Chapter 2

Probability

In the fields of observation chance favors only the prepared mind.
—Louis Pasteur (Lecture, University of Lille, December 7, 1854)

You have 22 songs on your mp3 player's playlist, and you set the player on shuffle mode, where songs are allowed to repeat, while you study. Country music is your favorite, but only 12 of the songs on this playlist are country with the rest being rock. What is the probability the first song will be country music? What is the probability the first song will not be country music? If 3 songs play, what is the probability that all 3 are country music? Or the probability that exactly 2 of them are country music? Or the probability that none of them are country music? If 3 songs play, what is the probability that only the first song is country, and is this different from the probability that exactly one song will be country? Why or why not?

2.1 Introduction

The probability of an event is a number between 0 and 1 (inclusive).[1] In Chapter 37, we further explore the idea of a probability of an event as the percentage of time that an event occurs, in the long run. For now, however, we begin with only three basic assumptions. We build the framework for our understanding from these 3 intuitive ideas:

[1] In more advanced courses, some events are so complicated that there is not a reasonable way to assign them a probability. We will not study such complications here. See, e.g., Billingsley [1] or Durrett [2].

Remark 2.1. Three Probability Axioms (Intuitive Statements)

1. Any event occurs a certain percentage of the time, so probabilities are always between 0 and 1.

2. With probability 1, some outcome in the sample space occurs.

3. If events have no outcomes in common, then the probability of their union is the sum of the probabilities of the individual events.

Definition 2.2. A pair of events A, B is **disjoint** (also called **mutually exclusive**) if they have no outcome in common, i.e., if their intersection is empty, $A \cap B = \emptyset$. A collection of events is **disjoint** if every pair of events is disjoint.

Definition 2.3. We always use "P" to denote a probability that is defined on the events associated with some random phenomenon.

Now we can state the fundamental ideas (mentioned earlier) in a precise way.

Axioms 2.4. Three Probability Axioms (Mathematical Statements)

1. For each event A,
$$0 \le P(A) \le 1.$$

2. For the sample space S,
$$P(S) = 1.$$

3. If $A_1, A_2, \ldots$ is a collection of *disjoint* events, then
$$P\Big(\bigcup_{j=1}^{\infty} A_j\Big) = \sum_{j=1}^{\infty} P(A_j).$$

Theorem 2.5. The probability of the empty set $\emptyset$ is always 0.

This theorem makes sense intuitively, but does it fit with our basic assumptions? Yes! Here is the reasoning:

$$\begin{aligned}
1 &= P(S) && \text{by Axiom 2.4.2} \\
&= P(S \cup \emptyset \cup \emptyset \cup \cdots) && \text{since } S = S \cup \emptyset \cup \emptyset \cdots \\
&= P(S) + P(\emptyset) + P(\emptyset) + \cdots && \text{by Axiom 2.4.3 } (A_1 = S;\ A_j = \emptyset \text{ for } j \ge 2)
\end{aligned}$$

By Axiom 2.4.2, $P(S) = 1$. The rest of the right hand side consists of nonnegative terms $P(\emptyset)$, which must therefore each be 0. So $P(\emptyset) = 0$.

Theorem 2.6. If $A_1, A_2, \ldots, A_n$ is a collection of finitely many *disjoint* events, then the probability of the union of the events equals the sum of the probabilities of the events:

$$P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j).$$

Again, this makes intuitive sense. To prove it, using our basic assumptions, define $A_j = \emptyset$ for all $j > n$. Then we use the probability theory axioms, as follows:

$$\begin{aligned} P\left(\bigcup_{j=1}^{n} A_j\right) &= P\left(\bigcup_{j=1}^{\infty} A_j\right) && \text{since } A_j = \emptyset \text{ for } j > n \\ &= \sum_{j=1}^{\infty} P(A_j) && \text{by Axiom 2.4.3} \\ &= \sum_{j=1}^{n} P(A_j) && \text{since } P(A_j) = 0 \text{ for } j > n \end{aligned}$$

2.2 Equally Likely Events

We begin by considering the probabilities assigned to some of the events that were discussed at the start of Chapter 1.

Example 2.7. When rolling a die, each of the six outcomes should be equally likely. This means that each single-outcome event should have the same probability.

The probability of each seems (intuitively) to be 1/6. Using our simple assumptions,

$$\begin{aligned} 1 &= P(S) \qquad \text{by Axiom 2.4.2} \\ &= P(\{1,2,3,4,5,6\}) \\ &= P(\{1\} \cup \{2\} \cup \{3\} \cup \{4\} \cup \{5\} \cup \{6\}) \\ &= P(\{1\}) + P(\{2\}) + P(\{3\}) + P(\{4\}) + P(\{5\}) + P(\{6\}) \quad \text{by Theorem 2.6} \end{aligned}$$

If all $P(\{j\})$'s are the same, then $1 = 6P(\{j\})$, so $P(\{j\}) = 1/6$ for each j. So our intuition is correct. Now we can compute any kind of probability associated with one roll of a die. For instance, the probability a die roll is odd is:

$$P(\{1,3,5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}.$$

Example 2.8. A pregnancy that yields exactly one baby would yield an outcome of either a boy or a girl, which are equally likely (as in the die example above).

www.cdc.gov/nchs/data/nvsr/nvsr61/nvsr61_01.pdf suggests that the odds are really closer to 51.17% for boys vs 48.83% for girls, but we assume a 50/50 ratio.

