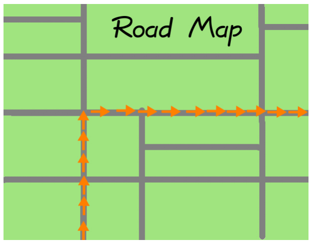
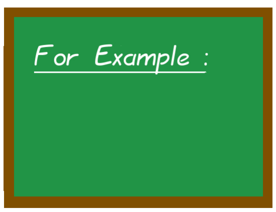
Module 7.6: You Can't Just Add Probabilities



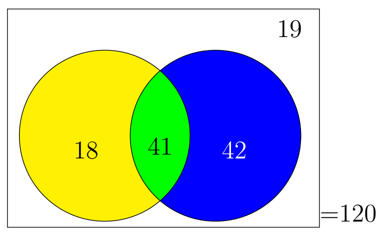
It has been my practice for many years to go over the final examinations of the middle quintile of my students. After the grades are submitted, I carefully re-examine what they wrote, not for the purposes of re-grading, but to try to identify any common misconceptions. Often, I can then warn future students about those misconceptions, so that they do not fall into the same trap.

One misconception that I discovered this way has to do with probabilities of the union of two events—problems centering around the word "or." This easy module will discuss how to avoid that pitfall.

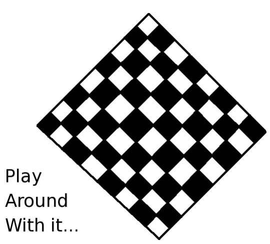


7-6-1

We're going to start with a very simple problem. Let's say that there is an office building in New Orleans where 120 people work. They are trying to decide what kind of music to play in the elevators and other public areas. After some conversation, it comes down to jazz or country music, and they decide to use an online survey to decide the matter. The following Venn Diagram summarizes the information:



The left circle represents jazz, and the right circle represents country music. I'd now like to ask you a few questions in the next box.



7-6-2

Let's look again at the data in the previous box.

- What is the probability that a random person in the building likes jazz? [Answer: $59/120 = 0.491\overline{6}$.]
- What is the probability that a random person in the building likes country music? [Answer: $83/120 = 0.6916\overline{6}$.]

By now, this is a phenomenally easy problem for us. You might be wondering where I'm going with all of this.

Continuing with the previous two boxes, let $Pr\{J\}$ represent the probability that someone likes jazz, and $Pr\{C\}$ represent the probability that someone likes country music. There are many students who understandably imagine that

$$\text{WRONG!} \rightarrow \qquad Pr\{J \cup C\} = Pr\{J\} + Pr\{C\} \qquad \leftarrow \text{WRONG!}$$

but this isn't true. We will see why, in the next box.

We can see that

$$\text{WRONG!} \rightarrow \qquad Pr\{J \cup C\} = Pr\{J\} + Pr\{C\} \qquad \leftarrow \text{WRONG!}$$

isn't true in this case by simply adding the probabilities:

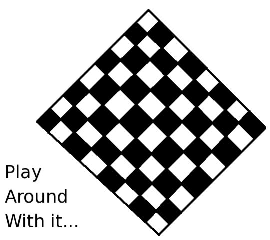
$$\frac{59}{120} + \frac{83}{120} = \frac{142}{120}$$

Surely 142/120 does not make sense as a probability! Probabilities must always be between zero and one, inclusively. For those students who prefer decimals,

$$0.491\overline{6} + 0.6916\overline{6} = 1.183\overline{3}$$

but surely no probability can be greater than 1.

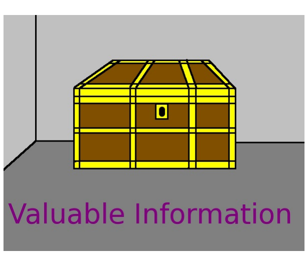
Let's now see one correct way of doing this problem in the next box.



7-6-3

Let's look at the Venn Diagram from the previous example. Let $\mathcal{J}$ be the set of people who like jazz, and let $\mathcal{C}$ be the set of people who like country music.

- How many people are in the set $\mathcal{C}$? [Answer: 83.]
- How many people are in the set $\mathcal{J}$? [Answer: 59.]
- How many people are in the set $\mathcal{C} \cup \mathcal{J}$? [Answer: 101.]
- What is the probability that a random person in the office building likes either country music or jazz? [Answer: $101/120 = 0.8416\overline{6}$.]



Now we know that for events A and B, it is not always true that

$$\text{WRONG!} \rightarrow \qquad Pr\{A \cup B\} = Pr\{A\} + Pr\{B\} \qquad \leftarrow \text{WRONG!}$$

- One strategy for calculating the probabilities of a union (an "or") is to draw a Venn Diagram, if one has not already been provided. Then you can see what the Venn Diagram says for the union of the two sets, and compute the correct probability from there.
- Another strategy involves "the inclusion-exclusion formula of probability," which is just a tiny variation of "the inclusion-exclusion formula of sets" that we first saw on Page 899. This formula works as a "shortcut" formula for problems involving a union (an "or"). We will learn this formula momentarily.
- There are also occasions where adding the probabilities will give the correct answer. These are not very rare. We will talk about these later in this module.

Let's see what happens when we manipulate a familiar formula, first given on Page 899, called the inclusion-exclusion principle for sets:

$$\#(\mathcal{A} \cup \mathcal{B}) = \#\mathcal{A} + \#\mathcal{B} - \#(\mathcal{A} \cap \mathcal{B})$$

We'll start by dividing every term by the "universal set" of the problem being studied, $\mathcal{U}$. We obtain

$$\frac{\#(\mathcal{A} \cup \mathcal{B})}{\#\mathcal{U}} = \frac{\#\mathcal{A}}{\#\mathcal{U}} + \frac{\#\mathcal{B}}{\#\mathcal{U}} - \frac{\#(\mathcal{A} \cap \mathcal{B})}{\#\mathcal{U}}$$

which looks frightening at first. However, we need only realize that

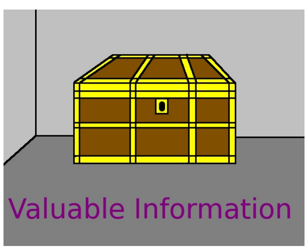
$$\frac{\#\mathcal{A}}{\#\mathcal{U}} = Pr\{\mathcal{A}\} \qquad \frac{\#\mathcal{B}}{\#\mathcal{U}} = Pr\{\mathcal{B}\}$$

and similarly

$$\frac{\#(\mathcal{A} \cup \mathcal{B})}{\#\mathcal{U}} = Pr\{\mathcal{A} \cup \mathcal{B}\} \qquad \frac{\#(\mathcal{A} \cap \mathcal{B})}{\#\mathcal{U}} = Pr\{\mathcal{A} \cap \mathcal{B}\}$$

The above probably looks like symbol soup. All I am doing is dividing various counts of objects in sets by the total number of objects in the entire problem. Those counts then become probabilities. Substituting all of that in there, we get

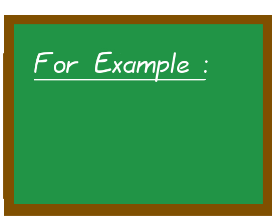
$$Pr\{\mathcal{A} \cup \mathcal{B}\} = Pr\{\mathcal{A}\} + Pr\{\mathcal{B}\} - Pr\{\mathcal{A} \cap \mathcal{B}\}$$



For any two events A and B from the same sample space, the following formula holds true.

$$Pr\{A \cup B\} = Pr\{A\} + Pr\{B\} - Pr\{A \cap B\}$$

This formula is called "the inclusion-exclusion principle for probability," and is very useful. In particular, it is great at converting situations where we know the probability of A as well as B, into knowledge of the probability of $A \cup B$, as well as the reverse direction. We'll practice with it now.



