Chapter 2

Probability (special topic)

Probability forms a foundation for statistics. You might already be familiar with many aspects of probability, however, formalization of the concepts is new for most. This chapter aims to introduce probability on familiar terms using processes most people have seen before.

2.1 Defining probability (special topic)

● **Example 2.1** A "die", the singular of dice, is a cube with six faces numbered 1, 2, 3, 4, 5, and 6. What is the chance of getting 1 when rolling a die?

If the die is fair, then the chance of a 1 is as good as the chance of any other number. Since there are six outcomes, the chance must be 1-in-6 or, equivalently, 1/6.

● **Example 2.2** What is the chance of getting a 1 or 2 in the next roll?

1 and 2 constitute two of the six equally likely possible outcomes, so the chance of getting one of these two outcomes must be $2/6 = 1/3$.

● **Example 2.3** What is the chance of getting either 1, 2, 3, 4, 5, or 6 on the next roll?

100%. The outcome must be one of these numbers.

● **Example 2.4** What is the chance of not rolling a 2?

Since the chance of rolling a 2 is 1/6 or $16.\bar{6}\%$, the chance of not rolling a 2 must be $100\% - 16.\bar{6}\% = 83.\bar{3}\%$ or 5/6.

Alternatively, we could have noticed that not rolling a 2 is the same as getting a 1, 3, 4, 5, or 6, which makes up five of the six equally likely outcomes and has probability 5/6.

● **Example 2.5** Consider rolling two dice. If $1/6^{th}$ of the time the first die is a 1 and $1/6^{th}$ of those times the second die is a 1, what is the chance of getting two 1s?

If $16.\bar{6}\%$ of the time the first die is a 1 and $1/6^{th}$ of *those* times the second die is also a 1, then the chance that both dice are 1 is $(1/6) \times (1/6)$ or 1/36.
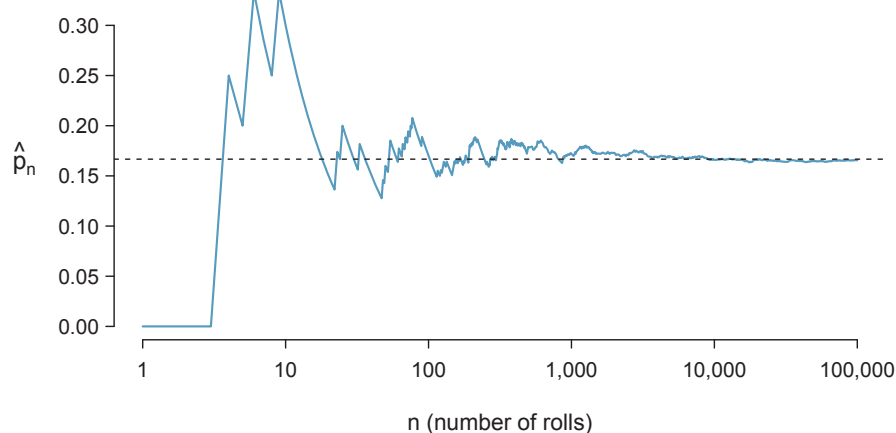
Figure 2.1: The fraction of die rolls that are 1 at each stage in a simulation. The proportion tends to get closer to the probability $1/6 \approx 0.167$ as the number of rolls increases.

2.1.1 Probability

We use probability to build tools to describe and understand apparent randomness. We often frame probability in terms of a **random process** giving rise to an **outcome**.

$$\text{Roll a die} \quad \rightarrow \quad \text{1, 2, 3, 4, 5, or 6}$$
$$\text{Flip a coin} \quad \rightarrow \quad \text{H or T}$$

Rolling a die or flipping a coin is a seemingly random process and each gives rise to an outcome.

> **Probability**
> The **probability** of an outcome is the proportion of times the outcome would occur if we observed the random process an infinite number of times.

Probability is defined as a proportion, and it always takes values between 0 and 1 (inclusively). It may also be displayed as a percentage between 0% and 100%.

Probability can be illustrated by rolling a die many times. Let $\hat{p}_n$ be the proportion of outcomes that are 1 after the first n rolls. As the number of rolls increases, $\hat{p}_n$ will converge to the probability of rolling a 1, $p = 1/6$. Figure 2.1 shows this convergence for 100,000 die rolls. The tendency of $\hat{p}_n$ to stabilize around p is described by the **Law of Large Numbers**.

