Jonathan Weisberg

Odds & Ends

Introducing Probability & Decision with a Visual Emphasis

v0.3 beta

An Open Access Publication

Contents

Preface

THIS textbook is awesome.

1. It is open access, hence free.
2. It's also open source, so other instructors can modify it to their liking.
3. It's available in both PDF and HTML. So it can be read comfortably on a range of devices, or printed.

If you teach from this book I'd love to know: email or tweet me.

A "cheat sheet" summarizing key definitions and formulas appears in [Appendix][Cheat Sheet] [A][Cheat Sheet]. Further appendices cover the axiomatic construction of probability theory, Hume's problem of induction, and Goodman's new ridle of induction.

I usually get a mix of students in my course, with different ideological inclinations and varying levels of background. For some the technical material is easy, even review. For others, a healthy skepticism about scientific methods and discourses comes naturally. My goal is to get these students all more or less on the same page.

If it weren't for these tools, I never would have written this book. It wouldn't have been possible to create one that does all the things this book is meant to do. I also owe inspiration to Kieran Healy's book *Data Visualization: A Practical Introduction*, which uses the same suite of tools. It gave me the idea to use those tools for an updated, open, and visually enhanced rendition of the classic material from Skyrms and Hacking.

Finally, I'm indebted to several teaching assistants and students who helped with earlier drafts. Thanks especially to Liang Zhou Koh and Meagan Phillips, who contributed several exercises; to Soroush Marouzi and Daniel Munro, who worked out the kinks during the first semester the book was piloted; and to the students in that course, who bore with us, and contributed several corrections of their own.

Part I

1   *The Monty Hall Problem*

When tackling a math problem,
I encourage students to draw a picture.
—Francis Su

IMAGINE you're on a game show. There are three doors, one with a prize behind it. You're allowed to pick any door, so you choose the first one at random, door A.

Now the rules of the game require the host to open one of the other doors and let you switch your choice if you want. Because the host doesn't want to give away the game, they always open an empty door.

In your case, the host opens door C: no prize, as expected. "Do you want to switch to door B?" the host asks.

Pause a moment to think about your answer before reading on.

WHAT did you decide? Did you conclude it doesn't matter whether you stick with door A or switch to door B?

If so, you're in good company. Most people find this answer sensible, including some professors of statistics and mathematics. They figure there are only two possibilities remaining, door A and door B, each with the same one-in-two chance of being the winner. So it doesn't matter which one you pick.

But the right answer is you should switch. Door B is now twice as likely to be the winner as door A. Why?

The reason is subtle. One way to think about it is that the host's choice of which door to open is a bit of a tell. Maybe they *had* to open door C, because the prize is behind door B and they didn't want to give that away. Of course, it could be behind door A instead, so maybe they just picked door C at random. But there was only a one-in-three chance the prize would be behind door A. Which means there's a two-in-three chance they didn't really have a choice, they had to open door C to avoid showing you the prize behind door B.

Here's another way to think about it. Imagine the game had a hundred doors instead of just three. And suppose again you start by picking the first door at random. Then the host opens *all the other doors*

Monty Hall was creator and host of the game show *Let's Make a Deal*.

Marilyn vos Savant made the Monty Hall problem famous when she solved it correctly in her magazine column. Read about it here.
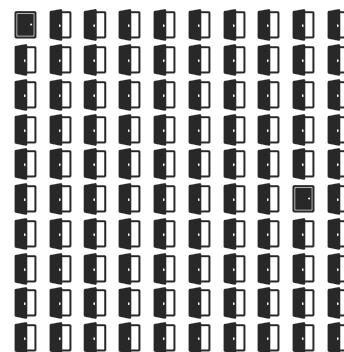

Figure 1.1: The hundred-door version of the Monty Hall problem, suggested by Marilyn vos Savant

but one, door 59 let's say. You have to ask yourself: why did they pick door 59 to leave closed?? Almost certainly because that's where the prize is hidden! Maybe you got really lucky and picked right with the first door at the beginning. But it's way more likely you didn't, and the host had to keep door 59 closed to avoid giving away the game.

1.1   Diagramming the Solution

A picture helps clarify things. At first the prize could be behind any of the three doors, with equal probability each way. So we draw a tree with three branches, each labeled with a probability of 1/3. Figure 1.2 shows the result.

Now, which door the host opens may depend on where the prize is, i.e. which branch we're on. If it's behind door C, they won't show you by opening that door. They would have to open door B in this case.

Likewise, if the prize is behind door B, then opening door C is their only option.

Only if the prize is behind door A do they have a choice: open either door B or door C. In that case it's a tossup which door they'll open, so each of those possibilities has a 1/2 chance. Check out Figure 1.3.

