An Introduction to Probability and Inductive Logic

ment of causes, and of means to ends. Therefore, our universe too must have a Creator.

A common objection: The emergence of a well-organized universe just by chance would be astonishing. But if we imagine that matter in motion has been around, if not forever, at least for a very long time, then of course sooner or later we would arrive at a well-organized universe, just by chance. Hence the argument from design is defective.

Is this a sound objection?

4 Elementary Probability Ideas

KEY WORDS FOR REVIEW

<table>
<thead>
<tr>
<th>Relative frequency</th>
<th>Gambling system</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chance setup</td>
<td>Complexity</td>
</tr>
<tr>
<td>Biased/unbiased</td>
<td>Independence</td>
</tr>
<tr>
<td>Random</td>
<td>Probability model</td>
</tr>
</tbody>
</table>

This chapter explains the usual notation for talking about probability, and then reminds you how to add and multiply with probabilities.

WHAT HAS A PROBABILITY?

Suppose you want to take out car insurance. The insurance company will want to know your age, sex, driving experience, make of car, and so forth. They do so because they have a question in mind:

What is the probability that you will have an automobile accident next year?

That asks about a proposition (statement, assertion, conjecture, etc.):

“You will have an automobile accident next year.”

The company wants to know: What is the probability that this proposition is true?

The insurers could ask the same question in a different way:

What is the probability of your having an automobile accident next year?

This asks about an event (something of a certain sort happening). Will there be “an automobile accident next year, in which you are driving one of the cars involved”?

The company wants to know: What is the probability of this event occurring?

Obviously these are two different ways of asking the same question.

PROPOSITIONS AND EVENTS

Logicians are interested in arguments from premises to conclusions. Premises and conclusions are propositions. So inductive logic textbooks usually talk about the probability of propositions.
Most statisticians and most textbooks of probability talk about the probability of events.
So there are two languages of probability, propositions and events.
Propositions are true or false.
Events occur or do not occur.
Most of what we say in terms of propositions can be translated into event-language, and most of what we say in terms of events can be translated into proposition-language.
To begin with, we will sometimes talk one way, sometimes the other.
The distinction between propositions and events is not an important one now.
It matters only in the second half of this book.

WHY LEARN TWO LANGUAGES WHEN ONE WILL DO?
Because some students will already talk the event language, and others will talk the proposition language.
Because some students will go on to learn more statistics, and talk the event language. Other students will follow logic, and talk the proposition language.
The important thing is to be able to understand anyone who has something useful to say.
There is a general moral here. Be very careful and very clear about what you say. But do not be dogmatic about your own language. Be prepared to express any careful thought in the language your audience will understand. And be prepared to learn from someone who talks a language with which you are not familiar.

NOTATION: LOGIC
Propositions or events are represented by capital letters: A, B, C . . .
Logical compounds will be represented as follows, no matter whether we have propositions or events in mind:
Disjunction (or): A ∨ B for (A, or B, or both). We read this “A or B.”
Conjunction (and): A ∧ B for (A and B).
Negation (not): ¬A for (not A).

Example in the proposition language:
Let Z be the proposition that the roulette wheel stops at a zero. Let B be the proposition that the wheel stops at black.
ZvB is the proposition that the wheel stops at a zero or black.

Example:
Let Z: the wheel stops at a zero.
Let B: the wheel stops at black.
Then: ZvB = one or the other of those events occurs = the wheel stops at black or a zero.
Let R: the wheel stops at red.
In roulette, the wheel stops at a zero or black (ZvB) if and only if it does not stop at red (¬R). So,
¬R is equivalent to (ZvB).
R is equivalent to ¬(ZvB).

NOTATION: SETS
Statisticians usually do not talk about propositions. They talk about events in terms of set theory. Here is a rough translation of proposition language into event language.
The disjunction of two propositions, A v B, corresponds to the union of two sets of events, A ∪ B.
The conjunction of two propositions, A ∧ B, corresponds to the intersection of two sets of events, A ∩ B.
The negation of a proposition, ¬A, corresponds to the complement of a set of events, often written A'.

NOTATION: PROBABILITY
In courses on probability and statistics, textbooks usually write P( ) for probability. But our notation for probability will be:
P( ).
In the roulette example (Z for zero, R for red, B for black), all these are probabilities:
P(Z)  P(ZvB)  P(¬(ZvB))  P(R)
Earlier on this page we said that ¬R is equivalent to (ZvB). So:
P(¬R)  P(ZvB).
That is, the probability of not stopping at a red segment is the same as the probability of stopping at a zero or a black segment.
TWO CONVENTIONS

All of us—whether we were ever taught any probability theory or not—have got into the habit of expressing probabilities by percentages or fractions. That is:

Probabilities lie between 0 and 1.

