January 12, 1999, he would have been under seven feet tall. I will say that if a net force had been applied to my harpsichord, it would have moved. But I will not say that if this penny had been in Goodman's pocket on VE day it would have been silver nor will I say that if Jupiter were made of pure gold it would have a mass of less than 100,000 kilograms.

Some metaphysicians have held that statements of what might have been are objective statements about parallel worlds or branches of time. Other thinkers hold that correct counterfactuals are fables constructed according to our contingent rules for changing beliefs. Reasoning about what might have been has value for them only as practice for reasoning about what might be.

It should be now clear that lawlike and accidental conditions are different, and you have some general indication of how they are different, but the specification of differences has not been precise. How exactly do laws function as contingent rules of inference? What are the rules for changing our beliefs about laws? Just what is needed for a law to support a given counterfactual? Despite an enormous amount of work there is, as yet, no generally satisfactory solution to these and related problems. They remain a major area of concern for the philosophy of science.

Suggested readings

Nelson Goodman, *Fact, Fiction and Forecast* (4th. ed.). (Cambridge, MA: Harvard University Press, 1983).

Douglas Stalker (ed.), *Grue! the new riddle of induction* (Chicago Open Court, 1994).

VI

The Probability Calculus

VI.1. INTRODUCTION. The theory of probability resulted from the cooperation of two eminent seventeenth-century mathematicians and a gambler. The gambler, Chevalier de Méré, had some theoretical problems with practical consequences at the dice tables. He took his problems to Blaise Pascal who in turn entered into correspondence with Pierre de Fermat, in order to discuss them. The mathematical theory of probability was born in the Pascal-Fermat correspondence.

We have used the word "probability" rather freely in the discussion so far, with only a rough, intuitive grasp of its meaning. In this chapter we will learn the mathematical rules that a quantity must satisfy in order to qualify as a probability.

VI.2. PROBABILITY, ARGUMENTS, STATEMENTS, AND PROPERTIES. The word "probability" is used for a number of distinct concepts. Earlier I pointed out the difference between inductive probability, which applies to arguments, and epistemic probability, which applies to statements. There is yet another type of probability, which applies to properties. When we speak of the probability *of* throwing a "natural" in dice, or the probability *of* living to age 65, we are ascribing probabilities to properties. When we speak of the probability *that* John Q. Jones will live to age 65, or the probability *that* the next throw of the dice will come up a natural, we are ascribing probabilities to statements. Thus, there are at least three different types of probability which apply to three different types of things: arguments, statements, and properties.

Luckily, there is a common core to these various concepts of probability: Each of these various types of probability obeys the rules of the mathematical theory of probability. Furthermore, the different types of probability are interrelated in other ways, some of which were brought out in the discussion of inductive and epistemic probability. In Chapter VI it will be shown how these different concepts of probability put flesh on the skeleton of the mathematical theory of probability. Here, however, we shall restrict ourselves to developing the mathematical theory.

The mathematical theory is often called the *probability calculus*. In order to facilitate the framing of examples we shall develop the probability calculus as it applies to *statements*. But we shall see later how it can also accommodate arguments and properties.

Remember that the truth tables for "~," "&," and "v" enable us to find out whether a complex statement is true or false if we know whether its simple constituent statements are true or false. However, truth tables tell us nothing about the truth or falsity of the simple constituent statements. In a similar manner, the rules of the probability calculus tell us how the probability of a complex statement is related to the probability of its simple constituent statements, but they do not tell us how to determine the probabilities of simple statements. The problem of determining the probability of simple statements (or properties or arguments) is a problem of inductive logic, but it is a problem that is not solved by the probability calculus.

Probability values assigned to complex statements range from 0 to 1. Although the probability calculus does not tell us how to determine the probabilities of simple statements, it does assign the extreme values of 0 and 1 to special kinds of complex statements. Previously we discussed complex statements that are *true* no matter what the facts are. These statements were called tautologies. Since a tautology is guaranteed to be true, no matter what the facts are, it is assigned the highest possible probability value.

> **Rule 1:** If a statement is a tautology, then its probability is equal to 1.

Thus, just as the complex statement *sv~s* is true no matter whether its simple constituent statement, *s*, is true or false, so its probability is 1 regardless of the probability of the simple constituent statement.

We also discussed another type of statement that is *false* no matter what the facts are. This type of statement, called the self-contradiction, is assigned the lowest possible probability value.

> **Rule 2:** If a statement is a self-contradiction, then its probability is equal to 0.

Thus, just as the complex statement *s&~s* is false no matter whether its simple constituent statement, *s*, is true or false, so its probability is 0 regardless of the simple constituent statement.

When two statements make the same factual claim, that is, when they are true in exactly the same circumstances, they are logically equivalent. Now if a statement that makes a factual claim has a certain probability, another statement that makes exactly the same claim in different words should be equally probable. The statement "My next throw of the dice will come up a natural" should have the same probability as "It is *not* the case that my next throw of the dice will *not* come up a natural." This fact is reflected in the following rule:

> **Rule 3:** If two statements are logically equivalent, then they have the same probability.

By the truth table method it is easy to show that the simple statement *p* is logically equivalent to the complex statement that is its double negation, *~~p*, since they are true in exactly the same cases.

	p	*~p*	*~~p*
Case 1:	T	F	T
Case 2:	F	T	F

Thus, the simple statement "My next throw of the dice will come up a natural" has, according to Rule 3, the same probability as its double negation, "It is *not* the case that my next throw of the dice will *not* come up a natural."

The first two rules cover certain special cases. They tell us the probability of a complex statement if it is either a tautology or a contradiction. The third rule tells us how to find the probability of a complex contingent statement from its simple constituent statements, if that complex statement is logically equivalent to one of its simple constituent statements. But there are many complex contingent statements that are not logically equivalent to any of their simple constituent statements, and more rules shall be introduced to cover them. The next two sections present rules for each of the logical connectives.

Exercises

Instead of writing "The probability of *p* is $\frac{1}{2}$," we shall write, for short "Pr(*p*) = $\frac{1}{2}$." Now suppose that Pr(*p*) = $\frac{1}{2}$ and Pr(*q*) = $\frac{1}{4}$. Find the probabilities of the following complex statements, using Rules 1 through 3 and the method of truth tables:

1. *pvp*.
2. *q&q*.
3. *q&~q*.
4. ~(*q&~q*).
5. ~(*pv~p*).
6. ~~(*pv~p*).
7. *pv*(*q&~q*).
8. *q&*(*pv~p*).

VI.3. DISJUNCTION AND NEGATION RULES. The probability of a disjunction *pvq* is most easily calculated when its disjuncts, *p* and *q*, are *mutually exclusive* or inconsistent with each other. In such a case the probability of the disjunction can be calculated from the probabilities of the disjuncts by means of the *special disjunction rule*. We shall use the notation introducing the exercises at the end of the previous section writing "The probability of *p* is *x*" as: "Pr(*p*) = *x*."

> **Rule 4:** If *p* and *q* are mutually exclusive, then Pr(*pvq*) = Pr(*p*) + Pr(*q*).

For example, the statements "Socrates is both bald and wise" and "Socrates is neither bald nor wise" are mutually exclusive. Thus, if the probability that Socrates is both bald and wise is $\frac{1}{2}$ and the probability that Socrates is neither bald nor wise is $\frac{1}{4}$, then the probability that Socrates is either both bald and wise *or* neither bald nor wise is $\frac{1}{2} + \frac{1}{4}$, or $\frac{3}{4}$.

