3

Conditional Credences

Chapter 2's discussion was confined to unconditional credence, an agent's outright degree of confidence that a particular proposition is true. This chapter takes up conditional credence, an agent's credence that one proposition is true on the supposition that another one is.

The main focus of this chapter is our fourth core normative Bayesian rule: the Ratio Formula. This rational constraint on conditional credences has a number of important consequences, including Bayes's Theorem (which gives Bayesianism its name).

Conditional credences are also central to the way Bayesians understand evidential relevance. I will define relevance as positive correlation, then explain how this notion has been used to investigate causal relations through the concept of screening off.

Having achieved a deeper understanding of the mathematics of conditional credences, I return at the end of the chapter to what exactly a conditional credence is. In particular, I discuss an argument by David Lewis that a conditional credence can't be understood as an unconditional credence in a conditional.

3.1 Conditional credences and the Ratio Formula

Arturo and Baxter know that two events will occur simultaneously in separate rooms: a fair coin will be flipped, and a clairvoyant will predict how it will land. Let H represent the proposition that the coin comes up heads, and C represent the proposition that the clairvoyant predicts heads. Suppose Arturo and Baxter each assign an unconditional credence of $1/2$ to H and an unconditional credence of $1/2$ to C.

Although Arturo and Baxter assign the same unconditional credences as each other to H and C, they still might take these propositions to be *related* in different ways. We could tease out those differences by saying to each agent, "I have no idea how the coin is going to come up or what the clairvoyant is going to say. But suppose for a moment the clairvoyant predicts heads. On this supposition, how confident are you that the coin will come up heads?" If

Fundamentals of Bayesian Epistemology 1: Introducing Credences. Michael G. Titelbaum, Oxford University Press.
 DOI: 10.1093/oso/9780198707608.003.0003

Arturo says 1/2 and Baxter says 99/100, that's a good indication that Baxter has more faith in the mystical than Arturo.

The quoted question in the previous paragraph elicits Arturo and Baxter's *conditional* credences, as opposed to the *unconditional* credences discussed in Chapter 2. An unconditional credence is a degree of belief assigned to a single proposition, indicating how confident the agent is that that proposition is true. A **conditional credence** is a degree of belief assigned to an ordered pair of propositions, indicating how confident the agent is that the first proposition is true on the supposition that the second is. We symbolize conditional credences as follows:

$$\mathrm{cr}(H \mid C) = 1/2 \tag{3.1}$$

This equation says that a particular agent (in this case, Arturo) has a 1/2 credence that the coin comes up heads conditional on the supposition that the clairvoyant predicts heads. The vertical bar indicates a conditional credence; to the right of the bar is the proposition supposed; to the left of the bar is the proposition evaluated in light of that supposition. The proposition to the right of the bar is sometimes called the **condition**; I am not aware of any generally accepted name for the proposition on the left.

To be clear: A real agent never assigns any credences *ex nihilo*, without assuming at least some background information. An agent's unconditional credences in various propositions (such as H) are informed by her background information at that time. To assign a conditional credence, the agent combines her stock of background information with a *further* supposition that the condition is true. She then evaluates the other proposition in light of this combination.

A conditional credence is assigned to an *ordered* pair of propositions. It makes a difference which proposition is supposed and which is evaluated. Consider a case in which I'm going to roll a fair die and you have various credences involving the proposition E that it comes up even and the proposition 6 that it comes up six. Compare:

$$\mathrm{cr}(6 \mid E) = 1/3 \tag{3.2}$$

$$\mathrm{cr}(E \mid 6) = 1 \tag{3.3}$$

3.1.1 The Ratio Formula

Section 2.2 described Kolmogorov's probability axioms, which Bayesians take to represent rational constraints on an agent's unconditional credences.

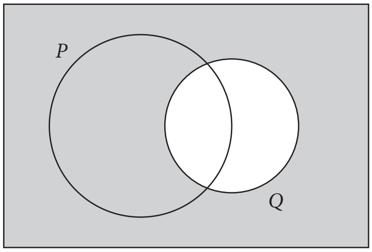

Figure 3.1 The region that dictates cr($P \mid Q$)

Bayesians then add a constraint relating conditional to unconditional credences:

Ratio Formula: For any P and Q in $\mathcal{L}$, if cr(Q) > 0 then

$$\mathrm{cr}(P \mid Q) = \frac{\mathrm{cr}(P \,\&\, Q)}{\mathrm{cr}(Q)}$$

Stated this way, the Ratio Formula remains silent on the value of cr($P \mid Q$) when cr(Q) $= 0$. There are various positions on how one should assign conditional credences when the condition has credence 0; we'll address some of them in our discussion of the infinite in Section 5.4.

Why should an agent's conditional credence equal the ratio of those unconditionals? Consider Figure 3.1. The rectangle represents all the possible worlds the agent entertains. The agent's unconditional credence in P is the fraction of that rectangle taken up by the P-circle. (The area of the rectangle is stipulated to be 1, so that fraction is the area of the P-circle divided by 1, which is just the area of the P-circle.) When we ask the agent to assign a credence conditional on the supposition that Q, she temporarily narrows her focus to just those possibilities that make Q true. In other words, she excludes from her attention the worlds I've shaded in the diagram, and considers only what's in the Q-circle. The agent's credence in P conditional on Q is the fraction of the Q-circle occupied by P-worlds. So it's the area of the PQ overlap divided by the area of the entire Q-circle, which is cr(P & Q)/cr(Q).

In the scenario in which I roll a fair die, your initial doxastic possibilities include all six outcomes of the die roll. I then ask for your credence that the die comes up six conditional on its coming up even—that is, cr($6 \mid E$). To assign this value, you exclude from consideration all the odd outcomes. You haven't actually *learned* that the die outcome is even; I've simply asked you to *suppose* that it comes up even and assign a confidence to other propositions in light of that supposition. You distribute your credence equally over the outcomes that remain under consideration (2, 4, and 6), so your credence in six conditional on even is 1/3.

We get the same result from the Ratio Formula:

$$\mathrm{cr}(6 \mid E) = \frac{\mathrm{cr}(6 \,\&\, E)}{\mathrm{cr}(E)} = \frac{1/6}{1/2} = \frac{1}{3} \tag{3.4}$$

The Ratio Formula allows us to calculate your conditional credences (confidences under a supposition) from your unconditional credences (confidences relative to no suppositions beyond your background information). Hopefully it's obvious why E gets an unconditional credence of 1/2 in this case; as for 6 & E, that's equivalent to just 6, so it gets an unconditional credence of 1/6.[1]

Warning

Mathematicians often treat the Ratio Formula as a *definition* of conditional probability. From their point of view, a conditional probability has the value it does *in virtue of* two unconditional probabilities' standing in a certain ratio. But I do not want to reduce the possession of a conditional credence to the possession of two unconditional credences standing in a particular relation. I take a conditional credence to be a genuine mental state (an attitude toward an ordered pair of propositions) capable of being elicited in various ways, such as by asking an agent her confidence in a proposition given a supposition. So I will interpret the Ratio Formula as a rational constraint on how an agent's conditional credences should relate to her unconditional credences. As a normative *constraint* (rather than a *definition*), it can be violated—by assigning a conditional credence that doesn't equal the specified ratio.

The point of the previous warning is that the Ratio Formula is a rational constraint, and agents don't always meet all the rational constraints on their credences. Yet for agents who do satisfy the Ratio Formula, there can be no difference between their conditional credences without some difference in their unconditional credences as well. If we're both rational and I assign a different $\mathrm{cr}(P \mid Q)$ value than you, we cannot assign the same values to both $\mathrm{cr}(P \,\&\, Q)$ and $\mathrm{cr}(Q)$. (A rational agent's conditional credences **supervene** on her unconditional credences.) Fully specifying a rational agent's unconditional credence distribution suffices to specify her conditional credences as well.[2] For instance, we might specify Arturo's and Baxter's credence distributions using the following probability table:

C	H	cr_A	cr_B
T	T	1/4	99/200
T	F	1/4	1/200
F	T	1/4	1/200
F	F	1/4	99/200

Here cr_A represents Arturo's credences and cr_B represents Baxter's. Arturo's unconditional credence in C is identical to Baxter's—the values on the first two rows sum to 1/2 for each of them. Similarly, Arturo and Baxter have the same unconditional credence in H (the sum of the first and third rows). Yet Arturo and Baxter disagree in their confidence that the coin will come up heads (H) given that the clairvoyant predicts heads (C). Using the Ratio Formula, we calculate this conditional credence by dividing the value on the first row of the table by the sum of the values on the first two rows. This yields:

$$\mathrm{cr}_A(H \mid C) = \frac{1/4}{1/2} = \frac{1}{2} \neq \frac{99}{100} = \frac{99/200}{100/200} = \mathrm{cr}_B(H \mid C) \tag{3.5}$$

Baxter has high confidence in the clairvoyant's abilities. So on the supposition that the clairvoyant predicts heads, Baxter is almost certain that the flip comes up heads. Arturo, on the other hand, is skeptical, so supposing that the clairvoyant predicts heads leaves his opinions about the flip outcome unchanged.