7-6-4

Let's suppose a medical company has a new drug, which will react badly for anyone who is on blood-pressure medicine or cholesterol medicine. With that in mind, it might not make sense to advertise in certain venues. For example, teenagers are often not on either blood-pressure medicine nor cholesterol medicine, while executives are often on both. A market survey is done online, and returns that 75% of executives are on cholesterol medication and 55% are on blood-pressure medication. This adds to more than 100%, which confuses the marketing people. An intern explains that some people are taking both medicines. The same survey audience is asked one more question: how many executives are taking both? The answer is 40%. Compute the probability that a random executive in the sample can take the new drug, and the probability that a random executive cannot take the new drug.

To solve this problem, let C be the event that a random executive is on cholesterol medicine, and B be the event that a random executive is on blood-pressure medicine. We now know that

$$Pr\{C\} = 0.75 \qquad Pr\{B\} = 0.55 \qquad Pr\{B \cap C\} = 0.4$$

We will continue the problem in the next box.

Continuing with the previous box, we know the inclusion-exclusion principle for probability

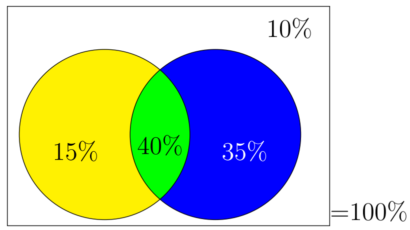
$$Pr\{C \cup B\} = Pr\{C\} + Pr\{B\} - Pr\{C \cap B\}$$

which becomes

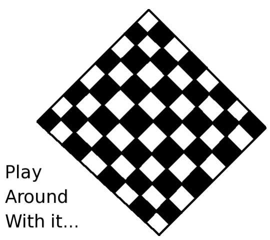
$$Pr\{C \cup B\} = 0.75 + 0.55 - 0.40 = 0.90$$

Therefore, our final conclusion is that a random executive from the survey can take the medicine with probability 10%, and cannot take it with probability 90%. The medical company should probably advertise in magazines of interest to different demographics, such as teenagers and athletes, and not to executives.

Here is a Venn Diagram that represents the previous example.



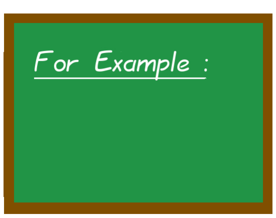
As you can see, the cholesterol-medicine takers are on the right, and the blood-pressure medicine takers are on the left.



7-6-5

Continuing with the previous example, suppose that a jogging magazine is contacted. They do a survey of their subscribers, and find out that 7% are taking blood-pressure medicine, 29% are taking cholesterol medicine, yet only 2% are taking both.

- What is the probability that a random survey-member can take the new drug? [Answer: 66%.]
- What is the probability that a random survey-member cannot take the new drug? [Answer: 34%.]



7-6-6

Let's suppose that you have an internship with a major worldwide airline. Further suppose that on this particular airline, you can achieve "gold frequent flyer status" either by flying six transatlantic routes in a year, or by flying 50,000 miles or more in a year. Your boss wants to know what percentage of customers achieved both of those requirements last year, but there's no command in the database to find this quickly. A query into the mileage tracking system of the airline reveals that 9% of customers flew more than 50,000 miles last year; furthermore, 5% of customers flew six or more times on a transatlantic flight. Moreover, 11% of the customers have gold status. What is the answer to your boss's question?

Let M be the probability that a customer flew more than 50,000 miles last year, and let S be the probability that a customer flew six or more transatlantic flights last year. Our data provides us with the following facts

$$Pr\{M\} = 0.09 \qquad Pr\{S\} = 0.05 \qquad Pr\{M \cup S\} = 0.11$$

which will enable us to solve the problem in the next box.

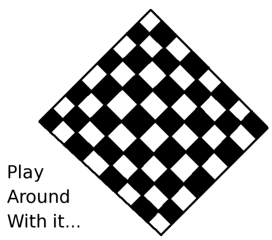
Continuing with the previous box, the inclusion-exclusion principle tells us that

$$Pr\{M \cup S\} = Pr\{M\} + Pr\{S\} - Pr\{M \cap S\}$$

which, after plugging in our data, becomes

$$\begin{aligned} 0.11 &= 0.09 + 0.05 - Pr\{M \cap S\} \\ 0.11 - 0.09 - 0.05 &= -Pr\{M \cap S\} \\ -0.03 &= -Pr\{M \cap S\} \\ 0.03 &= Pr\{M \cap S\} \end{aligned}$$

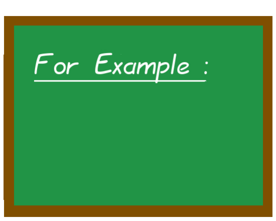
Therefore, you can tell your boss that 3% of customers achieved both criteria for gold frequent flyer status.



7-6-7

Let's suppose that you are the general manager at a large apartment complex. You have one of those terrible interns working for you: namely the type who does *almost exactly* what they are told, instead of *exactly* what they are told. Let's imagine that the health department wants to know what percentage of apartments in your complex have cats, what percentage have dogs, and what percentage have both.

First, you instruct your intern to perform an email survey of your renters, and he comes back a week later to say, "29% of residents have a cat or a dog." Of course, this fact is neither of the three facts that he was assigned to discover. You tell him to run another survey, and he comes back a week later to say, "21% of residents have a cat, and 10% of residents have a dog." Okay, now we're getting someplace—you have two of the three facts requested by the health department. The issue is that you don't want to survey your residents a third time, because it will annoy them. Find out what percentage of residents have both a cat and a dog from the given data. [Answer: 2%.]



7-6-8

Let's suppose that you are reading a newspaper article about drug and alcohol consumption on your college campus. The article mentions that after a large survey, it was found that 81% of students consumed alcohol during the past semester, and 19% used marijuana during the past semester, but 12% had used neither. Just out of curiosity, you'd like to know what percentage of students had consumed both during the past semester.

Let M be the probability that a student used marijuana in the previous semester, and let A be the probability that a student drank alcohol in the previous semester. Our data provides us with the following facts,

$$Pr\{M\} = 0.19 \qquad Pr\{A\} = 0.81 \qquad Pr\{A^c \cap M^c\} = 0.12$$

Recall that DeMorgan's Law (see Page 862) tells us that

$$A^c \cap M^c = (A \cup M)^c$$

which means that

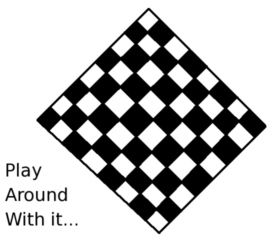
$$Pr\{A \cup M\} = 1 - 0.12 = 0.88$$

The inclusion-exclusion principle is now ready, and we will use it in the next box.

