 13, and 12. Again, there are ten of them.

I didn't have to list them in reverse order, but in doing so we notice something interesting: they correspond exactly to the 3-combinations! Choosing which two elements to include is equivalent to choosing which three to leave out, so we can line up the list of 2-combinations with the list of 3-combinations and see that there is a one-to-one correspondence. In mathematical parlance, we call this a **bijection.** *If you want to prove that two sets have the same size, finding a bijection between them is a great way to do it!*

Notation. Write $C(n,r)$ or $\binom{n}{r}$ for the number of r-combinations of an n-element set. The latter notation is read "n choose r", and is ubiquitous in mathematics. These numbers are also called 'binomial coefficients', because we have

$$(x+1)^n = \binom{n}{n}x^n + \binom{n}{n-1}x^{n-1} + \binom{n}{n-2}x^{n-2} + \cdots + \binom{n}{1}x + \binom{n}{0}.$$

For example, we have

$$(x+1)^10 = x^{10} + 10x^9 + 45x^8 + 120x^7 + 210x^6 + 252x^5 + 210x^4 + 120x^3 + 45x^2 + 10x + 1,$$

so we can FOIL without FOILing. If you think about it carefully, you can figure out why the first equation is true. But we still haven't explained how to actually *compute* these things. Here's the answer.

Theorem 4.10 *We have*

$$C(n,r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}.$$

To explain thiis we will be very careful and work backwards. First of all, note that it is enough to show that

$$P(n,r) = C(n,r) \cdot r! \tag{2}$$

We will first explain why (4.2) implies the theorem, and then we will explain why (4.2) is true. First, note that (4.2) implies that

$$C(n,r) = \frac{P(n,r)}{r!},$$

but remember that we showed that $P(n,r) = \frac{n!}{(n-r)!}$. Therefore,

$$C(n,r) = \frac{\frac{n!}{(n-r)!}}{r!} = \frac{n!}{(n-r)!r!},$$

as desired.

We are left to explain why is true. To do this, we explain why both sides of (4.2) count the number of r-permutations of a string of n elements:

- This is true of $P(n,r)$ by definition.
- Instead, we could first choose which r objects to make an r-permutation out of, without worrying about the order. By definition, there are $C(n,r)$ ways to do this. Now, we have to put these r symbols in some order – i.e., to write down a permutation of them. There are $r!$ ways to do this. So the total number of ways is $C(n,r) \cdot r!$

If you haven't seen this before, you probably didn't understand what just happened. That's okay. Read through it again.

5 (To be incorporated later)

Poker. Poker provides many interesting examples of expected value computations. We digress to explain the rules of poker.

A **poker hand** consists of five playing cards. From best to worst, they are ranked as follows:

- **A straight flush**, five consecutive cards of the same suit, e.g. $5\spadesuit 6\spadesuit 7\spadesuit 8\spadesuit 9\spadesuit$. An ace may be counted high or low but straights may not 'wrap around' (e.g. KA234 is not a straight).

  In case of a tie, the high card in the straight flush settles ties.
- **Four of a kind**, for example $K\spadesuit K\clubsuit K\diamondsuit K\heartsuit$ and any other card. (If two players have four of a kind, the highest set of four cards win.)
- **A full house**, i.e. three of a kind and a pair, $K\spadesuit K\clubsuit K\diamondsuit 7\heartsuit 7\diamondsuit$. (If two players have a full house, the highest set of three cards wins.)
- **A flush**, any five cards of the same suit, e.g. $Q\clubsuit 10\clubsuit 7\clubsuit 6\clubsuit 3\clubsuit$. The high card breaks ties (followed by the second highest, etc.)
- **A straight**, any five consecutive cards, e.g. $8\clubsuit 7\diamondsuit 6\diamondsuit 5\heartsuit 4\spadesuit$. The high card breaks ties.
- **Three of a kind**, e.g. $8\clubsuit 8\diamondsuit 8\spadesuit A\heartsuit 4\spadesuit$.

- **Two pair**, e.g. $8\clubsuit 8\diamondsuit 6\spadesuit 6\heartsuit A\spadesuit$.
- **One pair**, e.g. $8\clubsuit 8\diamondsuit 6\spadesuit 5\heartsuit A\spadesuit$.
- **High card**, e.g. none of the above. The value of your hand is determined by the highest card in it; then, ties are settled by the second highest card, and so on.

Poker is played for *chips*, which may or may not represent money. A **betting round** works as follows. The players (at least two) sit around a table. The first opens the betting, and chooses any amount to **bet** (including zero, which is referred to as a **check**). Then, proceeding clockwise, each other player may choose to **fold** (i.e., quit), **call** (i.e., match the bet), or **raise** (make a larger bet). This continues until everyone has either called or folded to the largest raise. If everyone folds to the last raiser, she wins and doesn't have to show her cards.

In Texas Hold'em, on the first round only, the first two players are required to make fixed, initial bets known as 'blind bets'. For example, in a low stakes game these might be 10 and 20 cents respectively.

There are multiple variants of poker, ... **finish later**