Now imagine playing the game over and over. A third of the time things will follow the top path; a third of the time they'll follow the middle one; and the remaining third they'll follow one of the two bottom paths.

When things follow the bottom branches, half of those times the host will open door B, and half the time they'll open door C. So one in every six plays will follow the *A-and-Open-B* path. And one in every six plays will follow the *A-and-Open-C* path. See Figure 1.4.

Now we can understand what happens when the host opens door C. Usually it's because the prize is behind door B. Sometimes they open door C because the prize is behind door A instead. But that's only a sixth of the time, compared to a third of the time where they open door C because the prize is behind door B.

So when you see the host open door C, you should think it's more likely you're on the middle branch, with the prize behind door B. Switch!

1.2   Lessons Learned

Tree diagrams are a handy tool for solving probability problems. They also illustrate some central concepts of probability.

Probabilities are numbers assigned to possibilities. In the Monty Hall problem, there are three possibilities for where the prize is: door A, door B, and door C. Each of these possibilities starts with the same
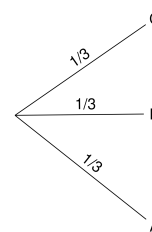

Figure 1.2: First stage of a tree diagram for the Monty Hall problem
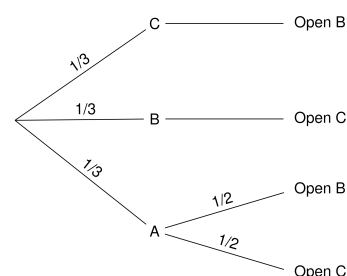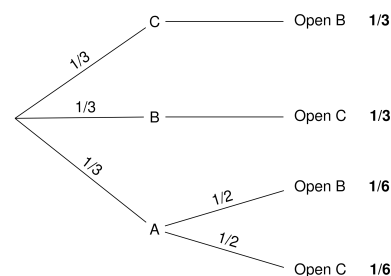

Figure 1.3: Second stage

Figure 1.4: Third and final stage

probability: 1/3.

Some possibilities are ***mutually exclusive***, meaning only one of them can obtain. The prize can't be behind door A and door B, for example. Here are more examples of mutually exclusive possibilities:

- A coin can land heads or tails, but it can't do both on the same toss.
- A card drawn from a standard deck could be either an ace or a queen, but it can't be both.
- The temperature at noon tomorrow could be 20 degrees, or it could be 25 degrees, but it can't be both.

When possibilities are mutually exclusive, their probabilities add up. For example, the initial probability the prize will be behind either door A or door B is $1/3 + 1/3 = 2/3$. And the probability a card drawn from a standard deck will be either an ace or a queen is $4/52 + 4/52 = 8/52 = 2/13$.

Another key concept is possibilities that are ***exhaustive***. In the Monty Hall problem, the prize has to be behind one of the three doors, so A, B, and C "exhaust" all the possibilities. Here are more examples of exhaustive possibilities:

- A card drawn from a standard deck must be either red or black.
- The temperature at noon tomorrow must be either above zero, below zero, or zero.

IN OUR tree diagrams, each branch-point always uses a set of possibilities that is *both* exclusive *and* exhaustive. The first split on the three doors covers all the possibilities for where the prize might be, and only one of those possibilities can be the actual location of the prize. Likewise for the second stage of the diagram. On the bottom branch for example, the host must open either door B or door C given the rules, but he will only open one or the other.

Figure 1.5: Three partitions for a card drawn from a standard deck

When a set of possibilities is both exclusive and exhaustive, it's called a ***partition***. A partition "carves up" the space of possibilities into distinct, non-overlapping units.

There can be more than one way to partition the space of possibilities. For example, a randomly drawn playing card could be black or red; it could be a face card or not; and it could be any of the four suits ($\heartsuit$, $\diamondsuit$, $\clubsuit$, $\spadesuit$).

When possibilities form a partition, their probabilities must add up to 1. Initially, the probability the prize will be behind one of the three doors is $1/3 + 1/3 + 1/3 = 1$. And the probability that a card drawn from a standard deck at random will be either red or black is $1/2 + 1/2 = 1$.

In a way, the fundamental principle of probability is that probabilities over a partition must add up to 1.