In symbols, for any $A$,

$$0 \leq \Pr(A) \leq 1$$

At the extremes we have 0 and 1.

* In the language of propositions, what is certainly true has probability 1.
* In the language of events, what is bound to happen has probability 1.
* In probability textbooks, the sure event or a proposition that is certainly true is often represented by the last letter in the Greek alphabet, omega, written as a capital letter: $\Omega$. So our convention is written:

$$\Pr(\Omega) = 1$$

The probability of a proposition that is certainly true, or of an event that is sure to happen, is 1.

MUTUALLY EXCLUSIVE

Two propositions are called mutually exclusive if they can’t both be true at once. An ordinary roulette wheel cannot both stop at a zero (the house wins) and, on the same spin, stop at red. Hence these two propositions cannot both be true. They are mutually exclusive:

* The wheel will stop at a zero on the next spin.
* The wheel will stop at red on the next spin.

Likewise, two events which cannot both occur at once are called mutually exclusive. They are also called disjoint.

ADDING PROBABILITIES

There are some things about probability that “everybody” in college seems to know. For the moment we will just use this common knowledge. “Everybody” knows how to add probabilities.

More carefully: the probabilities of mutually exclusive propositions or events add up.

If $A$ and $B$ are mutually exclusive, $\Pr(A\cup B) = \Pr(A) + \Pr(B)$.

Thus if the probability of zero, in roulette, is $1/19$, and the probability of red is $9/19$, the probability that one or the other happens is:

$$\Pr(Z\cup R) = \Pr(Z) + \Pr(R) = 1/19 + 9/19 = 10/19.$$  

Example: Take a fair die.

Let $E$ = the die falls with an even number up.

$$E = (\text{the die falls 2, 4, or 6}). \quad \Pr(E) = \frac{1}{2}$$

* Why? Because $\Pr(2) = 1/6$, $\Pr(4) = 1/6$, $\Pr(6) = 1/6$. The events 2, 4, and 6 are mutually exclusive.
* Add $1/6 + 1/6 + 1/6$. You get $\frac{1}{2}$.

People who roll dice call the one-spot on a die the ace.

Let $M = \text{the die falls either ace up, or with a prime number up}$.

$$M = (\text{the die falls 1, 2, 3, or 5}). \quad \Pr(M) = \frac{4}{6} = \frac{2}{3}$$

But you cannot add $\Pr(E)$ to $\Pr(M)$ to get

** $\Pr(E\cup M) = 7/6.$ (WRONG)**

(We already know probabilities lie between 0 and 1, so $7/6$ is impossible.)

* Why can’t we add them up? Because $E$ and $M$ overlap: 2 is in both $E$ and $M$.
* $E$ and $M$ are not mutually exclusive.

In fact, $E\cap M = \text{the die falls 1, 2, 3, 4, 5, or 6}$, so that,

$$\Pr(E\cap M) = 1.$$  

You cannot add if the events or propositions “overlap.”

Adding probabilities is for mutually exclusive events or propositions.

INDEPENDENCE

Intuitively:

Two events are independent when the occurrence of one does not influence the probability of the occurrence of the other.
Two propositions are independent when the truth of one does not make the truth of the other any more or less probable.

Many people—like Fallacious Gambler—don’t understand independence very well. All the same, “everybody” seems to know that probabilities can be multiplied. More carefully: the probabilities of independent events or propositions can be multiplied.

**MULTIPLYING**

If A and B are independent, \( Pr(A \& B) = Pr(A) \times Pr(B) \).

We are rolling two fair dice. The outcome of tossing one die is independent of the outcome of tossing the other, so the probability of getting

- a five on the first toss (Five<sub>1</sub>)
- and a six on the second toss (Six<sub>2</sub>) is:

\[
Pr(\text{Five}_1 \& \text{Six}_2) = Pr(\text{Five}_1) \times Pr(\text{Six}_2) = 1/6 \times 1/6 = 1/36.
\]

Independence matters! Here is a mistake:

The probability of getting an even number (E) with a fair die is 1/2.

We found that the probability of M, of getting either an ace or a prime number, is 2/3.

What is the probability that on a single toss a die comes up both E and M?

We cannot reason:

\[
** Pr(E \& M) = Pr(E) \times Pr(M) = 1/2 \times 2/3 = 1/3. \text{(WRONG)}
\]

The two events are not independent. In fact, only one outcome is both even and prime, namely 2. Hence:

\[
Pr(E \& M) = Pr(2) = 1/6.
\]

Sometimes the fallacy is not obvious. Suppose you decide that the probability of the Toronto Blue Jays playing in the next World Series is 0.3 [Pr(J)], and that the probability of the Los Angeles Dodgers playing in the next world series is 0.4 [Pr(D)]. You cannot conclude that

\[
** Pr(D \& J) = Pr(D) \times Pr(J) = 0.4 \times 0.3 = 0.12. \text{(WRONG)}
\]

This is because the two events may not be independent. Maybe they are. But maybe, because of various player trades and so on, the Dodgers will do well only if they trade some players with the Jays, in which case the Jays won’t do so well.