Continuing from the previous box, because we have 3 out of the 4 probabilities involved, we can now use the inclusion-exclusion principle to find the unknown probability.

$$\begin{aligned} Pr\{A \cup M\} &= Pr\{A\} + Pr\{M\} - Pr\{A \cap M\} \\ 0.88 &= 0.81 + 0.19 - Pr\{A \cap M\} \\ 0.88 - 0.81 - 0.19 &= -Pr\{A \cap M\} \\ -0.12 &= -Pr\{A \cap M\} \\ 0.12 &= Pr\{A \cap M\} \end{aligned}$$

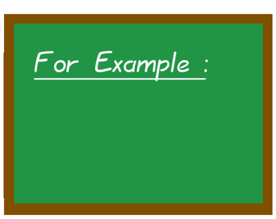
Thus only 12% of students consumed both marijuana and alcohol during the last semester.



7-6-9

Let's consider a factory that makes smartphones. Naturally it is important to test each smartphone as it is assembled for functionality. Suppose that 95.1% of the phones pass inspection for both the screen and the antenna. It turns out that 4.6% of phones have a defective screen and 1.9% of phones have a defective antenna.

What is the probability that a random smartphone has both defects? [Answer: 1.60%.]



7-6-10

Let's imagine a class on mathematical economics, which gives two midterm exams. An absent-minded professor recalls that last semester, 40% of the students got a "B-" or higher on the second midterm; however, he cannot find his records for the first midterm! Then, he remembers two useful facts about the class. Students who got a "B-" or higher on both midterms were given a free subscription to *The Economist*. Students who did not achieve a "B-" or higher on either exam were given a list of links to help them study for the final. He has these lists, and therefore knows that 22% of the students received the *Economist* subscription, and 45% of the students received the list of links to help them study for the final. Can we figure out how many students received a "B-" or higher on the first exam? We certainly can!

Let F be the probability that a student got a "B-" or higher on the first exam, and S be the probability that a student got a "B-" or higher on the second exam. Our data tells us that

$$Pr\{S\} = 0.40 \qquad Pr\{F \cap S\} = 0.22 \qquad Pr\{F^c \cap S^c\} = 0.45$$

Furthermore, DeMorgan's Law tells us that

$$F^c \cap S^c = (F \cup S)^c$$

and from the complement principle we know that

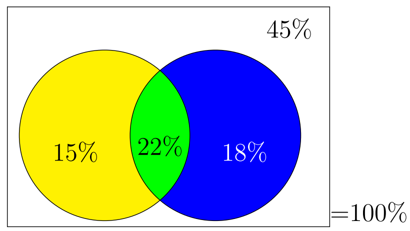
$$Pr\{F \cup S\} = 1 - Pr\{(F \cup S)^c\} = 1 - Pr\{F^c \cap S^c\} = 1 - 0.45 = 0.55$$

Finally, we can use the inclusion-exclusion formula, which we will do in the next box.

Continuing with the previous box, we can now use the inclusion-exclusion principle.

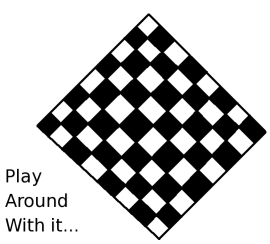
$$\begin{aligned} Pr\{F \cup S\} &= Pr\{F\} + Pr\{S\} - Pr\{F \cap S\} \\ 0.55 &= Pr\{F\} + 0.40 - 0.22 \\ 0.55 - 0.40 + 0.22 &= Pr\{F\} \\ 0.37 &= Pr\{F\} \end{aligned}$$

The previous example can be summarized by the following Venn Diagram:



where the left circle represents those who passed the first midterm exam, and the right circle represents those who passed the second midterm exam.

It is very fair to point out that some students will find it easier to just approach the problem using only Venn-Diagrams, and not bother with DeMorgan's Law nor the inclusion-exclusion principle. I think that's perfectly fine as an alternative method to solving this problem.



7-6-11

Let's suppose that your company's HR office does background checks on new employees, as most companies do. There are two criteria that they check: no candidate should have a felony conviction, and no candidate should have a low FICO credit rating, as this indicates financial irresponsibility. If a candidate has both of these faults, he or she is thrown out of the applicant pool—but if a candidate has only one fault, then they are forwarded to a panel which will decide their fate.

The CEO is curious about how many potential employees are turned away for having a low FICO score. He asks the HR director for this statistic, but citing privacy concerns, the HR director refuses to answer. However, the HR director will share that of the job candidates who successfully completed their interviews, and who were forwarded to HR for screening, 84% were hired, 10% were forwarded to the panel, and 6% were thrown out.

The COO then obtains a list of all of the job candidates who successfully completed their interviews, and has her interns type their names, one at a time, into an online database of criminal records. At the end of this laborious process, it is discovered that 9% of the job candidates who successfully completed their interviews had felony convictions.

Based on all of this information, compute for me the probability that a job candidate, who successfully completed his or her interviews, has a low FICO rating.

The answer will be given on Page 987, along with a full solution, because this question is fairly hard.

Play
Around
With it...

7-6-12

At many high schools in the USA, advanced classes in many subjects, including math, have the "AP" designation, which means "advanced placement."

Suppose that at a certain high school, both AP Statistics and AP Calculus are offered. While 5% of the seniors take both, 40% of seniors take AP Statistics and 30% of seniors take AP Calculus. A particular scholarship is available, but only seniors who are taking an AP math course are eligible to apply.

- What is the probability that a random high school senior is ineligible to apply? [Answer: 35%.]
- What is the probability that a random high school senior is eligible to apply? [Answer: 65%.]

Play
Around
With it...

7-6-13

Consider a university with a large computer engineering major. Suppose that among computer engineering majors there are 127 students who have taken a course in the `Java` programming language, and there are 289 students who have taken a course in the `Python` programming language. Furthermore, suppose that students who take either course are added to the "programming mailing list," and that list now has 358 students on it. Moreover, there are 159 students in the major who have taken neither programming course yet, presumably because they are just starting out.

What is the probability that a student (selected at random from the major) has taken both courses? [Answer: $11.2185\cdots\%$.]

Play
Around
With it...

7-6-14

Imagine a language-translation company which brags that 79% of their employees speak Spanish, while 65% of their employees speak French. Someone mentions to you that 19% of the employees know neither Spanish nor French, but focus on the Asian languages instead.

What percentage of employees speak both Spanish and French? [Answer: 63%.]

Play
Around
With it...

7-6-15

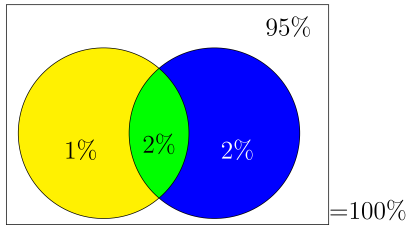
Let's imagine that you have an internship doing quality control in a factory that makes inkjet printers. The devices are tested for two known flaws: the paper-feeding mechanism and the ink-spraying mechanism. The way this factory's quality-control system works is that if one of the mechanisms fails the test, then the device gets refurbished. However, if both mechanisms fail their tests, then the printer is declared defective and is thrown out. The vast majority of printers pass both tests, are declared functional, and get shipped out to the warehouse.

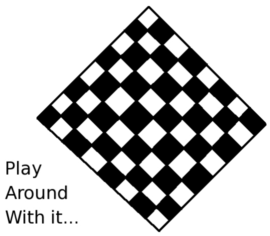
A new manufacturing process is proposed which will drastically cut costs, but the failure rate will increase slightly. The percentage of printers that have the paper-feeding mechanism flaw will rise to 4%, and the percentage of printers that have the ink-spraying mechanism flaw will rise to 3%. On the other hand, 95% of printers will be declared functional.