The four relevant probabilities are

1. $P(\emptyset) = 0$,
2. $P(\{\text{boy}\}) = 1/2$,
3. $P(\{\text{girl}\}) = 1/2$,
4. $P(\{\text{boy, girl}\}) = P(S) = 1$.

In the last case, $S = \{\text{boy, girl}\}$, so we are really just emphasizing the fact that $P(S) = 1$.

These observations about equally likely outcomes are handy and very general:

Theorem 2.9. If a sample space S has n **equally likely** outcomes, then each outcome has probability $1/n$ of occurring.

This is true for just the same reasons as in the die example. Let $x_1, x_2, \ldots, x_n$ be the n outcomes. Then

$$\begin{aligned} 1 &= P(S) \qquad \text{by Axiom 2.4.2} \\ &= P(\{x_1, x_2, \ldots, x_n\}) \\ &= P(\{x_1\} \cup \{x_2\} \cup \cdots \cup \{x_n\}) \\ &= P(\{x_1\}) + P(\{x_2\}) + \cdots + P(\{x_n\}) \qquad \text{by Axiom 2.4.3} \end{aligned}$$

If all $P(\{x_j\})$'s are the same, then $1 = nP(\{x_j\})$, so $P(\{x_j\}) = 1/n$ for each j.

Corollary 2.10. If sample space S has n **equally likely** outcomes, and A is an event with j outcomes, then event A has probability j/n of occurring, i.e., $P(A) = j/n$.

To prove this corollary, write $y_1, \ldots, y_j$ as the j outcomes in A. Then

$$\begin{aligned} P(A) &= P(\{y_1, y_2, \ldots, y_j\}) \\ &= P(\{y_1\} \cup \{y_2\} \cup \cdots \cup \{y_j\}) \\ &= P(\{y_1\}) + P(\{y_2\}) + \cdots + P(\{y_j\}) \qquad \text{by Axiom 2.4.3} \\ &= \frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n} \qquad \text{by Theorem 2.9} \\ &= j/n \end{aligned}$$

Definition 2.11. The number of outcomes in an event A, also called the size of A, is denoted as $|A|$.

Using the notation of $|S|$ and in $|A|$ as the number of items in S and A, respectively, Corollary 2.10 can be rewritten as follows:

Corollary 2.12. If sample space S has a finite number of equally likely outcomes, then event A has probability

$$P(A) = |A|/|S|,$$

where $|S|$ and $|A|$ denote the number of items in S and A, respectively.

We will study equally likely outcomes to a much greater extent, in Chapters 20 and 22. For now, however, we give a few examples.

Example 2.13. As in Example 1.9, consider a dartboard split into 20 regions. We treat each of the 20 regions as an outcome, and the possibility of a "miss" as a 21st potential outcome. For simplicity, suppose that the probability of a miss is 0. (Just allow the player to try again, if initially missing the board.) Also suppose the player is equally likely to hit any of the 20 regions on the board.

The sample space is $S = \{1, 2, 3, \ldots, 20\}$. We have 20 disjoint events, each with one outcome: $A_1 = \{1\}$, $A_2 = \{2\}$, and in general,

$$A_j = \{j\}, \qquad \text{for } 1 \leq j \leq 20.$$

(This is a gentle introduction to enumerating events. We will not always use $A_j = \{j\}$.) We assumed $P(\{\text{miss}\}) = 0$. The other 20 outcomes are each equally likely, and the jth event has just 1 outcome, so

$$P(A_j) = 1/20 \qquad \text{for each } j.$$

Example 2.13 (continued) We can now split the dartboard into four regions.

If a new event R_1 is constructed as the union of several of the A_j's, then the probability of R_1 is just the sum of the probabilities. E.g., if

$$R_1 = \{1, 18, 4, 13, 6\} = A_1 \cup A_{18} \cup A_4 \cup A_{13} \cup A_6,$$

(i.e., R_1 is the event that the dart lands in the northeast portion), then

$$P(R_1) = P(A_1) + P(A_{18}) + P(A_4) + P(A_{13}) + P(A_6) = 5/20 = 1/4.$$

This could also seen by using Corollary 2.10, since R_1 contains 5 of the 20 equally likely outcomes, so $P(R_1) = 5/20$.

