3 | PROBABILITY TOPICS

Figure 3.1 Meteor showers are rare, but the probability of them occurring can be calculated. (credit: Navicore/flickr)

Introduction

It is often necessary to "guess" about the outcome of an event in order to make a decision. Politicians study polls to guess their likelihood of winning an election. Teachers choose a particular course of study based on what they think students can comprehend. Doctors choose the treatments needed for various diseases based on their assessment of likely results. You may have visited a casino where people play games chosen because of the belief that the likelihood of winning is good. You may have chosen your course of study based on the probable availability of jobs.

You have, more than likely, used probability. In fact, you probably have an intuitive sense of probability. Probability deals with the chance of an event occurring. Whenever you weigh the odds of whether or not to do your homework or to study for an exam, you are using probability. In this chapter, you will learn how to solve probability problems using a systematic approach.

3.1 | Terminology

Probability is a measure that is associated with how certain we are of outcomes of a particular experiment or activity. An **experiment** is a planned operation carried out under controlled conditions. If the result is not predetermined, then the experiment is said to be a **chance** experiment. Flipping one fair coin twice is an example of an experiment.

A result of an experiment is called an **outcome**. The **sample space** of an experiment is the set of all possible outcomes. Three ways to represent a sample space are: to list the possible outcomes, to create a tree diagram, or to create a Venn diagram. The uppercase letter S is used to denote the sample space. For example, if you flip one fair coin, $S = \{H, T\}$ where H = heads and T = tails are the outcomes.

An **event** is any combination of outcomes. Upper case letters like A and B represent events. For example, if the experiment is to flip one fair coin, event A might be getting at most one head. The probability of an event A is written $P(A)$.

The **probability** of any outcome is the **long-term relative frequency** of that outcome. **Probabilities are between zero and one, inclusive** (that is, zero and one and all numbers between these values). $P(A) = 0$ means the event A can never happen. $P(A) = 1$ means the event A always happens. $P(A) = 0.5$ means the event A is equally likely to occur or not to occur. For example, if you flip one fair coin repeatedly (from 20 to 2,000 to 20,000 times) the relative frequency of heads approaches

---

Example 3.26

A litter of kittens available for adoption at the Humane Society has four tabby kittens and five black kittens. A family comes in and randomly selects two kittens (without replacement) for adoption.
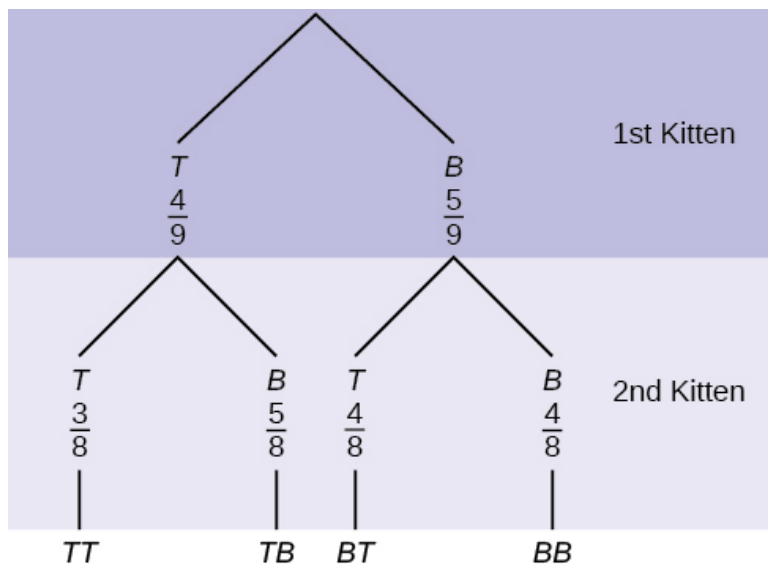
a. What is the probability that both kittens are tabby?

a. $\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)$ b. $\left(\frac{4}{9}\right)\left(\frac{4}{9}\right)$ c. $\left(\frac{4}{9}\right)\left(\frac{3}{8}\right)$ d. $\left(\frac{4}{9}\right)\left(\frac{5}{9}\right)$

b. What is the probability that one kitten of each coloring is selected?

a. $\left(\frac{4}{9}\right)\left(\frac{5}{9}\right)$ b. $\left(\frac{4}{9}\right)\left(\frac{5}{8}\right)$ c. $\left(\frac{4}{9}\right)\left(\frac{5}{9}\right)+\left(\frac{5}{9}\right)\left(\frac{4}{9}\right)$ d. $\left(\frac{4}{9}\right)\left(\frac{5}{8}\right)+\left(\frac{5}{9}\right)\left(\frac{4}{8}\right)$

c. What is the probability that a tabby is chosen as the second kitten when a black kitten was chosen as the first?

d. What is the probability of choosing two kittens of the same color?