Your boss wants to know,

- What percentage of printers will be thrown out, in the new scheme? [Answer: 2%.]
- What percentage of printers will be refurbished, in the new scheme? [Answer: 3%.]

The situation in the previous box can be summarized with the following Venn Diagram. The right circle represents printers that failed the paper-feeding mechanism test, and the left circle represents printers that failed the ink-spraying mechanism test.

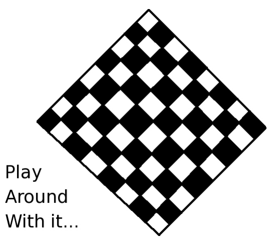




7-6-16

Let's imagine that you are working in a facility in Brazil for Doctors Without Borders, volunteering to help. It turns out that 34% of patients have Dengue Fever and 19% of patients have Malaria, but 54% of patients have neither.

What percentage of patients have both ailments? [Answer: 7%.]

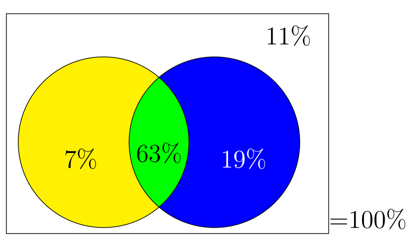


7-6-17

A salesperson sells machines that make brake pads, and there are only two automobile factories that would be interested in purchasing brake-pad machines at this time of year. One is in Flint, Michigan, and the other is in Detroit, Michigan. The salesperson believes that she will get a sale in Flint with probability 70%, and in Detroit with probability 82%. However, she also believes that the probability of not getting a sale at either factory is only 11%. With this in mind, please answer the following questions:

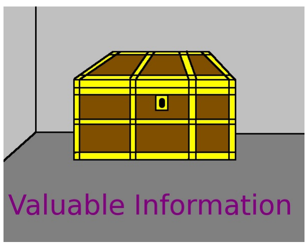
- What is the probability of getting a sale at both factories? [Answer: 63%.]
- What is the probability of getting a sale at either factory? (In other words, the probability of at least one sale?) [Answer: 89%.]
- What is the probability of getting a sale at exactly one factory? [Answer: 26%.]

For some students, a Venn Diagram is a great way to check your work. Other students prefer to solve the entire problem with only Venn Diagrams. Either way, this Venn Diagram summarizes the previous box. The left circle represents Flint, and the right circle represents Detroit.



We've now seen that we cannot just add two probabilities together to make a union of two events whenever we want to. Yet clearly there are some circumstances where it works, such as in the previous module, where we added separate outcomes (simple events) to make compound events.

Now we're going to explore precisely and exactly when we have the right to add the probabilities of two events. It turns out that we can add the probabilities whenever the events are *mutually exclusive*.



Sometimes two sets do not overlap at all, and have nothing in common. That means their intersection is the empty set. In this case, we call these sets *disjoint*. Formally,

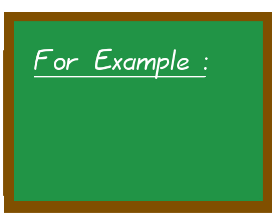
$$\text{A and B are disjoint sets if and only if } A \cap B = \{\}$$

which means that disjoint sets have the empty set as their intersection. (We actually defined the word *disjoint* back on Page 813.)

We have a different phrase for events that do not share any outcomes. If two events do not overlap, then we say that they are *mutually exclusive*, just as we did on Page 927, when we required mutual exclusivity as a criterion for being a sample space. Of course, if A and B are mutually exclusive then

$$Pr\{A \cap B\} = Pr\{\ \{\}\ \} = 0$$

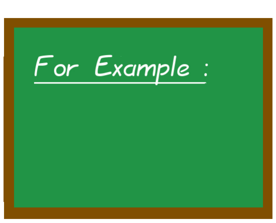
This is easier seen by examples.



\# 7-6-18

Here is an example and a non-example of mutually exclusive events:

- If you survey osteopaths about their ages, then because no doctor can be age 34 and age 36 simultaneously, the events "this osteopath is 34" and "this osteopath is 36" are mutually exclusive.
- However, the events "this osteopath is over 34" and "this osteopath is over 36" are not mutually exclusive. For example, a 40-year old osteopath would be in the intersection of both events.



\# 7-6-19

Let's suppose that a committee has to be formed at an accounting firm to audit something particularly boring. The committee consists of Alice, Bob, Charles, Diane, Edward, and Frank. They are going to choose the committee chairperson by putting their names (on slips of paper) into a hat, and drawing one slip out.

- The events "Bob is chosen" and "a female is chosen" are mutually exclusive. That is because Bob is not female.
- The events "Alice is chosen" and "a female is chosen" are not mutually exclusive. It is possible that both could happen, namely if Alice is chosen.

So what happens if two events are mutually exclusive?

$$\begin{aligned} Pr\{X \cup Y\} &= Pr\{X\} + Pr\{Y\} - Pr\{X \cap Y\} \\ Pr\{X \cup Y\} &= Pr\{X\} + Pr\{Y\} - 0 \qquad \leftarrow \text{only if } X \text{ and } Y \text{ are mutually exclusive.} \\ Pr\{X \cup Y\} &= Pr\{X\} + Pr\{Y\} \qquad \leftarrow \text{only if } X \text{ and } Y \text{ are mutually exclusive.} \end{aligned}$$

Therefore, we can say, for any two mutually exclusive events, the probability of their union is the sum of their probabilities.

How can we be certain, for sure, that there are not some other situations out there, no matter how obscure, where just adding the probabilities actually gives the correct answer?

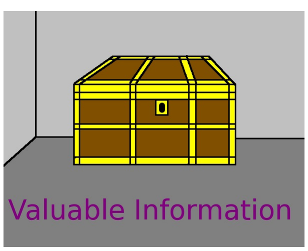
Suppose that for some events A and B, we know

$$Pr\{A\} = p_1 \qquad Pr\{B\} = p_2 \qquad Pr\{A \cup B\} = p_1 + p_2$$

Let's use the inclusion-exclusion formula and see what happens!

$$\begin{aligned} Pr\{A \cup B\} &= Pr\{A\} + Pr\{B\} - Pr\{A \cap B\} \\ p_1 + p_2 &= Pr\{A\} + Pr\{B\} - Pr\{A \cap B\} \\ p_1 + p_2 &= p_1 + Pr\{B\} - Pr\{A \cap B\} \\ p_1 + p_2 &= p_1 + p_2 - Pr\{A \cap B\} \\ p_1 + p_2 - p_1 - p_2 &= -Pr\{A \cap B\} \\ 0 &= Pr\{A \cap B\} \end{aligned}$$

As you can see, A and B are mutually exclusive because the probability of their intersection is zero. We've proven that if adding two probabilities together gives you the probability of their union then the events are mutually exclusive.