As in Figure 2.1, define northeast, southeast, southwest, and northwest regions as

$$R_1 = \{1, 18, 4, 13, 6\}, \qquad R_2 = \{10, 15, 2, 17, 3\},$$
$$R_3 = \{19, 7, 16, 8, 11\}, \qquad R_4 = \{14, 9, 12, 5, 20\}.$$

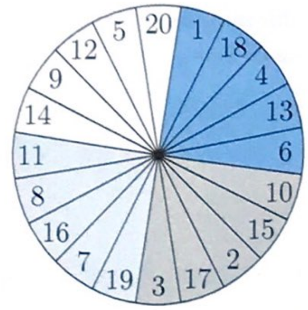

FIGURE 2.1: Twenty-one possible outcomes. Four colors for the northeast, southeast, southwest, and northwest regions.

A classification of the sample space, as we made in Example 2.13, when splitting the sample space into four regions, is called a **partition** of the outcomes in the sample space. More generally, in a partition, every outcome belongs in exactly one of the events, and the events in the partition are disjoint. Partitions will be very helpful in Chapter 5, on Bayes' Theorem.

Definition 2.14. If a collection of nonempty events is disjoint, and their union is the entire sample space, then the collection is called a **partition**. If $\bigcup_j B_j = S$ and the B_j's are disjoint events, then the collection of B_j's is called a **partition**.

In a partition, not all of the regions need to have the same size.

Example 2.15. For instance, a different partition of the dartboard could consist of three disjoint events:

$$\text{the upper portion, } T_1 = \{12, 5, 20, 1, 18\}$$
$$\text{the lower left portion, } T_2 = \{3, 19, 7, 16, 8, 11, 14, 9\}$$
$$\text{the lower right portion, } T_3 = \{4, 13, 6, 10, 15, 2, 17\}$$

See Figure 2.2. In this partition, $P(T_1) = 5/20$, $P(T_2) = 8/20$, and $P(T_3) = 7/20$. Notice $P(T_1) + P(T_2) + P(T_3) = 1$.

Remark 2.16. The probabilities of events in a partition always sum to 1.

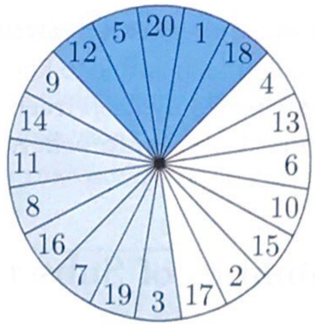

FIGURE 2.2: Twenty-one possible outcomes. Three colors for Example 2.15 regions.

To see this, in a partition consisting B_j's, every outcome is in one of the events, so $S = \bigcup_j B_j$. Also, each outcome is in exactly one of these events in the partition, so the B_j's are disjoint. Thus $P\big(\bigcup_j B_j\big) = \sum_j P(B_j)$. Putting these together, we get

$$1 = P(S) = P\Big(\bigcup_j B_j\Big) = \sum_j P(B_j).$$

So the sum of the probabilities of the events in a partition is always 1.

Example 2.17. As in Example 1.13, a student shuffles a deck of cards thoroughly (one time) and then selects cards from the deck *without replacement* until the ace of spades appears.

Let B_k denote the event that the ace of spades is drawn on exactly the kth draw:

$$B_k = \{(x_1, x_2, \ldots, x_k) \mid x_k = \mathbf{A}\spadesuit, \text{ and the } x_j\text{'s are distinct}\}.$$

We emphasize that $P(B_k) = 1/52$ for each k, since the initial placement of the ace of spades (i.e., during the initial shuffle) completely determines when the ace of spades will appear. Since the ace of spades is equally likely to be in any of the 52 places in the deck, then the ace of spades is equally likely to appear on any of the 52 draws.

The B_k's are disjoint events, since it is impossible for an outcome to simultaneously be in more than one of the B_k's. Also, every outcome is in exactly one of the events. So $B_1, B_2, \ldots, B_{52}$ form a partition of the sample space.

Not every set of outcomes is equally likely, so we must be careful when applying Corollary 2.10. For instance, when bowling, it is possible to knock down between 0 and 10 pins, so there are 11 outcomes (if we only keep track of the score, not the specific pins that fall down). We have no reason to believe that all of these 11 outcomes are equally likely.

2.3 Complements; Probabilities of Subsets

Example 2.18. As in Example 1.18, consider two events: A is the event that the amount of rainfall on a given day is strictly less than 2.8 inches, and B is the event that the amount of rainfall is 2.8 inches or more. Thus

$$A = [0, 2.8) \qquad B = [2.8, \infty).$$

The sample space consisting of all possible amounts of rain is $S = [0, \infty)$ so $S = A \cup B$. Also, A and B are disjoint. So

$$1 = P(S) = P(A \cup B) = P(A) + P(B),$$

so $P(B) = 1 - P(A)$. E.g., if the probability of "rainfall less than 2.8 inches" is 83%, then the probability of "rainfall 2.8 inches or more" must be $1 - 0.83 = 0.17$, i.e., 17%.