> **Law of Large Numbers**
> As more observations are collected, the proportion $\hat{p}_n$ of occurrences with a particular outcome converges to the probability p of that outcome.

Occasionally the proportion will veer off from the probability and appear to defy the Law of Large Numbers, as $\hat{p}_n$ does many times in Figure 2.1. However, these deviations become smaller as the number of rolls increases.

Above we write p as the probability of rolling a 1. We can also write this probability as

$P(A)$

Probability of
outcome A

$$P(\text{rolling a } 1)$$

As we become more comfortable with this notation, we will abbreviate it further. For instance, if it is clear that the process is "rolling a die", we could abbreviate $P(\text{rolling a } 1)$ as $P(1)$.

> ⊙ **Exercise 2.6**   Random processes include rolling a die and flipping a coin. (a) Think of another random process. (b) Describe all the possible outcomes of that process. For instance, rolling a die is a random process with potential outcomes 1, 2, ..., 6.[1]

What we think of as random processes are not necessarily random, but they may just be too difficult to understand exactly. The fourth example in the footnote solution to Exercise 2.6 suggests a roommate's behavior is a random process. However, even if a roommate's behavior is not truly random, modeling her behavior as a random process can still be useful.

> **TIP: Modeling a process as random**
> It can be helpful to model a process as random even if it is not truly random.

2.1.2   Disjoint or mutually exclusive outcomes

Two outcomes are called **disjoint** or **mutually exclusive** if they cannot both happen. For instance, if we roll a die, the outcomes 1 and 2 are disjoint since they cannot both occur. On the other hand, the outcomes 1 and "rolling an odd number" are not disjoint since both occur if the outcome of the roll is a 1. The terms *disjoint* and *mutually exclusive* are equivalent and interchangeable.

Calculating the probability of disjoint outcomes is easy. When rolling a die, the outcomes 1 and 2 are disjoint, and we compute the probability that one of these outcomes will occur by adding their separate probabilities:

$$P(1 \text{ or } 2) = P(1) + P(2) = 1/6 + 1/6 = 1/3$$

What about the probability of rolling a 1, 2, 3, 4, 5, or 6? Here again, all of the outcomes are disjoint so we add the probabilities:

$$P(1 \text{ or } 2 \text{ or } 3 \text{ or } 4 \text{ or } 5 \text{ or } 6)$$
$$= P(1) + P(2) + P(3) + P(4) + P(5) + P(6)$$
$$= 1/6 + 1/6 + 1/6 + 1/6 + 1/6 + 1/6 = 1.$$

The **Addition Rule** guarantees the accuracy of this approach when the outcomes are disjoint.

---

[1]Here are four examples. (i) Whether someone gets sick in the next month or not is an apparently random process with outcomes `sick` and `not`. (ii) We can *generate* a random process by randomly picking a person and measuring that person's height. The outcome of this process will be a positive number. (iii) Whether the stock market goes up or down next week is a seemingly random process with possible outcomes `up`, `down`, and `no_change`. Alternatively, we could have used the percent change in the stock market as a numerical outcome. (iv) Whether your roommate cleans her dishes tonight probably seems like a random process with possible outcomes `cleans_dishes` and `leaves_dishes`.

> **Addition Rule of disjoint outcomes**
> If A_1 and A_2 represent two disjoint outcomes, then the probability that one of them occurs is given by
>
> $$P(A_1 \text{ or } A_2) = P(A_1) + P(A_2)$$
>
> If there are many disjoint outcomes A_1, ..., A_k, then the probability that one of these outcomes will occur is
>
> $$P(A_1) + P(A_2) + \cdots + P(A_k) \tag{2.7}$$

⊙ **Exercise 2.8** We are interested in the probability of rolling a 1, 4, or 5. (a) Explain why the outcomes 1, 4, and 5 are disjoint. (b) Apply the Addition Rule for disjoint outcomes to determine $P(1 \text{ or } 4 \text{ or } 5)$.[2]