3.1.2 Consequences of the Ratio Formula

Combining the Ratio Formula with the probability axioms yields further useful probability rules. First we have the

Law of Total Probability: For any proposition P and finite partition $\{Q_1, Q_2, \ldots, Q_n\}$ in $\mathcal{L}$,

$$\mathrm{cr}(P) = \mathrm{cr}(P \mid Q_1) \cdot \mathrm{cr}(Q_1) + \mathrm{cr}(P \mid Q_2) \cdot \mathrm{cr}(Q_2) + \ldots + \mathrm{cr}(P \mid Q_n) \cdot \mathrm{cr}(Q_n)$$

Suppose you're trying to predict whether I will bike to work tomorrow, but you're unsure if the weather will rain, hail, or be clear. The Law of Total Probability allows you to systematically work through the possibilities in that partition. You multiply your confidence that it will rain by your confidence that I'll bike should it rain. Then you multiply your confidence that it'll hail by your confidence in my biking given hail. Finally you multiply your unconditional credence that it'll be clear by your conditional credence that I'll bike given that it's clear. Adding these three products together yields your unconditional credence that I'll bike. (In the formula, the proposition that I'll bike plays the role of P and the three weather possibilities are Q_1, Q_2, and Q_3.)

Next, the Ratio Formula connects conditional credences to Kolmogorov's axioms in a special way. Conditional credence is a two-place function, taking in an ordered pair of propositions and yielding a real number. Now suppose we designate some particular proposition R as our condition, and look at all of an agent's credences conditional on R. We now have a one-place function (because the second place has been filled by R) that we can think of as a distribution over the propositions in $\mathcal{L}$. Remarkably, if the agent's unconditional credences satisfy the probability axioms, then the Ratio Formula requires this conditional distribution $\text{cr}(\cdot \mid R)$ to satisfy those axioms as well. More formally, for any proposition R in $\mathcal{L}$ such that $\text{cr}(R) > 0$, the following will all be true:

- For any proposition P in $\mathcal{L}$, $\text{cr}(P \mid R) \geq 0$.
- For any tautology $\top$ in $\mathcal{L}$, $\text{cr}(\top \mid R) = 1$.
- For any mutually exclusive propositions P and Q in $\mathcal{L}$, $\text{cr}(P \vee Q \mid R) = \text{cr}(P \mid R) + \text{cr}(Q \mid R)$.

(You'll prove these three facts in Exercise 3.4.)

Knowing that a conditional credence distribution is a probability distribution can be a handy shortcut. (It also has a significance for updating credences that we'll discuss in Chapter 4.) Because it's a probability distribution, a conditional credence distribution must satisfy all the consequences of the probability axioms we saw in Section 2.2.1. For example, if I tell you that $\text{cr}(P \mid R) = 0.7$, you can immediately tell that $\text{cr}(\sim P \mid R) = 0.3$, by the following conditional implementation of the Negation rule:

$$\text{cr}(\sim P \mid R) = 1 - \text{cr}(P \mid R) \tag{3.6}$$

Similarly, Entailment tells us that if $P \vDash Q$, then $\text{cr}(P \mid R) \leq \text{cr}(Q \mid R)$.

One special conditional distribution is worth investigating at this point: What happens when the condition R is a tautology? Imagine I ask you to report your unconditional credences in a bunch of propositions. Then I ask you to assign credences to those same propositions conditional on the further supposition of... nothing. I give you nothing more to suppose. Clearly you'll just report back to me the same credences. Bayesians represent vacuous information as a tautology, so this means that a rational agent's credences conditional on a tautology equal her unconditional credences. In other words, for any P in $\mathcal{L}$,

$$\text{cr}(P \mid \top) = \text{cr}(P) \tag{3.7}$$

This fact (whose proof I'll leave to the reader) will be important to our theory of updating later on.[3]

3.1.3 Bayes's Theorem

The most famous consequence of the Ratio Formula and Kolmogorov's axioms is

Bayes's Theorem: For any H and E in $\mathcal{L}$,

$$\text{cr}(H \mid E) = \frac{\text{cr}(E \mid H) \cdot \text{cr}(H)}{\text{cr}(E)}$$

The first thing to say about Bayes's Theorem is *that it is a theorem*—it can be proven straightforwardly from the axioms and Ratio Formula. This is worth remembering, because there is a great deal of controversy about how Bayesians *apply* the theorem. (The significance they attach to this theorem is why Bayesians came to be called "Bayesians".)

What philosophical significance could attach to an equation that is, in the end, just a truth of mathematics? The theorem was first articulated by the Reverend Thomas Bayes in the 1700s.[4] Prior to Bayes, much of probability theory was concerned with problems of **direct inference**. Direct inference starts with the supposition of some probabilistic hypothesis, then asks how likely that hypothesis makes a particular experimental result. You probably learned to solve many direct inference problems in school, such as "Suppose I flip a fair coin 20 times; how likely am I to get exactly 19 heads?" Here the probabilistic hypothesis H says that the coin is fair, while the experimental result E is that 20 flips yield exactly 19 heads. Your credence that the experimental result will occur on the supposition that the hypothesis is true—$\mathrm{cr}(E \mid H)$—is called the **likelihood**.[5]

Yet Bayes was also interested in **inverse inference**. Instead of making suppositions about hypotheses and determining probabilities of courses of evidence, his theorem allows us to calculate probabilities of hypotheses from suppositions about evidence. Instead of calculating the likelihood $\mathrm{cr}(E \mid H)$, Bayes's Theorem shows us how to calculate $\mathrm{cr}(H \mid E)$. A problem of inverse inference might ask, "Suppose a coin comes up heads on exactly 19 of 20 flips; how probable is it that the coin is fair?"

Assessing the significance of Bayes's Theorem, Hans Reichenbach wrote:

> The *method of indirect evidence*, as this form of inquiry is called, consists of inferences that on closer analysis can be shown to follow the structure of the rule of Bayes. The physician's inferences, leading from the observed symptoms to the diagnosis of a specified disease, are of this type; so are the inferences of the historian determining the historical events that must be assumed for the explanation of recorded observations; and, likewise, the inferences of the detective concluding criminal actions from inconspicuous observable data. . . . Similarly, the general inductive inference from observational data to the validity of a given scientific theory must be regarded as an inference in terms of Bayes' rule. (Reichenbach 1935/1949, pp. 94–5)[6]

Here's an example of inverse inference: You're a biologist studying a particular species of fish, and you want to know whether the genetic allele coding for blue fins is dominant or recessive. Based on some other work you've done on fish, you're leaning toward recessive—initially you assign a 0.4 credence that the blue-fin allele is dominant. Given some background assumptions we won't worry about here,[7] a direct inference from the theory of genetics tells you that if the allele is dominant, roughly three out of four species members will have blue fins; if the allele is recessive blue fins will appear on roughly 25% of the fish. But you're going to perform an inverse inference, making experimental observations to decide between genetic hypotheses. You will capture fish from the species at random and examine their fins. How significant will your first observation be to your credences in dominant versus recessive? When you contemplate various ways that observation might turn out, how should supposing one outcome or the other affect your credences about the allele? Before we do the calculation, try estimating how confident you should be that the allele is dominant on the supposition that the first fish you observe has blue fins.

In this example our hypothesis H will be that the blue-fin allele is dominant. The evidence E to be supposed is that a randomly drawn fish has blue fins. We want to calculate the **posterior** value $\mathrm{cr}(H \mid E)$. This value is called the "posterior" because it's your credence in the hypothesis H *after* the evidence E has been supposed. In order to calculate this posterior, Bayes's Theorem requires the values of $\mathrm{cr}(E \mid H)$, $\mathrm{cr}(H)$, and $\mathrm{cr}(E)$.

$\mathrm{cr}(E \mid H)$ is the likelihood of drawing a blue-finned fish on the hypothesis that the allele is dominant. On the supposition that the allele is dominant, 75% of the fish have blue fins, so your $\mathrm{cr}(E \mid H)$ value should be 0.75. The other two values are known as **priors**; they are your unconditional credences in the hypothesis and the evidence *before* anything is supposed. We already said that your prior in the blue-fin dominant hypothesis H is 0.4. So $\mathrm{cr}(H)$ is 0.4. But what about the prior in the evidence? How confident are you before observing any fish that the first one you draw will have blue fins?

Here we can apply the Law of Total Probability to the partition containing H and $\sim H$. This yields:

$$\mathrm{cr}(E) = \mathrm{cr}(E \mid H) \cdot \mathrm{cr}(H) + \mathrm{cr}(E \mid {\sim}H) \cdot \mathrm{cr}({\sim}H) \tag{3.8}$$

The values on the right-hand side are all either likelihoods, or priors related to the hypothesis. These values we can easily calculate. So

$$\mathrm{cr}(E) = 0.75 \cdot 0.4 + 0.25 \cdot 0.6 = 0.45 \tag{3.9}$$

Plugging all these values into Bayes's Theorem gives us

$$\mathrm{cr}(H \mid E) = \frac{\mathrm{cr}(E \mid H) \cdot \mathrm{cr}(H)}{\mathrm{cr}(E)} = \frac{0.75 \cdot 0.4}{0.45} = 2/3 \tag{3.10}$$

Observing a single fish has the potential to change your credences substantially. On the supposition that the fish you draw has blue fins, your credence that the blue-fin allele is dominant goes from its prior value of 2/5 to a posterior of 2/3.