TREE diagrams follow a few simple rules based on these concepts. The parts of a tree are called *nodes*, *branches*, and *leaves*: see Figure 1.6.
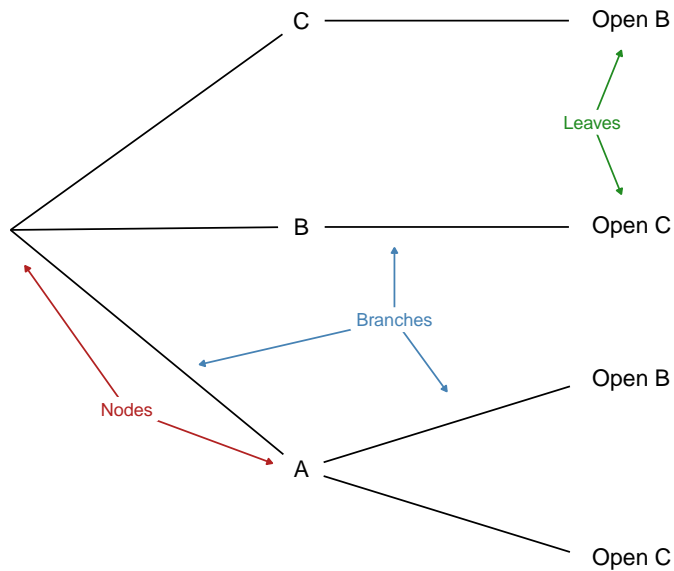

Figure 1.6: The parts of a tree diagram: nodes, branches, and leaves

The rules for a tree are as follows.

- *Rule 1.* Each node must branch into a partition. The subpossibilities that emerge from a node must be mutually exclusive, and they must include every way the possibility from which they emerge could obtain.

- *Rule 2.* The probabilities emerging from a node must add to 1. If we add up the numbers on the branches immediately coming out of a node, they should add to 1.

- *Rule 3.* The probability on a branch is *conditional* on the branches leading up to it. Consider the bottom path in the Monty Hall problem. The probability the host will open door C is 1/2 there because we're assuming the prize is behind door A.

- *Rule 4.* The probability of a leaf is calculated by multiplying across the branches on the path leading to it. This number represents the probability that all possibilities on that path occur.

Notice, Rule 4 is how we got the final probabilities we used to solve the Monty Hall problem (the numbers in bold).

Exercises

1. True or false: in the Monty Hall problem, it's essential to the puzzle that the host doesn't want to expose the prize. If they didn't care about giving away the location of the prize, there would be no reason to switch when they open door C.

2. In the version of the Monty Hall problem with a hundred doors, after the host opens every door except door 1 (your door) and door 59, the chance the prize is behind door 59 is:

   a. 1/100
   b. 1/99
   c. 1/2
   d. 99/100

3. Imagine three prisoners, A, B, and C, are condemned to die in the morning. But the king decides to pardon one of them first. He makes his choice at random and communicates it to the guard, who is sworn to secrecy. She can only tell the prisoners that one of them will be released at dawn, she can't say who.

   Prisoner A welcomes the news, as he now has a 1/3 chance of survival. Hoping to go even further, he says to the guard, "I know you can't tell me whether I am condemned or pardoned. But at least one other prisoner must still be condemned, so can you just name one who is?". The guard tells him that B is still condemned. "Ok," says A, "then it's either me or C who was pardoned. So my chance of survival has gone up to 1/2."

   Is prisoner A's reasoning correct? Use a probability tree to explain why/why not.

4. In a probability tree, each branch point should split into possibilities that are:

   a. Mutually exclusive.
   b. Exhaustive.
   c. Both mutually exclusive and exhaustive.
   d. None of the above.

5. Suppose you have two urns. The first has two black marbles and two white marbles. The second has three black marbles and one white marble. You are going to flip a fair coin to select one of the urns at random, and then draw one marble at random. What is the chance you will select a black marble?

Hint: draw a probability tree and ask yourself, "if I did this experiment over and over again, how often would I draw a black marble in the long run?"

a. 5/8
b. 3/8
c. 1/2
d. 1/4

6. An ice cream counter sells 4 different flavours of ice cream (chocolate, vanilla, strawberry, mint). There are 2 toppings (fudge, caramel), and 3 kinds of sprinkles (chocolate, rainbow, purple). You order by picking a flavour, a topping, and a kind of sprinkles.

a. How many possible orders are there?
b. If you make your three choices randomly, what is the probability your order will have strawberry ice cream but not rainbow sprinkles?

2   *The Problem of Induction*

> *It's tough to make predictions, especially about*
> *the future.*
> —*Yogi Berra*

MANY inductive arguments work by projecting an observed pattern onto as-yet unobserved instances. All the ravens we've observed have been black, so all ravens are. All the emeralds we've seen have been green, so all emeralds are.

The assumption that the unobserved will resemble the observed seems to be central to induction. Philosophers call this assumption the *Principle of Induction*.[1] But what justfies this assumption? Do we have any reason to think the parts of reality we've observed so far are a good representation of the parts we haven't seen yet?