⊙ **Exercise 2.9** In the email data set in Chapter 1, the number variable described whether no number (labeled none), only one or more small numbers (small), or whether at least one big number appeared in an email (big). Of the 3,921 emails, 549 had no numbers, 2,827 had only one or more small numbers, and 545 had at least one big number. (a) Are the outcomes none, small, and big disjoint? (b) Determine the proportion of emails with value small and big separately. (c) Use the Addition Rule for disjoint outcomes to compute the probability a randomly selected email from the data set has a number in it, small or big.[3]

Statisticians rarely work with individual outcomes and instead consider *sets* or *collections* of outcomes. Let A represent the event where a die roll results in 1 or 2 and B represent the event that the die roll is a 4 or a 6. We write A as the set of outcomes {1, 2} and $B = \{4, 6\}$. These sets are commonly called **events**. Because A and B have no elements in common, they are disjoint events. A and B are represented in Figure 2.2.
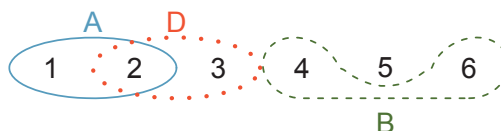

Figure 2.2: Three events, A, B, and D, consist of outcomes from rolling a die. A and B are disjoint since they do not have any outcomes in common.

The Addition Rule applies to both disjoint outcomes and disjoint events. The probability that one of the disjoint events A or B occurs is the sum of the separate probabilities:

$$P(A \text{ or } B) = P(A) + P(B) = 1/3 + 1/3 = 2/3$$

⊙ **Exercise 2.10** (a) Verify the probability of event A, $P(A)$, is 1/3 using the Addition Rule. (b) Do the same for event B.[4]

---

[2](a) The random process is a die roll, and at most one of these outcomes can come up. This means they are disjoint outcomes. (b) $P(1 \text{ or } 4 \text{ or } 5) = P(1) + P(4) + P(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$

[3](a) Yes. Each email is categorized in only one level of number. (b) Small: $\frac{2827}{3921} = 0.721$. Big: $\frac{545}{3921} = 0.139$. (c) $P(\text{small or big}) = P(\text{small}) + P(\text{big}) = 0.721 + 0.139 = 0.860$.

[4](a) $P(A) = P(1 \text{ or } 2) = P(1) + P(2) = \frac{1}{6} + \frac{1}{6} = \frac{2}{6} = \frac{1}{3}$. (b) Similarly, $P(B) = 1/3$.

2♣	3♣	4♣	5♣	6♣	7♣	8♣	9♣	10♣	J♣	Q♣	K♣	A♣
2♢	3♢	4♢	5♢	6♢	7♢	8♢	9♢	10♢	J♢	Q♢	K♢	A♢
2♡	3♡	4♡	5♡	6♡	7♡	8♡	9♡	10♡	J♡	Q♡	K♡	A♡
2♠	3♠	4♠	5♠	6♠	7♠	8♠	9♠	10♠	J♠	Q♠	K♠	A♠

Table 2.3: Representations of the 52 unique cards in a deck.

⊙ **Exercise 2.11**   (a) Using Figure 2.2 as a reference, what outcomes are represented by event D? (b) Are events B and D disjoint? (c) Are events A and D disjoint?[5]

⊙ **Exercise 2.12**   In Exercise 2.11, you confirmed B and D from Figure 2.2 are disjoint. Compute the probability that either event B or event D occurs.[6]

2.1.3   Probabilities when events are not disjoint

Let's consider calculations for two events that are not disjoint in the context of a regular deck of 52 cards, represented in Table 2.3. If you are unfamiliar with the cards in a regular deck, please see the footnote.[7]

⊙ **Exercise 2.13**   (a) What is the probability that a randomly selected card is a diamond? (b) What is the probability that a randomly selected card is a face card?[8]

Venn diagrams are useful when outcomes can be categorized as "in" or "out" for two or three variables, attributes, or random processes. The Venn diagram in Figure 2.4 uses a circle to represent diamonds and another to represent face cards. If a card is both a diamond and a face card, it falls into the intersection of the circles. If it is a diamond but not a face card, it will be in part of the left circle that is not in the right circle (and so on). The total number of cards that are diamonds is given by the total number of cards in the diamonds circle: $10 + 3 = 13$. The probabilities are also shown (e.g. $10/52 = 0.1923$).