Again, all of this is pure mathematics from a set of axioms that are rarely disputed. So why has Bayes's Theorem been the focus of controversy? One issue is the role Bayesians give the theorem in *updating* attitudes over time; we'll return to that application in Chapter 4. But the main idea Bayesians take from Bayes—the idea that has proven controversial—is that probabilistic inverse inference is the key to induction. Bayesians think the primary way we ought to draw conclusions from data—how we ought to reason about scientific hypotheses, say, on the basis of experimental evidence—is by calculating posterior credences using Bayes's Theorem. This view stands in direct conflict with other statistical methods, such as frequentism and likelihoodism. Advocates of those methods also have deep concerns about where Bayesians get the priors that Bayes's Theorem requires. Once we've considerably deepened our understanding of Bayesian epistemology, we will discuss those concerns in Chapter 13, and assess frequentism and likelihoodism as alternatives to Bayesianism.

Before moving on, I'd like to highlight two useful alternative forms of Bayes's Theorem. We've just seen that calculating the prior of the evidence—$\mathrm{cr}(E)$—can be easier if we break it up using the Law of Total Probability. Incorporating that manuever into Bayes's Theorem yields

$$\mathrm{cr}(H \,|\, E) = \frac{\mathrm{cr}(E \,|\, H) \cdot \mathrm{cr}(H)}{\mathrm{cr}(E \,|\, H) \cdot \mathrm{cr}(H) + \mathrm{cr}(E \,|\, {\sim}H) \cdot \mathrm{cr}({\sim}H)} \tag{3.11}$$

When a particular hypothesis H is under consideration, its negation ${\sim}H$ is known as the **catchall** hypothesis. So this form of Bayes's Theorem calculates the posterior in the hypothesis from the priors and likelihoods of the hypothesis and its catchall.

In other situations we have multiple hypotheses under consideration instead of just one. Given a finite partition of n hypotheses $\{H_1, H_2, \ldots, H_n\}$, the Law of Total Probability transforms the denominator of Bayes's Theorem to yield

$$\mathrm{cr}(H_i \,|\, E) = \frac{\mathrm{cr}(E \,|\, H_i) \cdot \mathrm{cr}(H_i)}{\mathrm{cr}(E \,|\, H_1) \cdot \mathrm{cr}(H_1) + \mathrm{cr}(E \,|\, H_2) \cdot \mathrm{cr}(H_2) + \ldots + \mathrm{cr}(E \,|\, H_n) \cdot \mathrm{cr}(H_n)} \tag{3.12}$$

This version allows you to calculate the posterior of one particular hypothesis H_i in the partition from the priors and likelihoods of all the hypotheses.

3.2 Relevance and independence

Arturo doesn't believe in hocus pocus; from his point of view, information about what a clairvoyant predicts is irrelevant to determining how a coin flip will come out. So supposing that a clairvoyant predicts heads makes no difference to Arturo's confidence in a heads outcome. If C says the clairvoyant predicts heads, H says the coin lands heads, and cr_A is Arturo's credence distribution, we have

$$\mathrm{cr}_A(H \mid C) = 1/2 = \mathrm{cr}_A(H) \tag{3.13}$$

Generalizing this idea yields a key definition: Proposition P is **probabilistically independent** of proposition Q relative to distribution cr just in case

$$\mathrm{cr}(P \mid Q) = \mathrm{cr}(P) \tag{3.14}$$

In this case Bayesians also say that Q is **irrelevant** to P. When Q is irrelevant to P, supposing Q leaves an agent's credence in P unchanged.[8]

Notice that probabilistic independence is always relative to a distribution cr. The very same propositions P and Q might be independent relative to one distribution but dependent relative to another. (Relative to Arturo's credences the clairvoyant's prediction is irrelevant to the flip outcome, but relative to the credences of his friend Baxter—who believes in psychic powers—it is not.) In what follows I may omit reference to a particular distribution when context makes it clear, but you should keep the relativity of independence to a probability distribution in the back of your mind.

While Equation (3.14) will be our official *definition* of probabilistic independence, there are many equivalent tests for independence. Given the probability axioms and Ratio Formula, the following equations are all true just when Equation (3.14) is:

$$\mathrm{cr}(P) = \mathrm{cr}(P \mid {\sim}Q) \tag{3.15}$$

$$\mathrm{cr}(P \mid Q) = \mathrm{cr}(P \mid {\sim}Q) \tag{3.16}$$

$$\mathrm{cr}(Q \mid P) = \mathrm{cr}(Q) = \mathrm{cr}(Q \mid {\sim}P) \tag{3.17}$$

$$\mathrm{cr}(P \,\&\, Q) = \mathrm{cr}(P) \cdot \mathrm{cr}(Q) \tag{3.18}$$

The equivalence of Equations (3.14) and (3.15) tells us that Q is probabilistically independent of P just in case ${\sim}Q$ is. The equivalence of (3.14) and (3.17) shows us that independence is symmetric: if supposing Q makes no difference to an agent's credence in P, then supposing P won't change that agent's attitude toward Q. Finally, Equation (3.18) embodies a useful probability rule:

Multiplication: P and Q are probabilistically independent relative to cr if and only if $\mathrm{cr}(P \,\&\, Q) = \mathrm{cr}(P) \cdot \mathrm{cr}(Q)$.

(Some authors define probabilistic independence using this biconditional, but we will define independence using Equation (3.14), then treat Multiplication as a consequence.)

We can calculate the probability of a conjunction by multiplying the probabilities of its conjuncts only when those conjuncts are *independent*. This trick will not work for any arbitrary propositions. The general formula for probability in a conjunction can be derived quickly from the Ratio Formula:

$$\mathrm{cr}(P \,\&\, Q) = \mathrm{cr}(P \mid Q) \cdot \mathrm{cr}(Q) \tag{3.19}$$

When P and Q are probabilistically independent, the first term on the right-hand side equals cr(P).

It's important not to get Multiplication and Finite Additivity confused. Finite Additivity says that the credence of a *disjunction* is the *sum* of the credences of its *mutually exclusive* disjuncts. Multiplication says that the credence of a *conjunction* is the *product* of the credences of its *independent* conjuncts. If I flip a fair coin twice in succession, heads on the first flip and heads on the second flip are independent, while heads on the first flip and tails on the first flip are mutually exclusive.

When two propositions fail to be probabilistically independent (relative to a particular distribution), we say those propositions are **relevant** to each other. Replace the "=" signs in Equations (3.14) through (3.18) with ">" signs and you have tests for Q's being **positively relevant** to P. Once more the tests are equivalent—if any of the resulting inequalities is true, all of them are. Q is positively relevant to P when assuming Q makes you more confident in P. For example, since Baxter believes in mysticism, he takes the clairvoyant's predictions to be highly relevant to the outcome of the coin flip—supposing that the clairvoyant has predicted heads takes him from equanimity to near-certainty in a heads outcome. Baxter assigns

$$\mathrm{cr}_B(H \mid C) = 99/100 > 1/2 = \mathrm{cr}_B(H) \tag{3.20}$$

Like independence, positive relevance is symmetric. Supposing that the coin came up heads will make Baxter highly confident that the clairvoyant predicted it would.

Similarly, replacing the "=" signs with "<" signs above yields tests for **negative relevance**. For Baxter, the clairvoyant's predicting heads is negatively relevant to the coin's coming up tails. Like positive correlation, negative correlation is symmetric (supposing a tails outcome makes Baxter less confident in a heads prediction). Note also that there are many synonyms in the statistics community for "relevance". Instead of finding "positively/negatively relevant" terminology, you'll sometimes find "positively/negatively dependent", "positively/negatively correlated", or even "correlated/anti-correlated".

The strongest forms of positive and negative relevance are entailment and refutation. Suppose a hypothesis H has nonextreme prior credence. If a particular piece of evidence E *entails* the hypothesis, the probability axioms and Ratio Formula tell us that

$$\mathrm{cr}(H \mid E) = 1 \tag{3.21}$$

Supposing E takes H from a middling credence to the highest credence allowed. Similarly, if E refutes H (what philosophers of science call **falsification**), then

$$\mathrm{cr}(H \mid E) = 0 \tag{3.22}$$

Relevance will be most important to us because of its connection to confirmation, the Bayesian notion of evidential support. A piece of evidence confirms a hypothesis only if it's relevant to that hypothesis. Put another way, learning a piece of evidence changes a rational agent's credence in a hypothesis only if that evidence is relevant to the hypothesis. (Much more on this later.)

3.2.1 Conditional independence and screening off

The definition of probabilistic independence compares an agent's conditional credence in a proposition to her unconditional credence in that proposition. But we can also compare conditional credences. When Baxter, who believes in the occult, hears a clairvoyant's prediction about the outcome of a fair coin flip, he takes it to be highly correlated with the true flip outcome. But what if we ask Baxter to suppose that this particular clairvoyant is an impostor? Once he supposes the clairvoyant is an impostor, Baxter may take the clairvoyant's predictions to be completely irrelevant to the flip outcome. Let C be the proposition that the clairvoyant predicts heads, H be the proposition that the coin comes up heads, and I be the proposition that the clairvoyant is an impostor. It's possible for Baxter's credences to satisfy both of the following equations at once:

$$\mathrm{cr}(H \mid C) > \mathrm{cr}(H) \tag{3.23}$$

$$\mathrm{cr}(H \mid C \,\&\, I) = \mathrm{cr}(H \mid I) \tag{3.24}$$

Inequality (3.23) tells us that unconditionally, Baxter takes C to be relevant to H. But conditional on the supposition of I, C becomes independent of H (Equation (3.24)); once Baxter has supposed I, adding C to his suppositions doesn't affect his confidence in H.

