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?

KEY WORDS FOR REVIEW

Relative frequency	Gambling system
Chance setup	Complexity
Biased/unbiased	Independence
Random	Probability model

4 Elementary Probability Ideas

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): AvB 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, AvB, corresponds to the union of two sets of events, $A \cup B$.

The conjunction of two propositions, A&B, corresponds to the intersection of two sets of events, $A \cap B$.

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:

Pr ( ).

In the roulette example (Z for zero, R for red, B for black), all these are probabilities:

Pr(Z) Pr(ZvB) Pr(~(ZvB)) Pr(R)

Earlier on this page we said that ~R is equivalent to (ZvB). So:

Pr(~R) = Pr(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: Ω. 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(AvB) = 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(ZvR) = 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 = (the die falls 2, 4, or 6). Pr(E) = ½

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 ½.

People who roll dice call the one-spot on a die the *ace*.
Let M = the die falls either ace up, or with a prime number up.

M = (the die falls 1, 2, 3, or 5). Pr(M) = 4/6 = 2/3

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

** Pr(EvM) = 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, (EvM) = (the die falls 1, 2, 3, 4, 5, or 6), so that,

Pr(EvM) = 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) × 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 ($Five_1$)
and a six on the second toss (Six_2) is:

$Pr(Five_1 \& Six_2) = Pr(Five_1) \times Pr(Six_2) = 1/6 \times 1/6 = 1/36.$

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) × Pr(M) = 1/2 × 2/3 = 1/3. (*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) × Pr(J) = 0.4 × 0.3 = 0.12. (*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.

Multiplying probabilities is for independent events or propositions.

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 a 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 a 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 with 2 dice) = 6/36 = 1/6. (*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 with 2 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 $Pr(R_1) = 3/4$.

With Urn 2, it is $Pr(R_2) = 1/4$.

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).

$$Pr(H\&R_1) = 1/2 \times 3/4 = 3/8.$$

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

$$Pr(T\&R_2) = 1/2 \times 1/4 = 1/8.$$

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.

$$Pr(Red) = Pr(H\&R_1) + Pr(T\&R_2) = 3/8 + 1/8 = 1/2.$$

So the probability of drawing red, in this set-up, is 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 setup?

We know the probability of getting one red is 1/2.

Is the probability of two reds (1/2)(1/2) = 1/4? 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 ½), and then getting red from Urn 1 (R_1, an event of probability 3/4) followed by replacing the ball and again drawing red from Urn 1 (another event of probability 3/4).

Y: By tossing tails (probability 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) = (1/2)(3/4)(3/4) = 9/32.
- Pr(Y) = (1/2)(1/4)(1/4) = 1/32.

Hence,

Pr(first ball drawn is red & second ball drawn is red)
= Pr(X) + Pr(Y) = 10/32 = 5/16.

UNDERSTANDING THE TRICK QUESTION

Did you think that the probability of two reds would be 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.

Results 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.)

Results 2. You get two red balls about 5/16 of the time, two green balls about 5/16 of the time, and a mix of one red and one green about 6/16 = 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:

(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:

Pr(hit the bull's-eye) = 0.1
Pr(hit first circle, but not bull's-eye) = 0.3.
Pr(hit second circle, but no better) = 0.2.
Pr(hit third circle, but no better) = 0.2.
Pr(hit fourth circle, but no better) = 0.1.

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

Events	Addition
Propositions	Independence
Mutually exclusive	Multiplication

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.