⊙ **Exercise 2.14**   Using the Venn diagram, verify $P(\text{face card}) = 12/52 = 3/13$.[9]

Let A represent the event that a randomly selected card is a diamond and B represent the event that it is a face card. How do we compute $P(A \text{ or } B)$? Events A and B are not disjoint – the cards $J\diamondsuit$, $Q\diamondsuit$, and $K\diamondsuit$ fall into both categories – so we cannot use the

---

[5](a) Outcomes 2 and 3. (b) Yes, events B and D are disjoint because they share no outcomes. (c) The events A and D share an outcome in common, 2, and so are not disjoint.

[6]Since B and D are disjoint events, use the Addition Rule: $P(B \text{ or } D) = P(B) + P(D) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$.

[7]The 52 cards are split into four **suits**: ♣ (club), ♢ (diamond), ♡ (heart), ♠ (spade). Each suit has its 13 cards labeled: 2, 3, ..., 10, J (jack), Q (queen), K (king), and A (ace). Thus, each card is a unique combination of a suit and a label, e.g. 4♡ and J♣. The 12 cards represented by the jacks, queens, and kings are called **face cards**. The cards that are ♢ or ♡ are typically colored red while the other two suits are typically colored black.

[8](a) There are 52 cards and 13 diamonds. If the cards are thoroughly shuffled, each card has an equal chance of being drawn, so the probability that a randomly selected card is a diamond is $P(\diamondsuit) = \frac{13}{52} = 0.250$. (b) Likewise, there are 12 face cards, so $P(\text{face card}) = \frac{12}{52} = \frac{3}{13} = 0.231$.

[9]The Venn diagram shows face cards split up into "face card but not ♢" and "face card and ♢". Since these correspond to disjoint events, $P(\text{face card})$ is found by adding the two corresponding probabilities: $\frac{3}{52} + \frac{9}{52} = \frac{12}{52} = \frac{3}{13}$.
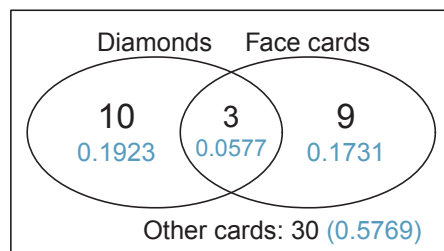
Figure 2.4: A Venn diagram for diamonds and face cards.

Addition Rule for disjoint events. Instead we use the Venn diagram. We start by adding the probabilities of the two events:

$$P(A) + P(B) = P(\diamondsuit) + P(\text{face card}) = 12/52 + 13/52$$

However, the three cards that are in both events were counted twice, once in each probability. We must correct this double counting:

$$
\begin{aligned}
P(A \text{ or } B) &= P(\text{face card or } \diamondsuit) \\
&= P(\text{face card}) + P(\diamondsuit) - P(\text{face card and } \diamondsuit) \qquad (2.15) \\
&= 12/52 + 13/52 - 3/52 \\
&= 22/52 = 11/26
\end{aligned}
$$

Equation (2.15) is an example of the **General Addition Rule**.

General Addition Rule

If A and B are any two events, disjoint or not, then the probability that at least one of them will occur is

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B) \qquad (2.16)$$

where $P(A \text{ and } B)$ is the probability that both events occur.

TIP: "or" is inclusive

When we write "or" in statistics, we mean "and/or" unless we explicitly state otherwise. Thus, A or B occurs means A, B, or both A and B occur.

⊙ **Exercise 2.17** (a) If A and B are disjoint, describe why this implies $P(A$ and $B) = 0$. (b) Using part (a), verify that the General Addition Rule simplifies to the simpler Addition Rule for disjoint events if A and B are disjoint.[10]

---

[10](a) If A and B are disjoint, A and B can never occur simultaneously. (b) If A and B are disjoint, then the last term of Equation (2.16) is 0 (see part (a)) and we are left with the Addition Rule for disjoint events.

⊙ **Exercise 2.18** In the `email` data set with 3,921 emails, 367 were spam, 2,827 contained some small numbers but no big numbers, and 168 had both characteristics. Create a Venn diagram for this setup.[11]

⊙ **Exercise 2.19** (a) Use your Venn diagram from Exercise 2.18 to determine the probability a randomly drawn email from the `email` data set is spam and had small numbers (but not big numbers). (b) What is the probability that the email had either of these attributes?[12]

2.1.4 Probability distributions

A **probability distribution** is a table of all disjoint outcomes and their associated probabilities. Table 2.5 shows the probability distribution for the sum of two dice.