Solution 3.26

a. c, b. d, c. $\frac{4}{8}$, d. $\frac{32}{72}$

Try It Σ

3.26 Suppose there are four red balls and three yellow balls in a box. Two balls are drawn from the box without replacement. What is the probability that one ball of each coloring is selected?

3.5 | Venn Diagrams
Venn Diagrams

A **Venn diagram** is a picture that represents the outcomes of an experiment. It generally consists of a box that represents the sample space S together with circles or ovals. The circles or ovals represent events. Venn diagrams also help us to convert common English words into mathematical terms that help add precision.

Venn diagrams are named for their inventor, John Venn, a mathematics professor at Cambridge and an Anglican minister. His main work was conducted during the late 1870's and gave rise to a whole branch of mathematics and a new way to approach issues of logic. We will develop the probability rules just covered using this powerful way to demonstrate the probability postulates including the Addition Rule, Multiplication Rule, Complement Rule, Independence, and Conditional Probability.

Example 3.27

Suppose an experiment has the outcomes 1, 2, 3, ... , 12 where each outcome has an equal chance of occurring. Let event $A = \{1, 2, 3, 4, 5, 6\}$ and event $B = \{6, 7, 8, 9\}$. Then A intersect $B = A \cap B = \{6\}$ and A union $B = A \cup B = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$. . The Venn diagram is as follows:
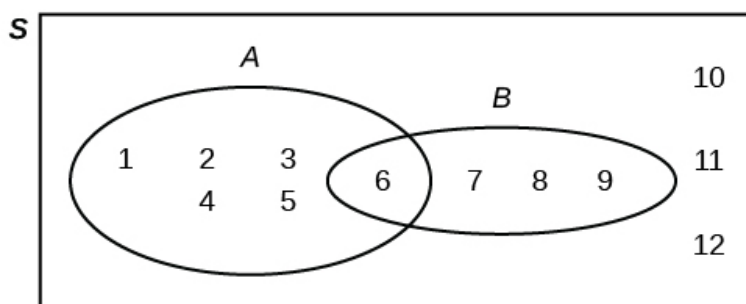

Figure 3.6

Figure 3.6 shows the most basic relationship among these numbers. First, the numbers are in groups called sets; set A and set B. Some number are in both sets; we say in set A $\cap$ in set B. The English word "and" means inclusive, meaning having the characteristics of both A and B, or in this case, being a part of both A and B. This condition is called the INTERSECTION of the two sets. All members that are part of both sets constitute the intersection of the two sets. The intersection is written as $A \cap B$ where $\cap$ is the mathematical symbol for intersection. The statement $A \cap B$ is read as "A intersect B." You can remember this by thinking of the intersection of two streets.

There are also those numbers that form a group that, for membership, the number must be in either one or the other group. The number does not have to be in BOTH groups, but instead only in either one of the two. These numbers are called the UNION of the two sets and in this case they are the numbers 1-5 (from A exclusively), 7-9 (from set B exclusively) and also 6, which is in both sets A and B. The symbol for the UNION is $\cup$, thus $A \cup B =$ numbers 1-9, but excludes number 10, 11, and 12. The values 10, 11, and 12 are part of the universe, but are not in either of the two sets.

Translating the English word "AND" into the mathematical logic symbol ∩ , intersection, and the word "OR" into the mathematical symbol ∪ , union, provides a very precise way to discuss the issues of probability and logic. The general terminology for the three areas of the Venn diagram in **Figure 3.6** is shown in **Figure 3.7**.

Try It Σ

3.27 Suppose an experiment has outcomes black, white, red, orange, yellow, green, blue, and purple, where each outcome has an equal chance of occurring. Let event C = {green, blue, purple} and event P = {red, yellow, blue}. Then $C \cap P$ = {blue} and $C \cup P$ = {green, blue, purple, red, yellow} . Draw a Venn diagram representing this situation.

Example 3.28

Flip two fair coins. Let A = tails on the first coin. Let B = tails on the second coin. Then A = {TT, TH} and B = {TT, HT}. Therefore, $A \cap B$ = {TT} . $A \cup B$ = {TH, TT, HT} .