We can do a little more with the special alternation rule in the following case: Suppose you are about to throw a single six-sided die and that each of the six outcomes is equally probable; that is:

$$\Pr(\text{the die will come up a } 1) = \tfrac{1}{6}$$
$$\Pr(\text{the die will come up a } 2) = \tfrac{1}{6}$$
$$\Pr(\text{the die will come up a } 3) = \tfrac{1}{6}$$
$$\Pr(\text{the die will come up a } 4) = \tfrac{1}{6}$$
$$\Pr(\text{the die will come up a } 5) = \tfrac{1}{6}$$
$$\Pr(\text{the die will come up a } 6) = \tfrac{1}{6}$$

Since the die can show only one face at a time, these six statements may be treated as being mutually exclusive.[1] Thus, the probability of getting a 1 or a 6 may be calculated by the special disjunction rule as follows:

$$\Pr(1\text{v}6) = \Pr(1) + \Pr(6) = \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{3}$$

The probability of getting an even number may be calculated as

$$\Pr(\text{even}) = \Pr(2\text{v}4\text{v}6) = \Pr(2) + \Pr(4) + \Pr(6) = \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{2}$$

The probability of getting an even number that is greater than 3 may be calculated as

$$\Pr(\text{even and greater than } 3) = \Pr(4\text{v}6) = \Pr(4) + \Pr(6) = \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{1}{3}$$

The probability of getting an even number or a 3 may be calculated as

$$\Pr(\text{even or } 3) = \Pr(2\text{v}4\text{v}6\text{v}3) = \tfrac{4}{6} = \tfrac{2}{3}$$

Finally, calculating the probability of getting either a 1, 2, 3, 4, 5, or 6 (that is, the probability that the die will show one face or another) gives $\frac{6}{6}$, or 1.

[1]Actually the statements are not mutually exclusive in the logical sense. We cannot show that they are inconsistent with each other by the method of truth tables, and it is logically possible that the die might change shape upon being thrown so as to display two faces simultaneously. To treat this case rigorously, we would have to use the general disjunction rule, along with a battery of assumptions: $\Pr(1\&2) = 0$, $\Pr(2\&3) = 0$, $\Pr(1\&3) = 0$, etc. However, we shall see that the result is the same as when we use the special disjunction rule, and treat these statements as if they were mutually exclusive.

We will now apply the special disjunction rule to a case of more general interest. It can be shown, by the method of truth tables, that any statement p is inconsistent with its negation, $\sim p$. Since p and $\sim p$ are therefore mutually exclusive, the special disjunction rule permits the conclusion that

$$\Pr(p\text{v}{\sim}p) = \Pr(p) + \Pr({\sim}p)$$

But the statement $p\text{v}{\sim}p$ is a tautology, so by Rule 1,

$$\Pr(p\text{v}{\sim}p) = 1$$

Putting these two conclusions together gives

$$\Pr(p) + \Pr({\sim}p) = 1$$

If the quantity $\Pr(p)$ is subtracted from both sides of the equation, the sides will remain equal, so we may conclude that

$$\Pr({\sim}p) = 1 - \Pr(p)$$

This conclusion holds good for any statement, since any statement is inconsistent with its negation, and for any statement p its disjunction with its negation, $p\text{v}{\sim}p$, is a tautology. This therefore establishes a general negation rule, which allows us to calculate the probability of a negation from the probability of its constituent statement:

Rule 5: $\Pr({\sim}p) = 1 - \Pr(p)$.

Suppose in the example using the die we wanted to know the probability of not getting a 3:

$$\Pr({\sim}3) = 1 - \Pr(3) = 1 - \tfrac{1}{6} = \tfrac{5}{6}$$

Note that we get the same answer as we would if we took the long road to solving the problem and confined ourselves to using the special disjunction rule:

$$\begin{aligned}\Pr({\sim}3) &= \Pr(1\text{v}2\text{v}4\text{v}5\text{v}6)\\ &= \Pr(1) + \Pr(2) + \Pr(4) + \Pr(5) + \Pr(6)\\ &= \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} + \tfrac{1}{6} = \tfrac{5}{6}\end{aligned}$$

We shall apply the special disjunction rule one more time in order to establish another generally useful rule. For any two statements, p, q, we can show by the truth table method that the complex statements $p\&q$, $p\&{\sim}q$, and ${\sim}p\&q$ are inconsistent with each other. As shown in the following table, there is no case in which two of them are true:

	p	q	$\sim p$	$\sim q$	$p\&q$	$p\&\sim q$	$\sim p\&q$
Case 1:	T	T	F	F	T	F	F
Case 2:	T	F	F	T	F	T	F
Case 3:	F	T	T	F	F	F	T
Case 4:	F	F	T	T	F	F	F

Since they are mutually exclusive, we can apply the special disjunction rule and conclude:

a. $\Pr[(p\&q)\mathrm{v}(p\&\sim q)] = \Pr(p\&q) + \Pr(p\&\sim q)$

b. $\Pr[(p\&q)\mathrm{v}(\sim p\&q)] = \Pr(p\&q) + \Pr(\sim p\&q)$

c. $\Pr[(p\&q)\mathrm{v}(p\&\sim q)\mathrm{v}(\sim p\&q)] = \Pr(p\&q) + \Pr(p\&\sim q) + \Pr(\sim p\&q)$

But the complex statement $(p\&q)\mathrm{v}(p\&\sim q)$ is logically equivalent to the simple statement p, as is shown by the following truth table:

	p	q	$\sim q$	$p\&q$	$p\&\sim q$	$(p\&q)\mathrm{v}(p\&\sim q)$
Case 1:	T	T	F	T	F	T
Case 2:	T	F	T	F	T	T
Case 3:	F	T	F	F	F	F
Case 4:	F	F	T	F	F	F

Since, according to Rule 3, logically equivalent statements have the same probability, equation (a) may be rewritten as

a′. $\Pr(p) = \Pr(p\&q) + \Pr(p\&\sim q)$

A similar truth table will show that the complex statement $(p\&q)\mathrm{v}(\sim p\&q)$ is logically equivalent to the simple statement q. Therefore, equation (b) may be rewritten as

b′. $\Pr(q) = \Pr(p\&q) + \Pr(\sim p\&q)$

Finally, a truth table will show that the complex statement $(p\&q)\mathrm{v}(p\&\sim q)\mathrm{v}(\sim p\&q)$ is logically equivalent to the complex statement $p\mathrm{v}q$, which enables us to rewrite equation (c) as

c′. $\Pr(p\mathrm{v}q) = \Pr(p\&q) + \Pr(p\&\sim q) + \Pr(\sim p\&q)$

Now let us add equations (a′) and (b′) together to get

d. $\Pr(p) + \Pr(q) = 2\Pr(p\&q) + \Pr(p\&\sim q) + \Pr(\sim p\&q)$

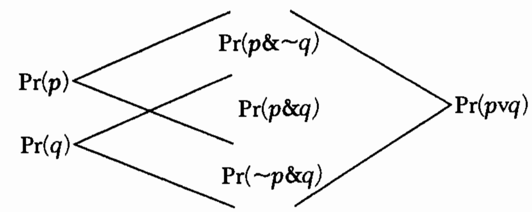
If we subtract the quantity $\Pr(p\&q)$ from both sides of the preceding equation, we get

d′. $\Pr(p) + \Pr(q) - \Pr(p\&q) = \Pr(p\&q) + \Pr(p\&\sim q) + \Pr(\sim p\&q)$

If equation (d′) is compared with equation (c′) we see that $\Pr(p\mathrm{v}q)$ is equal to the same thing as $\Pr(p) + \Pr(q) - \Pr(p\&q)$. This establishes a general disjunction rule that is good for all disjunctions, whether the disjuncts are mutually exclusive or not:

Rule 6: $\Pr(p\mathrm{v}q) = \Pr(p) + \Pr(q) - \Pr(p\&q)$.