Dice sum	2	3	4	5	6	7	8	9	10	11	12
Probability	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Table 2.5: Probability distribution for the sum of two dice.

Rules for probability distributions
A probability distribution is a list of the possible outcomes with corresponding probabilities that satisfies three rules:

1. The outcomes listed must be disjoint.

2. Each probability must be between 0 and 1.

3. The probabilities must total 1.

⊙ **Exercise 2.20** Table 2.6 suggests three distributions for household income in the United States. Only one is correct. Which one must it be? What is wrong with the other two?[13]

Chapter 1 emphasized the importance of plotting data to provide quick summaries. Probability distributions can also be summarized in a bar plot. For instance, the distribution of US household incomes is shown in Figure 2.7 as a bar plot.[14] The probability distribution for the sum of two dice is shown in Table 2.5 and plotted in Figure 2.8.

In these bar plots, the bar heights represent the probabilities of outcomes. If the outcomes are numerical and discrete, it is usually (visually) convenient to make a bar plot that resembles a histogram, as in the case of the sum of two dice. Another example of plotting the bars at their respective locations is shown in Figure 2.20 on page 96.

---

[11]Both the counts and corresponding probabilities (e.g. 2659/3921 = 0.678) are shown. Notice that the number of emails represented in the left circle corresponds to 2659 + 168 = 2827, and the number represented in the right circle is 168 + 199 = 367.

[12](a) The solution is represented by the intersection of the two circles: 0.043. (b) This is the sum of the three disjoint probabilities shown in the circles: 0.678 + 0.043 + 0.051 = 0.772.

[13]The probabilities of (a) do not sum to 1. The second probability in (b) is negative. This leaves (c), which sure enough satisfies the requirements of a distribution. One of the three was said to be the actual distribution of US household incomes, so it must be (c).

[14]It is also possible to construct a distribution plot when income is not artificially binned into four groups. *Continuous* distributions are considered in Section 2.5.
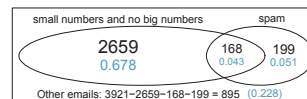
Income range ($1000s)	0-25	25-50	50-100	100+
(a)	0.18	0.39	0.33	0.16
(b)	0.38	-0.27	0.52	0.37
(c)	0.28	0.27	0.29	0.16

Table 2.6: Proposed distributions of US household incomes (Exercise 2.20).
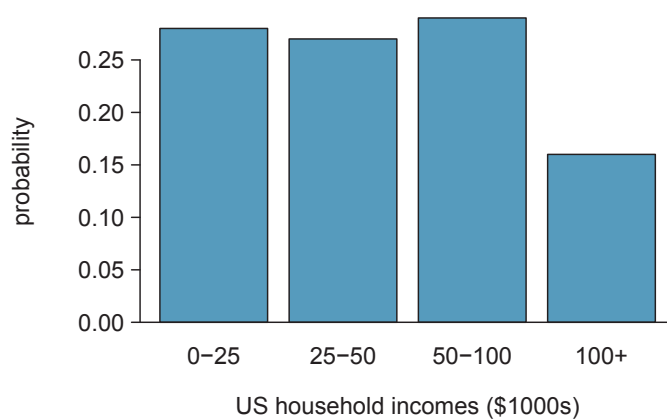

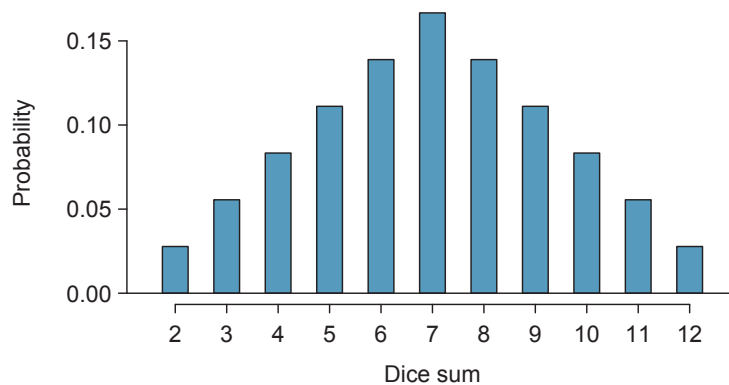Figure 2.7: The probability distribution of US household income.