---

**SIXES AND SEvens: ODD QUESTION 4**

Odd Question 4 went like this:

To throw a total of 7 with a pair of dice, you have to get a 1 and 6, or a 2 and a 5, or a 3 and a 4.

To throw a total of 6 with a pair of dice, you have to get a 1 and 5, or a 2 and a 4, or a 3 and another 3.

With two fair dice, you would expect:

- (a) To throw 7 more frequently than 6.
- (b) To throw 6 more frequently than 7.
- (c) To throw 6 and 7 equally often.

Many people think that 6 and 7 are equally probable. In fact, 7 is more probable than 6.

Look closely at what can happen in one roll of two dice, X and Y. We assume tosses are independent. There are 36 possible outcomes. In this table, (3,5), for example, means that X fell 3, while Y fell 5.

- (1,1) (2,1) (3,1) (4,1) (5,1) (6,1)
- (1,2) (2,2) (3,2) (4,2) (5,2) (6,2)
- (1,3) (2,3) (3,3) (4,3) (5,3) (6,3)
- (1,4) (2,4) (3,4) (4,4) (5,4) (6,4)
- (1,5) (2,5) (3,5) (4,5) (5,5) (6,5)
- (1,6) (2,6) (3,6) (4,6) (5,6) (6,6)

Circle the outcomes that add up to 6. How many?

Put a square around the outcomes that add up to 7. How many?

We can get a sum of seven in six ways: (1,6) or (2,5), or (3,4) or (4,3) or (5,2), or (6,1). Each of these outcomes has probability 1/36. (Independent tosses) So,

\[
Pr(7 \text{ with } 2 \text{ dice}) = 6/36 = 1/6. \text{(Mutually exclusive outcomes)}
\]

But we can get a sum of six in only five ways: (1,5) or (2,4) or (3,3) or (4,2) or (5,1). So,

\[
Pr(6 \text{ with } 2 \text{ dice}) = 5/36.
\]

**COMPOUNDING**

Throwing a 6 with one die is a single event. Throwing a sum of seven with two dice is a compound event. It involves two distinct outcomes, which are combined in the event "the sum of the dice equals 7."

A lot of simple probability reasoning involves compound events. Imagine a fair coin, and two urns, Urn 1 and Urn 2, made up as follows:

- Urn 1: 3 red balls, 1 green one.
- Urn 2: 1 red ball, 3 green ones.
In a fair drawing from Urn 1, the probability of getting a Red ball is \( Pr(R_1) = \frac{3}{4} \).

With Urn 2, it is \( Pr(R_2) = \frac{1}{4} \).

Now suppose we pick an urn by tossing a fair coin. If we get heads, we draw from Urn 1; if tails, from Urn 2. Assume independence, that is, that the toss of the coin has no effect on the urns.

What is the probability that we toss a coin and then draw a red ball from Urn 1? We first have to toss heads, and then draw a red ball from Urn 1 \((R_1)\).

\[ Pr(H\&R_1) = \frac{1}{2} \times \frac{3}{4} = \frac{3}{8} \]

The probability of tossing a coin and then drawing a red ball from Urn 2 \((R_2)\) is:

\[ Pr(T\&R_2) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8} \]

What is the probability of getting a red ball, using this setup? That is a compound event. We can get a red ball by getting heads with the coin, and then drawing red from Urn 1, \((H\&R_1)\), or by getting tails, and then drawing red from Urn 2 \((T\&R_2)\).

These are mutually exclusive events, and so can be added.

\[ Pr(\text{Red}) = Pr(H\&R_1) + Pr(T\&R_2) = \frac{3}{8} + \frac{1}{8} = \frac{1}{2} \]

So the probability of drawing red, in this set-up, is \( \frac{1}{2} \).

**A TRICK QUESTION**

Suppose we select one of those two urns by tossing a coin, and then make two draws from that urn with replacement. What is the probability of drawing two reds in a row in this set-up?

We know the probability of getting one red is \( \frac{1}{2} \).

Is the probability of two reds \( \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \)? \textbf{NO!}

The reason is that we can get two reds in a row in two different ways, which we'll call \( X \) and \( Y \):

\( X \): By tossing heads (an event of probability \( \frac{1}{2} \)), and then getting red from Urn 1 \((R_1)\), an event of probability \( \frac{3}{4} \) followed by replacing the ball and again drawing red from Urn 1 (another event of probability \( \frac{3}{4} \)).