Generalizing this idea yields the following definition of **conditional independence**: P is probabilistically independent of Q conditional on R just in case

$$\mathrm{cr}(P \mid Q \,\&\, R) = \mathrm{cr}(P \mid R) \tag{3.25}$$

If this equality fails to hold, we say that Q is relevant to P conditional on R.

One more piece of terminology: We will say that R **screens off** P from Q when Q is unconditionally relevant to P, but irrelevant to P conditional on each of R and $\sim R$. That is, R screens off P from Q just in case all three of the following are satisfied:

$$\mathrm{cr}(P \mid Q) \neq \mathrm{cr}(P) \tag{3.26}$$

$$\mathrm{cr}(P \mid Q \,\&\, R) = \mathrm{cr}(P \mid R) \tag{3.27}$$

$$\mathrm{cr}(P \mid Q \,\&\, {\sim}R) = \mathrm{cr}(P \mid {\sim}R) \tag{3.28}$$

When these equations are met, P and Q are correlated but supposing either R or $\sim R$ makes that correlation disappear.[9]

Conditional independence and screening off are both best understood through real-world examples. We'll see a number of those in the next few sections.

3.2.2 The Gambler's Fallacy

People often act as if future chancy events will "compensate" for unexpected past results. When a good hitter strikes out many times in a row, someone will say he's "due" for a hit. If a fair coin comes up heads nineteen times in a row, many people become more confident that the next outcome will be tails.

This mistake is known as the **Gambler's Fallacy**.[10] A person who makes the mistake is thinking along something like the following lines: In twenty flips of a fair coin, it's more probable to get nineteen heads and one tail than it is to get twenty heads. So having seen nineteen heads, it's much more likely that the twentieth flip will come up tails.

This person is providing the right answer to the wrong question. If the question is "When a fair coin is flipped twenty times, is it more likely that you'll get a *total* of nineteen heads and one tail than it is that you'll get twenty heads?", the answer to that question is "yes"—in fact, it's twenty times as likely! But that's the wrong question to ask in this case. Instead of wondering what sorts of total outcomes are probable when one flips a fair coin twenty times, in this case it's more appropriate to ask: *given* that the coin has already come up heads nineteen times, how confident are we that the twentieth flip will be tails? This is a question about our conditional credence

$$\text{cr}(\text{20th flip tails} \mid \text{first 19 flips heads}) \tag{3.29}$$

How should we calculate this conditional credence? Ironically, it might be more reasonable to make a mistake in the *opposite* direction from the Gambler's Fallacy. If I see a coin come up heads nineteen times, shouldn't that make me suspect that it's biased toward heads? If anything, shouldn't supposing nineteen consecutive heads make me *less* confident that the twentieth flip will come up tails?

This line of reasoning would be appropriate to the present case if we hadn't stipulated that the coin is fair. For a rational agent, the outcome of the twentieth flip is probabilistically independent of the outcomes of the first nineteen flips conditional on the fact that the coin is fair. That is,

$$\begin{aligned} \text{cr}(\text{20th flip tails} \mid \text{first 19 flips heads} \ \& \ \text{fair coin}) = \\ \text{cr}(\text{20th flip tails} \mid \text{fair coin}) \end{aligned} \tag{3.30}$$

We can justify this equation as follows: Typically, information that a coin came up heads nineteen times in a row would alter your opinion about whether it's a fair coin. Changing your opinion about whether it's a fair coin would then affect your prediction for the twentieth flip. So typically, information about the first nineteen flips alters your credences about the twentieth flip *by way of* your opinion about whether the coin is fair. But if you've already established that the coin is fair, information about the first nineteen flips has no further significance for your prediction about the twentieth. So conditional on the coin's being fair, the first nineteen flips' outcomes are irrelevant to the outcome of the twentieth flip.

The left-hand side of Equation (3.30) captures the correct question to ask about the Gambler's Fallacy case. The right-hand side is easy to calculate; it's 1/2. So after seeing a coin known to be fair come up heads nineteen times, we should be 1/2 confident that the twentieth flip will be tails.[11]

3.2.3 Probabilities are weird! Simpson's Paradox

Perhaps you're too much of a probabilistic sophisticate to ever commit the Gambler's Fallacy. Perhaps you successfully navigated Tversky and Kahneman's Conjunction Fallacy (Section 2.2.4) as well. But even probability experts sometimes have trouble with the counterintuitive relations that arise between conditional and unconditional dependence.

Here's an example of how odd things can get: In a famous case, the University of California, Berkeley's graduate departments were investigated for gender bias in admissions. The concern arose because in 1973 about 44% of overall male applicants were admitted to graduate school at Berkeley, while only 35% of female applicants were. Yet when the graduate departments (where admissions decisions are made) were studied one at a time, it turned out that individual departments either were admitting men and women at roughly equal rates, or in some cases were admitting a higher percentage of female applicants.

This is an example of **Simpson's Paradox**, in which probabilistic dependencies point in one direction conditional on each member of a partition, yet point the opposite way unconditionally. A Simpson's Paradox case involves a collection with a number of subgroups. Each of the subgroups displays the same correlation between two traits. Yet when we examine the collection as a whole, that correlation is reversed![12]

To see how this can happen, consider another example: Over the course of the 2016–2017 NBA season, Houston Rockets player James Harden made a higher percentage of his two-point shots than Toronto Raptors player DeMar DeRozan did. Harden also made a higher percentage of his three-point shots than DeRozan. Yet when we look at all the shots attempted (both two- and three-pointers), DeRozan made a higher percentage than Harden overall.[13]

Here are the data for the two players:

	Two-pointers		**Three-pointers**		**Combined**	
DeRozan	688/1421	48.4%	33/124	26.6%	721/1545	46.7%
Harden	412/777	53.0%	262/756	34.7%	674/1533	44.0%

The second number in each box is the number of attempts; the first is the number of makes; the third is the percentage (makes divided by attempts). Looking at the table, you can see how Harden managed to shoot better than DeRozan from each distance yet have a worse shooting percentage overall. Since three-pointers are more difficult to make than two-pointers, each player made a much higher percentage of his two-point attempts than his three-point attempts. DeRozan shot over 90% of his shots from the easier-to-make two-point distance. Harden, on the other hand, shot almost half of his shots from downtown. Harden was taking many more low-percentage shots than DeRozan, so even though Harden was better at those shots, his overall percentage suffered.

Scrutiny revealed a similar effect in Berkeley's 1973 admissions data. Bickel, Hammel, and O'Connell (1975) concluded, "The proportion of women applicants tends to be high in departments that are hard to get into and low in those that are easy to get into." Although individual departments were admitting women at comparable rates to men, female applications were less successful overall because more were directed at departments with low admission rates.[14]

Simpson's Paradox can be thought of entirely in terms of numerical proportions, as we've just done with the basketball and admissions examples. But these examples can also be analyzed using conditional probabilities. Suppose, for instance, that you are going to select a Harden or DeRozan shot attempt at random from the 2016–2017 season, making your selection so that each of the 3,078 attempts they put up together is equally likely to be selected. How confident should you be that the selected attempt will be a make? How should that confidence change on the supposition that a DeRozan attempt is selected, or a two-point attempt?

Below is a probability table for your credences, assembled from the real-life statistics above. Here D says that it's a DeRozan attempt; 2 says it's a two-pointer; and M says it's a make. (Given the pool from which we're sampling, $\sim D$ means a Harden attempt and ~ 2 means it's a three-pointer.)

D	2	M	cr
T	T	T	688/3078
T	T	F	733/3078
T	F	T	33/3078
T	F	F	91/3078
F	T	T	412/3078
F	T	F	365/3078
F	F	T	262/3078
F	F	F	494/3078

A bit of calculation with this probability table reveals the following:

$$\mathrm{cr}(M \mid D) > \mathrm{cr}(M \mid {\sim}D) \tag{3.31}$$

$$\mathrm{cr}(M \mid D \,\&\, 2) < \mathrm{cr}(M \mid {\sim}D \,\&\, 2) \tag{3.32}$$

$$\mathrm{cr}(M \mid D \,\&\, {\sim}2) < \mathrm{cr}(M \mid {\sim}D \,\&\, {\sim}2) \tag{3.33}$$

If you're selecting an attempt from the total sample, DeRozan is more likely to make it than Harden. Put another way, DeRozan's taking the attempt is unconditionally positively relevant to its being made (Equation (3.31)). But DeRozan's shooting is negatively relevant to a make conditional on each of the two distances (Equations (3.32) and (3.33)). If you're selecting from only the two-pointers, or from only the three-pointers, the shot is more likely to be made if it's attempted by Harden.

3.2.4 Correlation and causation

You may have heard the expression "correlation is not causation." People typically use this expression to point out that just because two events have both occurred—and maybe occurred in close spatio-temporal proximity—that doesn't mean they had anything to do with each other. But "correlation" is a technical term in probability discussions. The propositions describing two events may both be true, or you might have high credence in both of them, yet they still might not be probabilistically correlated. For the propositions to be correlated, supposing one to be true must *change* the probability of the other. I'm confident that I'm under 6 feet tall and that my eyes are blue, but that doesn't mean I take those facts to be correlated.