Figure 2.8: The probability distribution of the sum of two dice.

2.1.5 Complement of an event

S
Sample space

Rolling a die produces a value in the set $\{1, 2, 3, 4, 5, 6\}$. This set of all possible outcomes is called the **sample space** (S) for rolling a die. We often use the sample space to examine the scenario where an event does not occur.

A^c
Complement
of outcome A

Let $D = \{2, 3\}$ represent the event that the outcome of a die roll is 2 or 3. Then the **complement** of D represents all outcomes in our sample space that are not in D, which is denoted by $D^c = \{1, 4, 5, 6\}$. That is, D^c is the set of all possible outcomes not already included in D. Figure 2.9 shows the relationship between D, D^c, and the sample space S.
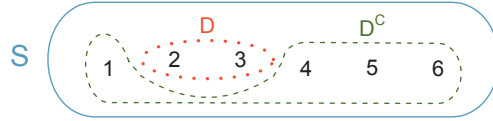

Figure 2.9: Event $D = \{2, 3\}$ and its complement, $D^c = \{1, 4, 5, 6\}$. S represents the sample space, which is the set of all possible events.

⊙ **Exercise 2.21** (a) Compute $P(D^c) = P(\text{rolling a 1, 4, 5, or 6})$. (b) What is $P(D) + P(D^c)$?[15]

⊙ **Exercise 2.22** Events $A = \{1, 2\}$ and $B = \{4, 6\}$ are shown in Figure 2.2 on page 71. (a) Write out what A^c and B^c represent. (b) Compute $P(A^c)$ and $P(B^c)$. (c) Compute $P(A) + P(A^c)$ and $P(B) + P(B^c)$.[16]

A complement of an event A is constructed to have two very important properties: (i) every possible outcome not in A is in A^c, and (ii) A and A^c are disjoint. Property (i) implies

$$P(A \text{ or } A^c) = 1 \tag{2.23}$$

That is, if the outcome is not in A, it must be represented in A^c. We use the Addition Rule for disjoint events to apply Property (ii):

$$P(A \text{ or } A^c) = P(A) + P(A^c) \tag{2.24}$$

Combining Equations (2.23) and (2.24) yields a very useful relationship between the probability of an event and its complement.

Complement

The complement of event A is denoted A^c, and A^c represents all outcomes not in A. A and A^c are mathematically related:

$$P(A) + P(A^c) = 1, \quad \text{i.e.} \quad P(A) = 1 - P(A^c) \tag{2.25}$$

In simple examples, computing A or A^c is feasible in a few steps. However, using the complement can save a lot of time as problems grow in complexity.

---

[15](a) The outcomes are disjoint and each has probability 1/6, so the total probability is $4/6 = 2/3$. (b) We can also see that $P(D) = \frac{1}{6} + \frac{1}{6} = 1/3$. Since D and D^c are disjoint, $P(D) + P(D^c) = 1$.
[16]Brief solutions: (a) $A^c = \{3, 4, 5, 6\}$ and $B^c = \{1, 2, 3, 5\}$. (b) Noting that each outcome is disjoint, add the individual outcome probabilities to get $P(A^c) = 2/3$ and $P(B^c) = 2/3$. (c) A and A^c are disjoint, and the same is true of B and B^c. Therefore, $P(A) + P(A^c) = 1$ and $P(B) + P(B^c) = 1$.

⊙ **Exercise 2.26**   Let A represent the event where we roll two dice and their total is less than 12. (a) What does the event A^c represent? (b) Determine $P(A^c)$ from Table 2.5 on page 74. (c) Determine $P(A)$.[17]

⊙ **Exercise 2.27**   Consider again the probabilities from Table 2.5 and rolling two dice. Find the following probabilities: (a) The sum of the dice is *not* 6. (b) The sum is at least 4. That is, determine the probability of the event $B = \{4, 5, ..., 12\}$. (c) The sum is no more than 10. That is, determine the probability of the event $D = \{2, 3, ..., 10\}$.[18]

2.1.6   Independence

Just as variables and observations can be independent, random processes can be independent, too. Two processes are **independent** if knowing the outcome of one provides no useful information about the outcome of the other. For instance, flipping a coin and rolling a die are two independent processes – knowing the coin was heads does not help determine the outcome of a die roll. On the other hand, stock prices usually move up or down together, so they are not independent.