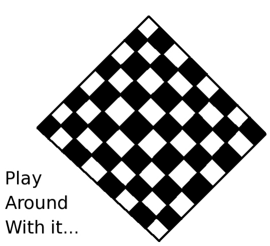


For any two events A and B, from the same sample space, the following formula holds true.

$$Pr\{A \cup B\} = Pr\{A\} + Pr\{B\} - Pr\{A \cap B\}$$

However, if A and B are mutually exclusive, then we can use a simpler formula:

$$Pr\{A \cup B\} = Pr\{A\} + Pr\{B\} \qquad \leftarrow \text{only if } A \text{ and } B \text{ are mutually exclusive.}$$



7-6-20

Abstract concepts, such as mutually exclusive events, are often much easier to understand by looking at concrete examples. Let's return to the committee of six accountants, Alice, Bob, Charles, Diane, Edward, and Frank, from earlier. We're going to define the following events:

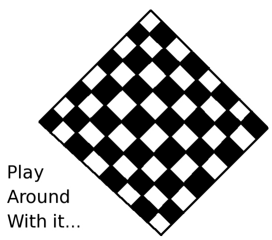
- The event M means that the chairperson is male.
- The event F means that the chairperson is female.
- The event C means that the chairperson's name begins with a consonant.
- The event V means that the chairperson's name begins with a vowel.

This question continues into the next box.

Continuing with the previous box, just imagine that you're studying with a classmate. First, try to figure out whether each of the pairs of events in the following box is mutually exclusive or not. Second, if the answer is "mutually exclusive," produce proof of this fact by naming someone in the intersection of the events.

- Are M and F mutually exclusive?
- Are M and C mutually exclusive?
- Are M and V mutually exclusive?
- Are F and C mutually exclusive?
- Are F and V mutually exclusive?
- Are C and V mutually exclusive?

The answers to these questions can be found on Page 988.



7-6-21

Before continuing from the problem of the previous box, please check your work on Page 988.

Now that you've checked your work, let's compute the following probabilities:

- What is $Pr\{M\}$? [Answer: 2/3.]
- What is $Pr\{V\}$? [Answer: 1/3.]
- What is $Pr\{M \cap V\}$? [Answer: 1/6.]
- What is $Pr\{M \cup V\}$? [Answer: 5/6.]

Let's take a closer look at the problem of the last two boxes. Observe that the formula

$$Pr\{M \cup V\} = Pr\{M\} + Pr\{V\} - Pr\{M \cap V\}$$

will, after plugging in the data from the previous box, give us the statement

$$5/6 = 2/3 + 1/3 - 1/6$$

which is clearly true. Contrastingly, the formula

$$Pr\{M \cup V\} = Pr\{M\} + Pr\{V\} \qquad \leftarrow \text{Unauthorized !!!}$$

will, after plugging in the data from the previous box, give us the statement

$$5/6 = 2/3 + 1/3 \qquad \leftarrow \text{Very False !!!}$$

which is clearly false. That's because the first formula works all the time, but the second formula only works for mutually exclusive events. Since M and V are not mutually exclusive, the formula fails, and we get an incorrect equation.

In summary: do not use the shortened formula unless you know that the events are indeed mutually exclusive. Otherwise, you risk getting a wrong answer.

Play
Around
With it...

7-6-22

Let's imagine a smartphone factory. The phones are tested after manufacture for flaws in their speaker, their antenna, and their screen. After the phones with multiple flaws have been thrown out, it is found that 1% have a flaw with the speaker, 2% have a flaw with the antenna, and 4% have a flaw with the screen.

After the phones with multiple flaws have been thrown out, the boss grabs one phone to demonstrate it to a group of investors who are touring the factory. They are mostly older folks, so their idea of testing a smartphone is to make an ordinary phone call from it. Accordingly, they will decide that the phone isn't working only if either the antenna or the speaker is flawed.

What is the probability that they decide the phone works? [Answer: 97%.]

Play
Around
With it...

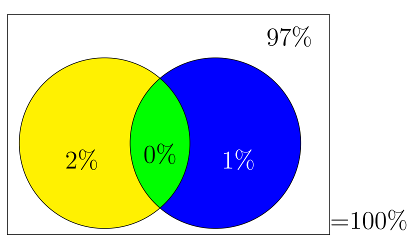
7-6-23

Let's now imagine that you are interning for a company that imports electronics. You come across the quality-control manual of an important supplier. The manual is written in Japanese—for the sake of this problem, we will imagine that you cannot read Japanese, but please accept my apologies if you can. However, the cover letter on the manual is written in English and partway through it says that 97% of keyboards tested last month were tagged green (meaning fully functional), while 2% of keyboards were tagged orange (indicating some flaw, but you don't know what yet), and 1% of keyboards were tagged red (indicating some other flaw, which you also do not know).

- Using this information, can we determine whether the red tag flaw and the orange tag flaw are mutually exclusive? [Answer: Yes, we can.]
- Are they mutually exclusive? [Answer: Yes, they are.]

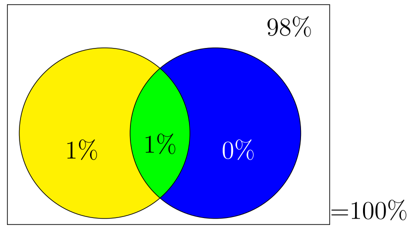
Check Your Work !!

One way to visualize what is happening in the previous box is to look at the Venn Diagram:



As you can see, the orange-tagged keyboards are in the left circle, and the red-tagged keyboards are in the right circle. The football-shaped region, indicating the overlap, has probability 0%. Therefore, the two flaws are indeed mutually exclusive.

Interestingly, if we were to change the problem of the previous checkerboard to make the 97% into 98%, then we would have the following Venn Diagram:



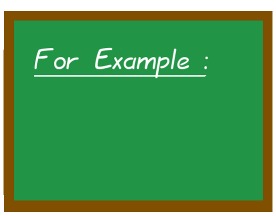
Again, the orange-tagged keyboards are in the left circle, and the red-tagged keyboards are in the right circle. The football-shaped region, indicating the overlap, has probability 1%. Therefore, in this modified problem, the two flaws are not mutually exclusive. Interestingly, in this case, the red tag never occurs alone.

It is interesting how sensitive probability problems, combinatorics problems, and other problems in this chapter, are to changes in their initial conditions. We saw that in the previous box, and we'll explore this further on Page 1003 and Page 1019.

On Page 936, we stated that probability distributions come from four possible sources:

1. The "uniform distribution assumption" or "equally likely assumption."
2. A survey or experiment.
3. Using calculations to make one distribution from another distribution.
4. Making arbitrary assumptions based on experience.

We're now going to explore the third case in more detail—it is straight forward.



7-6-24

Back when we were looking at the malfunctioning printers on Page 974 of this module, we saw the following probability distribution, given below on the left. Recall that printers with no flaws were to be shipped to the warehouse for sale, printers with both flaws were to be thrown out, and printers with just one flaw were to be refurbished. We can combine the two middle cases to get the probability distribution on the right.

No flaws = 95%	→	Ship for Sale = 95%
Only an ink-spraying flaw = 1% Only a paper-feeding flaw = 2%	} →	Refurbish = 3%
Both flaws = 2%	→	Thrown Out = 2%

As you can see in the previous box, each outcome (simple event) in the new sample space consists of a subset of the outcomes of the original sample space. We must take care that no old outcome ends up as part of two (or more) new outcomes.

Instead, the new outcomes must be disjoint subsets of the set of old outcomes. That will guarantee that the new sample space will be mutually exclusive.

The easiest way to understand this is to think of "gluing" outcomes in the old sample space together, to form outcomes in the new sample space.

For example, in the previous box, we glued together the event "only an ink-spraying flaw" and the event "only a paper-feeding flaw."