Theorem 2.19. The complement A^c of event A has probability $P(A^c) = 1 - P(A)$.

The argument is straightforward: An event and its complement are disjoint, and their union is the whole sample space:

$$A \cup A^c = S$$

So

$$1 = P(S) = P(A \cup A^c) = P(A) + P(A^c),$$

and the theorem follows, $P(A^c) = 1 - P(A)$.

Now we consider a bound for the probability of an event that is the subset of another event. Recall that A is a subset of B if every outcome of event A is contained in B too (denoted by $A \subset B$). Intuitively, B seems more likely to occur than A in this case. This intuition is correct:

Theorem 2.20. If $A \subset B$ then $P(A) \leq P(B)$.

The event B can be expressed as a disjoint union $B = A \cup (B \setminus A)$. So

$$P(B) = P(A) + P(B \setminus A),$$

but probability is always positive, so $P(B \setminus A) \geq 0$, and thus $P(B) \geq P(A)$.

2.4 Inclusion-Exclusion

The method of inclusion-exclusion allows us to relate overlaps among subsets to unions and intersections. This decomposition enables us to calculate probabilities for events that are overlapping.

Example 2.21. Consider an observer who randomly chooses a car without knowing its color. There are exactly 10 cars available, one from each of these colors: red, blue, yellow, green, lime, teal, orange, silver, brown, or black. Let A denote the event that the car is red, blue, yellow, orange, or silver; let B denote the event that the car is orange, silver, brown, or black. So $A \cap B$ is the event that the car is orange or silver. Thus

$$P(A) = 5/10, \qquad P(B) = 4/10, \qquad P(A \cap B) = 2/10, \qquad P(A \cup B) = 7/10.$$

These probabilities correspond to the scenario in Figure 2.3.

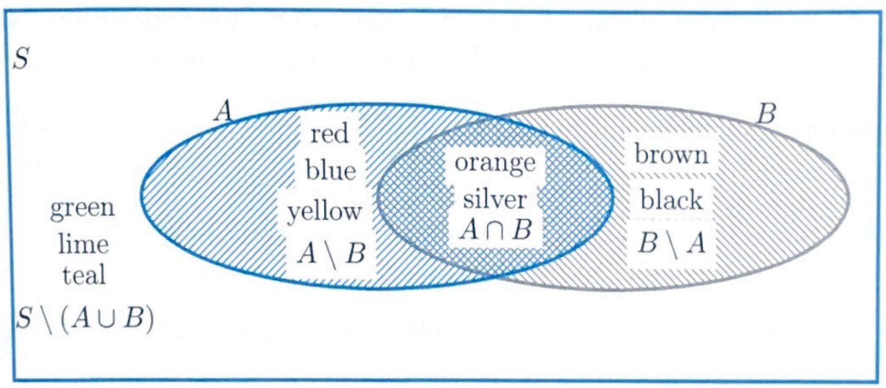

FIGURE 2.3: Observing ten colors of cars.

Notice

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

This might seem intuitively clear because $A \cup B$ accounts for each of the colors red, blue, yellow, orange, silver one time, while

1. A accounts for red, blue, yellow, orange, silver
2. B accounts for orange, silver, brown, black
3. $A \cap B$ accounts for orange, silver

So we are just removing the duplication, i.e., correcting for the fact that the orange and silver were accounted for twice.

The very same argument works much more generally:

Theorem 2.22. For any two events A and B,

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

To prove this, write

$$\begin{aligned} P(A \cup B) &= P(A \setminus B) + P(A \cap B) + P(B \setminus A) \\ &= P(A \setminus B) + P(A \cap B) + P(B \setminus A) + P(A \cap B) - P(A \cap B) \\ &= P(A) + P(B) - P(A \cap B) \end{aligned}$$

Similar nice things happen if we try to characterize $P(A \cup B \cup C)$ by taking each set into account the proper amount of times. Intuitively, we might first guess that $P(A \cup B \cup C)$ and $P(A) + P(B) + P(C)$ are close in value, but we have double-counted all of the contributions from the outcomes in $A \cap B$ and $A \cap C$ and $B \cap C$, so we remove those. Then $P(A \cup B \cup C)$ is close to $P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C)$, but the contribution from $A \cap B \cap C$ was originally accounted for three times and then removed three times, so we must add it back on. (We ask for a proof in the exercises.)