The sample space when you flip two fair coins is X = {HH, HT, TH, TT}. The outcome HH is in NEITHER A NOR B. The Venn diagram is as follows:
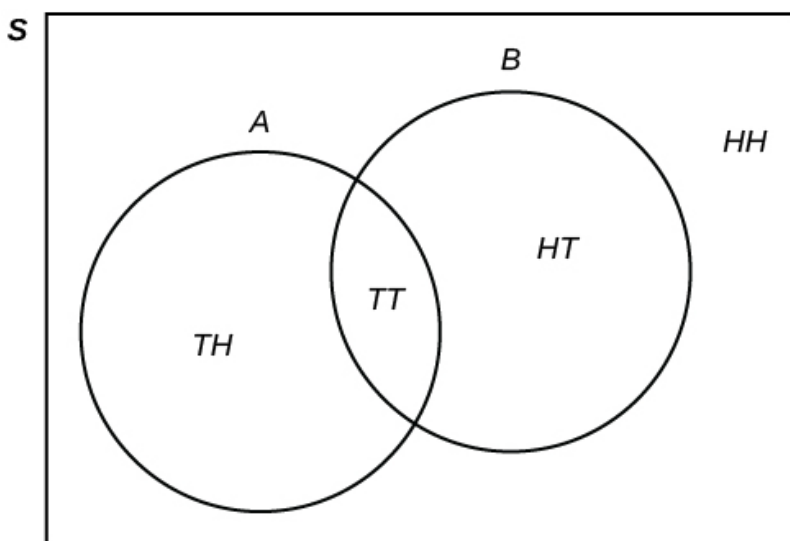

Figure 3.7

Try It Σ

3.28 Roll a fair, six-sided die. Let A = a prime number of dots is rolled. Let B = an odd number of dots is rolled. Then A = {2, 3, 5} and B = {1, 3, 5}. Therefore, $A \cap B$ = {3, 5} . $A \cup B$ = {1, 2, 3, 5} . The sample space for rolling a fair die is S = {1, 2, 3, 4, 5, 6}. Draw a Venn diagram representing this situation.

Example 3.29

A person with type O blood and a negative Rh factor (Rh-) can donate blood to any person with any blood type. Four percent of African Americans have type O blood and a negative RH factor, 5−10% of African Americans have the Rh- factor, and 51% have type O blood.
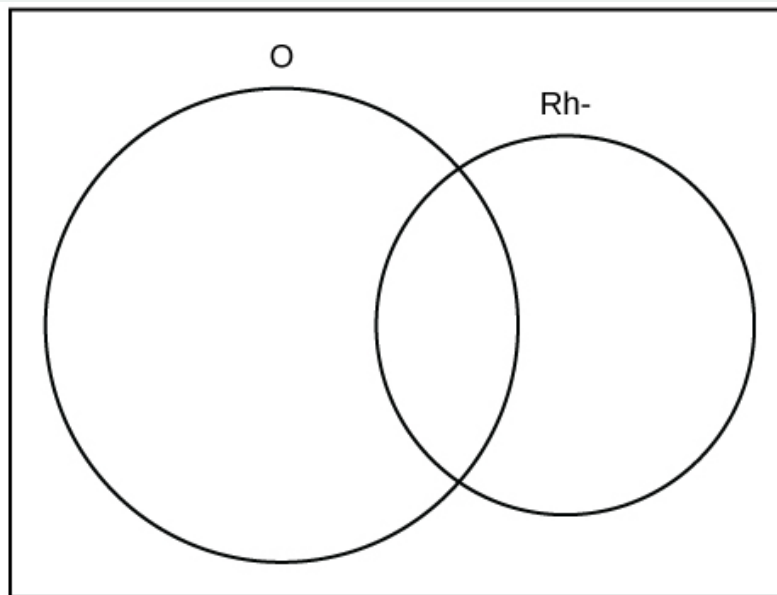

Figure 3.8

The "O" circle represents the African Americans with type O blood. The "Rh-" oval represents the African Americans with the Rh- factor.

We will take the average of 5% and 10% and use 7.5% as the percent of African Americans who have the Rh- factor. Let O = African American with Type O blood and R = African American with Rh- factor.

   a. $P(O) = $ _____

   b. $P(R) = $ _____

   c. $P(O \cap R) = $ _____

   d. $P(O \cup R) = $ _____

   e. In the Venn Diagram, describe the overlapping area using a complete sentence.

   f. In the Venn Diagram, describe the area in the rectangle but outside both the circle and the oval using a complete sentence.