If some of the algebra used to establish the general disjunction rule has left you behind, the following diagram may help to make the reasoning clear:



When $\Pr(p)$ is added to $\Pr(q)$, then $\Pr(p\&q)$ is counted twice. But to get $\Pr(p\mathrm{v}q)$, it should be counted only once. Thus, to get $\Pr(p\mathrm{v}q)$, we add $\Pr(p)$ and $\Pr(q)$ and then subtract $\Pr(p\&q)$ to make up for having counted it twice. In the case in which p and q are mutually exclusive, this makes no difference, because when p and q are mutually exclusive, $\Pr(p\&q) = 0$. No matter how many times 0 is counted, we will always get the same result. For example, by the general disjunction rule, $\Pr(p\mathrm{v}\sim p) = \Pr(p) + \Pr(\sim p) - \Pr(p\&\sim p)$. But the statement $p\&\sim p$ is a self-contradiction, so its probability is zero. Thus, we get the same result as if we had used the special disjunction rule. Counting $\Pr(p\&q)$ twice *does* make a difference when p and q are *not* mutually exclusive. Suppose we use the general disjunction rule to calculate the probability of the complex statement $p\mathrm{v}p$:

$$\Pr(p\mathrm{v}p) = \Pr(p) + \Pr(p) - \Pr(p\&p)$$

But since the complex statement $p\&p$ is logically equivalent to the simple statement p, $\Pr(p\&p) = \Pr(p)$, we get

$$\Pr(p\mathrm{v}p) = \Pr(p) + \Pr(p) - \Pr(p) = \Pr(p)$$

We know this is the correct answer, because the complex statement $p\mathrm{v}p$ is also logically equivalent to the simple statement p.

The example with the die shall be used to give one more illustration of the use of the general disjunction rule. Suppose that we want to know the probability that the die will come up an even number or a number less than 3. There is a way to calculate this probability using only the special disjunction rule:

$$\begin{aligned}\Pr(\text{even v less than 3}) &= \Pr(1\text{v}2\text{v}4\text{v}6)\\ &= \Pr(1) + \Pr(2) + \Pr(4) + \Pr(6) = \tfrac{4}{6} = \tfrac{2}{3}\end{aligned}$$

We may use the special disjunction rule because the outcomes 1, 2, 4, and 6 are mutually exclusive. However, the outcomes "even" and "less than 3" are not mutually exclusive, since the die might come up 2. Thus, we may apply the general disjunction rule as follows:

$$\begin{aligned}&\Pr(\text{even v less than 3})\\ &\quad = \Pr(\text{even}) + \Pr(\text{less than 3}) - \Pr(\text{even\&less than 3})\end{aligned}$$

Now we may calculate Pr(even) as Pr(2v4v6) by the special disjunction rule; it is equal to $\frac{1}{2}$. We may calculate Pr(less than 3) as Pr(1v2) by the special disjunction rule; it is equal to $\frac{1}{3}$. And we may calculate Pr(even&less than 3) as Pr(2), which is equal to $\frac{1}{6}$. So, by this method,

$$\Pr(\text{even v less than 3}) = \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{6} = \tfrac{2}{3}$$

The role of the subtraction term can be seen clearly in this example. What we have done is to calculate Pr(even v less than 3) as

$$\Pr(2\text{v}4\text{v}6) + \Pr(1\text{v}2) - \Pr(2)$$

so the subtraction term compensates for adding in Pr(2) twice when we add Pr(even) and Pr(less than 3). In this example use of the general disjunction rule was the long way of solving the problem. But in some cases it is necessary to use the general disjunction rule. Suppose you are told that

$$\Pr(p) = \tfrac{1}{2}$$

$$\Pr(q) = \tfrac{1}{3}$$

$$\Pr(p\&q) = \tfrac{1}{4}$$

You are asked to calculate $\Pr(p\text{v}q)$. Now you cannot use the special disjunction rule since you know that p and q are not mutually exclusive. If they were, $\Pr(p\&q)$ would be 0, and you are told that it is $\frac{1}{4}$. Therefore, you must use the general disjunction rule in the following way:

$$\begin{aligned}\Pr(p\text{v}q) &= \Pr(p) + \Pr(q) - \Pr(p\&q)\\ &= \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{4} = \tfrac{7}{12}\end{aligned}$$

In Section VI.2, we compared the rules of the probability calculus to the way in which the truth tables for the logical connectives relate the truth or falsity of a complex statement to the truth or falsity of its simple constituent statements. We are now at the point where we must qualify this comparison. We can always determine the truth or falsity of a complex statement if we know whether its simple constituent statements are true or false. But we cannot always calculate the probability of a complex statement from the probabilities of its simple constituent statements. Sometimes, as in the example above, in order to calculate the probability of the complex statement $p\text{v}q$, we need not only know the probabilities of its simple constituent statements, p and q, we also need to know the probability of another complex statement, $p\&q$. We shall discuss the rules that govern the probabilities of such conjunctions in the next section. However, we shall find that it is not always possible to calculate the probability of a conjunction simply from the probabilities of its constituent statements.

Exercises

1. Suppose you have an ordinary deck of 52 playing cards. You are to draw one card. Assume that each card has a probability of 1/52 of being drawn. What is the probability that you will draw:
   a. The ace of spades?
   b. The queen of hearts?
   c. The ace of spades or the queen of hearts?
   d. An ace?
   e. A heart?
   f. A face card (king, queen, or jack)?
   g. A card that is not a face card?
   h. An ace or a spade?
   i. A queen or a heart?
   j. A queen or a non-spade?
2. $\Pr(p) = \frac{1}{2}$, $\Pr(q) = \frac{1}{2}$, $\Pr(p\&q) = \frac{1}{8}$. What is $\Pr(p\text{v}q)$?
3. $\Pr(r) = \frac{1}{2}$, $\Pr(s) = \frac{1}{4}$, $\Pr(r\text{v}s) = \frac{3}{4}$. What is $\Pr(r\&s)$?
4. $\Pr(u) = \frac{1}{2}$, $\Pr(t) = \frac{3}{4}$, $\Pr(u\&{\sim}t) = \frac{1}{8}$. What is $\Pr(u\text{v}{\sim}t)$?

VI.4. CONJUNCTION RULES AND CONDITIONAL PROBABILITY. Before the rules that govern the probability of conjunctions are discussed, it is necessary to introduce the notion of *conditional probability*. We may write $\Pr(q \text{ given } p)$ as the probability of q *on the condition that* p.

This probability may or may not be different from $\Pr(q)$. We shall deal with the concept of conditional probability on the intuitive level before a precise definition for it is introduced.

In the example with the die, we found that the probability of throwing an even number was $\frac{1}{2}$. However, the probability of getting an even number *given that* a 2 or a 4 is thrown is not $\frac{1}{2}$ but 1. And the probability of casting an even number *given* that a 1 or a 3 is thrown is 0. To take a little more complicated example, suppose that the die remains unchanged and you are to bet on whether it will come up even, with a special agreement that if it comes up 5 all bets will be off and it will be thrown again. In such a situation you would be interested in the probability that it will come up even *given that* it will be either a 1, 2, 3, 4, or 6. This probability should be greater than $\frac{1}{2}$ since the condition excludes one of the ways in which the die could come up odd. It is, in fact, $\frac{3}{5}$. Thus, the probabilities of "even," given three different conditions, are each different from the probability of "even" by itself:

a. $\Pr(\text{even}) = \frac{1}{2}$

b. $\Pr(\text{even } \textit{given } 2\text{v}4) = 1$

c. $\Pr(\text{even } \textit{given } 1\text{v}3) = 0$

d. $\Pr(\text{even } \textit{given } 1\text{v}2\text{v}3\text{v}4\text{v}6) = \frac{3}{5}$

Conditional probabilities allow for the fact that if a certain statement, p, is known to be true, this may affect the probability to be assigned to another statement, q. The most striking cases occur when there is a deductively valid argument from p to q:

p = The next throw of the die will come up 2
v
the next throw of the die will come up 4.

---

q = The next throw of the die will come up even.

In this case, $\Pr(q \textit{ given } p) = 1$:[2]

$\Pr(\text{even } \textit{given } 2\text{v}4) = 1$

Suppose there is a deductively valid argument from p to $\sim q$:

[2] We must make one qualification to this statement. When p is a self-contradiction, then for any statement q there is a deductively valid argument from p to q *and* a deductively valid argument from p to $\sim q$. In such a case, $\Pr(q \textit{ given } p)$ has no value.

p = The next throw of the die will come up 1
v
the next throw of the die will come up 3.