[1] See Section **??** and Appendix **??** for previous discussions of the Principle of Induction.

Actually there are strong reasons to doubt whether this assumption can be justified. It may be impossible to give any good argument for expecting the unobserved to resemble the observed.

The Dilemma

We noted in [Chapter 2][Logic] that there are two kinds of argument, inductive and deductive. Some arguments establish their conclusions necessarily, others only support them with high probability. If there is an argument for the Principle of Induction, it must be one of these two kinds. Let's consider each in turn.

Could we give an inductive argument for the Principle of Induction? At first it seems we could. Scientists have been using inductive reasoning for millenia, often with great success. Indeed, it seems humans, and other creatures too, have relied on it for much longer, and could not have survived without it. So the Principle of Induction has a very strong track record. Isn't that a good argument for believing it's correct?

No, because the argument is circular. It uses the Principle of Induction to justify believing in the Principle of Induction. Consider that the argument we are attempting looks like this:

Figure 2.1: David Hume (1711–1776) raised the problem of induction in 1739. Our presentation of it here is somewhat modernized from his original argument.

> The principle has worked well when we've used it in the past. Therefore it will work well in future instances.

This is an inductive argument, an argument from observed instances to ones as yet unobserved. So, under the hood, it appeals to the Principle of Induction. But that's exactly the conclusion we're trying to establish. And one can't use a principle to justify itself.

What about our second option: could a deductive argument establish the Principle of Induction? Well, by definition, a deductive argument establishes its conclusion with necessity. Is it necessary that the unobserved will be like the observed? It doesn't look like it. It seems perfectly possible that tomorrow the world will go haywire, randomly switching from pattern to pattern, or even to no pattern at all.

Maybe tomorrow the sun will fail to rise. Maybe gravity will push apart instead of pull together, and all the other laws of physics will reverse too. And just as soon as we get used to those patterns and start expecting them to continue, another pattern will arise. And then another. And then, just as we give up and come to have no expectation at all about what will come next, everything will return to normal. Until we get comfortable and everything changes again.

Thankfully, our universe hasn't been so mischievous. We get surprised now and again, but for the most part inductive reasoning is pretty reliable, when we do it carefully. But we're lucky in this respect, is the point.

Nature *could* have been mischievous, totally unpredictable. It is not a necessary truth that the unobserved must resemble the observed. And so it seems there cannot be a deductive argument for the Principle of Induction. Because such an argument would establish the principle as a necessary truth.

The Problem of Induction vs. the Grue Paradox

If you read Appendix **??**, you know of another famous problem with the Principle of Induction: the grue paradox. (If you haven't read that chapter, you might want to skip this section.)

The two problems are quite different, but it's easy to get them confused. The problem we're discussing here is about justifying the Principle of Induction. Is there any reason to believe it's true? Whereas the grue paradox points out that we don't even really know what the principle says, in a way. It says that what we've observed is a good indicator of what we haven't yet obsered. But in what respects? Will unobserved emeralds be green, or will they be grue?

So the challenge posed by grue is to spell out, precisely, what the Principle of Induction says. But even if we can meet that challenge,

this challenge will remain. Why should we believe the principle, once it's been spelled out? Neither a deductive argument nor an inductive argument seems possible.

Probability Theory to the Rescue?

The Problem of Induction is centuries old. Isn't it out of date? Hasn't the modern, mathematical theory of probability solved the problem for us?

Not at all, unfortunately. One thing we learn in this book is that the laws of probability are very weak in a way. They don't tell us much, without us first telling them what the prior probabilities are. And as we've seen over and again throughout Part III, the problem of priors is very much unsolved.

For example, suppose we're going to flip a mystery coin three times. We don't know whether the coin is fair or biased, but we hope to have some idea after a few flips.

Now suppose we've done the first two flips, both heads. The Principle of Induction says we should expect the next flip to be heads too. At least, that outcome should now be more probable.

Do the laws of probability agree? Well, we need to calculate the quantity:

$$Pr(H_3 \mid H_1 \,\&\, H_2).$$

The definition of conditional probability tells us:

$$Pr(H_3 \mid H_2 \,\&\, H_1) = \frac{Pr(H_3 \,\&\, H_2 \,\&\, H_1)}{Pr(H_2 \,\&\, H_1)}.$$

But the laws of probability don't tell us what numbers go in the numerator and the denominator.

The numbers have to be between 0 and 1. And we have to be sure mutually exclusive propositions have probabilities that add up, according to the Additivity rule. But that still leaves things wide open.