Once we've understood probabilistic correlation correctly, does its presence always indicate a causal connection? When two propositions about empirical events are correlated, must the event described by one cause the event described by the other? Hans Reichenbach offered a classic counterexample to this proposal. He wrote:

> Suppose two geysers which are not far apart spout irregularly, but throw up their columns of water always at the same time. The existence of a subterranean connection of the two geysers with a common reservoir of hot water is then practically certain. (1956, p. 158)

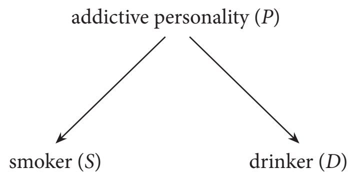

Figure 3.2 A causal fork

If you've noticed that two nearby geysers always spout simultaneously, seeing one spout will increase your confidence that the other is spouting as well. So your credences about the geysers are correlated. But you don't think one geyser's spouting *causes* the other to spout. Instead, you hypothesize an unobserved reservoir of hot water that is the **common cause** of both spouts.

Reichenbach proposed a famous principle about empirically correlated events:

Principle of the Common Cause: When event outcomes are probabilistically correlated, either one causes the other or they have a common cause.[15]

In addition to this principle, he offered a key mathematical insight about causation: a common cause screens its effects off from each other.

Let's work through an example of this insight concerning causation and screening off. Suppose the proposition that a particular individual is a drinker is positively relevant to the proposition that she's a smoker. According to the Principle of the Common Cause, there must be some causal link between these propositions. Perhaps drinking causes smoking—drinking creates situations in which one is more likely to smoke—or vice versa. Or they may be linked through a common cause: maybe being a smoker and being a drinker are both promoted by an addictive personality, which we can imagine results from a genetic endowment unaffected by one's behavior. (See Figure 3.2; the arrows indicate causal influence.)

Imagine the latter explanation is true, and moreover is the *only* true explanation of the correlation between drinking and smoking. That is, being a smoker and being a drinker are positively correlated only due to their both being caused by an addictive personality. Given this assumption, let's take a particular subject whose personality you're unsure about, and consider what happens to your credences when you make various suppositions about her.

If you begin by supposing that the subject drinks, this will make you more confident that she smokes—*but only because it makes you more confident*

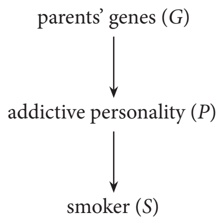

Figure 3.3 A causal chain

that the subject has an addictive personality. On the other hand, you might start by supposing that the subject has an addictive personality. That will certainly make you more confident that she's a smoker. But once you've made that adjustment, going on to suppose that she's a drinker won't affect your confidence in smoking. Information about drinking affects your smoking opinions only *by way of* helping you figure out whether she has an addictive personality, and the answer to the personality question was filled in by your initial supposition. Once an addictive personality is supposed, drinking has no further relevance to smoking. (Compare: Once a coin is supposed to be fair, the outcomes of its first nineteen flips have no relevance to the outcome of the twentieth.) Drinking becomes probabilistically independent of smoking conditional on suppositions about whether the subject has an addictive personality. That is,

$$\mathrm{cr}(S \mid D) > \mathrm{cr}(S) \tag{3.34}$$

$$\mathrm{cr}(S \mid D \,\&\, P) = \mathrm{cr}(S \mid P) \tag{3.35}$$

$$\mathrm{cr}(S \mid D \,\&\, {\sim}P) = \mathrm{cr}(S \mid {\sim}P) \tag{3.36}$$

Causal forks (as in Figure 3.2) give rise to screening off. P is a common cause of S and D, so P screens off S from D.

But that's not the only way screening off can occur. Consider Figure 3.3. Here we've focused on a different portion of the causal structure. Imagine that the subject's parents' genes causally influence whether she has an addictive personality, which in turn causally promotes smoking. Now her parents' genetics are probabilistically relevant to the subject's smoking, but that correlation is screened off by facts about her personality. Again, if you're uncertain whether the subject's personality is addictive, facts about her parents' genes will affect your opinion of whether she's a smoker. But once you've made a firm supposition about the subject's personality, suppositions about her parents' genetics have no further influence on your smoking opinions. In equation form:

$$\mathrm{cr}(S \mid G) > \mathrm{cr}(S) \tag{3.37}$$

$$\mathrm{cr}(S \mid G \,\&\, P) = \mathrm{cr}(S \mid P) \tag{3.38}$$

$$\mathrm{cr}(S \mid G \,\&\, {\sim}P) = \mathrm{cr}(S \mid {\sim}P) \tag{3.39}$$

P screens off *S* from *G*.[16]

Relevance, conditional relevance, and causation can interact in very complex ways.[17] My goal here has been to introduce the main ideas and terminology employed in their analysis. The state of the art in this field has come a long way since Reichenbach; computational tools now available can look at statistical correlations among a large number of variables and hypothesize a causal structure lying beneath them. The resulting causal diagrams are known as **Bayes Nets**, and have practical applications from satellites to health care to car insurance to college admissions.

These causal methods all start from Reichenbach's insight that common causes screen off their effects. And what of his more metaphysically radical Principle of the Common Cause? It remains highly controversial.

3.3 Conditional credences and conditionals

I now want to circle back and get clearer on the nature of conditional credence. First, it's important to note that the conditional credences we've been discussing are indicative, not subjunctive. This distinction is familiar from the theory of conditional propositions. Compare:

If Shakespeare didn't write *Hamlet*, then someone else did.
If Shakespeare hadn't written *Hamlet*, then someone else would have.

The former conditional is indicative, while the latter is subjunctive. The traditional distinction between these two types of conditional begins with the assumption that a conditional is evaluated by considering possible worlds in which the antecedent is satisfied, then checking whether the consequent is true in those worlds as well. In evaluating an indicative conditional, the antecedent worlds are restricted to those among the agent's doxastic possibilities.[18] Evaluating a subjunctive conditional, on the other hand, permits *counterfactual* reasoning involving antecedent worlds considered non-actual. So when you assess the subjunctive conditional above, you are allowed to consider worlds that make the antecedent true by making *Hamlet* never exist. But when evaluating the indicative conditional, you have to take into account that *Hamlet* actually does exist, and entertain only worlds in which that's true. So you consider bizarre "author-conspiracy" worlds that, while far-fetched, satisfy the antecedent and are among your current doxastic possibilities. In the end, I'm guessing you take the indicative conditional to be true but the subjunctive to be false.

Now suppose I ask for your credence in the proposition that someone wrote *Hamlet*, conditional on the supposition that Shakespeare didn't. This value will be high, again because you take *Hamlet* to exist. In assigning this conditional credence, you aren't bringing into consideration possible worlds you'd otherwise ruled out (such as *Hamlet*-free worlds). Instead, you're focusing in on the narrow set of author-conspiracy worlds you currently entertain. As we saw in Figure 3.1, assigning a conditional credence strictly narrows the worlds under consideration; it's doesn't expand your attention to worlds previously ruled out. Thus the conditional credences discussed in this book—and typically discussed in the Bayesian literature—are indicative rather than subjunctive.[19]

Are there more features of conditional propositions that can help us understand conditional credences? Might we understand conditional credences *in terms of* conditionals? Initiating his study of conditional degrees of belief, F.P. Ramsey warned against assimilating them to conditional propositions:

> We are also able to define a very useful new idea—"the degree of belief in p given q". This does not mean the degree of belief in "If p then q", or that in "p entails q", or that which the subject would have in p if he knew q, or that which he ought to have. (1931, p. 82)[20]

Yet many authors failed to heed Ramsey's warning. It's very tempting to equate conditional credences with some simple combination of conditional propositions and unconditional credences. For example, when I ask, "How confident are you in P given Q?", it's easy to hear that as "Given Q, how confident are you in P?" or just "If Q is true, how confident are you in P?" This simple slide might suggest that for any real number r and propositions P and Q,

$$\text{“cr}(P \mid Q) = r\text{” is equivalent to “}Q \rightarrow \text{cr}(P) = r\text{”} \tag{3.40}$$

Here I'm using the symbol "→" to represent some kind of conditional. For the reasons discussed above, it should be an indicative conditional. But it need not be the material conditional symbolized by "⊃"; many authors think the material conditional's truth-function fails to accurately represent the meaning of natural-language indicative conditionals.

Endorsing the equivalence in (3.40) would require serious changes to the traditional logic of conditionals. We can demonstrate this in two ways. First, we usually take indicative conditionals to satisfy the disjunctive syllogism rule. (The material conditional certainly does!) This rule tells us that

$$\text{“}X \rightarrow Z\text{” and “}Y \rightarrow Z\text{” jointly entail “}(X \vee Y) \rightarrow Z\text{”} \tag{3.41}$$

for any propositions X, Y, and Z. Thus for any propositions A, B, and C and constant k we have

$$\text{“}A \rightarrow [\mathrm{cr}(C) = k]\text{” and “}B \rightarrow [\mathrm{cr}(C) = k]\text{” entail “}(A \vee B) \rightarrow [\mathrm{cr}(C) = k]\text{”} \tag{3.42}$$

Combining (3.40) and (3.42) yields

$$\text{“}\mathrm{cr}(C \mid A) = k\text{” and “}\mathrm{cr}(C \mid B) = k\text{” entail “}\mathrm{cr}(C \mid A \vee B) = k\text{”} \tag{3.43}$$

(3.43) may look appealing, as a sort of probabilistic analog of disjunctive syllogism. But it's false. Not only can one design a credence distribution satisfying the probability axioms and Ratio Formula such that $\mathrm{cr}(C \mid A) = k$ and $\mathrm{cr}(C \mid B) = k$ but $\mathrm{cr}(C \mid A \vee B) \neq k$; one can even describe real-life examples in which it's rational for an agent to assign such a distribution. (See Exercise 3.14.) The failure of (3.43) is another case in which credences confound expectations developed by our experiences with classificatory states.