$\sim q$ = The next throw of the die will *not* come up even.

In this case, $\Pr(q \textit{ given } p) = 0$:

$\Pr(\text{even } \textit{given } 1\text{v}3) = 0$.[3]

There are, however, important cases where neither the argument from p to q nor the argument from p to $\sim q$ is deductively valid and yet $\Pr(q \textit{ given } p)$ differs from $\Pr(q)$, as in the previous example with the die:

$\Pr(\text{even } \textit{given } 1\text{v}2\text{v}3\text{v}4\text{v}6) = \frac{3}{5}$

$\Pr(\text{even}) = \frac{1}{2}$

There are other cases where the knowledge that p is true may be completely irrelevant to the probability to be assigned to q. For example, it was said that the probability that the next throw of the die will come up even is $\frac{1}{2}$. We could say that the probability that the next throw of the die will come up even, given that the President of the United States sneezes simultaneously with our throw, is still $\frac{1}{2}$. The President's sneeze is irrelevant to the probability assigned to "even." Thus, the two statements "The next throw of the die will come up even" and "The President of the United States will sneeze simultaneously with the next throw of the die" are *independent*.[4]

We can now give substance to the intuitive notions of conditional probability and independence by defining them in terms of pure statement probabilities. First we will define conditional probability:

Definition 12: *Conditional probability:*[5]

$$\Pr(q \textit{ given } p) = \frac{\Pr(p\&q)}{\Pr(p)}$$

Let us see how this definition works out in the example of the die:

[3] We must make one qualification to this statement. When p is a self-contradiction, then for any statement q there is a deductively valid argument from p to q *and* a deductively valid argument from p to $\sim q$. In such a case, $\Pr(q \textit{ given } p)$ has no value.

[4] This type of independence is called probabilistic or stochastic independence. It should not be confused with the mutual logical independence discussed in deductive logic. Stochastic independence of two statements is neither a necessary nor a sufficient condition for their mutual logical independence.

[5] When $\Pr(p) = 0$ the quotient is not defined. In this case there is no $\Pr(q \textit{ given } p)$.

a. $\Pr(\text{even } \textit{given } 2\text{v}4) = \dfrac{\Pr[\text{even}\&(2\text{v}4)]}{\Pr(2\text{v}4)} = \dfrac{\Pr(2\text{v}4)}{\Pr(2\text{v}4)} = 1$

b. $\Pr(\text{even } \textit{given } 1\text{v}3) = \dfrac{\Pr[\text{even}\&(1\text{v}3)]}{\Pr(1\text{v}3)} = \dfrac{0}{\frac{1}{3}} = 0$

c. $\Pr(\text{even } \textit{given } 1\text{v}2\text{v}3\text{v}4\text{v}6) = \dfrac{\Pr[\text{even}\&(1\text{v}2\text{v}3\text{v}4\text{v}6)]}{\Pr(1\text{v}2\text{v}3\text{v}4\text{v}6)}$

$$= \frac{\Pr(2\text{v}4\text{v}6)}{\Pr(1\text{v}2\text{v}3\text{v}4\text{v}6)} = \frac{\frac{3}{6}}{\frac{5}{6}} = \frac{3}{5}$$

Notice that the conditional probabilities computed by using the definition accord with the intuitive judgments as to conditional probabilities in the die example. We may test the definition in another way. Consider the special case of $\Pr(q \text{ given } p)$, where p is a tautology and q is a contingent statement. Since a tautology makes no factual claim, we would not expect knowledge of its truth to influence the probability that we would assign to the contingent statement, q. The probability that the die will come up even given that it will come up either even or odd should be simply the probability that it will come up even. In general, if we let T stand for an arbitrary tautology, we should expect $\Pr(q$ *given* $T)$ to be equal to $\Pr(q)$. Let us work out $\Pr(q$ *given* $T)$, using the definition of conditional probability:

$$\Pr(q \textit{ given } T) = \frac{\Pr(T\&q)}{\Pr(T)}$$

But the probability of a tautology is always equal to 1. This gives

$$\Pr(q \textit{ given } T) = \Pr(T\&q)$$

When T is a tautology and q is any statement whatsoever, the complex statement $T\&q$ is logically equivalent to the simple statement q. This can always be shown by truth tables. Since logically equivalent statements have the same probability, $\Pr(q$ *given* $T) = \Pr(q)$.[6] Again, the definition of conditional probability gives the expected result.

Now that conditional probability has been defined, that concept can be used to define independence:

[6] We could have constructed the probability calculus by taking conditional probabilities as basic, and then defining pure statement probabilities as follows: The probability of a statement is defined as its probability *given* a tautology. Instead we have taken statement probabilities as basic, and defined conditional probabilities. The choice of starting point makes no difference to the system as a whole. The systems are equivalent.

Definition 13: *Independence*: Two statements p and q are independent if and only if $\Pr(q$ *given* $p) = \Pr(q)$.

We talk of two statements p and q being independent, rather than p being independent of q and q being independent of p. We can do this because we can prove that $\Pr(q$ *given* $p) = \Pr(q)$ if and only if $\Pr(p$ *given* $q) = \Pr(p)$. If $\Pr(q$ *given* $p) = \Pr(q)$, then, by the definition of conditional probability,

$$\frac{\Pr(p\&q)}{\Pr(p)} = \Pr(q)$$

Multiplying both sides of the equation by $\Pr(p)$ and dividing both sides by $\Pr(q)$, we have

$$\frac{\Pr(p\&q)}{\Pr(q)} = \Pr(p)$$

But by the definition of conditional probability, this means $\Pr(p$ *given* $q) = \Pr(p)$.

This proof only works if neither of the two statements has 0 probability. Otherwise, one of the relevant quotients would not be defined. To take care of this eventuality, we may add an additional clause to the definition and say that two statements are also independent if at least one of them has probability 0. It is important to realize the difference between independence and mutual exclusiveness. The statement about the outcome of the throw of the die and the statement about the President's sneeze are independent, but they are not mutually exclusive. They can very well be true together. On the other hand, the statements "The next throw of the die will come up an even number" and "The next throw of the die will come up a 5" are mutually exclusive, but they are not independent. $\Pr(\text{even}) = \frac{1}{2}$, but $\Pr(\text{even given } 5) = 0$. $\Pr(5) = \frac{1}{6}$, but $\Pr(5 \text{ given even}) = 0$. In general, if p and q are mutually exclusive they are not independent, and if they are independent they are not mutually exclusive.[7]

Having specified the definitions of conditional probability and independence, the rules for conjunctions can now be introduced. The *general conjunction rule* follows directly from the definition of conditional probability:

Rule 7: $\Pr(p\&q) = \Pr(p) \times \Pr(q$ *given* $p)$.

The proof is simple. Take the definition of conditional probability:

$$\Pr(q \textit{ given } p) = \frac{\Pr(p\&q)}{\Pr(p)}$$

[7] The exception is when at least one of the statements is a self-contradiction and thus has probability 0.

Multiply both sides of the equation by Pr(p) to get

$$\Pr(p) \times \Pr(q \textit{ given } p) = \Pr(p\&q)$$

which is the general conjunction rule. When p and q are independent, Pr(q *given* p) = Pr(q), and we may substitute Pr(q) for Pr(q *given* p) in the general conjunction rule, thus obtaining

$$\Pr(p) \times \Pr(q) = \Pr(p\&q)$$

Of course, the substitution may only be made in the special case when p and q are independent. This result constitutes the *special conjunction rule:*

Rule 8: If p and q are independent, then $\Pr(p\&q) = \Pr(p) \times \Pr(q)$.