Example 3.30

Forty percent of the students at a local college belong to a club and **50%** work part time. **Five percent** of the students work part time and belong to a club. Draw a Venn diagram showing the relationships. Let C = student belongs to a club and PT = student works part time.
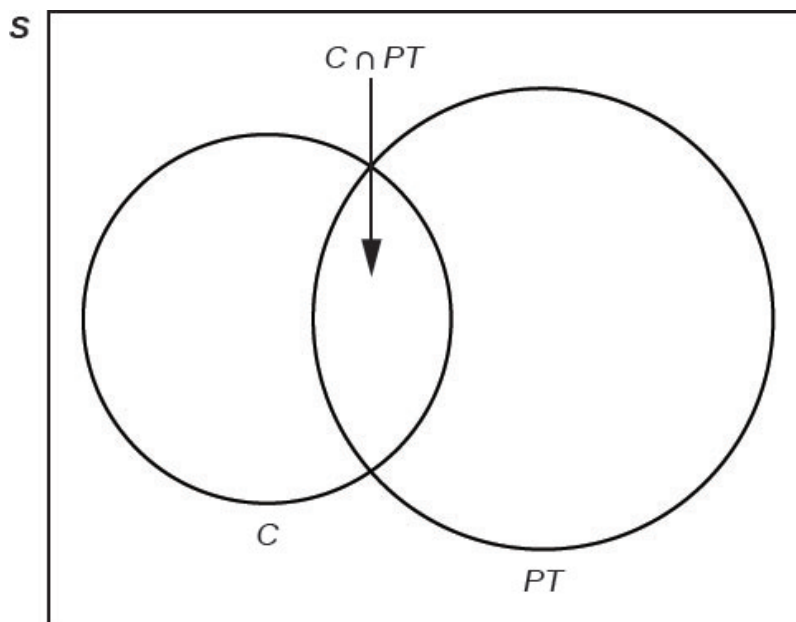
Figure 3.9

If a student is selected at random, find

- the probability that the student belongs to a club. $P(C) = 0.40$

- the probability that the student works part time. $P(PT) = 0.50$

- the probability that the student belongs to a club AND works part time. $P(C \cap \text{PT}) = 0.05$

- the probability that the student belongs to a club **given** that the student works part time.
$$P(C|PT) = \frac{P(C \cap PT)}{P(PT)} = \frac{0.05}{0.50} = 0.1$$

- the probability that the student belongs to a club **OR** works part time.
$$P(C \cup \text{PT}) = P(C) + P(\text{PT}) - P(C \cap \text{PT}) = 0.40 + 0.50 - 0.05 = 0.85$$

In order to solve **Example 3.30** we had to draw upon the concept of conditional probability from the previous section. There we used tree diagrams to track the changes in the probabilities, because the sample space changed as we drew without replacement. In short, conditional probability is the chance that something will happen given that some other event has already happened. Put another way, the probability that something will happen conditioned upon the situation that something else is also true. In **Example 3.30** the probability $P(C|PT)$ is the conditional probability that the randomly drawn student is a member of the club, conditioned upon the fact that the student also is working part time. This allows us to see the relationship between Venn diagrams and the probability postulates.

Try It Σ

3.30 Fifty percent of the workers at a factory work a second job, 25% have a spouse who also works, 5% work a second job and have a spouse who also works. Draw a Venn diagram showing the relationships. Let W = works a second job and S = spouse also works.

Try It Σ

3.30 In a bookstore, the probability that the customer buys a novel is 0.6, and the probability that the customer buys a non-fiction book is 0.4. Suppose that the probability that the customer buys both is 0.2.

a. Draw a Venn diagram representing the situation.

b. Find the probability that the customer buys either a novel or a non-fiction book.

c. In the Venn diagram, describe the overlapping area using a complete sentence.

d. Suppose that some customers buy only compact disks. Draw an oval in your Venn diagram representing this event.

Example 3.31

A set of 20 German Shepherd dogs is observed. 12 are male, 8 are female, 10 have some brown coloring, and 5 have some white sections of fur. Answer the following using Venn Diagrams.

Draw a Venn diagram simply showing the sets of male and female dogs.