For example, we could assume that all possible sequences of heads and tails are equally likely. In other words:

$$Pr(HHH) = Pr(THH) = Pr(HTH) = \ldots = Pr(TTT) = 1/8.$$

Given that assumption, we get the result that $Pr(H_3 \mid H_2 \,\&\, H_1) = 1/2$.

$$
\begin{aligned}
Pr(H_3 \mid H_2 \,\&\, H_1) &= \frac{Pr(H_3 \,\&\, H_2 \,\&\, H_1)}{Pr(H_2 \,\&\, H_1)} \\
&= \frac{1/8}{1/8 + 1/8} \\
&= 1/2.
\end{aligned}
$$

But that means the first two flips didn't tell us anything about the third! Even after we got two heads, the chance of another heads is still stuck at 1/2, same as it was to start with.

WE DIDN'T *have* to assume all possible sequences are equally likely. We can make a different assumption, and get a different result.

Let's try assuming instead that all possible *frequencies* of heads are equally probable. In other words, the probability of getting 0 heads is the same as the probability of getting 1 head, which is also the same as the probability of getting 2 heads, and likewise for 3 heads. So we're grouping the possible sequences like so:

- 0 heads: TTT
- 1 head: HTT, THT, TTH
- 2 heads: HHT, HTH, THH
- 3 heads: HHH

Each grouping has the same probability, 1/4. And for the groups in the middle, which have multiple members, we divide that evenly between the members. So $Pr(HTH) = 1/12$, for example, but $Pr(TTT) = 1/4$.

This might seem a funny way of assigning prior probabilities. But it actually leads to very sensible results; the same results as the Rule of Succession in fact! For example, we get $Pr(H_3 \mid H_2 \& H_1) = 3/4$:

$$
\begin{aligned}
Pr(H_3 \mid H_1 \& H_2) &= \frac{Pr(H_3 \& H_2 \& H_1)}{Pr(H_2 \& H_1)} \\
&= \frac{1/4}{1/4 + 1/12} \\
&= 3/4.
\end{aligned}
$$

So, on this analysis, the first two tosses do tell us what to expect on the next toss.

WE'VE seen two different ways of assigning prior probabilities, which lead to very different results. The first way, where we assume all possible sequences are equally likely, disagrees with the Principle of Induction. Our observations of the first two flips tell us nothing about the next one. But the second way, where we assume all possible frequencies are equally likely, agrees with with the Principle of Induction. Observing heads on the first two does tell us to expect another heads on the next one.

Both assumptions are consistent with the laws of probability. So those laws don't, by themselves, tell us what to expect. The laws of probability only tell us what to expect once we've specified the prior

probabilities. The problem of induction challenges us to justify one choice of prior probabilities over the alternatives.

In the 280 years since this challenge was first raised by David Hume, no answer has gained general acceptance.

Exercises

1. Suppose we do 100 draws, with replacement, from an urn containing an unknown mixture of black and white balls. All 100 draws come out black. Which of the following is correct?

   a. According to the laws of probability, the next draw is more likely to be black than white.
   b. According to the laws of probability, the next draw is more likely to be white than black.
   c. According to the laws of probability, the next draw is equally likely to be black vs. white.
   d. The laws of probability are consistent with any of the above conclusions; it depends on the prior probabilities.
   e. None of the above.

2. Write a short essay (3–4 paragraphs) explaining the problem of induction. Your essay should include all of the following:

   • a clear, accurate explanation of the Principle of Induction,
   • a clear, accurate explanation of the dilemma we face in justifying the Principle of Induction,
   • a clear, accurate explanation of the challenge for the deductive horn of this dilemma, and
   • a clear, accurate explanation of the challenge for the inductive horn of the dilemma.

3. A coin will be tossed 3 times, so the number of heads could be 0, 1, 2, or 3. Suppose all 4 of these possibilities are equally likely. Moreover, any two sequences with the same number of heads are also equally likely. For example, $Pr(H_1 \& H_2 \& T_3) = Pr(T_1 \& H_2 \& H_3)$. Answer each of the following.

   a. What is $Pr(H_2 \mid H_1)$?
   b. What is $Pr(H_3 \mid H_1 \& H_2)$?

   Now suppose we do 4 tosses instead of 3. The prior probabilities follow the same rules: all possible numbers of heads are equally likely, and any two sequences with the same number of heads are equally likely.

c. If the first n tosses come up heads, what is the probability the next toss will come up heads? In other words, give a formula for $Pr(H_{n+1} \mid H_1 \& \ldots \& H_n)$ in terms of n.

d. What if only k out of the first n tosses land heads, then what formula gives the probability of heads on the next toss?

4. Suppose a computer program prints out a stream of A's and B's. After observing the sequence A, A, B, we want to know the probability of an A next.