The general conjunction rule is more basic than the special conjunction rule. But since the special conjunction rule is simpler, its application will be illustrated first. Suppose that two dice are thrown simultaneously. The basic probabilities are as follows:

Die *A*	Die *B*
$\Pr(1) = \frac{1}{6}$	$\Pr(1) = \frac{1}{6}$
$\Pr(2) = \frac{1}{6}$	$\Pr(2) = \frac{1}{6}$
$\Pr(3) = \frac{1}{6}$	$\Pr(3) = \frac{1}{6}$
$\Pr(4) = \frac{1}{6}$	$\Pr(4) = \frac{1}{6}$
$\Pr(5) = \frac{1}{6}$	$\Pr(5) = \frac{1}{6}$
$\Pr(6) = \frac{1}{6}$	$\Pr(6) = \frac{1}{6}$

Since the face shown by die *A* presumably does not influence the face shown by die *B*, or vice versa, it shall be assumed that all statements claiming various outcomes for die *A* are independent of all the statements claiming various outcomes for die *B*. That is, the statements "Die *A* will come up a 3" and "Die *B* will come up a 5" are independent, as are the statements "Die *A* will come up a 6" and "Die *B* will come up a 6." The statements "Die *A* will come up a 5" and "Die *A* will come up a 3" are not independent; they are mutually exclusive (when made in regard to the same throw).

Now suppose we wish to calculate the probability of throwing a 1 on die *A* and a 6 on die *B*. The special conjunction rule can now be used:

$$\Pr(1 \text{ on } A \,\&\, 6 \text{ on } B) = \Pr(1 \text{ on } A) \times \Pr(6 \text{ on } B)$$
$$= \tfrac{1}{6} \times \tfrac{1}{6} = \tfrac{1}{36}$$

In the same way, the probability of each of the 36 possible combinations of results of die *A* and die *B* may be calculated as $\frac{1}{36}$, as shown in Table VI.1. Note that each of the cases in the table is mutually exclusive of each other case.

Thus, by the special alternation rule, the probability of case 1 v case 3 is equal to the probability of case 1 plus the probability of case 3.

Table VI.1

Possible results when throwing two dice

Case	Die *A*	Die *B*	Case	Die *A*	Die *B*
1	1	1	19	4	1
2	1	2	20	4	2
3	1	3	21	4	3
4	1	4	22	4	4
5	1	5	23	4	5
6	1	6	24	4	6
7	2	1	25	5	1
8	2	2	26	5	2
9	2	3	27	5	3
10	2	4	28	5	4
11	2	5	29	5	5
12	2	6	30	5	6
13	3	1	31	6	1
14	3	2	32	6	2
15	3	3	33	6	3
16	3	4	34	6	4
17	3	5	35	6	5
18	3	6	36	6	6

Suppose now that we wish to calculate the probability that the dice will come up showing a 1 and a 6. There are two ways this can happen: a 1 on die *A* and a 6 on die *B* (case 31). The probability of this combination appearing may be calculated as follows:

$$\Pr(1 \,\&\, 6) = \Pr[(1 \text{ on } A \,\&\, 6 \text{ on } B) \vee (1 \text{ on } B \,\&\, 6 \text{ on } A)]$$

Since the cases are mutually exclusive, the special disjunction rule may be used to get

$$\Pr[(1 \text{ on } A \,\&\, 6 \text{ on } B) \vee (1 \text{ on } B \,\&\, 6 \text{ on } A)]$$
$$= \Pr(1 \text{ on } A \,\&\, 6 \text{ on } B) + \Pr(1 \text{ on } B \,\&\, 6 \text{ on } A)$$

But it has already been shown, by the special conjunction rule, that

$$\Pr(1 \text{ on } A \,\&\, 6 \text{ on } B) = \tfrac{1}{36}$$
$$\Pr(1 \text{ on } B \,\&\, 6 \text{ on } A) = \tfrac{1}{36}$$

so the answer is $\frac{1}{36} + \frac{1}{36}$, or $\frac{1}{18}$.

The same sort of reasoning can be used to solve more complicated problems. Suppose we want to know the probability that the sum of spots showing on both dice will equal 7. This happens only in cases 6, 11, 16, 21, 26, and 31. Therefore

$$\begin{aligned}\Pr(\text{total of } 7) = \Pr[&(1 \text{ on } A \ \&\ 6 \text{ on } B)\\ &\text{v } (2 \text{ on } A \ \&\ 5 \text{ on } B)\\ &\text{v } (3 \text{ on } A \ \&\ 4 \text{ on } B)\\ &\text{v } (4 \text{ on } A \ \&\ 3 \text{ on } B)\\ &\text{v } (5 \text{ on } A \ \&\ 2 \text{ on } B)\\ &\text{v } (6 \text{ on } A \ \&\ 1 \text{ on } B)]\end{aligned}$$

Using the special disjunction rule and the special conjunction rule Pr(total of 7) = $\frac{6}{36}$, or $\frac{1}{6}$.

In solving a particular problem, there are often several ways to apply the rules. Suppose we wanted to calculate the probability that both dice will come up even. We could determine in which cases both dice are showing even numbers, and proceed as before, but this is the long way to solve the problem. Instead, we can calculate the probability of getting an even number on die A as $\frac{1}{2}$ by the special disjunction rule:

$$\begin{aligned}\Pr(\text{even on } A) &= \Pr(2 \text{ on } A \text{ v } 4 \text{ on } A \text{ v } 6 \text{ on } A)\\ &= \Pr(2 \text{ on } A) + \Pr(4 \text{ on } A) + \Pr(6 \text{ on } A)\\ &= \tfrac{3}{6} = \tfrac{1}{2}\end{aligned}$$

and calculate the probability of getting an even number on die B as $\frac{1}{2}$ by the same method. Then, by the special conjunction rule,[8]

$$\begin{aligned}\Pr(\text{even on } A \ \&\ \text{even on } B) &= \Pr(\text{even on } A) \times \Pr(\text{even on } B)\\ &= \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{4}\end{aligned}$$

We apply the *general conjunction rule* when two statements are not independent. Such is the case in the following example. Suppose you are presented with a bag containing ten gumdrops, five red and five black. You are to shake the bag, close your eyes and draw out a gumdrop, look at it, eat it, and then repeat the process once more. We shall assume that, at the time of each draw, each gumdrop in the bag has an equal probability of being drawn. The problem is to find the probability of drawing two red gumdrops.

[8]It can be shown that the statements "Die A will come up even" and "Die B will come up even" are independent, on the basis of the independence assumptions made in setting up this example.

To solve this problem we must find the probability of the conjunction Pr(red on 1 & red on 2). We will first find Pr(red on 1). We will designate each of the gumdrops by a letter: $A, B, C, D, E, F, G, H, I, J$. We know that we will draw one of these on the first draw, so

$$\Pr(A \text{ on } 1 \text{ v } B \text{ on } 1 \text{ v } C \text{ on } 1 \text{ v} \ldots \text{v } J \text{ on } 1) = 1$$

Now, by the special disjunction rule,

$$\Pr(A \text{ on } 1) + \Pr(B \text{ on } 1) + \Pr(C \text{ on } 1) + \ldots + (\Pr(J \text{ on } 1) = 1$$

Since each of the gumdrops has an equal chance of being drawn, and there are 10 gumdrops, therefore

$$\Pr(A \text{ on } 1) = \tfrac{1}{10}$$
$$\Pr(B \text{ on } 1) = \tfrac{1}{10}$$
$$\vdots$$
$$\Pr(J \text{ on } 1) = \tfrac{1}{10}$$

We said that there were five red ones. We will use the letters A, B, C, D, and E to designate the red gumdrops and the remaining letters to designate the black ones. By the special disjunction rule, the probability of getting a red gumdrop on draw 1 is

$$\begin{aligned}&\Pr(A \text{ on } 1 \text{ v } B \text{ on } 1 \text{ v } C \text{ on } 1 \text{ v } D \text{ on } 1 \text{ v } E \text{ on } 1)\\ &= \Pr(A \text{ on } 1) + \Pr(B \text{ on } 1) + \Pr(C \text{ on } 1) + \Pr(D \text{ on } 1) + \Pr(E \text{ on } 1)\\ &= \tfrac{5}{10} = \tfrac{1}{2}\end{aligned}$$

