MATH 2510-001: PROBABILITY IN-DEPTH

DIEZ, ÇETINKAYA-RUNDEL, BARR (EDITED BY COLTON)
SEPTEMBER 17, 2019

This was prepared for University of Colorado's MATH 2510-001 (Introduction to Statistics) in Fall 2019. Please send questions, comments, complaints, and corrections to colton.grainger@colorado.edu. Any mistakes are my own. These notes are scraped from the "Probability" chapter of OpenIntro Statistics v4.0.

CONTENTS

Lecture 2: Probability notions

A **probability space** consists of a set Ω called **laboratory** a set of subsets of Ω called **events** and a function P from events to the interval $[0, 1]$ so that probabilities of disjoint sets add up and such that the entire laboratory has probability 1. Every point in Ω is an **experiment**. Think of an event A as a collection of experiments and of $P[A]$ as the likelihood that A occurs.

Examples:

1. We turn a wheel of fortune and assume it is fair in the sense that every angle range $[a, b]$ appears with probability $(b-a)/2\pi$. What is the chance that the wheel stops with an angle between 30 and 90 degrees?
   Answer: The laboratory here is the circle $[0, 2\pi)$. Every point in this circle is a possible experiment. The event that the wheel stops between 30 and 90 degrees is the interval $[\pi/6, \pi/2]$. Assuming that all angles are equally probable, the answer is $1/6$.

2. This example is called **Bertrand's paradox**. Assume we throw randomly a line into the unit disc. What is the probability that its length is larger than the length of the inscribed triangle?
   Answer: Interestingly, the answer depends as we will see in the lecture.

3. Lets look at the DowJonesIndustrial average DJI from the start. What is the probability that the index will double in the next 50 years?
   Answer: This is a strange question because we have **only one** data set. How can we talk about probability in this situation? One way is to see this graph as a sample of a larger probability space. A simple model would be to fit the data with some polynomial, then add random noise to it. The real DJI graph now looks very similar to a typical graph of those.

4. Lets look at the digits of π. What is the probability that the digit 5 appears? **Answer:** Also this is a strange example since the digits are not randomly generated. They are given by nature. There is no randomness involved. Still, one observes that the digits behave like a random number and that the number is "normal": every digit appears with the same frequency. This is independent of the base.

   Here is a more precise list of conditions which need to be satisfied for events.

   1. The entire laboratory Ω and the empty set $\emptyset$ are events.
   2. If A_j is a sequence of events, then $\bigcup_j A_j$ and $\bigcap_j A_j$ are events.

   It follows that also the complement of an event is an event.

   Here are the conditions which need to be satisfied for the **probability function** P:

   1. $0 \le P[A] \le 1$ and $P[\Omega] = 1$.
   2. A_j are disjoint events, then $P[\bigcup_{j=1}^{\infty} A_j] = \sum_{j=1}^{\infty} P[A_j]$.

FIGURE 1. 2011 Lecture notes from my hero, Oliver Knill.

Index of Terms and Keywords

WHAT'S THIS NEXT SECTION ALL ABOUT?   Introductory examples

Before we get into technical ideas, let's walk through some basic examples that may feel more familiar.

Example 1.1. 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 1.2. 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 1.3. 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 1.4. 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 1.5. Consider rolling two dice. If $1/6$ of the time the first die is a 1 and $1/6$ 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 $1/6$ 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 $1/36$. ◄

WHAT'S THIS NEXT SECTION ALL ABOUT?   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.

Definition 1.6 (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 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.