Theorem 2.23. For any three events A, B, C,

$$\begin{aligned} P(A \cup B \cup C) = P(A) + P(B) + P(C) \\ - P(A \cap B) - P(A \cap C) - P(B \cap C) \\ + P(A \cap B \cap C). \end{aligned}$$

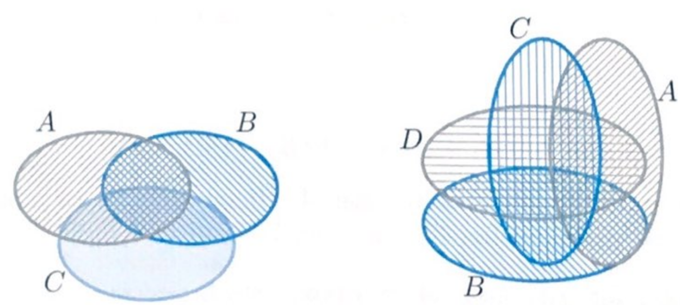

FIGURE 2.4: Left: Overlaps among three events A, B, C. Right: Overlaps among four events A, B, C, D.

The overlaps among A, B, C can be visualized on the left side of Figure 2.4. The overlaps among four events, A, B, C, D can be visualized on the right side of the same figure. It is true, furthermore, by similar reasoning, that

Theorem 2.24. For any four events A, B, C, D,

$$\begin{aligned}P(A \cup B \cup C \cup D) = P(A) + P(B) + P(C) + P(D) \\ - P(A \cap B) - P(A \cap C) - P(A \cap D) \\ - P(B \cap C) - P(B \cap D) - P(C \cap D) \\ + P(A \cap B \cap C) + P(A \cap B \cap D) \\ + P(A \cap C \cap D) + P(B \cap C \cap D) \\ - P(A \cap B \cap C \cap D).\end{aligned}$$

This same kind of reasoning can continue, and we get a general inclusion-exclusion formula.

Theorem 2.25. Inclusion-Exclusion Rule
For any finite sequence of events $A_1, A_2, \ldots, A_n$,

$$\begin{aligned}P\left(\bigcup_{j=1}^{n} A_j\right) = \sum_{j=1}^{n} P(A_j) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \\ - \sum_{i<j<k<l} P(A_i \cap A_j \cap A_k \cap A_l) \\ \pm \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n)\end{aligned}$$

2.5 More Examples of Probabilities of Events

Example 2.26. As in Example 1.14, consider a student who draws cards from a deck but this time he always replaces the card after each selection and then reshuffles the deck. He only stops if he reaches the ace of spades.

Let B_k be the set of outcomes in which the ace of spades is discovered for the first time on the kth draw:

$$B_k = \{(x_1, \ldots, x_k) \mid \text{only } x_k \text{ is } \mathbf{A}\spadesuit\}$$

Then the set of all possibilities in which the student actually finds the ace of spades is $\bigcup_{k=1}^{\infty} B_k$. Also let

$$C = \{(x_1, x_2, x_3, \ldots) \mid \text{none of the } x_k\text{'s is } \mathbf{A}\spadesuit\}$$

So the entire sample space can be partitioned using the B_k's and C, i.e.,

$$S = \left(\bigcup_{k \geq 1} B_k\right) \cup C.$$

The probability that a single draw does not contain the ace of spades is always $51/52$. Since the draws do not affect each other (a phenomenon that we will explore much further in Chapter 3 on independence), it follows that

$$P(B_k) = \overbrace{\left(\frac{51}{52}\right)\left(\frac{51}{52}\right)\cdots\left(\frac{51}{52}\right)}^{k-1}\left(\frac{1}{52}\right) = \left(\frac{51}{52}\right)^{k-1}\left(\frac{1}{52}\right).$$

$\sum_{j=0}^{\infty}\left(\frac{51}{52}\right)^j = \frac{1}{1-\frac{51}{52}}$ is a geometric sum; see the Math Review.

We know that the B_k's and C are all disjoint, and also

$$S = \left(\bigcup_{k=1}^{\infty} B_k\right) \cup C.$$

So

$$1 = P(S) = P(C \cup B_1 \cup B_2 \cup B_3 \cup \cdots) = P(C) + \sum_{k=1}^{\infty} P(B_k).$$

We already computed $P(B_k) = (51/52)^{k-1}(1/52)$, so

$$\sum_{k=1}^{\infty} P(B_k) = \sum_{k=1}^{\infty}\left(\frac{51}{52}\right)^{k-1}\left(\frac{1}{52}\right) = \left(\frac{1}{1-\frac{51}{52}}\right)\left(\frac{1}{52}\right) = 1.$$

In summary, $1 = P(C) + 1$, so $P(C)$ must be 0. In other words, the probability that the ace of spades never appears is 0.

Example 2.27. **(continued from Example 1.17)** A student hears ten songs (in a random shuffle mode) on her music player, paying special attention to how many of these songs belong to her favorite type of music. We assume that the songs are picked independently of each other and that each song has probability p of being a song of the student's favorite type.

As in Chapter 1, use F and N to denote when a song belongs to her favorite, or not to her favorite, type of music. So, for each song, the probability that the song is one of her favorite type can be called p, and the probability that the song is not one of her favorite type is $1 - p$.