Solution 3.31

The Venn diagram below demonstrates the situation of mutually exclusive events where the outcomes are independent events. If a dog cannot be both male and female, then there is no intersection. Being male precludes being female and being female precludes being male: in this case, the characteristic gender is therefore mutually exclusive. A Venn diagram shows this as two sets with no intersection. The intersection is said to be the null set using the mathematical symbol $\varnothing$.
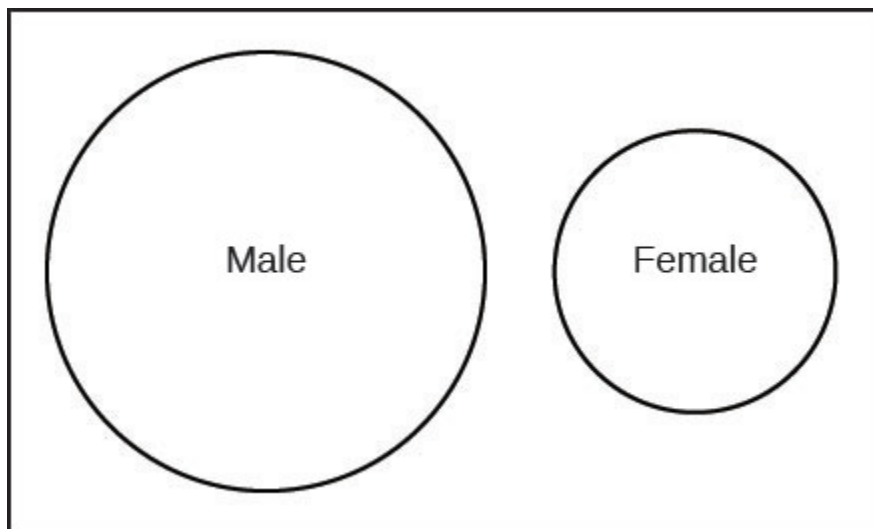
Figure 3.10

Draw a second Venn diagram illustrating that 10 of the male dogs have brown coloring.

Solution 3.31

The Venn diagram below shows the overlap between male and brown where the number 10 is placed in it. This represents Male ∩ Brown : both male and brown. This is the intersection of these two characteristics. To get the union of Male and Brown, then it is simply the two circled areas minus the overlap. In proper terms, Male ∪ Brown = Male + Brown − Male ∩ Brown will give us the number of dogs in the union of these two sets. If we did not subtract the intersection, we would have double counted some of the dogs.
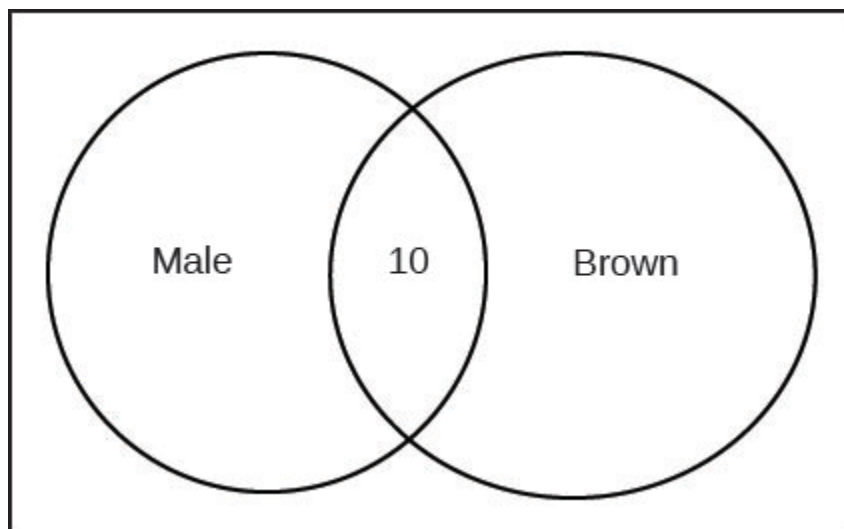

Figure 3.11

Now draw a situation depicting a scenario in which the non-shaded region represents "No white fur and female," or *White fur′ ∩ Female.* the prime above "fur" indicates "not white fur." The prime above a set means not in that set, e.g. A' means not A. Sometimes, the notation used is a line above the letter. For example, $\bar{A} = A'$.
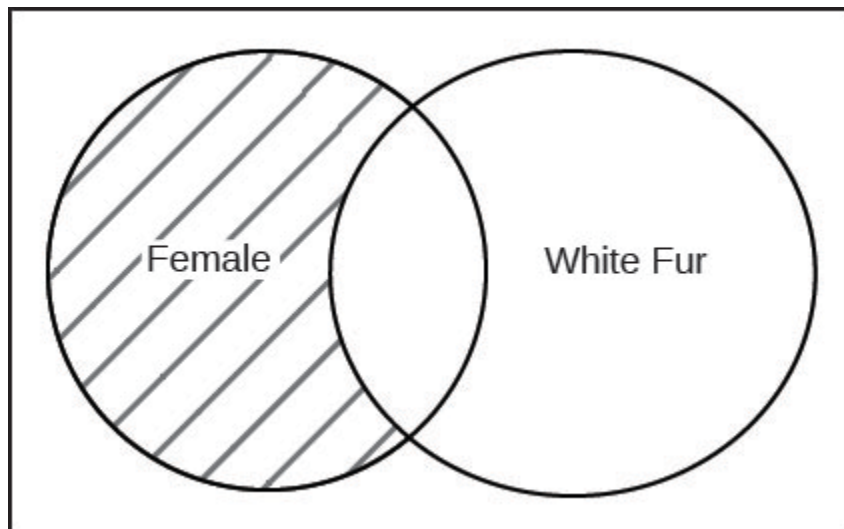
Solution 3.31