Our friend Charlie suggests we reason as follows. We are going to observe 4 characters total. Before we observed any characters, there were 5 possibilities: the total number of As could turn out to be 0, 1, 2, 3, or 4. And all of these possibilities are equally likely. So each has prior probability 1/5.

Some of these possibilities can be subdivided. For example, there are 4 ways to get 3 A's:

A, A, A, B
A, A, B, A
A, B, A, A
B, A, A, A

So each of these sequences gets 1/4 of 1/5, in other words prior probability 1/20.

According to Charlie's way of reasoning, what is the probability the 4th character will be an A, given that the first 3 were A, A, B?

5. In this chapter we considered a coin to be flipped 3 times, with the same prior probability for every possible sequence of heads/tails. Now suppose the coin will be flipped some very large number of times, n. And again suppose the prior probability is the same for every possible sequence of heads/tails. Prove that no matter how many times the coin lands heads, the probability of heads on the next toss is still 1/2. In other words, prove that $Pr(H_{k+1} \mid H_1 \& \ldots \& H_k) = 1/2$ no matter how large k gets.

6. Suppose a coin will be flipped 3 times. There are $2^3 = 8$ possible sequences of heads/tails that we might get. Find a way of assigning prior probabilities to these 8 sequences so that, the more heads we observe, the *less* likely it becomes we'll get heads on the next toss. In other words, assign a prior probability to each sequence so that $Pr(H_2 \mid H_1) < Pr(H_2)$ and $Pr(H_3 \mid H_2 \& H_1) < Pr(H_3 \mid H_1)$.

7. Suppose a coin will be flipped 4 times. Recall the two different ways of assigning a probability to each possible sequence of heads and tails discussed in the chapter:

- Scheme 1: all possible sequences have the same probability.
- Scheme 2: all possible *frequencies* (numbers of heads) have the same probability, and all sequences that share the same frequency have the same probability.

According to each scheme, how probable is each of the following propositions?

- $A =$ All tosses will land the same way.

- $B =$ The number of heads and tails will be the same.

Now answer the same question with 10 tosses instead of 4.

3   *Solutions to Selected Exercises*

Chapter 1

Exercise 3. Prisoner A is incorrect. His chances of survival are still only 1/3. To see why, we can use the same tree as we did for the Monty Hall problem, and just change the labels on the leaves from "Monty Opens X" to "Guard Names X."

Exercise 5. Option (a): 5/8.

Chapter 2

Exercise 1. (a) Invalid, (b) Valid, (c) Valid, (d) Valid, (e) Invalid, (f) Valid.

Exercise 2. (a) Compatible, (b) Compatible, (c) Mutually Exclusive.

Exercise 3. False.

Exercise 5. Yes, it is possible.

Chapter 3

Exercise 1. (a) $\sim A$, (b) $A \& B$, (c) $A \& \sim B$, (d) $\sim A \& \sim B$.

Exercise 2a. These are mutually exclusive: there's only one row where $A \& B$ is true, and $A \& \sim B$ is false in that row.

Exercise 3b. These are logically equivalent: their final columns are identical, T T F F.

Exercise 4. The column for $B \& C$ is T F F F T F F F. The column for $A \vee (B \& C)$ is T T T T T F F F.

Chapter 4

Exercise 1. (a) Independent, (b) Not Independent, (c) Independent, (d) Not Independent.

Exercise 2. Option (d): All of the above.

Exercise 3. Option (e): None of the above.

Exercise 4. (a) Not a gambler's fallacy, (b) Gambler's fallacy.

Chapter 5

Exercise 1. This is a contradiction so its probability is 0.

Exercise 4. (a) 8/15, (b) 8/15, (c) No.

Exercise 5. (a) 0, (b) 1/6, (c) No.

Exercise 7. (a) Yes, (b) No, (c) No.

Exercise 8. 1/6.

Chapter 6

Exercise 2. (a) 1/2, (b) 2/7.

Exercise 5. (a) 9/70, (b) 9/140, (c) 1/30, (d) 1/60, (e) 17/210, (f) 27/34.

Exercise 6. 1/9.