For Example :

\# 7-6-25

On Page 949, we saw the data for the age groups of osteopaths back in 2001. After some calculation, we got the following probability distribution, found below. A manufacturer of recreational boats wants to target the "under 45" age group. That company will make the assumption that the age distribution of osteopaths attending the national osteopathic convention will be the same as the age distribution of osteopaths in the USA overall—this is a common and safe assumption. We want to collapse this probability distribution into only two outcomes: "inside target age group" and "outside target age group."

- Under age 35 $= 9866/46,962 = 21.0084\cdots\%$.
- Age 35–44 $= 14,798/46,962 = 31.5105\cdots\%$.
- Age 45–54 $= 12,754/46,962 = 27.1581\cdots\%$.
- Age 55–64 $= 4706/46,962 = 10.0208\cdots\%$.
- Age 65 and over $= 4838/46,962 = 10.3019\cdots\%$.

We will solve this example in the next box.

Continuing from the previous box, all we have to do is add the relevant probabilities:

- Inside Target Age Group:
$$21.0084\% + 31.5105\% = 52.5190\%$$
- Outside Target Age Group:
$$27.1581\% + 10.0208\% + 10.3019\% = 47.4809\%$$

Since the majority of osteopaths are in the target age group, this is a good convention for the company to visit, with the hope of selling recreational boats and yachts.

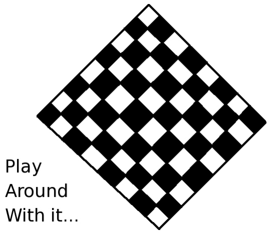
For Example :

\# 7-6-26

Recall that on the Roulette wheel, we have 38 spots. Because the wheel spins quickly and the little white ball jumps around, all 38 spots are equally likely. There are 18 red spots, 18 black spots, and 2 green spots on most wheels. Since there are 38 spots, each spot has probability 1/38 of being the spot where the little white ball will land.

In this case, the old probability distribution is 38 spots, each with probability 1/38. We will glue together the 18 red spots, for a probability of $18/38 = 9/19$. Similarly, we will glue together the 18 black spots, for a probability of $18/38 = 9/19$. Finally, we will glue together the 2 green spots, for a probability of $2/38 = 1/19$. Our final distribution, as we saw on Page 1077, is the following:

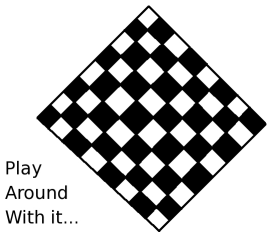
- Red: $18/38 = 9/19$
- Black: $18/38 = 9/19$
- Green: $2/38 = 1/19$



7-6-27

We return to our earlier problem (from Page 979 of this module) about tagging keyboards with orange and red tags, for quality control. We did this problem twice, once with 97% of keyboards receiving zero tags, and once with 98% of keyboards receiving zero tags. In each of those cases, I'd like you to construct a probability distribution for keyboards receiving 0 tags, 1 tag, or 2 tags.

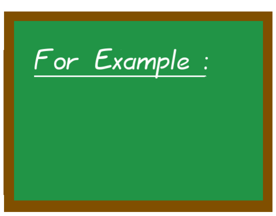
- For the 97% case?
  [Answer: 0 tags = 97%, 1 tag = 3%, 2 tags = 0%.]
- For the 98% case?
  [Answer: 0 tags = 98%, 1 tag = 1%, 2 tags = 1%.]



7-6-28

Let's re-examine the probability distribution of osteopaths, but this time from the year 2010. You can find the distribution on Page 965 of the previous module. We are still interested in knowing the probability distribution of osteopaths being inside or outside the target market group of "under age 45."

[Answer: $51.9423\cdots\%$ are inside the target age group, and $48.0576\cdots\%$ are outside the target age group.]



7-6-29

Let's say that someone is analyzing a situation involving two events, E and F. They believe the following three statements:

$$Pr\{E\} = 0.4 \qquad Pr\{F\} = 0.3 \qquad Pr\{E \cup F\} = 0.8$$

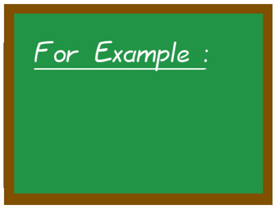
Believe it or not, we already have enough information to conclude that the above is absolutely impossible, and therefore one (or more) of the three statements must be false.

I will show why this data is impossible in the next box.

In the previous box, I made the bold claim that the given data was impossible. One way to see this is to check the inclusion-exclusion formula.

$$\begin{aligned} Pr\{E \cup F\} &= Pr\{E\} + Pr\{F\} - Pr\{E \cap F\} \\ 0.8 &= 0.4 + 0.3 - Pr\{E \cap F\} \\ 0.4 &= 0.3 - Pr\{E \cap F\} \\ 0.1 &= -Pr\{E \cap F\} \\ -0.1 &= Pr\{E \cap F\} \end{aligned}$$

Clearly, no probability can ever be negative, so we know that $Pr\{E \cap F\} = -0.1$ is impossible. Thus the given data is impossible as well.



7-6-30

Here is another impossible situation. Imagine someone is analyzing a situation involving two events, X and Y. They believe the following three statements:

$$Pr\{X\} = 0.7 \qquad Pr\{Y\} = 0.6 \qquad Pr\{X \cap Y\} = 0.2$$

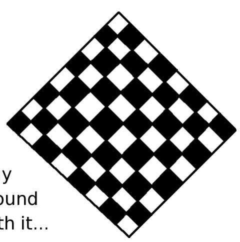
As before, we already have enough information to conclude that the above is impossible, and therefore one (or more) of the three statements must be false. We will see this using the inclusion-exclusion formula.

$$\begin{aligned} Pr\{X \cup Y\} &= Pr\{X\} + Pr\{Y\} - Pr\{X \cap Y\} \\ Pr\{X \cup Y\} &= 0.7 + 0.6 - 0.2 \\ Pr\{X \cup Y\} &= 1.1 \end{aligned}$$

Clearly, no probability can ever be more than 1, so we know that $Pr\{X \cup Y\} = 1.1$ is impossible. One (or more) of the given statements must be false.

My usual strategy for the above kind of problem is to use the inclusion-exclusion formula, as I have shown you. However, some students like to make a two-circle Venn Diagram and see what happens. If you have a region with a negative probability, or a probability above one, then the given data represents an impossible situation.

Whether you like to use Venn Diagrams or algebra, try the following situations out and determine for each whether it is possible or not.



Play
Around
With it...

7-6-31

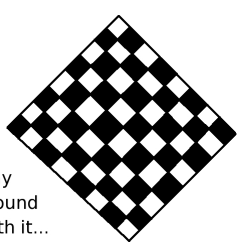
Consider the following data, and for each situation tell me whether it is possible or impossible.

(1) $Pr\{K\} = 0.3, Pr\{L\} = 0.4, Pr\{K \cup L\} = 0.7$

(2) $Pr\{M\} = 0.6, Pr\{N\} = 0.5, Pr\{M \cap N\} = 0.1$

(3) $Pr\{A\} = 0.2, Pr\{B\} = 0.4, Pr\{A \cup B\} = 0.7$

The answers are on Page 989.



Play
Around
With it...

7-6-32

Consider the following data, and for each situation tell me whether it is possible or impossible.