Second, we usually take indicative conditionals to satisfy *modus tollens*. But consider the following facts about me: Unconditionally, I am highly confident that I will be alive tomorrow. But conditional on the proposition that the sun just exploded, my confidence that I will be alive tomorrow is very low. Given these facts, *modus tollens*, and (3.40), I could run the following argument:

cr(alive tomorrow | sun exploded) is low.	[given]	(3.44)
If the sun exploded, cr(alive tomorrow) is low.	[(3.44), (3.40)]	(3.45)
cr(alive tomorrow) is high.	[given]	(3.46)
The sun did not explode.	[*modus tollens*]	(3.47)

While intriguing for its promise of astronomy by introspection, this argument is unsound. So I conclude that as long as indicative conditionals satisfy classical logical rules such as disjunctive syllogism and *modus tollens*, any analysis of conditional credences in terms of conditionals that uses (3.40) must be false.[21]

Perhaps we've mangled the transition from conditional credences to conditional propositions. Perhaps we should hear "How confident are you in P given Q?" as "How confident are you in 'P, given Q'?", which is in turn "How confident are you in 'If Q, then P'?" Maybe a conditional credence is a credence in a conditional. Or perhaps more weakly: an agent assigns a particular conditional credence value whenever she unconditionally assigns that value to a conditional. In symbols, the proposal is that

$$\text{“cr}(P \mid Q) = r\text{” is equivalent to “cr}(Q \rightarrow P) = r\text{”} \tag{3.48}$$

for any real r, any propositions P and Q, any credence distribution cr, and some indicative conditional $\rightarrow$. If true, this equivalence would offer another possibility for analyzing conditional credences in terms of unconditional credences and conditionals.

We can quickly show that (3.48) fails if "$\rightarrow$" is read as the material conditional $\supset$. Under the material reading, the proposal entails that

$$\text{cr}(P \mid Q) = \text{cr}(Q \supset P) \tag{3.49}$$

Using the probability calculus and Ratio Formula, we can show that Equation (3.49) holds only when $\text{cr}(Q) = 1$ or $\text{cr}(Q \supset P) = 1$. (See Exercise 3.15.) This is a *triviality result*: It shows that Equation (3.49) can hold only in trivial cases, namely over the narrow range of conditionals for which the agent is either certain of the antecedent or certain of the conditional itself. Equation (3.49) does not express a truth that holds for *all* conditional credences in *all* propositions; nor does (3.48) when "$\rightarrow$" is read materially.

Perhaps the equivalence in (3.48) can be saved from this objection by construing its "$\rightarrow$" as something other than a material conditional. But David Lewis (1976) provided a clever objection that works whichever conditional $\rightarrow$ we choose. Begin by selecting arbitrary propositions P and Q. We then derive the following from the proposal in Equation (3.48):

$$\text{cr}(Q \rightarrow P) = \text{cr}(P \mid Q) \qquad \text{[from (3.48)]} \tag{3.50}$$

$$\text{cr}(Q \rightarrow P \mid P) = \text{cr}(P \mid Q \,\&\, P) \qquad \text{[see below]} \tag{3.51}$$

$$\text{cr}(Q \rightarrow P \mid P) = 1 \qquad [Q \,\&\, P \text{ entails } P] \tag{3.52}$$

$$\text{cr}(Q \rightarrow P \mid {\sim}P) = \text{cr}(P \mid Q \,\&\, {\sim}P) \qquad \text{[see below]} \tag{3.53}$$

$$\text{cr}(Q \rightarrow P \mid {\sim}P) = 0 \qquad [Q \,\&\, {\sim}P \text{ refutes } P] \tag{3.54}$$

$$\mathrm{cr}(Q \rightarrow P) = \mathrm{cr}(Q \rightarrow P \mid P) \cdot \mathrm{cr}(P) + \mathrm{cr}(Q \rightarrow P \mid {\sim}P) \cdot \mathrm{cr}({\sim}P) \quad \text{[Law of Total Prob.]} \tag{3.55}$$

$$\mathrm{cr}(Q \rightarrow P) = 1 \cdot \mathrm{cr}(P) + 0 \cdot \mathrm{cr}({\sim}P) \quad [(3.52), (3.54), (3.55)] \tag{3.56}$$

$$\mathrm{cr}(Q \rightarrow P) = \mathrm{cr}(P) \tag{3.57}$$

$$\mathrm{cr}(P \mid Q) = \mathrm{cr}(P) \quad [(3.50)] \tag{3.58}$$

Some of these lines require explanation. The idea of lines (3.51) and (3.53) is this: We've already seen that a credence distribution conditional on a particular proposition satisfies the probability axioms. This suggests that we should think of a distribution conditional on a proposition as being just like any other credence distribution. (We'll see more reason to think this in Chapter 4, note 3.) So a distribution conditional on a proposition should satisfy the proposal of (3.48) as well. If you conditionally suppose X, then under that supposition you should assign $Y \rightarrow Z$ the same credence you would assign Z were you to *further* suppose Y. In other words, anyone who maintains (3.48) should also maintain that for any X, Y, and Z,

$$\mathrm{cr}(Y \rightarrow Z \mid X) = \mathrm{cr}(Z \mid Y \,\&\, X) \tag{3.59}$$

In line (3.51) the roles of X, Y, and Z are played by P, Q, and P; in line (3.53) it's ${\sim}P$, Q, and P.

Lewis has given us another triviality result. Assuming the probability axioms and Ratio Formula, the proposal in (3.48) can hold only for propositions P and Q such that $\mathrm{cr}(P \mid Q) = \mathrm{cr}(P)$. In other words, it can hold only for propositions the agent takes to be independent.[22] Or (taking things from the other end), the proposed equivalence can hold for all the conditionals in an agent's language only if the agent treats every pair of propositions in $\mathcal{L}$ as independent![23]

So a rational agent's conditional credence will not in general equal her unconditional credence in a conditional. This is not to say that conditional credences have *nothing* to do with conditionals. A popular idea now usually called "Adams's Thesis" (Adams 1965) holds that an indicative conditional $Q \rightarrow P$ is *acceptable* to a degree equal to $\mathrm{cr}(P \mid Q)$.[24] But we cannot maintain that an agent's conditional credence is always equal to her credence that some conditional is *true*.

This brings us back to a proposal I discussed in Chapter 1. One might try to relate degrees of belief to binary beliefs by suggesting that whenever an agent has an r-valued credence, she has a binary belief in a traditional proposition with r as part of its content. Working out this proposal for conditional credences reveals how hopeless it is. Suppose an agent assigns $\mathrm{cr}(P \mid Q) = r$. Would we suggest that the agent believes that if Q, then the probability of P is r? This proposal mangles the logic of conditional credences. Perhaps the agent believes that the probability of "if Q, then P" is r? Lewis's argument dooms this idea.

I said in Chapter 1 that the numerical value of an unconditional degree of belief is an attribute of the *attitude taken* toward a proposition, not a *constituent of* that proposition itself. As for conditional credences, $\mathrm{cr}(P \mid Q) = r$ does not say that an agent takes some attitude toward a conditional proposition with a probability value in its consequent. Nor does it say that the agent takes some attitude toward a single, conditional proposition composed of P and Q. $\mathrm{cr}(P \mid Q) = r$ says that the agent takes an r-valued attitude toward an *ordered pair* of propositions—neither of which need involve the number r.

3.4 Exercises

Unless otherwise noted, you should assume when completing these exercises that the credence distributions under discussion satisfy the probability axioms and Ratio Formula. You may also assume that whenever a conditional credence expression occurs, the condition has a nonzero unconditional credence so that the conditional credence is well defined.

Problem 3.1. A family has two children of different ages. Assume that each child has a probability of 1/2 of being a girl, and that the probability that the elder is a girl is independent of the probability that the younger is.

(a) Conditional on the older child's being a girl, what's the probability that the younger one is?

(b) Conditional on at least one child's being a girl, what's the probability that they both are?

Problem 3.2. Flip and Flop are playing a game. They have a fair coin that they are going to keep flipping until one of two things happens: either the coin comes up heads twice in a row, or it comes up tails followed by heads. The first time one of these things happens, the game ends—if it ended with HH, Flip wins; if it ended with TH, Flop wins.

(a) What's the probability that Flip wins after the first two tosses of the coin? What's the probability that Flop wins after the first two tosses of the coin?

(b) Flip and Flop play their game until it ends (at which point one of them wins). What's the probability that Flop is the winner?[25]

Problem 3.3. 🌶🌶 One might think that real humans only ever assign credences that are rational numbers—and perhaps only rational numbers involving relatively small whole-number numerators and denominators. But we can construct simple conditions that *require* an irrational-valued credence distribution. For example, consider the scenario below.

You have a biased coin that you are going to flip twice in a row. Suppose your credence distribution satisfies all of the following conditions:

(i) You are equally confident that the first flip will come up heads and that the second flip will come up heads.

(ii) You treat the outcomes of the two flips as probabilistically independent.

(iii) Given what you know about the bias, your confidence that the two flips will *both* come up heads equals your confidence in all of the other outcomes put together.