\( Y \): By tossing tails (probability \( \frac{1}{2} \)), and then getting red from Urn 2 \((R_2)\), followed by replacing the ball and again drawing red from Urn 2 (another \( R_2 \)).

The probabilities are:

- \( Pr(X) = \frac{1}{2} \times \frac{3}{4} \times \frac{3}{4} = \frac{9}{32} \).
- \( Pr(Y) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{32} \).

Hence,

\[ Pr(\text{first ball drawn is red} \& \text{second ball drawn is red}) = Pr(X) + Pr(Y) = \frac{9}{32} + \frac{1}{32} = \frac{5}{16} \]

**UNDERSTANDING THE TRICK QUESTION**

Did you think that the probability of two reds would be \( \frac{1}{4} \)? Here is one way to understand why not. Think of doing two different experiments over and over again.

**Experiment 1.** Choose an urn by tossing a coin, and then draw a ball.

**Result 1.** You draw a red ball about half the time.

**Experiment 2.** Choose an urn by tossing a coin, and then draw two balls with replacement. (After you have drawn a ball, you put it back in the urn.)

**Result 2.** You get two red balls about \( \frac{5}{16} \) of the time, two green balls about \( \frac{5}{16} \) of the time, and a mix of one red and one green about \( \frac{6}{16} = \frac{3}{8} \) of the time.

**Explanation.** Once you have picked an urn with a "bias" for a given color, it is more probable that both balls will be of that color, than that you will get one majority and one minority color.

**LAPLACE**

This example was a great favorite with P.S. de Laplace (1749–1827), a truly major figure in the development of probability theory. He wrote the very first introductory college textbook about probability, *A Philosophical Essay on Probabilities*. He wrote this text for a class he taught at the polytechnic school in Paris in 1795, between the French Revolution and the rule of Napoleon.

Laplace was one of the finest mathematicians of his day. His *Analytic Theory of Probabilities* is still a rich source of ideas. His *Celestial Mechanics*—the mathematics of gravitation and astronomy—was equally important. He was very popular with the army, because he used mathematics to improve the French artillery.

He used to go to Napoleon's vegetarian lunches, where he gave informal talks about probability theory.

**EXERCISES**

1. **Galileo.** Don't feel bad if you gave the wrong answer to Odd Question 4, about rolling dice. A long time ago someone asked a similar question, about throwing three dice. Galileo (1564–1642), one of the greatest astronomers and physicists ever, took the time to explain the right and wrong answers.

   Explain why you might (wrongly) expect three fair dice to yield a sum of 9 as often as they yield a sum of 10. Why is it wrong to think that 9 is as probable as 10?

2. **One card.** A card is drawn from a standard deck of fifty-two cards which have been well shuffled seven times. What is the probability that the card is:
An Introduction to Probability and Inductive Logic

(a) Either a face card (jack, queen, king) or a ten?
(b) Either a spade or a face card?

3 Two cards. When two cards are drawn in succession from a standard pack of cards, what are the probabilities of drawing:
(a) Two hearts in a row, with replacement, and (b) without replacement.
(c) Two cards, neither of which is a heart, with replacement, and (d) without replacement.

4 Archery. An archer’s target has four concentric circles around a bull’s-eye. For a certain archer, the probabilities of scoring are as follows:

\[
\begin{align*}
Pr(\text{hit the bull's-eye}) &= 0.1 \\
Pr(\text{hit first circle, but not bull's-eye}) &= 0.3 \\
Pr(\text{hit second circle, but no better}) &= 0.2 \\
Pr(\text{hit third circle, but no better}) &= 0.2 \\
Pr(\text{hit fourth circle, but no better}) &= 0.1
\end{align*}
\]

Her shots are independent.
(a) What is the probability that in two shots she scores a bull’s-eye on the first shot, and the third circle on the second shot?
(b) What is the probability that in two shots she hits the bull’s-eye once, and the third circle once?
(c) What is the probability that on any one shot she misses the target entirely?

5 Polio from diapers (a news story).

Southampton, England: A man contracted polio from the soiled diaper of his niece, who had been vaccinated against the disease just days before, doctors said yesterday. "The probability of a person contracting polio from soiled diapers is literally one in three million," said consultant Martin Wale. What did Dr. Wale mean?

6 Languages. We distinguish between a “proposition-language” and an “event-language” for probability. Which language was used in:
(a) Question 2. (b) Question 3. (c) Question 4. (d) Dr. Wale’s statement in question 5?