Definition 1.7 (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. However, these deviations become smaller as the number of rolls increases.
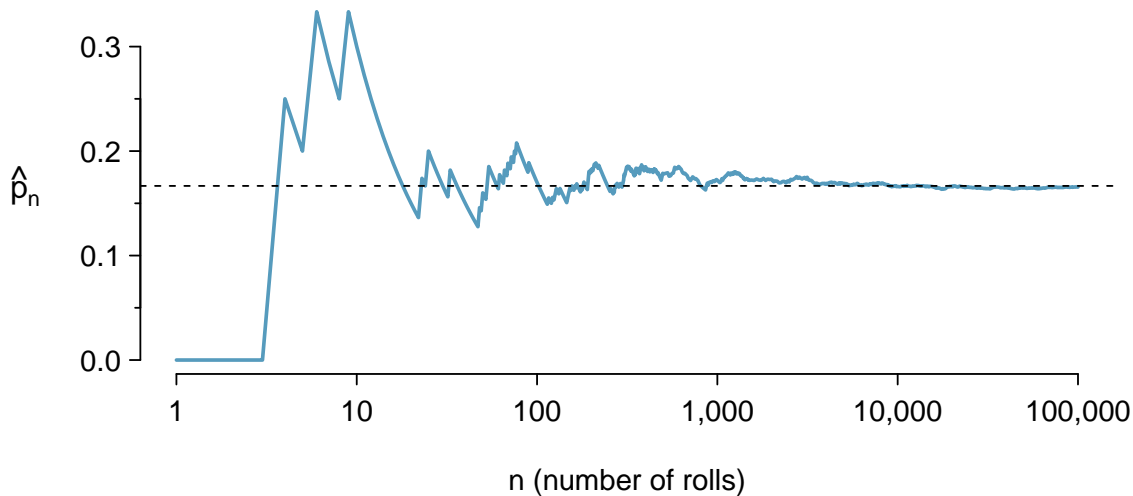
FIGURE 2. The fraction of die rolls that are 1 at each stage in a simulation. The proportion tends to get closer to the probability $1/6 \approx 0.167$ as the number of rolls increases.

Above we write p as the probability of rolling a 1. We can also write this probability as $\mathbb{P}(\text{rolling a 1})$. 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 $\mathbb{P}(\text{rolling a 1})$ as $\mathbb{P}(1)$.

Exercise 1.8. 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 possible 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 1.8 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.

WHAT'S THIS NEXT SECTION ALL ABOUT? 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:

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

---

[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`.

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:

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

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

Definition 1.9 (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

$$\mathbb{P}(A_1 \text{ or } A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2)$$

If there are many disjoint outcomes A_1, ..., A_k, then the probability that one of these outcomes will occur is

$$\mathbb{P}(A_1) + \mathbb{P}(A_2) + \cdots + \mathbb{P}(A_k)$$

Exercise 1.10. 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 $P(\texttt{1 or 4 or 5})$.[2]

Exercise 1.11. In the `loans` data set in "Summarizing Data" (OpenIntro Statistics) the `homeownership` variable described whether the borrower rents, has a mortgage, or owns her property. Of the 10,000 borrowers, 3858 rented, 4789 had a mortgage, and 1353 owned their home.[3]

(1) Are the outcomes `rent`, `mortgage`, and `own` disjoint?
(2) Determine the proportion of loans with value `mortgage` and `own` separately.
(3) Use the Addition Rule for disjoint outcomes to compute the probability a randomly selected loan from the data set is for someone who has a mortgage or owns her home.

Data scientists 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 $\{\texttt{1, 2}\}$ and $B = \{\texttt{4, 6}\}$. 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 3.
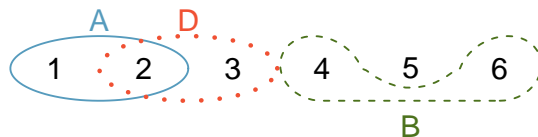

FIGURE 3. Three events, A, B, and D, consist of outcomes from rolling a die. A and B are disjoint since they do not have any outcomes in common.

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:

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

---

[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) $\mathbb{P}(\texttt{1 or 4 or 5}) = \mathbb{P}(\texttt{1}) + \mathbb{P}(\texttt{4}) + \mathbb{P}(\texttt{5}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}$

[3](a) Yes. Each loan is categorized in only one level of `homeownership`. (b) Mortgage: $\frac{4789}{10000} = 0.479$. Own: $\frac{1353}{10000} = 0.135$. (c) $\mathbb{P}(\texttt{mortgage or own}) = \mathbb{P}(\texttt{mortgage}) + \mathbb{P}(\texttt{own}) = 0.479 + 0.135 = 0.614$.