Assuming your credence distribution satisfies the three conditions above, how confident are you that the first flip will come up heads?[26]

Problem 3.4. 🌶🌶 Prove that credences conditional on a particular proposition form a probability distribution. That is, prove that for any proposition R in $\mathcal{L}$ such that $\text{cr}(R) > 0$, the following three conditions hold:

(a) For any proposition P in $\mathcal{L}$, $\text{cr}(P \mid R) \geq 0$.

(b) For any tautology T in $\mathcal{L}$, $\text{cr}(\mathsf{T} \mid R) = 1$.

(c) For any mutually exclusive propositions P and Q in $\mathcal{L}$, $\text{cr}(P \vee Q \mid R) = \text{cr}(P \mid R) + \text{cr}(Q \mid R)$.

Problem 3.5. 🌶 Pink gumballs always make my sister sick. (They remind her of Pepto Bismol.) Blue gumballs make her sick half of the time (they just look unnatural), while white gumballs make her sick only one-tenth of the time. Yesterday, my sister bought a single gumball randomly selected from a machine that's one-third pink gumballs, one-third blue, and one-third white. Applying the version of Bayes's Theorem in Equation (3.12), how confident should I be that my sister's gumball was pink conditional on the supposition that it made her sick?

Problem 3.6. 🌶🌶

(a) Prove Bayes's Theorem from the probability axioms and Ratio Formula. (Hint: Start by using the Ratio Formula to write down expressions involving $\text{cr}(H \,\&\, E)$ and $\text{cr}(E \,\&\, H)$.)

(b) Exactly which unconditional credences must we assume to be positive in order for your proof to go through?

(c) Where exactly does your proof rely on the probability axioms (and not just the Ratio Formula)?

Problem 3.7. Once more, consider the probabilistic credence distribution specified by this probability table (from Exercise 2.9):

P	Q	R	cr
T	T	T	0.1
T	T	F	0.2
T	F	T	0
T	F	F	0.3
F	T	T	0.1
F	T	F	0.2
F	F	T	0
F	F	F	0.1

Answer the following questions about this distribution:

(a) What is $\mathrm{cr}(P \mid Q)$?

(b) Relative to this distribution, is Q positively relevant to P, negatively relevant to P, or probabilistically independent of P?

(c) What is $\mathrm{cr}(P \mid R)$?

(d) What is $\mathrm{cr}(P \mid Q \,\&\, R)$?

(e) Conditional on R, is Q positively relevant to P, negatively relevant to P, or probabilistically independent of P?

(f) Does R screen off P from Q? Explain why or why not.

Problem 3.8. Prove that all the alternative statements of probabilistic independence in Equations (3.15) through (3.18) follow from our original independence definition. That is, prove that each equation (3.15) through (3.18) follows from Equation (3.14), the probability axioms, and the Ratio Formula. (Hint: Once you prove that a particular equation follows from Equation (3.14), you may use it in subsequent proofs.)

Problem 3.9. Show that probabilistic independence is not transitive. That is, provide a single probability distribution on which all of the following are true: X is independent of Y, and Y is independent of Z, but X is not independent of Z. Show that your distribution satisfies all three conditions. (For an added chili pepper of difficulty, have your distribution assign every state-description a nonzero unconditional credence.)

Problem 3.10. In the text we discussed what makes a *pair* of propositions probabilistically independent. If we have a larger collection of propositions, what does it take to make them all independent of each other? You might think all that's necessary is *pairwise independence*—for each pair within the set of propositions to be independent. But pairwise independence doesn't guarantee that each proposition will be independent of *combinations* of the others.

To demonstrate this fact, describe a real-world example (spelling out the propositions represented by X, Y, and Z) in which it would be rational for an agent to assign credences meeting all four of the following conditions:

(i) $\text{cr}(X \mid Y) = \text{cr}(X)$
(ii) $\text{cr}(X \mid Z) = \text{cr}(X)$
(iii) $\text{cr}(Y \mid Z) = \text{cr}(Y)$
(iv) $\text{cr}(X \mid Y \,\&\, Z) \neq \text{cr}(X)$

Show that your example satisfies all four conditions.

Problem 3.11. Using the program PrSAT referenced in the Further Readings for Chapter 2, find a probability distribution satisfying all the conditions in Exercise 3.10, plus the following *additional* condition: Every state-description expressible in terms of X, Y, and Z must have a nonzero unconditional cr-value.

Problem 3.12.

(a) The 2016–2017 NBA season has just ended, and you're standing on a basketball court. Suddenly aliens appear, point to a spot on the court, and say that unless someone makes the next shot attempted from that spot, they will end your life. You're highly interested in self-preservation but terrible at basketball; luckily James Harden and DeMar DeRozan are standing right there. DeRozan says, "I had a better overall shooting percentage than Harden this year, so I should attempt the shot." Given the statistics on page 70, explain why DeRozan's argument is unconvincing.

(b) Suppose the aliens pointed to a spot that would yield a three-point attempt. You're about to hand the ball to Harden, when DeRozan says, "I know that's a three-pointer, and Harden shot better from three-point range in general than I did this year. But from *that particular spot*, I had a better percentage than him." Is DeRozan's claim consistent with the statistics on page 70? (That is, could DeRozan's claim be true while those statistics are also accurate?) Explain why or why not.

Problem 3.13. After laying down probabilistic conditions for a causal fork, Reichenbach demonstrated that a causal fork induces correlation. Consider the following four conditions:

(i) $\text{cr}(A \mid C) > \text{cr}(A \mid {\sim}C)$
(ii) $\text{cr}(B \mid C) > \text{cr}(B \mid {\sim}C)$
(iii) $\text{cr}(A \,\&\, B \mid C) = \text{cr}(A \mid C) \cdot \text{cr}(B \mid C)$
(iv) $\text{cr}(A \,\&\, B \mid {\sim}C) = \text{cr}(A \mid {\sim}C) \cdot \text{cr}(B \mid {\sim}C)$

(a) Assuming C is the common cause of A and B, explain what each of the four conditions means in terms of relevance, independence, conditional relevance, or conditional independence.
(b) Prove that if all four conditions hold, then $\text{cr}(A \,\&\, B) > \text{cr}(A) \cdot \text{cr}(B)$.

Problem 3.14. In Section 3.3 I pointed out that the following statement (labeled Equation (3.43) there) does not hold for every constant k and propositions A, B, and C:

$$\text{"cr}(C \mid A) = k\text{" and "cr}(C \mid B) = k\text{" entail "cr}(C \mid A \vee B) = k\text{"}$$

(a) Describe a real-world example (involving dice, or cards, or something more interesting) in which it's rational for an agent to assign $\text{cr}(C \mid A) = k$ and $\text{cr}(C \mid B) = k$ but $\text{cr}(C \mid A \vee B) \neq k$. Show that your example meets this description.
(b) Prove that if A and B are mutually exclusive, then whenever $\text{cr}(C \mid A) = k$ and $\text{cr}(C \mid B) = k$ it's also the case that $\text{cr}(C \mid A \vee B) = k$.

Problem 3.15. Fact: For any propositions P and Q, if $\text{cr}(Q) > 0$ then $\text{cr}(Q \supset P) \geq \text{cr}(P \mid Q)$.

(a) Starting from a language $\mathcal{L}$ with atomic propositions P and Q, build a probability table on its state-descriptions and use that table to prove the fact above.
(b) Show that Equation (3.49) in Section 3.3 entails that either $\text{cr}(Q) = 1$ or $\text{cr}(Q \supset P) = 1$.

3.5 Further reading

INTRODUCTIONS AND OVERVIEWS

Alan Hájek (2011a). Conditional Probability. In: *Philosophy of Statistics*. Ed. by Prasanta S. Bandyopadhyay and Malcolm R. Forster. Vol. 7. Handbook of the Philosophy of Science. Amsterdam: Elsevier, pp. 99–136

Describes the Ratio Formula and its motivations. Then works through a number of philosophical applications of conditional probability, and a number of objections to the Ratio Formula. Also discusses conditional-probability-first formalizations (as described in note 3 of this chapter).

Todd A. Stephenson (2000). *An Introduction to Bayesian Network Theory and Usage*. Tech. rep. 03. IDIAP

Section 1 provides a nice, concise overview of what a Bayes Net is and how it interacts with conditional probabilities. (Note that the author uses A, B to express the *conjunction* of A and B.) Things get fairly technical after that as he covers algorithms for creating and using Bayes Nets. Sections 6 and 7, though, contain real-life examples of Bayes Nets for speech recognition, Microsoft Windows troubleshooting, and medical diagnosis.

Christopher R. Hitchcock (2021). Probabilistic Causation. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2021

While this entry is primarily about *analyses* of the concept of causation using probability theory, along the way Hitchcock includes impressive coverage of the Principle of the Common Cause, Simpson's Paradox, causal modeling with Bayes Nets, and related material.

CLASSIC TEXTS

Hans Reichenbach (1956). The Principle of Common Cause. In: *The Direction of Time*. Berkeley: University of California Press, pp. 157–60

Text in which Reichenbach introduces his account of common causes in terms of screening off. (Note that Reichenbach uses a period to express conjunction, and a comma rather than a vertical bar for conditional probabilities—what we would write as $\mathrm{cr}(A \mid B)$ he writes as $\mathrm{P}(B, A)$.)

David David Lewis (1976). Probabilities of Conditionals and Conditional Probabilities. *The Philosophical Review* 85, pp. 297–315

Article in which Lewis presents his triviality argument concerning probabilities of conditionals.