KEY WORDS FOR REVIEW

<table>
<thead>
<tr>
<th>Events</th>
<th>Addition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propositions</td>
<td>Independence</td>
</tr>
<tr>
<td>Mutually exclusive</td>
<td>Multiplication</td>
</tr>
</tbody>
</table>

5 Conditional Probability

The most important new idea about probability is the probability that something happens, on condition that something else happens. This is called conditional probability.

CATEGORICAL AND CONDITIONAL

We express probabilities in numbers. Here is a story I read in the newspaper. The old tennis pro Ivan was discussing the probability that the rising young star Stefan would beat the established player Boris in the semifinals. Ivan was set to play Pete in the other semifinals match. He said,

The probability that Stefan will beat Boris is 40%.

Or he could have said,

The chance of Stefan’s winning is 0.4.

These are categorical statements, no ifs and buts about them. Ivan might also have this opinion:

Of course I’m going to win my semifinal match, but if I were to lose, then Stefan would not be so scared of meeting me in the finals, and he would play better; there would then be a 50-50 chance that Stefan would beat Boris.

This is the probability of Stefan’s winning in his semifinal match, conditional on Ivan losing the other semifinal. We call it the conditional probability. Here are other examples:

Categorical: The probability that there will be a bumper grain crop on the prairies next summer.
Conditional: The probability that there will be a bumper grain crop next summer, given that there has been very heavy snowfall the previous winter.

Categorical: The probability of dealing an ace as the second card from a standard pack of well-shuffled cards (regardless of what card is dealt first).

There are 4 aces and 52 cards, any one of which may come up as the second card. So the probability of getting an ace as the second card should be 4/52 = 1/13.

Conditional: The probability of dealing an ace as the second card on condition that the first card dealt was a king.

If a king was dealt first, there are 51 cards remaining. There are 4 aces still in the pack, so the conditional probability is 4/51.

Conditional: The probability of dealing an ace as the second card on condition that the first card dealt was also an ace.

When an ace is dealt first, there are 51 cards remaining, but only 3 aces, so the conditional probability is 3/51.

NOTATION

Categorical probability is represented:

Pr( )

Conditional probability is represented:

Pr( / ).

Examples of categorical probability:

Pr(S wins the final) = 0.4.
Pr(second card dealt is an ace) = 1/13.

Examples of conditional probability:

Pr(S wins his semifinal/1 loses his semifinal) = 0.5.
Pr(second card dealt is an ace/first card dealt is a king) = 4/51.

BINGO

Bingo players know about conditional probability.

In a game of bingo, you have a 5×5 card with 25 squares. Each square is marked with a different number from 1 to 99. The master of ceremonies draws numbered balls from a bag. Each time a number on your board is drawn, you fill in the corresponding square. You win (BINGO!) when you fill in a complete column, row, or diagonal.

Bingo players are fairly relaxed when they start the game. The probability that they will soon complete a line is small. But as they begin to fill in a line they get very excited, because the conditional probability of their winning is not so small.

PARKING TICKETS

If you park overnight near my home, and don't live on the block, you may be ticketed for not having a permit for overnight parking. The fine will be $20. But the street is only patrolled on average about once a week.

What is the probability of being fined?

Apparently the street is never patrolled on two consecutive nights. What is the probability of being ticketed tonight, conditional on having been ticketed on this street last night?

DEFINITION OF CONDITIONAL PROBABILITY

There is a very handy definition of conditional probability. We first state it, and then illustrate how it works.

\[
\text{When } Pr(B) > 0 \quad Pr(A / B) = Pr(A \& B) / Pr(B)
\]

Pr(B) must be a positive number, because we cannot divide by zero. But why is the rest of this definition sensible? Some examples will suggest why.

CONDITIONAL DICE

Think of a fair die. We say the outcome of a toss is even if it falls 2, 4, or 6 face up.

Here is conditional probability:

Pr(6/even)

In ordinary English:

The probability that we roll a 6, on condition that we rolled an even number.

The conditional probability of sixes, given evens.

With a fair die, we roll 2, 4, and 6 equally often. So 6 comes up a third of the time that we get an even outcome.

Pr(6/even) = 1/3.
This fits our definition, because,

\[
\Pr(6 \text{ & even}) = \Pr(6) = 1/6. \\
\Pr(\text{even}) = 1/2. \\
\Pr(6/\text{even}) = (1/6)/(1/2) = 1/3.
\]

**OVERLAPS**

Now ask a more complicated question, which involves overlapping outcomes. Let M mean that the die either falls 1 up or falls with a prime number up (2, 3, 5). Thus M happens when the die falls 1, 2, 3, or 5 uppermost. What is

\[
\Pr(\text{even}/M)\?
\]

The only prime even number is 2. There are 4 ways to throw M (1, 2, 3, 5). Hence, if the die is fair,

\[
\Pr(\text{even}/M) = 1/4.
\]

This fits our definition, because

\[
\Pr(\text{even} \& M) = 1/6. \\
\Pr(M) = 4/6. \\
\Pr(\text{even}/M) = (1/6)/(4/6) = 1/4.
\]

**WELL-SHUFFLED CARDS**

Think of a well-shuffled standard pack of 52 cards, from which the dealer deals the top card. He tells you that it is either red, or clubs. But not which. Call this information RvC.