Figure 3.12

The Addition Rule of Probability

We met the addition rule earlier but without the help of Venn diagrams. Venn diagrams help visualize the counting process that is inherent in the calculation of probability. To restate the Addition Rule of Probability:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Remember that probability is simply the proportion of the objects we are interested in relative to the total number of objects. This is why we can see the usefulness of the Venn diagrams. **Example 3.31** shows how we can use Venn diagrams to count the number of dogs in the union of brown and male by reminding us to subtract the intersection of brown and male. We can see the effect of this directly on probabilities in the addition rule.

Example 3.32

Let's sample 50 students who are in a statistics class. 20 are freshmen and 30 are sophomores. 15 students get a "B" in the course, and 5 students both get a "B" and are freshmen.

Find the probability of selecting a student who either earns a "B" OR is a freshmen. We are translating the word OR to the mathematical symbol for the addition rule, which is the union of the two sets.

Solution 3.32

We know that there are 50 students in our sample, so we know the denominator of our fraction to give us probability. We need only to find the number of students that meet the characteristics we are interested in, i.e. any freshman and any student who earned a grade of "B." With the Addition Rule of probability, we can skip directly to probabilities.

Let "A" = the number of freshmen, and let "B" = the grade of "B." Below we can see the process for using Venn diagrams to solve this.

The $P(A) = \frac{20}{50} = 0.40$, $P(B) = \frac{15}{50} = 0.30$, and $P(A \cap B) = \frac{5}{50} = 0.10$.

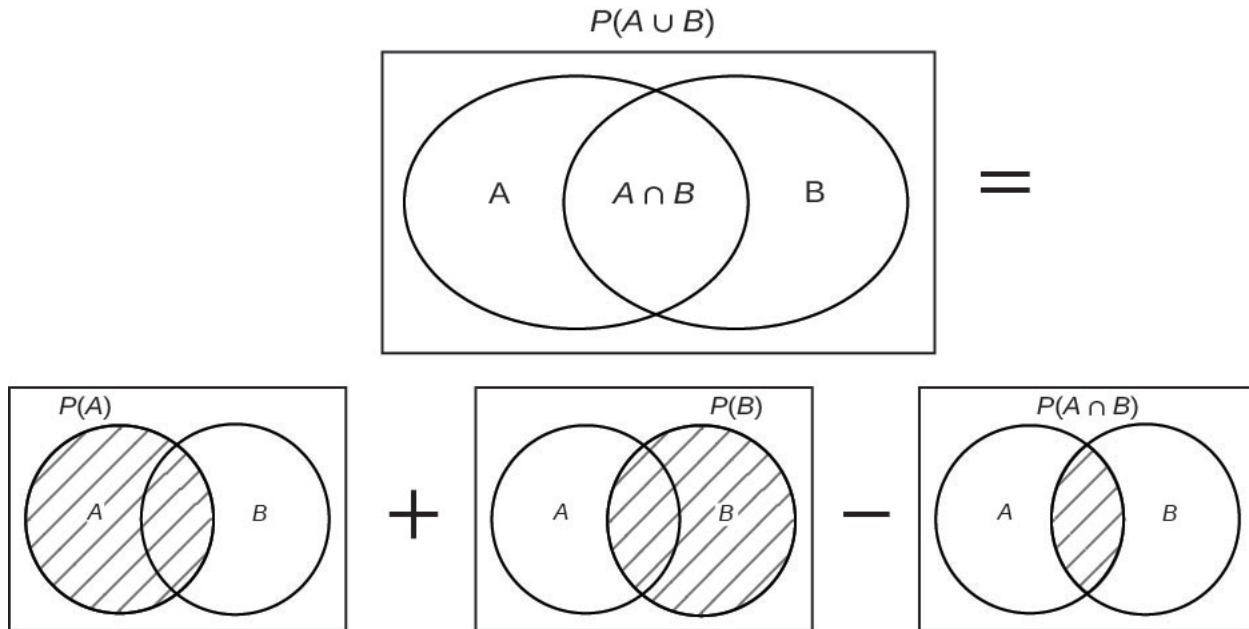Therefore, $P(A \cap B) = 0.40 + 0.30 - 0.10 = 0.60$.