EXTENDED DISCUSSION

Bas C. van Fraassen (1982). Rational Belief and the Common Cause Principle. In: *What? Where? When? Why?* Ed. by Robert McLaughlin. Dordrecht: Reidel, pp. 193–209

Frank Arntzenius (1993). The Common Cause Principle. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2, pp. 227–37

Discuss the meaning and significance of Reichenbach's Principle of the Common Cause, then present possible counterexamples (including counterexamples from quantum mechanics).

Alan Hájek (2011b). Triviality Pursuit. *Topoi* 30, pp. 3–15

Explains the plausibility and significance of the claim that probabilities of conditionals are conditional probabilities. Then canvasses a variety of Lewis-style triviality arguments against that claim.

Notes

1. Here's a good way to double-check that 6 & E is equivalent to 6: Remember that equivalence is mutual entailment. Clearly 6 & E entails 6. Going in the other direction, 6 entails 6, but 6 also entails E. So 6 entails 6&E. When evaluating conditional credences using the Ratio Formula, we'll often find ourselves simplifying a conjunction down to just one or two of its conjuncts. For this to work, the conjunct that remains has to entail each of the conjuncts that was removed.
2. When I refer to an agent's "credence distribution" going forward, I will often be referring to both her unconditional and conditional credences. Strictly speaking this extends our definition of a "distribution", but since conditional credences supervene on unconditional for rational agents, not much damage will be done.
3. Some authors take advantage of Equation (3.7) to formalize probability theory in exactly the opposite order from the way I've been proceeding. They begin by introducing conditional credences and subject them to a number of constraints somewhat like

Kolmogorov's axioms. The desired rules for *un*conditional credences are then obtained by introducing the single constraint that for all P in $\mathcal{L}$, $\text{cr}(P) = \text{cr}(P \mid \top)$. For more on this approach, its advocates, and its motivations, see Section 5.4.

4. Bayes never published the theorem; Richard Price found it in Bayes's notes and published it after Bayes's death in 1761. Pierre-Simon Laplace independently rediscovered the theorem later on and was responsible for much of its early popularization.
5. In everyday English "likely" is a synonym for "probable". Yet R.A. Fisher introduced the technical term "likelihood" to represent a particular *kind* of probability—the probability of some evidence given a hypothesis. This somewhat peculiar terminology has stuck.
6. Quoted in Galavotti (2005, p. 51).
7. For instance, we have to assume that the base rates of the alleles are equal in the population, none of the relevant phenotypes is fitter than any of the others, and the blue-finned fish don't assortatively mate with other blue-finned fish. (Thanks to Hayley Clatterbuck for discussion.)
8. Throughout this section and Section 3.2.1, I will assume that any proposition appearing in the condition of a conditional cr-expression has a nonzero cr-value. Defining probabilistic independence for propositions with probability 0 can get complicated. (See e.g. Fitelson and Hájek (2014))
9. One will sometimes see "screening off" defined without Equation (3.28) or its analogue. (That is, some authors define screening off in terms of R's making the correlation between P and Q disappear, without worrying whether $\sim R$ has the same effect.) Equation (3.28) makes an important difference to our definition: in the Baxter example I does not screen off H from C according to our definition because when $\sim I$ is supposed, C becomes very relevant to H.

   I have included Equation (3.28) in our definition because it connects our approach to the more general notion of screening off used in the statistics community. In statistics one often works with continuous random variables, and the idea is that random variable Z screens off X from Y if X and Y become independent conditional on each possible value of Z. Understanding proposition R as a dichotomous random variable (Chapter 2, note 7) converts this general definition of screening off into the particular definition I've given in the text.

   Many authors also leave Equation (3.26) (or its analogue) implicit in their definitions of "screening off". But since examples of screening off always involve unconditional correlations that disappear under conditioning, I've made this feature explicit in my definition.
10. Not to be confused with the Rambler's Fallacy: I've said so many false things in a row, the next one must be true!
11. Twenty flips of a fair coin provide a good example of what statisticians call **IID trials**. "IID" stands for "independent, identically distributed." The flips are "independent" because each is probabilistically independent of all the others; information about the outcomes of other coin flips doesn't change the probability that a particular flip will come up heads. The flips are "identically distributed" because each flip has the same probability of producing heads (as contrasted with a case in which some of the flips are of a fair coin while others are flips of a biased coin).

12. This paradoxical phenomenon is named after E.H. Simpson because of a number of striking examples he gave in his (1951). Yet the phenomenon had been known to statisticians as early as Pearson, Lee, and Bramley-Moore (1899) and Yule (1903).
13. I learned about the Harden/DeRozan example from Reuben Stern, who in turn learned about it from a reddit post by a user named Jerome Williams. (I copied the specific data for the two players from stats.nba.com.) The UC Berkeley example was brought to the attention of philosophers by Cartwright (1979).
14. Notice that these findings only address one potential form of bias that might have been present in Berkeley's graduate admissions. For instance, they're consistent with the possibility that women were being actively discouraged from applying to the less selective departments.
15. I'm playing a bit fast and loose with the objects of discussion here. Throughout this chapter we're considering correlations in an agent's credence distribution. Reichenbach was concerned not with probabilistic correlations in an agent's credences but instead with correlations in objective frequencies or chance distributions (about which more in Chapter 5). But presumably if the Principle of the Common Cause holds for objective probability distributions, that provides an agent who views particular propositions as empirically correlated with some reason to suppose that the events described in those propositions either stand as cause to effect or share a common cause.
16. You might worry that Figure 3.3 presents a counterexample to Reichenbach's Principle of the Common Cause, because G and S are unconditionally correlated yet G doesn't cause S and they have no common cause. It's important to the principle that the causal relations need not be *direct*; for Reichenbach's purposes G counts as a cause of S even though it's not the immediate cause of S.
17. Just to indicate a few more complexities that may arise: While our discussion in the text concerns "direct" common causes, one can have an "indirect" common cause that doesn't screen off its effects from each other. For example, if we imagine merging Figures 3.2 and 3.3 to show how the subject's parents' genes are a common cause of both smoking and drinking by way of her addictive personality, it is possible to arrange the numbers so that her parents' genetics don't screen off smoker from drinker. Even more complications arise if some causal arrows do end-arounds past others—what if in addition to the causal structure just described, the parents' genetics tend to make *them* smokers, which in turn directly influences the subject's smoking behavior?
18. Here I assume that a rational agent will entertain an indicative conditional only if she takes its antecedent to be possible. For arguments in favor of this position, and citations to the relevant literature, see Moss (2018, Sect. 4.3) and Titelbaum (2013a, Sect. 5.3.2). The analogous assumption for conditional credences is that an agent assigns a conditional credence only when its condition is true in at least one of her doxastically possible worlds.
19. One *could* study a kind of attitude different from the conditional credences considered in this book—something like a subjunctive degree of belief. Joyce (1999) does exactly that, but is careful to distinguish his analysandum from standard conditional degrees of belief. (For instance, the arguments for the Ratio Formula given earlier in this chapter do not extend to Joycean subjunctive credences.) Schwarz (2018) then evaluates triviality arguments for subjunctive conditional credences much like the triviality arguments for indicatives I will go on to consider in this section.

20. I realize some of the "p"s and "q"s in this quote are flipped around from what one might expect, but that's how it's printed in my copy of Ramsey (1931). Context makes clear that the ordering is *not* why Ramsey rejects the proposed equivalents to "the degree of belief in p given q"; he'd still reject them were the order inverted.
21. A variety of recent positions in linguistics and the philosophy of language suggest that indicative conditionals with modal expressions in their consequents do not obey classical logical rules. Yalcin (2012), among others, classes probability locutions with these modals and so argues that, inter alia, indicative conditionals with probabilistic consequents do not keep *modus tollens* truth-preserving. (His argument could easily be extended to disjunctive syllogism as well.) Yet the alternative positive theory of indicative conditionals Yalcin offers does not analyze conditional credences in terms of conditionals either, so even if he's correct, we would still need an independent understanding of what conditional credences are. (Thanks to Fabrizio Cariani for discussion of these points.)
22. A careful reader will note that the proof given fails when $\mathrm{cr}(P \mid Q)$ takes an extreme value. If $\mathrm{cr}(P \mid Q) = 0$, the condition in $\mathrm{cr}(P \mid Q \,\&\, P)$ will have unconditional credence 0, while if $\mathrm{cr}(P \mid Q) = 1$, the condition in $\mathrm{cr}(P \mid Q \,\&\, {\sim}P)$ will have credence 0. So strictly speaking the triviality result is that (3.48) can hold only when the agent takes P and Q to be independent, or dependent in the strongest possible fashion. This is no more plausible than the less careful version I've given in the text. (Thanks to Glauber de Bona for being the careful reader who caught this for me!)
23. Fitelson (2015) proves a triviality result like Lewis's using probability tables (instead of proceeding axiomatically). Moreover, he traces the triviality specifically to the combination of (3.48) with the assumption that the conditional $\rightarrow$ satisfies what's known as the "import-export" condition.
24. Interestingly, this idea is often traced back to a suggestion in Ramsey, known as "Ramsey's test" (Ramsey 1929/1990, p. 155n).
25. Thanks to Irving Lubliner for inspiring this problem.
26. This problem was inspired by a problem of Branden Fitelson's. Thanks to Catrin Campbell-Moore for devising this particularly elegant set of conditions.