Now consider the event A that none of the first three songs is her favorite type of music. This event is

$$A = \{(N, N, N, x_4, \ldots, x_{10}) \mid x_j \in \{F, N\}\},$$

and the probability of the event is $P(A) = (1 - p)^3$, because all that the event requires is that none of the first three songs is from her favorite type of music. We do not impose any restrictions on songs 4, 5, ..., 10, so these do not affect the probability of event A occurring.

Now consider the event that the even-numbered songs are from her favorite type of music. We write this event as

$$B = \{(x_1, F, x_3, F, x_5, F, x_7, F, x_9, F) \mid x_j \in \{F, N\}\},$$

and the probability of B is $P(B) = p^5$, because we only require that five specific songs are from her favorite type of music.

Similarly, if

$$C = \{(x_1, x_2, x_3, x_4, x_5, F, F, F, F, F) \mid x_j \in \{F, N\}\},$$

then $P(C) = p^5$ too.

An outcome is in $B \cap C$ if and only if the 2nd, 4th, 6th, 7th, 8th, 9th, and 10th songs are favorites, i.e., of type "F." So the event $B \cap C$ is

$$B \cap C = \{(x_1, F, x_3, F, x_5, F, F, F, F, F) \mid x_j \in \{F, N\}\},$$

which has probability p^7.

As one final event, let A_j denote the event that exactly j of the 10 songs are from her favorite type of music. The probability of A_j is

$$P(A_j) = \binom{10}{j} p^j (1-p)^{10-j},$$

where $\binom{10}{j} = \frac{10!}{j!(10-j)!}$ is the number of ways to pick exactly j out of 10 songs. We will explore this idea in greater depth in Chapter 15, on Binomial random variables.

2.6 Exercises

2.6.1 Practice

Exercise 2.1. Songs by genre. A song is chosen at random from a person's mp3 player. The student makes a partition of the sample space, according to genre of music. The table below gives the number of outcomes in each part of the partition. There are 27,333 songs altogether.

1032	Alternative	83	Electronic	56	Metal
330	Blues	508	Folk	2718	Other
275	Books & Spoken	183	Gospel	1786	Pop
1468	Children's Music	82	Hip-Hop	403	R&B
921	Classical	564	Holiday	8286	Rock
6169	Country	537	Jazz	1432	Soundtrack
178	Easy Listening	106	Latin	216	World

Let A be the event that a song is either blues, jazz, or rock, i.e.,

$$A = \{x \mid x \text{ is a blues, jazz, or rock song}\},$$

when one song is chosen at random. Assume that all songs are equally likely to appear.

a. What is the probability of randomly selecting a blues song?

b. What is $P(A)$?

c. Did you need to use inclusion-exclusion for part b? (Explain very briefly; one sentence will do.)

d. What is $P(A^c)$?

Exercise 2.2. Rock climbing. I am out rock climbing, and the rock face has 4 easy, 7 challenging, and 3 extreme routes to get to the top. The routes are poorly marked, so I just choose one at random, with all routes equally likely. What is the probability that I do not choose an extreme route?

Exercise 2.3. Student interests. A student is chosen at random. Let A, B, C be the events that the student is an Aeronautics major, a Basketball player, or a Co-op student. The events are not disjoint; we are told

$$P(A) = P(B) = P(C) = 0.38,$$

and

$$P(A \cap B) = P(A \cap C) = P(B \cap C) = 0.12,$$

and

$$P(A \cap B \cap C) = 0.05.$$

Find the probability that the student participates in at least one of these three programs, i.e., find $P(A \cup B \cup C)$.

Exercise 2.4. Dining with Dad. At a random meal during a parent weekend in the dining hall, a student notices the food chosen by her father. Let A, B, C be the events that his meal include Artichokes, Broccoli, or Cauliflower. These events have the property that

$$\begin{aligned} P(B) &= 0.39 \\ P(C) &= 0.44 \\ P(A \cap B) &= 0.13 \\ P(A \cap C) &= 0.12 \\ P(B \cap C) &= 0.13 \\ P(A \cap B \cap C) &= 0.09 \\ P(A \cup B \cup C) &= 0.89 \end{aligned}$$

What is the probability that the father includes Artichokes in his meal, i.e., what is $P(A)$?

Exercise 2.5. Selecting a pair of shoes. A woman has 3 pairs of sneakers, 8 pairs of flip flops, 6 pairs of flats, 4 pairs of wedges, and 9 pairs of high heels. (Hint for those of you unfamiliar with women's footwear: The wedges and high heels will make her look taller.)

a. What is the probability that she selects a pair of shoes that makes her taller if she pulls a pair from her closet without looking?

b. What is the probability that she selects a pair of shoes that does not make her look taller?

c. Create another type of partition for this woman's collection of shoes.