We shall have to use the general conjunction rule to find Pr(red on 1 & red on 2), since the statements "A red gumdrop will be drawn the first time" and "A red gumdrop will be drawn the second time" are not independent. If a red gumdrop is drawn the first time, this will leave four red and five black gumdrops in the bag with equal chances of being drawn on the second draw. But if a black gumdrop is drawn the first time, this will leave five red and four black gumdrops awaiting the second draw. Thus, the knowledge that a red one is drawn the first time will influence the probability we assign to a red one being drawn the second time, and the two statements are not independent. Applying the general conjunction rule, we get

$$\Pr(\text{red on } 1 \ \&\ \text{red on } 2) = \Pr(\text{red on } 1) \times \Pr(\text{red on } 2 \text{ given red on } 1)$$

We have already found Pr(red on 1). Now we must calculate Pr(red on 2 given red on 1). Given that we draw a red gumdrop on the first draw, there will be nine gumdrops remaining: four red and five black. We must draw one of them, and they each have an equal chance of being drawn. By reasoning similar to that used above, each has a probability of $\frac{1}{9}$ of being drawn, and the probability of drawing a red one is $\frac{4}{9}$. Therefore

$$\text{Pr(red on 2 given red on 1)} = \tfrac{4}{9}$$

We can now complete our calculations:

$$\text{Pr(red on 1 \& red on 2)} = \tfrac{1}{2} \times \tfrac{4}{9} = \tfrac{2}{9}$$

We can calculate Pr(black on 1 & red on 2) in the same way:

$$\text{Pr(black on 1)} = \tfrac{1}{2}$$

$$\text{Pr(red on 2 given black on 1)} = \tfrac{5}{9}$$

Therefore, by the general conjunction rule,

$$\text{Pr(black on 1 \& red on 2)} = \tfrac{1}{2} \times \tfrac{5}{9} = \tfrac{5}{18}$$

At this point the question arises as to what the Pr(red on 2) is. We know Pr(red on 2 given red on 1) $= \frac{4}{9}$. We know Pr(red on 2 given black on 1) $= \frac{5}{9}$. But what we want to know now is the probability of getting a red gumdrop on the second draw before we have made the first draw. We can get the answer if we realize that *red on 2* is logically equivalent to

(red on 1 & red on 2) v (not-red on 1 & red on 2)

Remember that the simple statement q is logically equivalent to the complex statement $(p\&q)\text{v}(\sim p\&q)$. Therefore

$$\text{Pr(red on 2)} = \text{Pr[(red on 1 \& red on 2) v (not-red on 1 \& red on 2)]}$$

By the special disjunction rule,

$$\text{Pr(red on 2)} = \text{Pr(red on 1 \& red on 2)} + \text{Pr(not-red on 1 \& red on 2)}$$

We have calculated Pr(red on 1 & red on 2) as $\frac{2}{9}$. We have also calculated

$$\text{Pr(not-red on 1 \& red on 2)} = \text{Pr(black on 1 \& red on 2)} = \tfrac{5}{18}$$

Therefore

$$\text{Pr(red on 2)} = \tfrac{2}{9} + \tfrac{5}{18} + \tfrac{4}{18} + \tfrac{5}{18} = \tfrac{9}{18} = \tfrac{1}{2}$$

The same sort of applications of conditional probability and the general conjunction rule would apply to card games where the cards that have been played are placed in a discard pile rather than being returned to the deck. Such considerations are treated very carefully in manuals on poker and blackjack. In fact, some gambling houses have resorted to using a new deck for each hand of blackjack in order to keep astute students of probability from gaining an advantage over the house.

Exercises

1. $\text{Pr}(p) = \frac{1}{2}$, $\text{Pr}(q) = \frac{1}{2}$, p and q are independent.
   a. What is $\text{Pr}(p\&q)$?
   b. Are p and q mutually exclusive?
   c. What is $\text{Pr}(p\text{v}q)$?
2. Suppose two dice are rolled, as in the example above.
   a. What is the probability of both dice showing a 1?
   b. What is the probability of both dice showing a 6?
   c. What is the probability that the total number of spots showing on both dice will be either 7 or 11?
3. A coin is flipped three times. Assume that on each toss Pr(heads) $= \frac{1}{2}$ and Pr(tails) $= \frac{1}{2}$. Assume that the tosses are independent.
   a. What is Pr(3 heads)?
   b. What is Pr(2 heads and 1 tail)?
   c. What is Pr(1 head and 2 tails)?
   d. What is Pr(head on toss 1 & tail on toss 2 & head on toss 3)?
   e. What is Pr(at least 1 tail)?
   f. What is Pr(no heads)?
   g. What is Pr(either 3 heads or 3 tails)?
4. Suppose you have an ordinary deck of 52 cards. A card is drawn and is not replaced, then another card is drawn. Assume that on each draw each of the cards then in the deck has an equal chance of being drawn.
   a. What is Pr(ace on draw 1)?
   b. What is Pr(10 on draw 2 given ace on draw 1)?
   c. What is Pr(ace on draw 1 & 10 on draw 2)?
   d. What is Pr(10 on draw 1 & ace on draw 2)?
   e. What is Pr(an ace and a 10)?
   f. What is Pr(2 aces)?
5. The probability that George will study for the test is $\frac{4}{5}$. The probability that he will pass the test given that he studies is $\frac{3}{5}$. The probability that he will pass the test given that he does not study is $\frac{1}{10}$. What is the probability that George will pass the test? Hint: The simple statement "George will pass the test" is logically equivalent to the complex statement "Either George will study and pass the test or George will not study and pass the test."

VI.5. EXPECTED VALUE OF A GAMBLE. The attractiveness of a wager depends not only on the probabilities involved, but also on the odds given. The probability of getting a head and a tail on two independent tosses of a fair coin is $\frac{1}{2}$, while the probability of getting two heads is only $\frac{1}{4}$. But if someone were to offer either to bet me even money that I will not get a head and a tail or give 100 to 1 odds against my getting two heads, I would be well advised to take the second wager. The probability that I will win the second wager is less, but this is more than compensated for by the fact that if I win, I will win a great deal, and if I lose, I will lose much less. The attractiveness of a wager can be measured by calculating its ***expected value***. To calculate the expected value of a gamble, first list all the possible outcomes, along with their probabilities and the amount won in each case. A loss is listed as a negative amount. Then for each outcome multiply the probability by the amount won or lost. Finally, add these products to obtain the expected value. To illustrate, suppose someone bets me 10 dollars that I will not get a head and a tail on two tosses of a fair coin. The expected value of this wager for me can be calculated as follows:

Possible outcomes				
Toss 1	Toss 2	Probability	Gain	Probability × Gain
H	H	$\frac{1}{4}$	−$10	−$2.50
H	T	$\frac{1}{4}$	10	2.50
T	H	$\frac{1}{4}$	10	2.50
T	T	$\frac{1}{4}$	−10	−2.50
			Expected value:	$0.00

Thus, the expected value of the wager for me is $0, and since my opponent wins what I lose and loses what I win, the expected value for him is also $0. Such a wager is called a *fair bet*. Now let us calculate the expected value for me of a wager where my opponent will give me 100 dollars if I get two heads, and I will give him one dollar if I do not.

Possible outcomes				
Toss 1	Toss 2	Probability	Gain	Probability × Gain
H	H	$\frac{1}{4}$	$100	$25.00
H	T	$\frac{1}{4}$	−1	−0.25
T	H	$\frac{1}{4}$	−1	−0.25
T	T	$\frac{1}{4}$	−1	−0.25
			Expected value:	$24.25

The expected value of this wager for me is $24.25. Since my opponent loses what I win, the expected value for him is −$24.25. This is not a fair bet, since it is favorable to me and unfavorable to him.

The procedure for calculating expected value and the rationale behind it are clear, but let us try to attach some meaning to the numerical answer. This can be done in the following way. Suppose that I make the foregoing wager many times. And suppose that over these many times the distribution of results corresponds to the probabilities; that is, I get two heads one-fourth of the time; a head and then a tail one-fourth of the time; a tail and then a head one-fourth of the time; and two tails one-fourth of the time. Then the expected value will be equal to my average winnings on a wager (that is, my total winnings divided by the number of wagers I have made).