Clubs are black. There are 13 clubs in the pack, and 26 other cards that are red. We were told that the first card is RvC. What is the probability that it is an ace? What is \(\Pr(A/RvC)\)? A & (RvC) is equivalent to ace of clubs, or a red ace, diamonds or hearts. For a total of 3. Hence 3 cards out of the 39 RvC cards are aces.

Hence the conditional probability is:

\[
\Pr(A/RvC) = 1/13.
\]

This agrees with our definition:

\[
\Pr(A \& (RvC)] = 3/52. \\
\Pr(RvC) = 39/52. \\
\Pr(A/(RvC)] = \frac{\Pr(A \& (RvC)]}{\Pr(RvC)} = 3/39 = 1/13.
\]

**URNS**

Imagine two urns, each containing red and green balls. Urn A has 80% red balls, 20% green, and Urn B has 60% green, 40% red. You pick an urn at random. Is it A or B? Let's draw balls from the urn and use this information to guess which urn it is. After each draw, the ball drawn is replaced. Hence for any draw, the probability of getting red from urn A is 0.8, and from urn B, the probability of getting red is 0.4.

\[
\Pr(R/A) = 0.8 \\
\Pr(R/B) = 0.4 \\
\Pr(A) = \Pr(B) = 0.5
\]

You draw a red ball. If you are like Alert Learner, that may lead you to suspect that this is urn A (which has more red balls than green ones). That is just a hunch. Let's be more exact.

We want to find \(\Pr(A/R)\), which is \(\Pr(A\&R)/[\Pr(R)]\).

You can get a red ball from either urn A or urn B. You get a red ball either when the event A&R happens, or when the event B&R happens. Event R is thus identical to (A&R)&(B&R).

The two alternatives (A&R) and (B&R) are mutually exclusive, so we can add up the probabilities.

\[
\Pr(R) = \Pr(A&R) + \Pr(B&R) \quad [1]
\]

The probability of getting urn B is 0.5; the probability of getting a red ball from it is 0.4, so that the probability of both happening is

\[
\Pr(B&R) = \Pr(R/B) = \Pr(R/B)\Pr(B) = 0.4 \times 0.5 = 0.2.
\]

Likewise,

\[
\Pr(A&R) = 0.8 \times 0.5 = 0.4.
\]

Putting these into formula [1] above,

\[
\Pr(R) = \Pr(A&R) + \Pr(B&R) = 0.4 + 0.2 = 0.6. \\
\text{Hence,} \Pr(A/R) = \Pr(A&R)/\Pr(R) = (0.4)/(0.6) = 2/3.
\]

**DRAWING THE CALCULATION TO CHECK IT**

You may find it helpful to visualize the calculation as a branching tree. We start out with our coin and the two urns. How can we get to a red ball? There are two routes. We can toss a heads (probability 0.5), giving us urn A. Then we can draw a red ball (probability 0.8). That is the route shown here on the top branch.
We can also get an R by tossing tails, going to urn B, and then drawing a red ball, as shown on the bottom branch.

We get to R on one of the two branches. So the total probability of ending up with R is the sum of the probabilities at the end of each branch. Here it is: 0.4 + 0.2 = 0.6.

The probability of getting to an R following an A branch is 0.4.

Thus that part of the probability that gets you to R by A, namely Pr(A/R), is A/.6 = 2/3.

MODELS

All our examples up to now have been dice, cards, urns. Now we turn to more interesting cases, more like real life. In each we make a model of a situation, and say that the real-life story is modeled by a standard ball-and-urn example.

SHOCK ABSORBERS

An automobile plant contracted to buy shock absorbers from two local suppliers, Bolt & Co. and Acme Inc. Bolt supplies 40% and Acme 60%. All shocks are subject to quality control. The ones that pass are called reliable.

Of Acme's shocks, 96% test reliable. But Bolt has been having some problems on the line, and recently only 72% of Bolt's shock absorbers have tested reliable.

What is the probability that a randomly chosen shock absorber will test reliable?

Intuitive guess: the probability will be lower than 0.96, because Acme's product is diluted by a proportion of shock absorbers from Bolt. The probability must be between 0.96 and 0.72, and nearer to 0.96. But by how much?