Figure 3.13

If two events are mutually exclusive, then, like the example where we diagram the male and female dogs, the addition rule is simplified to just $P(A \cup B) = P(A) + P(B) - 0$. This is true because, as we saw earlier, the union of mutually exclusive events is the null set, $\emptyset$. The diagrams below demonstrate this.
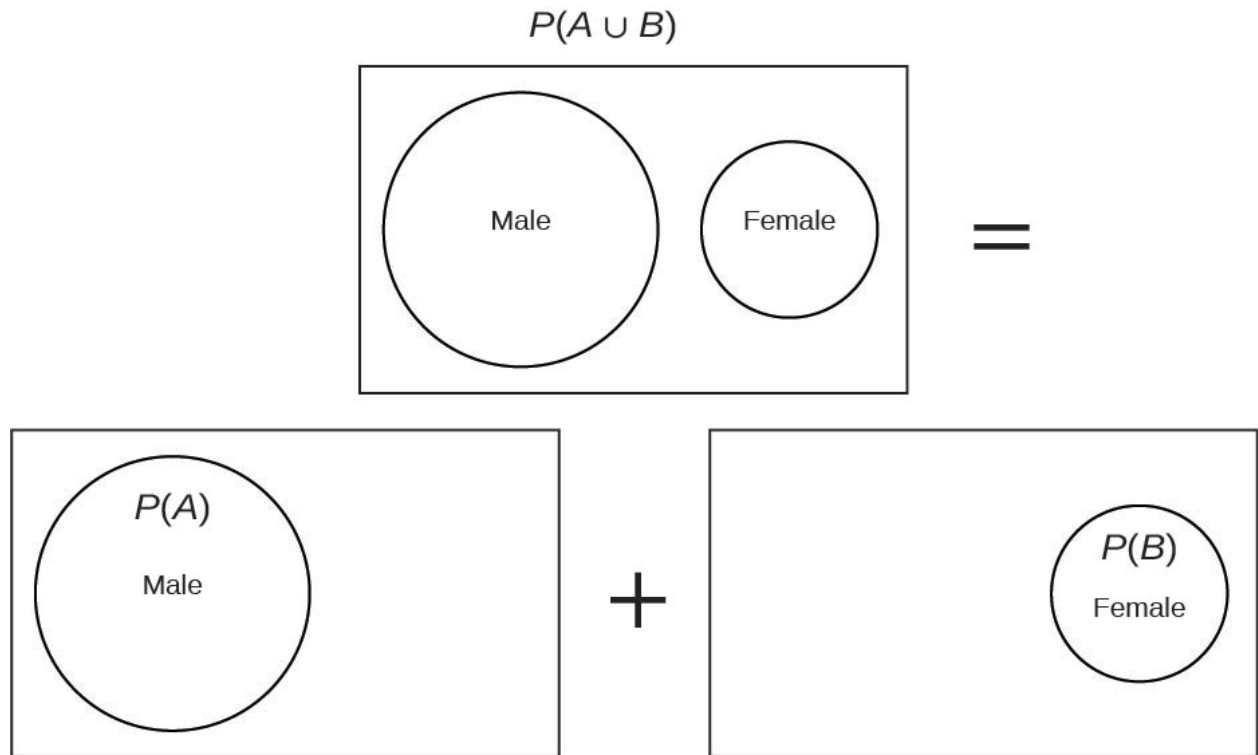
$$P(A \cup B)$$

Figure 3.14

The Multiplication Rule of Probability

Restating the Multiplication Rule of Probability using the notation of Venn diagrams, we have:

$$P(A \cap B) = P(A|B) \cdot P(B)$$

The multiplication rule can be modified with a bit of algebra into the following conditional rule. Then Venn diagrams can then be used to demonstrate the process.

The conditional rule: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

Using the same facts from **Example 3.32** above, find the probability that someone will earn a "B" if they are a "freshman."