Exercise 2.6. Lollipops and licorice. In a kindergarten class, there are 30 children. Altogether, 19 of them like lollipops, and 10 of them like licorice (some like both). There are 8 students who don't like either of these. A child is chosen at random. What is the probability that the child likes both lollipops and licorice?

Exercise 2.7. Application to weather. Measure the amount of rainfall that occurs in your city for a year.

a. List 5 possible outcomes.

b. What is the sample space?

c. Devise a useful way to partition the sample space so that the partition sheds some insight for the general public about the annual rainfall.

d. Explain how the answer to part c meets the definition of a partition; see Definition 2.14.

Exercise 2.8. Coin flips. You flip a coin 5 times. What is the probability the first 4 are heads and the last one is a tail?

Exercise 2.9. Pizza meat. The guys on one floor of a college dorm all decide to get pizzas to share. They get 3 pepperoni pizzas, 2 bacon pizzas, 1 cheese pizza, 3 sausage pepperoni pizzas, and 3 meat lovers pizzas with sausage, pepperoni, and bacon. What is the probability of a randomly selected slice of pizza containing:

a. Bacon?

b. Pepperoni?

c. Sausage?

Exercise 2.10. Math and physics. In a class of 100 people, 60 of them are math majors, and 75 of them are physics majors. There are no students majoring in anything else in this class. This means that some of the students are double-majoring in both math and physics. A student is picked at random. What is the probability the student is double-majoring?

2.6.2 Extensions

Exercise 2.11. Monkey keystrokes. A monkey is let loose in a computer lab and starts playing with a keyboard. What is the probability that the monkey, without any comprehension or intention, types out the word "bananas" if he types exactly 7 keys? The typical keyboard has 101 keys, and the monkey only presses one key at a time.

Exercise 2.12. Apples. There are 6 apples in a basket. Two of them are red, and four are green.

a. What is the probability of selecting a red apple when choosing at random?

b. What is the probability that, if one apple is randomly chosen per day (and then eaten, not replaced), red apples are chosen on the first two days and green apples are chosen on the last four days?

Exercise 2.13. Shuffling and star ratings. I have 20 five-star songs and 200 four-star songs on my iPod, which has 2000 songs total.

a. What is the probability of shuffling to a five-star song?

b. What is the probability of shuffling to a four-star song?

c. What is the probability of shuffling to either a five- or four-star song?

d. What is the probability of shuffling to a song with fewer than four stars?

Exercise 2.14. Pizza toppings. Consider the following preferences: 35% of people like olive pizza, 54% like sausage pizza, and 12% like both olive and sausage pizza. What is the probability that a randomly chosen person likes neither olive pizza nor sausage pizza?

Exercise 2.15. Roll a die. If you roll a die, event A contains outcomes 1, 3, and 6; event B contains outcomes 1 and 6, and event C contains outcomes 4 and 6.

a. What is the probability of having $A \cup B \cup C$ occur?

b. What is the probability of having $A \cap B \cap C$ occur?

Exercise 2.16. DeMorgan's first law. For three events, A, B, C, consider the following probabilities: A, B, C have probability 20% of occurring, the probability of any two of these events occurring is 3% for each pair, and the probability of all three occurring is 1%. Using DeMorgan's first law, find the probability of $(A \cup B \cup C)^c$.

Exercise 2.17. Abstract art. A painter has three different jars of paint colors available, in colors green, yellow, and purple. She wants to paint something abstract, so she blindfolds herself, randomly dips her brush, and paints on the canvas. She continues trying paint jars until she finally gets some purple onto the canvas (her assistant will tell her when this happens) and then she stops.

Assume that she does not repeat any of the jars because her assistant removes a jar once it has been used. So the sample space is

$$S = \{(P), (G, P), (Y, P), (Y, G, P), (G, Y, P)\}.$$

Find the probabilities of each of the following events:

$$\{(P)\}, \quad \{(G, P), (Y, P)\}, \quad \{(G, P), (G, Y, P)\},$$
$$\{(Y, G, P), (G, Y, P)\}, \quad \{(P), (Y, P)\}$$

Exercise 2.18. Yahtzee. In the game Yahtzee, there are 5 dice with 6 possible numbers on each. What is the probability for a Yahtzee on a player's first roll? (In other words, what is the probability that all 5 dice show the same number the first time that they are rolled)?

Exercise 2.19. Locker combinations. You just forgot your locker combination and are too embarrassed to ask for it. You know for sure that the first number is 22, or was it 32? It's one of those. You're certain that the middle number is a one-digit number (0–9), and the last number could be anything between 0 and 45. If the lock is a 3-number lock with numbers 0 through 45, what is the maximum number of tries needed to open it, assuming you don't repeat any combinations?