Exercise 1.12. (a) Verify the probability of event A, $P(A)$, is 1/3 using the Addition Rule. (b) Do the same for event B.[4]

Exercise 1.13. (a) Using Figure 3 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 1.14. In Exercise 1.13, you confirmed B and D from Figure 3 are disjoint. Compute the probability that event B or event D occurs.[6]

WHAT'S THIS NEXT SECTION ALL ABOUT? | 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 Figure 4. If you are unfamiliar with the cards in a regular deck, please see the footnote.[7]

$$\begin{array}{llllllllllllll} 2\clubsuit & 3\clubsuit & 4\clubsuit & 5\clubsuit & 6\clubsuit & 7\clubsuit & 8\clubsuit & 9\clubsuit & 10\clubsuit & J\clubsuit & Q\clubsuit & K\clubsuit & A\clubsuit \\ 2\diamondsuit & 3\diamondsuit & 4\diamondsuit & 5\diamondsuit & 6\diamondsuit & 7\diamondsuit & 8\diamondsuit & 9\diamondsuit & 10\diamondsuit & J\diamondsuit & Q\diamondsuit & K\diamondsuit & A\diamondsuit \\ 2\heartsuit & 3\heartsuit & 4\heartsuit & 5\heartsuit & 6\heartsuit & 7\heartsuit & 8\heartsuit & 9\heartsuit & 10\heartsuit & J\heartsuit & Q\heartsuit & K\heartsuit & A\heartsuit \\ 2\spadesuit & 3\spadesuit & 4\spadesuit & 5\spadesuit & 6\spadesuit & 7\spadesuit & 8\spadesuit & 9\spadesuit & 10\spadesuit & J\spadesuit & Q\spadesuit & K\spadesuit & A\spadesuit \end{array}$$

FIGURE 4. Representations of the 52 unique cards in a deck.

Exercise 1.15. (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 5 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. $10/52 = 0.1923$).

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 $\mathbb{P}(A \text{ or } B)$? 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 Addition Rule for disjoint events. Instead we use the Venn diagram. We start by | adding the probabilities of the two events |:

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

---

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

[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*: $\clubsuit$ (club), $\diamondsuit$ (diamond), $\heartsuit$ (heart), $\spadesuit$ (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. $4\heartsuit$ and $J\clubsuit$. The 12 cards represented by the jacks, queens, and kings are called | face cards |. The cards that are $\diamondsuit$ or $\heartsuit$ 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 $P(\text{face card}) = \frac{12}{52} = \frac{3}{13} = 0.231$.
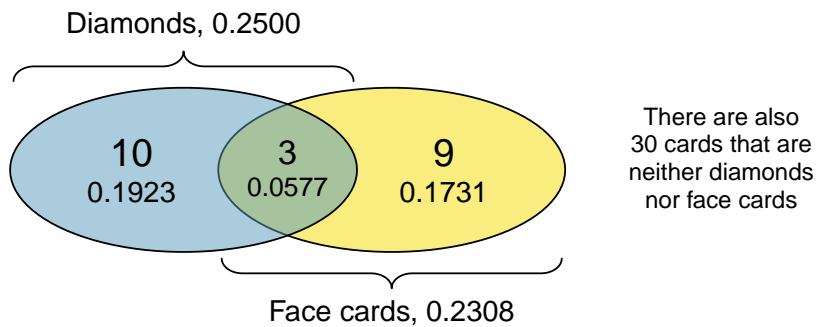
Figure 5. A Venn diagram for diamonds and face cards.

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

$$\mathbb{P}(A \text{ or } B) = \mathbb{P}(\diamondsuit \text{ or face card})$$
$$= \mathbb{P}(\diamondsuit) + \mathbb{P}(\text{face card}) - \mathbb{P}(\diamondsuit \text{ and face card})$$
$$= 13/52 + 12/52 - 3/52$$
$$= 22/52 = 11/26$$

This equation is an example of the *General Addition Rule*.