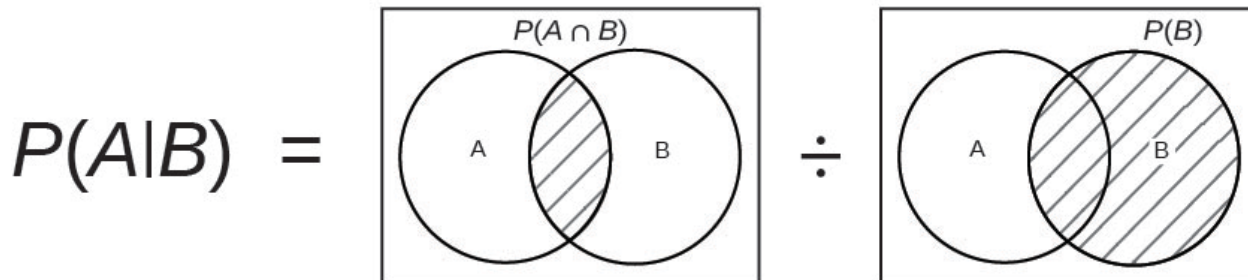$$P(A|B) = \frac{0.10}{0.30} = \frac{1}{3}$$

Figure 3.15

The multiplication rule must also be altered if the two events are independent. Independent events are defined as a situation where the conditional probability is simply the probability of the event of interest. Formally, independence of events is defined as $P(A|B) = P(A)$ or $P(B|A) = P(B)$. When flipping coins, the outcome of the second flip is independent of the outcome of the first flip; coins do not have memory. The Multiplication Rule of Probability for independent events thus becomes:

$$P(A \cap B) = P(A) \cdot P(B)$$

One easy way to remember this is to consider what we mean by the word "and." We see that the Multiplication Rule has translated the word "and" to the Venn notation for intersection. Therefore, the outcome must meet the two conditions of freshmen and grade of "B" in the above example. It is harder, less probable, to meet two conditions than just one or some other one. We can attempt to see the logic of the Multiplication Rule of probability due to the fact that fractions multiplied times each other become smaller.

The development of the Rules of Probability with the use of Venn diagrams can be shown to help as we wish to calculate probabilities from data arranged in a contingency table.

Example 3.33

Table 3.11 is from a sample of 200 people who were asked how much education they completed. The columns represent the highest education they completed, and the rows separate the individuals by male and female.

	Less than High School Grad	High School Grad	Some College	College Grad	**Total**
Male	5	15	40	60	120
Female	8	12	30	30	80
Total	13	27	70	90	**200**

Table 3.11

Now, we can use this table to answer probability questions. The following examples are designed to help understand the format above while connecting the knowledge to both Venn diagrams and the probability rules.

What is the probability that a selected person both finished college and is female?

Solution 3.33

This is a simple task of finding the value where the two characteristics intersect on the table, and then applying the postulate of probability, which states that the probability of an event is the proportion of outcomes that match the event in which we are interested as a proportion of all total possible outcomes.

$P(College\ Grad\ \cap\ Female)$ = $\frac{30}{200}$ = 0.15

What is the probability of selecting either a female or someone who finished college?

Solution 3.33

This task involves the use of the addition rule to solve for this probability.

$P(College\ Grad\ \cup\ Female)$ = $P(F) + P(CG) - P(F\ \cap\ CG)$

$P(College\ Grad\ \cup\ Female)$ = $\frac{80}{200} + \frac{90}{200} - \frac{30}{200} = \frac{140}{200}$ = 0.70

What is the probability of selecting a high school graduate if we only select from the group of males?

Solution 3.33

Here we must use the conditional probability rule (the modified multiplication rule) to solve for this probability.

$P(HS\ Grad\ |\ Male$ = $\dfrac{P(\text{HS Grad} \cap \text{Male})}{P(\text{Male})} = \dfrac{\left(\frac{15}{200}\right)}{\left(\frac{120}{200}\right)} = \dfrac{15}{120}$ = 0.125

Can we conclude that the level of education attained by these 200 people is independent of the gender of the person?

Solution 3.33

There are two ways to approach this test. The first method seeks to test if the intersection of two events equals the product of the events separately remembering that if two events are independent than $P(A)*P(B) = P(A\ \cap\ B)$. For simplicity's sake, we can use calculated values from above.

Does P(College Grad ∩ Female) = P(CG) · P(F)?

$\frac{30}{200} \neq \frac{90}{200} \cdot \frac{80}{200}$ *because* 0.15 ≠ 0.18.

Therefore, gender and education here are **not** independent.

The second method is to test if the conditional probability of *A given B* is equal to the probability of *A*. Again for simplicity, we can use an already calculated value from above.

Does P(HS Grad | Male) = P(HS Grad)?

$\frac{15}{120} \neq \frac{27}{200}$ *because* 0.125 ≠ 0.135.

Therefore, again gender and education here are **not** independent.