Chapter 7

Exercise 1. (a) 3/13, (b) 10/13, (c) 4/13, (d) 4/13, (e) 10/13

Exercise 4. 17/32

Exercise 6. Yes:

$$
\begin{aligned}
Pr(A) &= Pr((A \,\&\, B) \vee (A \,\&\, C) \vee (A \,\&\, D)) && \text{by Equivalence} \\
&= Pr(A \,\&\, B) + Pr(A \,\&\, C) + Pr(A \,\&\, D) && \text{by Addition} \\
&= Pr(A \mid B)Pr(B) + Pr(A \mid C)Pr(C) + Pr(A \mid D)Pr(D) && \text{by General Multiplication.}
\end{aligned}
$$

Exercise 8. (a) 1/13, (b) 4/51, (c) 4/663, (d) 4/663, (e) 8/663, (f) 1/221

Exercise 25. First observe that $A \,\&\, (A \vee B)$ is equivalent to A. This can be verified by a truth table or Euler diagram. We then reason as

follows:

$$Pr(A \mid A \vee B) = \frac{Pr(A \,\&\, (A \vee B))}{Pr(A \vee B)} \qquad \text{by definition}$$

$$= \frac{Pr(A)}{Pr(A \vee B)} \qquad \text{by Equivalence}$$

$$= \frac{Pr(A)}{Pr(A) + Pr(B)} \qquad \text{by Addition.}$$

Chapter 8

Exercise 1. 1/4

Exercise 3. 2/3

Exercise 5. 9/11

Exercise 7. The full formula is:

$$Pr(X \mid B) = \frac{Pr(X)Pr(B \mid X)}{Pr(X)Pr(B \mid X) + Pr(Y)Pr(B \mid Y) + Pr(Z)Pr(B \mid Z)}.$$

To derive this formula, start with the short form of Bayes' theorem for $Pr(X \mid B)$. Then apply the version of LTP from Exercise 7.6 to the denominator, $Pr(B)$.

Exercise 8. 1/57

Exercise 11. 1/4

Chapter 9

Exercise 1. (a) 1/6, (b) 1/2, (c) 1/2.

Exercise 2. 81/85.

Exercise 5.

$$Pr(A \mid B \,\&\, C) = \frac{Pr(A \,\&\, (B \,\&\, C))}{Pr(B \,\&\, C)} \qquad \text{by definition}$$

$$= \frac{Pr(A \,\&\, (C \,\&\, B))}{Pr(C \,\&\, B)} \qquad \text{by Equivalence}$$

$$= Pr(A \mid C \,\&\, B) \qquad \text{by definition.}$$

Exercise 7. First note that $Pr(C) = Pr((A \,\&\, C) \vee Pr(\sim A \,\&\, C))$ by Equivalence, and thus by Addition we have $Pr(\sim A \,\&\, C) = Pr(C) - Pr(A \,\&\,$

C). We then reason as follows:

$$Pr(\sim A \mid C) = \frac{Pr(\sim A \,\&\, C)}{Pr(C)} \qquad \text{by definition}$$

$$= \frac{Pr(C) - Pr(A \,\&\, C)}{Pr(C)} \qquad \text{by above}$$

$$= \frac{Pr(C)}{Pr(C)} - \frac{Pr(A \,\&\, C)}{Pr(C)} \qquad \text{by algebra}$$

$$= 1 - \frac{Pr(A \,\&\, C)}{Pr(C)} \qquad \text{by algebra}$$

$$= 1 - Pr(A \mid C) \qquad \text{by definition.}$$

Chapter 11

Exercise 2. $-\$80$.

Exercise 4. $-\$0.79$.

Exercise 9. (a) \$460 million, (b) \$580 million, (c) \$220 million, (d) no, they won't conduct the study because it would be a waste of $\$5,000$. (The EMV of enacting the tax will be positive regardless of the study's findings. So doing the study won't help them make their decision.)

Exercise 11. (a) $-\$60$, (b) $-\$52$.

Exercise 16. $x = 888$.

Exercise 22. Suppose that $E(A) = \$x$. Let the possible payoffs of A be $x_1, \ldots, x_n$. Then:

$$
\begin{aligned}
E(\text{Pay } \$x \text{ for } A) &= Pr(\$x_1) \cdot (\$x_1 - \$x) + \ldots + Pr(\$x_n) \cdot (\$x_n - \$x) \\
&= [Pr(\$x_1) \cdot \$x_1 - Pr(\$x_1) \cdot \$x] + \ldots + [Pr(\$x_n) \cdot \$x_n - Pr(\$x_n) \cdot \$x] \\
&= E(A) - [Pr(\$x_1) \cdot \$x + \ldots + Pr(\$x_n) \cdot \$x] \\
&= E(A) - \$x \, [Pr(\$x_1) + \ldots + Pr(\$x_n)] \\
&= E(A) - \$x \\
&= 0.
\end{aligned}
$$

Chapter 12

Exercise 2. (d)

Exercise 3. 3/5

Exercise 5. (a) 3/5, (b) 2/3.

Exercise 8. (a) 103/2, (b) 490/9, (c) 45/49.