Solution

Let A = The shock chosen at random was made by Acme.

Let B = The shock chosen at random was made by Bolt.

Let R = The shock chosen at random is reliable.

Pr(A) = 0.6
Pr(B) = 0.4

Pr(R/A) = 0.96
Pr(R/B) = 0.72

So, Pr(R&A) = 0.576
So, Pr(R&B) = 0.288

R = Pr(R&A)+Pr(R&B)

Answer: Pr(R) = (0.6 × 0.96) + (0.4 × 0.72) = 0.576 + 0.288 = 0.864.

We can ask a more interesting question.

What is the conditional probability that a randomly chosen shock absorber, which is tested and found to be reliable, is made by Bolt?

Intuitive guess: look at the numbers. The automobile plant buys more shocks from Acme than Bolt. And Bolt's shocks are much less reliable than Acme's. Both these pieces of evidence count against a reliable shock, chosen at random, being made by Bolt. We expect that the probability that the shock is from Bolt is less than 0.4. But by how much?

Solution

We require Pr(B/R).

By definition, Pr(B/R) = Pr(B&R)/Pr(R) = 0.288/0.864 = 1/3.

Actually, you may like to do this without any multiplying, because almost all the numbers cancel:

= 0.4 × 0.72
= 0.6 × 0.96

Answer: Pr(B/R) = 1/3.

DRAWING TO CHECK

Pr(R) = 0.576 + 0.288 = 0.864. Pr(B/R) = 1/3.
WEIGHTLIFTERS

You learn that a certain country has two teams of weightlifters, either of which it may send to an international competition. Of the members of one team (the Steroid team), 80% have been regularly using steroids, but only 20% of the members of the other team are regular users (the Cleaner team). The head coach flips a fair coin to decide which team will be sent.

One member of the competing team is tested at random. He has been using steroids.

What is the conditional probability that the team in competition is the Steroid team, given that a member was found by a urine test to be using steroids? That is, what is \( \Pr(S/U) \)?

**Solution**

Let \( S \) = The coach sent the Steroid team.

Let \( C \) = The coach sent the Cleaner team.

Let \( U \) = A member selected at random uses steroids.

\[
\begin{align*}
\Pr(S) &= 0.5 \\
\Pr(U/S) &= 0.8 \\
\Pr(U/C) &= 0.2 \\
\Pr(U) &= 0.5 \\
\Pr(S|U) &= \frac{\Pr(S\cap U)}{\Pr(U)} = \frac{0.5 \times 0.8}{0.5} = 0.8
\end{align*}
\]

**Answer:** \( \Pr(S|U) \) = 0.8.

So the fact that we randomly selected a team member who uses steroids, is pretty good evidence that this is the Steroid team.

TWO IN A ROW: WITH REPLACEMENT

Back to the urns on page 51. Suppose you pick an urn at random, and make two draws, with replacement. You get a red, and then a red again. What is \( \Pr(A/R_1\&R_2) \)?

Let \( R_1 \) be the event that the first ball drawn is red, and \( R_2 \) the event that the second ball drawn is red. Then you can work out \( \Pr(A/R_1\&R_2) \) as:

\[
\begin{align*}
\Pr(A/R_1\&R_2) &= \Pr(A/R_1\&R_2) \Pr(R_1\&R_2) \\
&= \Pr(R_1\&R_2) \Pr(A/R_1\&R_2) \\
&= 0.8 \times 0.8 = 0.64
\end{align*}
\]

Now we know \( \Pr(A/R_1\&R_2) = \Pr(R_1\&R_2) \Pr(A/R_1\&R_2) = 0.8 \times 0.8 = 0.32 \).

Likewise, \( \Pr(B/R_1\&R_2) = 0.8 \times 0.8 = 0.32 \).

\[
\begin{align*}
\Pr(R_1\&R_2) &= \Pr(A\&R_1\&R_2) + \Pr(B\&R_1\&R_2) \\
&= 0.32 + 0.32 = 0.64 \\
\Pr(A/R_1\&R_2) &= \frac{0.32}{0.64} = 0.5 \\
\Pr(B/R_1\&R_2) &= \frac{0.32}{0.64} = 0.5
\end{align*}
\]

Thus a second red ball “increases the conditional probability” that this is urn A. The extra red ball may be taken as more evidence.

This suggests how we learn by experience by obtaining more evidence.

THE GAMBLER’S FALLOACY ONCE AGAIN

Fallacious Gambler thought that he could “learn from experience” when he saw that a fair (unbiased, independent trials) roulette wheel came up 12 blacks in a row. That is, he thought that:

\[
\Pr(\text{red on 13th trial}/12 \text{ blacks in a row}) > \frac{1}{2}.
\]

But if trials are independent, this probability is

\[
\frac{\Pr(BB\ldots B)}{\Pr(BBB\ldots BB)} = \frac{\left(\frac{1}{2}\right)^{12}}{\left(\frac{1}{2}\right)^{12}} = \frac{1}{2}.
\]

This is a new way to understand the gambler’s fallacy.