(4) $Pr\{C\} = 0.5, Pr\{D\} = 0.7, Pr\{C \cap D\} = 0.1$

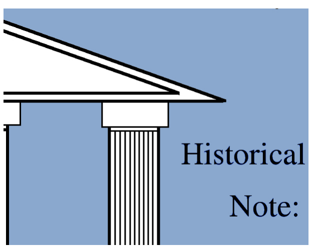
(5) $Pr\{E\} = 0.8, Pr\{E \cup F\} = 0.9, Pr\{E \cap F\} = 0.2$

(6) $Pr\{G\} = 0.8, Pr\{G \cap H\} = 0.9, Pr\{G \cup H\} = 0.2$

The answers are on Page 989.

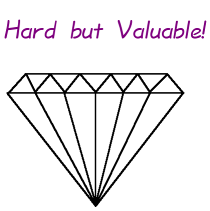
The inclusion-exclusion principle has a disputed history. It appears in an 1854-paper by Daniel da Silva (1814–1878), and also in an 1883-paper by James Joseph Sylvester (1814–1897). Yet the idea is so basic it seems hard to believe that it wasn't known by 1700. Some even attribute it to Abraham de Moivre (1667–1754).

James Joseph Sylvester had a fascinating career. He studied under Augustus DeMorgan, who invented DeMorgan's laws. Sylvester studied mathematics at the University of London and at Cambridge. In 1837, at *the tripos*, the annual contest of the senior undergraduates in mathematics at Cambridge, Sylvester ranked as "junior wrangler"—second place. Despite this accomplishment, he was not awarded a degree and could not compete for a fellowship for his graduate studies, because he was Jewish. At the time, degrees from Oxford and Cambridge were restricted only to Anglicans. (Cambridge did not give Sylvester his degree until thirty-five years later, in 1872.)



As a result of the difficulties mentioned in the previous box, Sylvester moved between the USA and the UK several times. He helped found the mathematics department at Johns Hopkins University, and had a variety of accomplishments. He influenced several branches of mathematics and even wrote a book about poetry.

I've even used his work in my own research, using something called a "Sylvester matrix." It can be used to find simultaneous solutions to polynomial systems of equations, which is an otherwise very difficult task. Polynomial systems of equations come up in a variety of areas, such as molecular chemistry, robotics, and cryptography.



This module is now essentially complete, but there are three additional uses of what we have learned that I would like to share. As it comes to pass, two very basic tools of the previous module can be rigorously proven correct with what we've learned in this module. That can be very important, because mathematics is not a subject where we should accept things on faith. We should want to see reasons to believe that something is true. Moreover, these are extremely compact and short proofs, easier to swallow than most. Moreover, we can complete our discussion of why the probabilities in a probability distribution always add up to one.

However, there's not really a way to put questions about this sort of thing on a quiz, test, or examination. Therefore, I'm going to guess that most instructors will skip the following material. Nonetheless, I would ask you to read it anyway, so that you can proceed through the rest of the chapter with greater confidence. I want you to be reassured that the theory of probability rests on extremely solid ground.

Believe it or not, we now have enough information about the laws of probability to prove "the complement principle of probability." Namely that

$$Pr\{A^c\} = 1 - Pr\{A\} \qquad \text{and similarly} \qquad Pr\{A\} = 1 - Pr\{A^c\}$$

Let's take any sample space whatsoever, and consider any event A whatsoever. We begin with some simple observations:

1. Clearly, A and A^c are mutually exclusive events. Either A happens or A^c happens. That's what being a complement means.
2. For that reason, we can be sure that $Pr\{A \cap A^c\} = 0$.
3. Likewise, $A \cup A^c$ is going to include everything possible. For that reason, we can be sure that $Pr\{A \cup A^c\} = 1$.

Now we can prove the principle in the next box, using the include-exclusion formula

We can start with the principle, plug in our information, and solve for $Pr\{A\}$. We'll write the general formula in terms of X and Y, where $X = A$ and $Y = A^c$.

$$\begin{aligned} Pr\{X \cup Y\} &= Pr\{X\} + Pr\{Y\} - Pr\{X \cap Y\} \\ Pr\{A \cup A^c\} &= Pr\{A\} + Pr\{A^c\} - Pr\{A \cap A^c\} \\ 1 &= Pr\{A\} + Pr\{A^c\} - 0 \\ 1 &= Pr\{A\} + Pr\{A^c\} \\ 1 - Pr\{A\} &= Pr\{A^c\} \end{aligned}$$

That's one of the two pieces that we needed. From the second-to-last line we could have subtracted $Pr\{A^c\}$ from both sides to instead get

$$1 - Pr\{A^c\} = Pr\{A\}$$

The way that we computed the probabilities of compound events in the previous module was by adding up the probabilities of the outcomes (simple events) that comprised them. It turns out that we do not have to accept this technique on faith. We can prove it true, rigorously. Because this textbook is not about proof, we will just consider the case of three outcomes A, B, and C, which can be any outcomes whatsoever in any sample space whatsoever, that has three or more outcomes in it. To accomplish this, we're going to make the three-variable version of the "inclusion-exclusion formula for probability."

First, let's recall the three-variable version of the "inclusion-exclusion formula for sets." We saw this on Page 907. The formula is

$$\#(A \cup B \cup C) = \#A + \#B + \#C - \#(A \cap B) - \#(B \cap C) - \#(A \cap C) + \#(A \cap B \cap C)$$

and as you might guess, for probability this becomes

$$Pr\{A \cup B \cup C\} = Pr\{A\} + Pr\{B\} + Pr\{C\} - Pr\{A \cap B\} - Pr\{B \cap C\} - Pr\{A \cap C\} + Pr\{A \cap B \cap C\}$$

by dividing the equation by the number of objects in the universal set of the problem.

We will continue in the next box.

Continuing from the previous box, because A, B, and C are outcomes in a sample space, they are mutually exclusive. That means the last four probabilities in the previous equation, containing the $\cap$s, must be zero. We have

$$Pr\{A \cup B \cup C\} = Pr\{A\} + Pr\{B\} + Pr\{C\} - 0 - 0 - 0 + 0$$

which clearly simplifies to

$$Pr\{A \cup B \cup C\} = Pr\{A\} + Pr\{B\} + Pr\{C\}$$

We have rigorously proven that for three outcomes, the probability of the compound event equal to their union is just the sum of the probabilities of the outcomes. As it turns out, this is true for two outcomes, four outcomes, five outcomes, six outcomes, or any other number of outcomes. That was the basis of how we computed the probabilities of compound events throughout the previous module. At least now we know that those computations were on solid ground.

By the way, the formula

$$Pr\{A \cup B \cup C\} = Pr\{A\} + Pr\{B\} + Pr\{C\} - Pr\{A \cap B\} - Pr\{B \cap C\} - Pr\{A \cap C\} + Pr\{A \cap B \cap C\}$$

could, in theory, be used in other contexts as a legitimate formula in probability. You can even write

$$\begin{aligned} Pr\{A \cup B \cup C \cup D\} \quad = \quad & Pr\{A\} + Pr\{B\} + Pr\{C\} + Pr\{D\} \\ & - Pr\{A \cap B\} - Pr\{B \cap C\} - Pr\{A \cap C\} - Pr\{A \cap D\} - Pr\{B \cap D\} - Pr\{C \cap D\} \\ & + Pr\{A \cap B \cap C\} + Pr\{A \cap B \cap D\} + Pr\{A \cap C \cap D\} + Pr\{B \cap C \cap D\} \\ & - Pr\{A \cap B \cap C \cap D\} \end{aligned}$$

However, I have yet to encounter a problem that needed these, except for extremely artificial examples. Therefore, we will not use these longer versions of the formula in this textbook.