I said that expected value was a measure of the attractiveness of a wager. Generally, it seems reasonable to accept a wager with a positive expected gain and reject a wager with a negative expected gain. Furthermore, if you are offered a choice of wagers, it seems reasonable to choose the wager with the highest expected value. These conclusions, however, are oversimplifications. They assume that there is no positive or negative value associated with risk itself, and that gains or losses of equal amounts of money represent gains or losses of equal amounts of money represent gains or losses of equal amount of value to the individual involved. Let us examine the first assumption.

Suppose that you are compelled to choose an even-money wager either for 1 dollar or for 100 dollars. The expected value of both wagers is 0. But if you wish to avoid risks as much as possible, you would choose the smaller wager. You would, then, assign a negative value to risk itself. However, if you enjoy taking larger risks for their own sake, you would choose the larger wager. Thus, although expected value is a major factor in determining the attractiveness of wagers, it is not the only factor. The positive or negative values assigned to the magnitude of the risk itself must also be taken into account.

We make a second assumption when we calculate expected value in terms of money. We assume that gains or losses of equal amounts of money represent gains or losses of equal amounts of value to the individual involved. In the language of the economist this is said to be the assumption that money has a constant marginal utility. This assumption is quite often false. For a poor man, the loss of 1000 dollars might mean he would starve, while the gain of 1000 dollars might mean he would merely live somewhat more comfortably. In this situation, the real loss accompanying a monetary loss of 1000 dollars is much greater than the real gain accompanying a monetary gain of 1000 dollars. A man in these circumstances would be foolish to accept an even money bet of 1000 dollars on the flip of a coin. In terms of money, the wager has an expected value of 0. But in terms of real value, the wager has a negative expected value.

Suppose you are in a part of the city far from home. You have lost your wallet and only have a quarter in change. Since the bus fare home is 35 cents, it looks as though you will have to walk. Now someone offers to flip you for a dime. If you win, you can ride home. If you lose, you are hardly any worse off than before. Thus, although the expected value of the wager in monetary terms is 0, in terms of real value, the wager has a positive expected value. In assessing the attractiveness of wagers by calculating their expected value, we must always be careful to see whether the monetary gains and losses accurately mirror the real gains and losses to the individual involved.

Exercises

1. What is the expected value of the following gamble? You are to roll a pair of dice. If the dice come up a natural, 7 or 11, you win 10 dollars. If the dice come up snake-eyes, 2, or boxcars, 12, you lose 20 dollars. Otherwise the bet is off.
2. What is the expected value of the following gamble? You are to flip a fair coin. If it comes up heads you win 1 dollar, and the wager is over. If it comes up tails you lose 1 dollar, but you flip again for 2 dollars. If the coin comes up heads this time you win 2 dollars. If it comes up tails you lose 2 dollars, but flip again for 4 dollars. If it comes up heads you win 4 dollars. If it comes up tails you lose 4 dollars. But in either case the wager is over.

   Hint: The possible outcomes are:

Toss 1	Toss 2	Toss 3
H	None	None
T	H	None
T	T	H
T	T	T

3. Suppose you extended the doubling strategy of Exercise 2 to four tosses. Would this change the expected value?
4. Suppose that you tripled your stakes instead of doubling them. Would this change the expected value?

VI.6. BAYES' THEOREM. You may wonder what the relation is between a conditional probability Pr(q *given* p) and its converse Pr(p *given* q). They need not be equal. The probability that Ezekial is an ape, given that he is a gorilla, is 1. But the probability that Ezekial is a gorilla, given that he is an ape, is less than 1. The value of a conditional probability is not determined by the value of its converse alone. But the value of a conditional probability can

Table VI.2

Step	Justification
1. $\Pr(q \textit{ given } p) = \dfrac{\Pr(p\&q)}{\Pr(p)}$	Definition of conditional probability
2. $\Pr(q \textit{ given } p) = \dfrac{\Pr(p\&q)}{\Pr(p\&q)\vee(p\&{\sim}q)}$	p is logically equivalent to $(p\&q)\vee(p\&{\sim}q)$
3. $\Pr(q \textit{ given } p) = \dfrac{\Pr(p\&q)}{\Pr(p\&q) + \Pr(p\&{\sim}q)}$	Special disjunction rule
4. $\Pr(q \textit{ given } p) = \dfrac{\Pr(q) \times \Pr(p \textit{ given } q)}{[\Pr(q) \times \Pr(p \textit{ given } q)] + [\Pr({\sim}q) \times \Pr(p \textit{ given } {\sim}q)]}$	General conjunction rule

be calculated from the value of its converse, together with certain other probability values. The basis of this calculation is set forth in *Bayes' theorem.* A simplified version of a proof of Bayes' theorem is presented in Table VI.2. Step 4 of this table states the simplified version of Bayes' theorem.[9] Note that it allows us to compute conditional probabilities going in one direction—that is, Pr(q *given* p)—from conditional probabilities going in the opposite direction—that is, Pr(p *given* q) and Pr(p *given* $\sim q$)—together with certain statement probabilities—that is, Pr(q) and Pr($\sim q$). Let us see how this theorem is applied in a concrete example.

Suppose we have two urns. Urn 1 contains eight red balls and two black balls. Urn 2 contains two red balls and eight black balls. Someone has selected an urn by flipping a fair coin. He then has drawn a ball from the urn he

[9] The general form of Bayes' theorem arises as follows: Suppose that instead of simply the two statements q and $\sim q$ we consider a set of n mutually exclusive statements, $q_1, q_2, \ldots q_n$, which is *exhaustive*. That is, the complex statement, $q_1 \vee q_2 \vee \ldots \vee q_n$, is a tautology. Then it can be proven that the simple statement p is logically equivalent to the complex statement $(p\&q_1)\vee(p\&q_2)\vee \ldots \vee(p\&q_n)$. This substitution is made in step 2, and the rest of the proof follows the model of the proof given. The result is

$$\Pr(q_1 \textit{ given } p) = \frac{\Pr(q_1) \times \Pr(p \textit{ given } q_1)}{[(\Pr(q_1) \times \Pr(p \textit{ given } q_1)] + [\Pr(q_2) \times \Pr(p \textit{ given } q_2)] + \cdots + [\Pr(q_n) \times \Pr(p \textit{ given } q_n)]}$$

selected. Assume that each ball in the urn he selected had an equal chance of being drawn. What is the probability that he selected urn 1, given that he drew a red ball? Bayes' theorem tells us the Pr(urn 1 given red) is equal to

$$\frac{\Pr(\text{urn } 1) \times \Pr(\text{red given urn } 1)}{[\Pr(\text{urn } 1) \times \Pr(\text{red given urn } 1)] + [\Pr(\sim\text{urn } 1) \times \Pr(\text{red given} \sim\text{urn } 1)]}$$

The probabilities needed may be calculated from the information given in the problem:

$$\Pr(\text{urn } 1) = \tfrac{1}{2}$$
$$\Pr(\sim\text{urn } 1) = \Pr(\text{urn } 2) = \tfrac{1}{2}$$
$$\Pr(\text{red given urn } 1) = \tfrac{8}{10}$$
$$\Pr(\text{red given} \sim\text{urn } 1) = \Pr(\text{red given urn } 2) = \tfrac{2}{10}$$

If these values are substituted into the formula, they give

$$\Pr(\text{urn 1 given red}) = \frac{\frac{1}{2} \times \frac{8}{10}}{(\frac{1}{2} \times \frac{8}{10}) + (\frac{1}{2} \times \frac{2}{10})} = \frac{\frac{4}{10}}{\frac{4}{10} + \frac{1}{10}} = \frac{4}{5}$$

A similar calculation will show that Pr(urn 2 given red) $= \frac{1}{5}$. Thus, the application of Bayes' theorem confirms our intuition that a red ball is more likely to have come from urn 1 than urn 2, and it tells us how much more likely.