Example 2.5 provides a basic example of two independent processes: rolling two dice. We want to determine the probability that both will be 1. Suppose one of the dice is red and the other white. If the outcome of the red die is a 1, it provides no information about the outcome of the white die. We first encountered this same question in Example 2.5 (page 68), where we calculated the probability using the following reasoning: $1/6^{th}$ of the time the red die is a 1, and $1/6^{th}$ of *those* times the white die will also be 1. This is illustrated in Figure 2.10. Because the rolls are independent, the probabilities of the corresponding outcomes can be multiplied to get the final answer: $(1/6) \times (1/6) = 1/36$. This can be generalized to many independent processes.
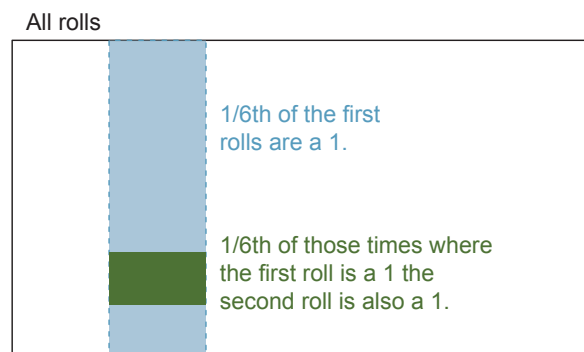

Figure 2.10: $1/6^{th}$ of the time, the first roll is a 1. Then $1/6^{th}$ of *those* times, the second roll will also be a 1.

---

[17](a) The complement of A: when the total is equal to 12. (b) $P(A^c) = 1/36$. (c) Use the probability of the complement from part (b), $P(A^c) = 1/36$, and Equation (2.25): $P(\text{less than } 12) = 1 - P(12) = 1 - 1/36 = 35/36$.

[18](a) First find $P(6) = 5/36$, then use the complement: $P(\text{not } 6) = 1 - P(6) = 31/36$.

(b) First find the complement, which requires much less effort: $P(2 \text{ or } 3) = 1/36 + 2/36 = 1/12$. Then calculate $P(B) = 1 - P(B^c) = 1 - 1/12 = 11/12$.

(c) As before, finding the complement is the clever way to determine $P(D)$. First find $P(D^c) = P(11 \text{ or } 12) = 2/36 + 1/36 = 1/12$. Then calculate $P(D) = 1 - P(D^c) = 11/12$.

● **Example 2.28** What if there was also a blue die independent of the other two? What is the probability of rolling the three dice and getting all 1s?

The same logic applies from Example 2.5. If $1/36^{th}$ of the time the white and red dice are both 1, then $1/6^{th}$ of *those* times the blue die will also be 1, so multiply:

$$P(white = 1 \text{ and } red = 1 \text{ and } blue = 1) = P(white = 1) \times P(red = 1) \times P(blue = 1)$$
$$= (1/6) \times (1/6) \times (1/6) = 1/216$$

Examples 2.5 and 2.28 illustrate what is called the Multiplication Rule for independent processes.

Multiplication Rule for independent processes
If A and B represent events from two different and independent processes, then the probability that both A and B occur can be calculated as the product of their separate probabilities:

$$P(A \text{ and } B) = P(A) \times P(B) \tag{2.29}$$

Similarly, if there are k events A_1, ..., A_k from k independent processes, then the probability they all occur is

$$P(A_1) \times P(A_2) \times \cdots \times P(A_k)$$

⊙ **Exercise 2.30** About 9% of people are left-handed. Suppose 2 people are selected at random from the U.S. population. Because the sample size of 2 is very small relative to the population, it is reasonable to assume these two people are independent. (a) What is the probability that both are left-handed? (b) What is the probability that both are right-handed?[19]

⊙ **Exercise 2.31** Suppose 5 people are selected at random.[20]

(a) What is the probability that all are right-handed?
(b) What is the probability that all are left-handed?
(c) What is the probability that not all of the people are right-handed?