On Page 957, we showed that for three of our four ways of making a probability distribution, the probabilities always add to one. Now we can complete that theory by showing it will work for the case that we skipped.

We had good arguments for the "equally likely assumption" case, the survey/experiment case, and the "assumptions based off experience" case. We are ready to show that the probabilities will always add to one, including in the case of gluing together a few outcomes from an old sample space to make a new one.

We will now attempt to show that when you glue outcomes together, adding their probabilities to make a new probability distribution out of an old one, that the probabilities will still add to one. First, we need a particular tool.

There's this law of algebra which you've probably heard of, but perhaps so long ago that you've forgotten it. When we have a list of numbers and we want to add them up, we can group them in whatever order we want. That's called the "Associative Law of Addition." Basically, it means that if you have a bunch of numbers being added together, you can put the parentheses wherever you like, and the total will come out the same.

This is much easier to see by example, which I will do in the next box.

Continuing with the previous box, I'd like to show you an example of the "Associative Law of Addition." Here's an addition where I've put parentheses in all sorts of arbitrary positions, and of course, the sum is unchanged.

$$\begin{aligned}
1/16 + 1/2 + 1/4 + 1/16 + 1/8 &= 1 \\
(1/16 + 1/2 + 1/4) + 1/16 + 1/8 &= 13/16 + 1/16 + 1/8 = 1 \\
1/16 + (1/2 + 1/4) + (1/16 + 1/8) &= 1/16 + 3/4 + 3/16 = 1 \\
1/16 + 1/2 + (1/4 + 1/16 + 1/8) &= 1/16 + 1/2 + 7/16 = 1 \\
1/16 + 1/2 + (1/4 + 1/16) + 1/8 &= 1/16 + 1/2 + 5/16 + 1/8 = 1 \\
1/16 + (1/2 + 1/4 + 1/16) + 1/8 &= 1
\end{aligned}$$

Moreover, you can even change the ordering however you might like. That's called the "Commutative Law of Addition." For example,

$$1/16 + 1/2 + 1/4 + 1/16 + 1/8 = 1/2 + 1/4 + 1/8 + 1/16 + 1/16 = 1$$

The previous box has told us that you can reorder and reorganize a bunch of additions putting the terms into any ordering and by putting parentheses wherever you might like.

Now let's consider how this affects a probability distribution. If we start with a probability distribution, we know that the probabilities will add to one. Then, if we reorder the simple events (the outcomes), it will still add to one, because of the commutative law of addition. Next, if we glue events together, adding their probabilities, that's like inserting some parentheses. The parentheses will group together the simple events from the old probability distribution, to make the simple events of the new probability distribution. Yet the sum never changes—it is still one.

Because of the associative law of addition, the total is still one, or 100%, no matter how we might choose to place the parentheses.

You have now completed this module. Here are the answers to some questions that appeared earlier.

Here is the answer to the question on Page 973, where the CEO wanted to know the probability that a job candidate who successfully completed his or her interviews had a low FICO rating.

Let C be the event that a candidate has a felony conviction, and let L be the event that a candidate has a low FICO rating. Let's summarize what we know:

- Because 6% of candidates were thrown out, we know that $Pr\{L \cap C\} = 0.06$.
- While 84% were hired without further fuss, we know that $Pr\{L^c \cap C^c\} = 0.84$, but we do not need this fact, strictly speaking.
- However, DeMorgan's Law (see Page 862) tells us that

  $$L^c \cap C^c = (L \cup C)^c$$

  which means that $Pr\{(L \cup C)^c\} = 0.84$ and therefore

  $$Pr\{L \cup C\} = 1 - 0.84 = 0.16$$

Continuing with the previous box,

- Alternatively, because 10% of candidates were forwarded to the panel, and 6% were thrown out directly,

$$Pr\{L \cup C\} = 0.10 + 0.06 = 0.16$$

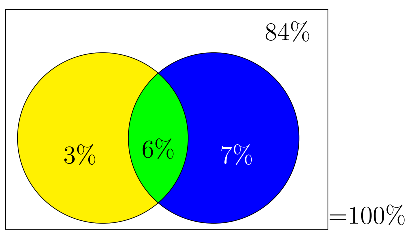
which is good for those students who do not remember DeMorgan's laws.

- Because 9% of candidates had a felony conviction, we know that $Pr\{C\} = 0.09$.

We are now ready to use the inclusion-exclusion principle:

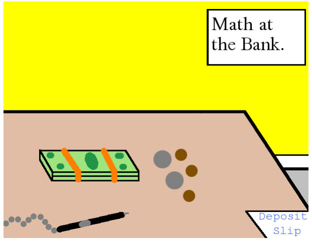
$$\begin{aligned} Pr\{L \cup C\} &= Pr\{L\} + Pr\{C\} - Pr\{L \cap C\} \\ 0.16 &= Pr\{L\} + 0.09 - 0.06 \\ 0.16 &= Pr\{L\} + 0.03 \\ 0.13 &= Pr\{L\} \end{aligned}$$

Therefore, we now know for sure that 13% of the job candidates who successfully completed their interviews had a low FICO score. We can summarize this with the following Venn Diagram, where low FICO score candidates are in the right circle, and convicted felons are in the left circle.



Here are the solutions to the questions about the mutual exclusivity of events, given on Page 977.

- Are M and F mutually exclusive? [Answer: Yes.]
- Are M and C mutually exclusive?
  [Answer: No. Your chosen counter-example can be Bob, Charles, or Frank.]
- Are M and V mutually exclusive?
  [Answer: No. Your chosen counter-example must be Edward.]
- Are F and C mutually exclusive?
  [Answer: No. Your chosen counter-example must be Diane.]
- Are F and V mutually exclusive?
  [Answer: No. Your chosen counter-example must be Alice.]
- Are C and V mutually exclusive? [Answer: Yes.]



By the way, the problem of the previous box does have a certain "20th century" feel to it, because it ignores the possibility that one or more of the committee members is transgendered. However, forward-looking businesses that welcome transgendered employees will allow their employees to select which gender, which pronouns, and whatever first name the employee wants.

We must hope that the hypothetical company of Alice, Bob, Charles, Diane, Edward, and Frank would be equally modern and accommodating.

Here are the answers to the two checkerboard boxes starting on Page 983, where we had to classify certain situations as either possible or impossible.

(1) Possible, because we get $Pr\{K \cap L\} = 0$, and there is nothing wrong with that.

(2) Possible, because we get $Pr\{M \cup N\} = 1$, and there is nothing wrong with that.

(3) Impossible, because we get $Pr\{A \cap B\} = -0.1$, and no probability can be negative.

Note: Another reason that (3) is impossible is that $Pr\{A \cup B\} > Pr\{A\} + Pr\{B\}$, which can never actually happen.

(4) Impossible, because we get $Pr\{C \cup D\} = 1.1$, but no probability can exceed 1.

(5) Possible, because we get $Pr\{F\} = 0.3$, and there is nothing wrong with that.

(6) Impossible, because $Pr\{G \cap H\} > Pr\{G \cup H\}$ cannot occur.