TWO WEIGHTLIFTERS: WITHOUT REPLACEMENT

Back to the weightlifters. Suppose we test two weightlifters chosen at random from a team that the coach selected by tossing a fair coin. We think: if both weightlifters test positive, that is pretty strong evidence that this is the Steroid team. Probability confirms this hunch.

We are sampling the team without replacement. So say there are ten members to a team. We randomly test two members.

Let \( S \) = The coach sent the Steroid team.

Let \( C \) = The coach sent the Cleaner team.

Let \( U_1 \) = The first member selected at random uses steroids.

Let \( U_2 \) = The second member selected at random uses steroids.

If we have the Steroid team, the probability that the first person tested uses steroids is 0.8 (on page 54 we had \( \Pr(U/S) = 0.8 \)). What is the probability of selecting two users?

There is a 4/5 probability of selecting one user. After the first person is selected, and turns out to be a user, there are 9 team members left, 7 of whom use steroids. So there is a 7/9 probability of getting a user for the next test. Hence the probability that the first two persons chosen from the Steroid team use steroids is \( 4/5 \times 7/9 = 28/45 \).

Likewise, the probability that the first two persons chosen from the Cleaner team use steroids is \( 1/5 \times 1/9 = 1/45 \).

The probability that the coach sent the Steroid team, when both team members selected at random are users, is,
\[
\text{Pr}(S \cap U \cap U_2) = 0.5(28/45) = 28/90.
\]
Likewise, \(\text{Pr}(C \cap U \cap U_2) = 0.5(1/45) = 1/90\).
\[
\text{Pr}(U_1 \cap U_2) = \text{Pr}(S \cap U \cap U_1) + \text{Pr}(C \cap U \cap U_2) = 29/90.
\]
\[
\text{Pr}(S/U \cap U_2) = \frac{\text{Pr}(S \cap U \cap U_2)}{\text{Pr}(U \cap U_2)} = 28/29 > 0.96
\]

Conditional probability that this is the Steroid team, given that we randomly selected:
- one weightlifter who was a user, is 0.8
- a second weightlifter who was also a user, is > 0.96

Getting two members who use steroids seems to be powerful evidence that the coach picked the Steroid team.

**EXERCISES**

1. **Phony precision about tennis.** In real life, the newspaper story about tennis quoted Ivan as stating a probability not of "40%" but:

   The probability that Stefan will beat Boris in the semifinals is only 37.325%.

   Can you make any sense out of this precise fraction?

2. **Heat lamps.** Three percent of production batches of Tropicana heat lamps fall below quality standards. Six percent of the batches of Florida heat lamps are below quality standards. A hardware store buys 40\% of its heat lamps from Tropicana, and 60\% from Florida.

   (a) What is the probability that a lamp taken at random in the store is made by Tropicana and is below quality standards?

   (b) What is the probability that a lamp taken at random in the store is below quality standards?

   (c) What is the probability that a lamp from this store, and found to be below quality standards, is made by Tropicana?

3. **The Triangle.** An unhealthy triangular-shaped region in an old industrial city once had a lot of chemical industry. Two percent of the children in the city live in the triangle. Fourteen percent of these test positive for excessive presence of toxic metals in the tissues. The rate of positive tests for children in the city, not living in the triangle, is only 1\%.

   (a) What is the probability that a child who lives in the city, and who is chosen at random, both lives in the Triangle and tests positive?

   (b) What is the probability that a child living in the city, chosen at random, tests positive?

   (c) What is the probability that a child chosen at random, who tests positive, lives in the Triangle?

4. **Taxicabs.** Draw a tree diagram for the taxicab problem. Odd Question 5.

5. **Laplace's trick question.** Look back at Laplace's question (page 44). An experiment consists of tossing a coin to select an urn, then drawing a ball, noting its color, replacing it, and drawing another ball and noting its color. Find, \(\text{Pr}\) (second ball drawn is red/first ball drawn is red).

6. **Understanding the question.** On page 45 we ended Chapter 4 with a way to understand Laplace's trick question. How does your answer to question 6 help with understanding Laplace's question?

**KEY WORDS FOR REVIEW**

Categorical
<table>
<thead>
<tr>
<th>Conditional probability</th>
<th>Calculating conditional probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Definition of conditional probability</td>
<td>Models</td>
</tr>
<tr>
<td>Learning from experience</td>
<td></td>
</tr>
</tbody>
</table>