Definition 1.16. 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

$$\mathbb{P}(A \text{ or } B) = \mathbb{P}(A) + \mathbb{P}(B) - \boxed{\mathbb{P}(A \text{ and } B)}$$

where $\mathbb{P}(A \text{ and } B)$ 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 1.17. (a) If A and B are disjoint, describe why this implies $P(A \text{ and } B) = 0$. (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.[9]

Exercise 1.18. In the `loans` data set describing 10,000 loans, 1495 loans were from joint applications (e.g. a couple applied together), 4789 applicants had a mortgage, and 950 had both of these characteristics. Create a Venn diagram for this setup.[10]

Exercise 1.19. (a) Use your Venn diagram from Exercise 1.18 to determine the probability a randomly drawn loan from the `loans` data set is from a joint application where the couple had a mortgage. (b) What is the probability that the loan had either of these attributes?[11]

---

[9](a) If A and B are disjoint, A and B can never occur simultaneously. (b) If A and B are disjoint, then the last $\mathbb{P}(A \text{ and } B)$ term of in the General Addition Rule formula is 0 (see part (a)) and we are left with the Addition Rule for disjoint events.

[10]Both the counts and corresponding probabilities (e.g. $3839/10000 = 0.384$) are shown. Notice that the number of loans represented in the left circle corresponds to $3839 + 950 = 4789$, and the number represented in the right circle is $950 + 545 = 1495$.
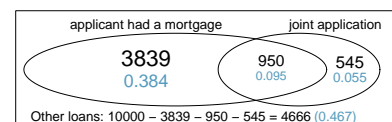
WHAT'S THIS NEXT SECTION ALL ABOUT? | Probability distributions |

A | probability distribution | is a table of all disjoint outcomes and their associated probabilities. Figure 6 shows the probability distribution for the sum of two dice.

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

FIGURE 6. Probability distribution for the sum of two dice.

Definition 1.20. 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 1.21. Figure 7 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?[12]

Income Range	$0-25k	$25k-50k	$50k-100k	$100k+
(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

FIGURE 7. Proposed distributions of US household incomes (Exercise 1.21).

The first three weeks of our course 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 8 as a bar plot. The probability distribution for the sum of two dice is shown in Figure 6 and plotted in Figure 9.

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.

WHAT'S THIS NEXT SECTION ALL ABOUT? | Complement of an event |

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 | (denoted by | S or Ω |) for rolling a die. We often use the sample space to examine the scenario where an event does not occur.

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 10 shows the relationship between D, D^c, and the sample space S.

Exercise 1.22. (a) Compute $\mathbb{P}(D^c) = \mathbb{P}(\text{rolling a } 1, 4, 5, \text{ or } 6)$. (b) What is $\mathbb{P}(D) + \mathbb{P}(D^c)$?[13]

---

[11](a) The solution is represented by the intersection of the two circles: 0.095. (b) This is the sum of the three disjoint probabilities shown in the circles: $0.384 + 0.095 + 0.055 = 0.534$ (off by 0.001 due to a rounding error).