---

[19](a) The probability the first person is left-handed is 0.09, which is the same for the second person. We apply the Multiplication Rule for independent processes to determine the probability that both will be left-handed: $0.09 \times 0.09 = 0.0081$.
(b) It is reasonable to assume the proportion of people who are ambidextrous (both right and left handed) is nearly 0, which results in $P(\text{right-handed}) = 1 - 0.09 = 0.91$. Using the same reasoning as in part (a), the probability that both will be right-handed is $0.91 \times 0.91 = 0.8281$.
[20](a) The abbreviations RH and LH are used for right-handed and left-handed, respectively. Since each are independent, we apply the Multiplication Rule for independent processes:

$$P(\text{all five are RH}) = P(\text{first} = \text{RH, second} = \text{RH, ..., fifth} = \text{RH})$$
$$= P(\text{first} = \text{RH}) \times P(\text{second} = \text{RH}) \times \cdots \times P(\text{fifth} = \text{RH})$$
$$= 0.91 \times 0.91 \times 0.91 \times 0.91 \times 0.91 = 0.624$$

(b) Using the same reasoning as in (a), $0.09 \times 0.09 \times 0.09 \times 0.09 \times 0.09 = 0.0000059$
(c) Use the complement, $P(\text{all five are RH})$, to answer this question:

$$P(\text{not all RH}) = 1 - P(\text{all RH}) = 1 - 0.624 = 0.376$$

Suppose the variables `handedness` and `gender` are independent, i.e. knowing some-one's `gender` provides no useful information about their `handedness` and vice-versa. Then we can compute whether a randomly selected person is right-handed and female[21] using the Multiplication Rule:

$$
\begin{aligned}
P(\text{right-handed and female}) &= P(\text{right-handed}) \times P(\text{female}) \\
&= 0.91 \times 0.50 = 0.455
\end{aligned}
$$

⊙ **Exercise 2.32** Three people are selected at random.[22]
  (a) What is the probability that the first person is male and right-handed?
  (b) What is the probability that the first two people are male and right-handed?.
  (c) What is the probability that the third person is female and left-handed?
  (d) What is the probability that the first two people are male and right-handed and the third person is female and left-handed?

Sometimes we wonder if one outcome provides useful information about another out-come. The question we are asking is, are the occurrences of the two events independent? We say that two events A and B are independent if they satisfy Equation (2.29).

● **Example 2.33** If we shuffle up a deck of cards and draw one, is the event that the card is a heart independent of the event that the card is an ace?

The probability the card is a heart is 1/4 and the probability that it is an ace is 1/13. The probability the card is the ace of hearts is 1/52. We check whether Equation 2.29 is satisfied:

$$
P(\heartsuit) \times P(\text{ace}) = \frac{1}{4} \times \frac{1}{13} = \frac{1}{52} = P(\heartsuit \text{ and ace})
$$

Because the equation holds, the event that the card is a heart and the event that the card is an ace are independent events.

2.2 Conditional probability (special topic)

Are students more likely to use marijuana when their parents used drugs? The `drug_use` data set contains a sample of 445 cases with two variables, `student` and `parents`, and is summarized in Table 2.11.[23] The `student` variable is either `uses` or `not`, where a student is labeled as `uses` if she has recently used marijuana. The `parents` variable takes the value `used` if at least one of the parents used drugs, including alcohol.

● **Example 2.34** If at least one parent used drugs, what is the chance their child (`student`) uses?

We will estimate this probability using the data. Of the 210 cases in this data set where `parents = used`, 125 represent cases where `student = uses`:

$$
P(\texttt{student = uses given parents = used}) = \frac{125}{210} = 0.60
$$

---

[21]The actual proportion of the U.S. population that is `female` is about 50%, and so we use 0.5 for the probability of sampling a woman. However, this probability does differ in other countries.

[22]Brief answers are provided. (a) This can be written in probability notation as P(a randomly selected person is male and right-handed) = 0.455. (b) 0.207. (c) 0.045. (d) 0.0093.

[23]Ellis GJ and Stone LH. 1979. Marijuana Use in College: An Evaluation of a Modeling Explanation. Youth and Society 10:323-334.