It is important to emphasize the importance of the pure statement probabilities $\Pr(q)$ and $\Pr(\sim q)$ in Bayes' theorem. If we had not known that the urn to be drawn from had been selected by flipping a fair coin, if we had just been told that it was selected some way or other, we could not have computed Pr(urn 1 given red). Indeed if Pr(urn 1) and Pr(~urn 1) had been different, then our answer would have been different. Suppose that the urn had been selected by throwing a pair of dice. If the dice came up "snake-eyes" (a 1 on each die), urn 1 would be selected; otherwise urn 2 would be selected. If this were the case, then Pr(urn 1) $= \frac{1}{36}$ and Pr(~urn 1) = Pr(urn 2) $= \frac{35}{36}$. Keeping the rest of the example the same, Bayes' theorem gives

$$\Pr(\text{urn 1 given red}) = \frac{\frac{1}{36} \times \frac{8}{10}}{(\frac{1}{36} \times \frac{8}{10}) + (\frac{35}{36} \times \frac{2}{10})} = \frac{\frac{8}{360}}{\frac{8}{360} + \frac{70}{360}} = \frac{8}{78} = \frac{4}{39}$$

This is quite a different answer from the one we got when urns 1 and 2 had an equal chance of being selected. In each case Pr(urn 1 given red) is higher than Pr(urn 1). This can be interpreted as saying that in both cases the additional information that a red ball was drawn would raise confidence that urn 1 was selected. But the initial level of confidence that urn 1 was selected is different in the two cases, and consequently the final level is also.

Exercises

1. The probability that George will study for the test is $\frac{4}{10}$. The probability that he will pass, given that he studies, is $\frac{9}{10}$. The probability that he passes, given that he does not study, is $\frac{3}{10}$. What is the probability that he has studied, given that he passes?
2. Suppose there are three urns. Urn 1 contains six red balls and four black balls. Urn two contains nine red balls and one black ball. Urn 3 contains five red balls and five black balls. A ball is drawn at random from urn 1. If it is black a second ball is drawn at random from urn 2, but if it is red the second ball is drawn at random from urn 3.
   a. What is the probability of the second ball being drawn from urn 2?
   b. What is the probability of the second ball being drawn from urn 3?
   c. What is the probability that the second ball drawn is black, given that it is drawn from urn 2?
   d. What is the probability that the second ball drawn is black, given that it is drawn from urn 3?
   e. What is the probability that the second ball is black?
   f. What is the probability that the second ball was drawn from urn 2, given that it is black?
   g. What is the probability that the second ball was drawn from urn 3, given that it is black?
   h. What is the probability that the second ball drawn was drawn from urn 2, given that it is red?
   i. What is the probability that the second ball drawn was drawn from urn 3, given that it is red?
   j. What is the probability that the first ball drawn was red, given that the second ball drawn is black?
   k. What is the probability that the first ball is black, given that the second ball is black?
   l. What is the probability that both balls drawn are black?
3. A fair coin is flipped twice. The two tosses are independent. What is the probability of a heads on the first toss given a heads on the second toss?
4. Their captors have decided that two of three prisoners—Smith, Jones, and Fitch—will be executed tomorrow. The choice has been made at random, but the identity of the unfortunate selectees is to be kept from the prisoners until the final hour. The prisoners, who are held in separate cells, unable to communicate with each other, know this. Fitch asks a guard to tell the name of one of the other prisoners who will be executed. Regardless of whether Fitch was chosen or not, one of the others will be executed, so the guard reasons that he is not giving Fitch any illicit information by answering truthfully. He says: "Jones will be executed." Fitch is heartened by the news for he reasons that his probability of being the one who escapes execution has risen from $\frac{1}{3}$ to $\frac{1}{2}$. Has Fitch made a mistake? Has the guard? Use Bayes' theorem to analyze the reasoning involved. (Hint: Calculate

the probability that Fitch will not be executed given that *the guard tells him that* Jones will be executed, not the probability that Fitch will not be executed given that Jones will be. What assumptions are possible about the probability that the guard tells Fitch that Jones will be executed given that Fitch escapes execution?)

VI.7. PROBABILITY AND CAUSALITY. What is meant when it is said that smoking causes lung cancer? Not that smoking is a *sufficient* condition for contraction of lung cancer, for many people smoke and never contract the disease. Not that smoking is a *necessary* condition for lung cancer, for some who never smoke nevertheless develop lung cancer. What is meant is something probabilistic: that smoking increases one's chances of getting lung cancer.

We might say that smoking has a tendency in the direction of sufficientness if Pr(cancer given smoking) is greater than Pr(cancer given ~smoking)—that is, if smoking is *positively statistically relevant* to cancer. We might say that smoking has a tendency in the direction of necessaryness for lung cancer if Pr(having smoked given cancer) is greater than Pr(having smoked given no cancer)—that is, if cancer is positively statistically relevant to smoking. But we can show from the probability calculus that for any two statements, P,Q,[10] P is positively statistically relevant to Q if and only if Q is positively statistically relevant to P. By Bayes' theorem:

$$\Pr(Q \text{ given } P) = \frac{\Pr(P \text{ given } Q)\ \Pr(Q)}{\Pr(P)}$$

$$\text{So: } \frac{\Pr(Q \text{ given } P)}{\Pr(Q)} = \frac{\Pr(P \text{ given } Q)}{\Pr(P)}$$

P is positively relevant to Q just in case the left-hand side of the equation is greater than one; Q is positively relevant to P just in case the right-hand side of the equation is greater than one. So the probabilistic notions of being a tendency toward a sufficient condition, and having a tendency toward being a necessary condition come to the same thing! Considerations appear to be simpler in this way in a probabilistic setting than in a deterministic one.

But there is a complication that we must now discuss. Suppose that smoking itself did not cause the cancer, but that desire to smoke and cancer were both effects of some underlying genetically determined biological condition. Then smoking would still be positively statistically relevant to cancer, but as a symptom of having the bad gene rather than as a cause of cancer. If this hypothesis were correct, we would not say that smoking raised one's *chances of* getting lung cancer. If someone, say *you*, had the bad genes, then your chances of contracting cancer would be already high and smoking would not make them worse; if you didn't have the bad genes, your chances of contracting cancer would be lower and smoking wouldn't make them worse. That is, the positive statistical relevance of smoking to cancer would disappear if we looked at probabilities *conditional* on having the bad genes; likewise if we looked at probabilities conditional on not having the bad genes:

Pr(cancer given smoking and bad genes) = Pr(cancer given bad genes)

Pr(cancer given smoking and good genes) = Pr(cancer given good genes)

To support the claim that smoking is a probabilistic cause of lung cancer, the foregoing hypothesis (and others like it) must be ruled out. Perhaps identical twins can be found such that one of each pair is a long-time smoker, and more of the smokers develop cancer. Perhaps subjects who don't want to smoke but are forced to inhale smoke anyway (certain laboratory mice, cocktail waitresses, and so on) have a higher incidence of lung cancer.

If we believe that a certain constellation of factors determines the chance of getting lung cancer, then we consider smoking a probabilistic cause of lung cancer if, when we hold all the other preexisting factors *fixed*, smoking increases the chance of lung cancer. That is, if:

Pr(cancer given background factors and smoking) is greater than
Pr(cancer given background factors and no smoking)

Whether X is a probabilistic cause of Y for individual a may depend on just what constellation of background factors is present for a. Some lucky people have a biochemistry such that for them, contact with poison oak is not a probabilistic cause of skin eruptions and intense itching, but for most of us it unfortunately is.

[10] With positive probability.

Exercises

1. Discuss the following argument: Most heroin users have previously smoked marijuana. Therefore, marijuana use causes heroin use.
2. How would you go about finding out whether *for you* exposure to ragweed pollen is a cause of a stuffed-up nose, runny eyes, and so on?
3. Some studies have found that, on average, convicted criminals exhibit vitamin deficiencies. This suggests to some researchers that vitamin deficiencies might