[12]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).
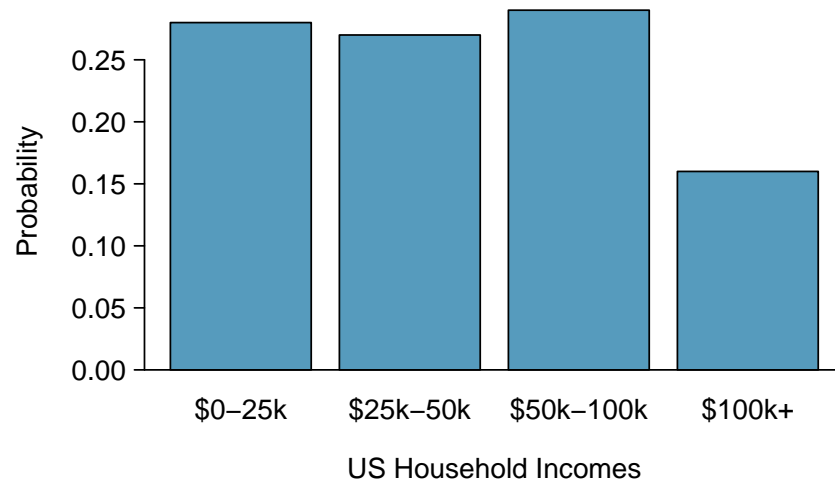
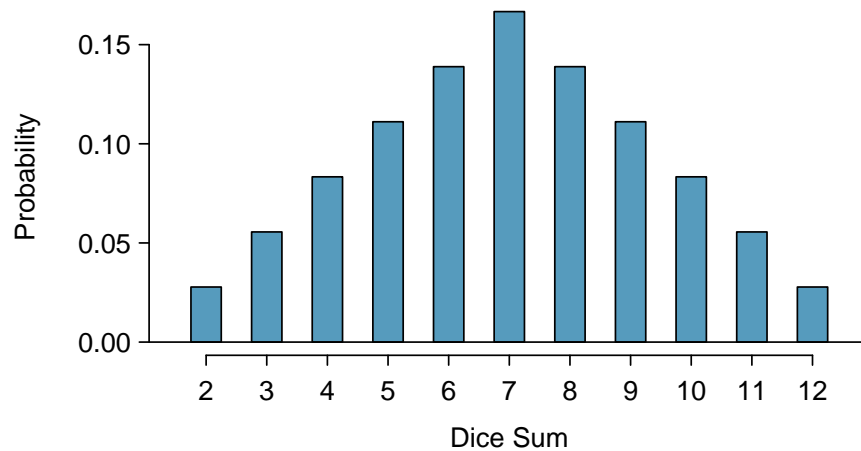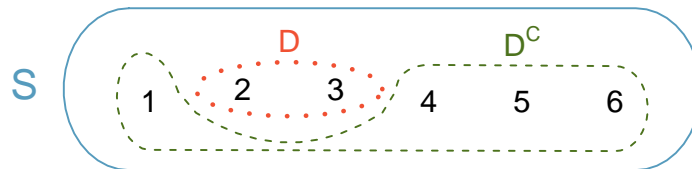FIGURE 8. The probability distribution of US household income.

FIGURE 9. The probability distribution of the sum of two dice.

FIGURE 10. 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 outcomes.

Exercise 1.23. Events $A = \{1, 2\}$ and $B = \{4, 6\}$ are shown in Figure 3 on page 5. (a) Write out what A^c and B^c represent. (b) Compute $\mathbb{P}(A^c)$ and $\mathbb{P}(B^c)$. (c) Compute $\mathbb{P}(A) + \mathbb{P}(A^c)$ and $\mathbb{P}(B) + \mathbb{P}(B^c)$.[14]

---

[13](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 $\mathbb{P}(D) = \frac{1}{6} + \frac{1}{6} = 1/3$. Since D and D^c are disjoint, $\mathbb{P}(D) + \mathbb{P}(D^c) = 1$.

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

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

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):

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

Combining the last two equations yields a very useful relationship between the probability of an event and its complement.

Definition 1.24 (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:

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

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.

Exercise 1.25. 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 $\mathbb{P}(A^c)$ from Figure 6 on page 8. (c) Determine $\mathbb{P}(A)$.[15]

Exercise 1.26. Find the following probabilities for rolling two dice:[16]
  (1) The sum of the dice is *not* 6.
  (2) The sum is at least 4. That is, determine the probability of the event $B = \{4, 5, ..., 12\}$.
  (3) The sum is no more than 10. That is, determine the probability of the event $D = \{2, 3, ..., 10\}$.

WHAT'S THIS NEXT SECTION ALL ABOUT? 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 (or events) — 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 1.5 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 1.5 (page 3), where we calculated the probability

---

[14]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 $\mathbb{P}(A^c) = 2/3$ and $\mathbb{P}(B^c) = 2/3$. (c) A and A^c are disjoint, and the same is true of B and B^c. Therefore, $\mathbb{P}(A) + \mathbb{P}(A^c) = 1$ and $\mathbb{P}(B) + \mathbb{P}(B^c) = 1$.