Exercise 10. Suppose action A has only two possible consequences, C_1 and C_2, such that $Pr(C_1) = Pr(C_2)$ and $U(C_1) = -U(C_2)$. Since

$Pr(C_1) = Pr(C_2) = 1/2$, we have:

$$
\begin{aligned}
EU(A) &= Pr(C_1)U(C_1) + Pr(C_2)U(C_2) \\
&= \frac{1}{2}U(C_1) + \frac{1}{2}U(C_2) \\
&= -\frac{1}{2}U(C_2) + \frac{1}{2}U(C_2) \\
&= 0.
\end{aligned}
$$

Chapter 13

Exercise 1. (d)

Exercise 2. (a)

Exercise 3. (b)

Exercise 5. (c)

Exercise 6. (c)

Exercise 8. (d)

Chapter 14

Exercise 1. (d)

Exercise 2. (b)

Exercise 5. $f(y) = e^y$. If we plug e^y into $\log(x)$ we get y back: $\log(e^y) = y$.

Exercise 8. (b)

Exercise 9. (c)

Chapter 16

Exercise 1. a) 7/10, b) 4/5, c) 2/3.

Exercise 3. (a) 5/6, (b) 1/3, (c) 2/5

Exercise 4. (a) 1/2, (b) 1/3, (c) 2/3, (d) 2/9

Chapter 17

Exercise 1. We make the following deals with Ronnie.

- He pays us $.40; we pay him $1 if A is true.
- He pays us $.70; we pay him $1 if A is false.

Each of these deals is fair according to Ronnie because the betting rates match his personal probabilities. For example, the expected value of the first bet is 0:

$$(\$1 - \$.40)Pr(A) - \$.40(1 - Pr(A)) = (\$.60)(4/10) - \$.40(6/10) = 0.$$

But Ronnie must pay us $1.10 for these bets, and he will only get $1 in return. Whether A is true or false, only one of the bets will pay off for Ronnie. So his net "gain" will be $\$1 - \$1.10 = -\$.10$.

Exercise 2. We make the following deals with Marco.

- We pay him $0.30; he pays us $1 if X is true.
- We pay him $0.20; he pays us $1 if Y is true.
- He pays us $0.60; we pay him $1 if $X \vee Y$ is true.

The explanation is similar to 17.1: each bet is fair according to Marco's personal probabilities, as can be checked by doing the expected value calculations which come out to 0 for each one. But he pays us $.10 more up front than we pay him, and he can't possibly win it back. If he wins the third bet, we win one or both of the first two bets.

Exercise 8. We make the following deals with Piz.

- We pay him $1/4; he pays us $1 if Pia is a basketball player.
- He pays us $1/3; we pay him $1 if Pia is a good basketball player.

Again the deals are fair given his personal probabilities. And again he pays us more up front than we pay him, money which he can't win back. If he wins the second bet, we win the first.

Chapter 18

Exercise 1. (a) 7/10, (b) 3/10, (c) same answers as before, (d) different answers: 217/300 and 11/36.

Exercise 2. a) 1/2, b) 5/8, c) yes, different answers here: 2/3 and 21/32, d) different again: 49/62 and 713/992.

Exercise 6. a) 2/3, b) 2/1, c) 1/2, d) different: 1/3 now

Chapter 19

Exercise 1. (a) $\mu = 20, \sigma = 4$, (b) a bell curve centered at 20 and reaching $y \approx 0$ around $x \approx 8$, (c) $a = 16, b = 24$, (d) $a = 12, b = 28$, (e) $a = 8, b = 32$, (f) 7 or fewer, (g) 33 or more, (h) answers will vary.

Exercise 3. (a) $\mu = 144, \sigma = 6$, (b) $132, 156$, (c) $126, 162$, (d) not significant at either level, (e) false (more accurately, not enough information is given to draw a conclusion either way).

Exercise 8. (a) $\mu = 10, \sigma = 3$, (b) no, (c) false, (d) false.

Exercise 9. (a) $\mu = 80, \sigma = 4$, (b) yes, (c) if the null hypothesis is true, and we repeated the experiment over and over, we would get a result this far from the mean less than 1% of the time.

Exercise 11. If the null hypothesis is true, the probability of getting a result this "extreme" (i.e. as improbable as this one, or even less probable) is below .05.

Chapter 20

Exercise 1. (a) and (b)

Exercise 2. (a) 25, (b) 400, (c) 16/17 or about 94%.

Exercise 3. (a) 225, (b) 675/676 or about 99.9%.

Exercise 5. (a) $\mu = 5, \sigma = 2$, (b) yes, (c) $\mu = 5/2, \sigma = 3/2$, (d) yes.