Exercise 2.20. Mathematical Science majors. There are 89 students in a volunteer group for majors in math, statistics, or actuarial science. Thirteen students have a double-major in math and statistics. Twelve students have a double major in math and actuarial science. Thirteen students have a double major in statistics and actuarial science. There are 9 students doing a triple major. Thirty-nine students are majoring in at least statistics, and forty-four are majoring in at least actuarial science. How many students are majoring in at least mathematics?

Exercise 2.21. Postal customers. A sequence of seven people walk into a post office and only their sexes are noted (in sequence) as they enter.

a. How many outcomes are in the sample space?

b. Let A_2 be the event that exactly two of the people are females. How many outcomes are in A_2?

c. Let A_j be the event that exactly j of the people are females. How many outcomes are in A_j?

d. If each of the seven customers is equally likely to be male or female, what probability should be associated with event A_j?

Exercise 2.22. Deducing a probability. Events A, B, C are to be considered with the following properties:

$$P(A) = 0.17$$
$$P(B) = 0.37$$
$$P(C) = 0.19$$
$$P(A \cap B) = 0.07$$
$$P(B \cap C) = 0.11$$
$$P(A \cap B \cap C) = 0.03$$
$$P(A \cup B \cup C) = 0.48$$

Find the probability of $A \cap C$.

Exercise 2.23. Mystery probability. Suppose there are 3 events such that

$$P(A) = 0.20$$
$$P(B) = 0.10$$
$$P(C) = 0.40$$
$$P(A \cap B) = 0.05$$
$$P(A \cap C) = 0.10$$
$$P(B \cap C) = 0.03$$
$$P(A \cap B \cap C) = 0.01$$

What is the probability that none of the events happens?

Exercise 2.24. Prove Theorem 2.23.

Exercise 2.25. Prove Theorem 2.24.

2.6.3 Advanced

Exercise 2.26. Prove Theorem 2.25.

Exercise 2.27. Is the whole smaller than the sum of the parts?

a. It is always true, for any events A, B, that $P(A \cup B) \leq P(A) + P(B)$. Why? Explain briefly with words or a very clear picture.

b. Is it always true that $P(A \cup B \cup C) \leq P(A) + P(B) + P(C)$? If so, explain why, either using words or a very clear picture. If not, please give a counterexample.

Exercise 2.28. Grabbing a pen. You find a container of 27 old pens in your school supplies and continue to test them (without replacement), until you find one that works. If each individual pen works 25% of the time (regardless of the other pens), what is the probability that you find one that works within the first four tries?

Exercise 2.29. Die rolls. You roll a die three times. What is the probability the sum of the first two rolls is equal to the third roll?

Exercise 2.30. Cookies. Consider a jar of 9 chocolate chip and 11 peanut butter cookies. You randomly select 2 cookies to eat. All possible choices are equally likely.

a. What is the probability that the 2 you select will both will be chocolate chip?

b. What is the probability that at least one of your cookies will be peanut butter?

c. What is the probability that last 2 cookies left in the jar (after 18 have been eaten) will be chocolate chip? (Is this answer the same or different than part a? Why or why not?)

Exercise 2.31. Seating arrangements. Alice, Bob, Catherine, Doug, and Edna are randomly assigned seats at a circular table in a perfectly circular room. Assume that rotations of the table do not matter, so there are exactly 24 possible outcomes in the sample space.

Bob and Catherine are married. Doug and Edna are married.

Let A_j denote the event that exactly j of the married couples are happy because they are sitting together. Find $P(A_0)$ and $P(A_1)$ and $P(A_2)$.

Exercise 2.32. Socks. In your drawer you have 10 white socks, 6 black socks, 4 red socks, and 2 purple socks. Your roommate is still asleep, and you can't turn the light on while you're getting dressed. You reach in blindly and grab two socks. What is the probability of pulling out a matching pair of purple socks?

Exercise 2.33. Maximum of three dice. Roll three distinguishable dice (e.g., assume that there is a way to tell them apart, for instance, that the dice are three different colors). There are $6 \times 6 \times 6 = 216$ possible outcomes.

Let B_k be the event that the maximum value that appears on all three dice when they are rolled is less than or equal to k. Find $P(B_1)$, $P(B_2)$, $P(B_3)$, $P(B_4)$, $P(B_5)$, and $P(B_6)$. If you prefer, you are welcome to just give a general formula that covers all six of these cases, i.e., you are welcome to just give a formula for $P(B_k)$ itself.

Exercise 2.34. If $A_1, A_2, \ldots, A_n$ is a collection of events, is it always true that

$$P\Big(\bigcup_{k=1}^{n} A_k\Big) \le \sum_{k=1}^{n} P(A_k)?$$

Prove this inequality, or give a counterexample.

Exercise 2.35. Tickets. Consider n people who drop a ticket into a box. The box of tickets is thoroughly shaken and randomized. Each person then draws one ticket, without replacement. A person is a winner if she selects her own ticket.

a. What is the probability that nobody is a winner?

b. Find the limit of the probability in part a, as $n \to \infty$.