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

[16](a) First find $\mathbb{P}(6) = 5/36$, then use the complement: $\mathbb{P}(\text{not } 6) = 1 - \mathbb{P}(6) = 31/36$.
  (b) First find the complement, which requires much less effort: $\mathbb{P}(2 \text{ or } 3) = 1/36 + 2/36 = 1/12$. Then calculate $\mathbb{P}(B) = 1 - \mathbb{P}(B^c) = 1 - 1/12 = 11/12$.
  (c) As before, finding the complement is the clever way to determine $\mathbb{P}(D)$. First find $\mathbb{P}(D^c) = \mathbb{P}(11 \text{ or } 12) = 2/36 + 1/36 = 1/12$. Then calculate $\mathbb{P}(D) = 1 - \mathbb{P}(D^c) = 11/12$.

using the following reasoning: 1/6 of the time the red die is a 1, and 1/6 of *those* times the white die will also be 1. This is illustrated in Figure 11. 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.
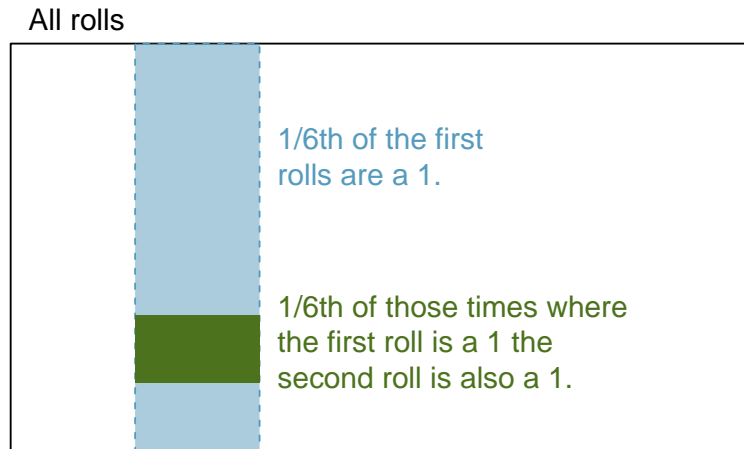
All rolls

1/6th of the first rolls are a 1.

1/6th of those times where the first roll is a 1 the second roll is also a 1.

FIGURE 11. 1/6 of the time, the first roll is a 1. Then 1/6 of *those* times, the second roll will also be a 1.

Example 1.27. 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 1.5. If 1/36 of the time the white and red dice are both 1, then 1/6 of *those* times the blue die will also be 1, so multiply:

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

Example 1.27 illustrates what is called the Multiplication Rule for independent processes or independent events.

Definition 1.28 (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:

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

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

$$\mathbb{P}(A_1) \times \mathbb{P}(A_2) \times \cdots \times \mathbb{P}(A_k)$$

Exercise 1.29. 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?[17]

Exercise 1.30. Suppose 5 people are selected at random.[18]

[17](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: $0.09 \times 0.09 = 0.0081$.

(1) 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?

Suppose the variables `handedness` and `sex` are independent, i.e. knowing someone's `sex` provides no useful information about their `handedness` and vice-versa. Then we can compute whether a randomly selected person is right-handed and female[19] using the Multiplication Rule:

$$\mathbb{P}(\text{right-handed and female}) = \mathbb{P}(\text{right-handed}) \times \mathbb{P}(\text{female})$$
$$= 0.91 \times 0.50 = 0.455$$

Exercise 1.31. Three people are selected at random.[20]
(1) 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 outcome. 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 $\mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \times \mathbb{P}(B)$.

Example 1.32. 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 $\mathbb{P}(A \text{ and } B) = \mathbb{P}(A) \times \mathbb{P}(B)$ is satisfied:

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

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

> *"You know, the most amazing thing happened to me tonight. I was coming here, on the way to the lecture, and I came in through the parking lot. And you won't believe what happened. I saw a car with the license plate ARW 357. Can you imagine? Of all the millions of license plates in the state, what was the chance that I would see that particular one tonight? Amazing!"*

(b) It is reasonable to assume the proportion of people who are ambidextrous (both right- and left-handed) is nearly 0, which results in $\mathbb{P}(\text{right-handed}) = 1 - 0.09 = 0.91$. Using the same reasoning as in part (a), the probability that both will be right-handed is $0.91 \times 0.91 = 0.8281$.

[18](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:

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

(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, $\mathbb{P}(\text{all five are RH})$, to answer this question:

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

[19]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.

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