3 | PROBABILITY TOPICS

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

Introduction

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

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

3.1 | Terminology

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

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

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

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

0.5 (the probability of heads).

Equally likely means that each outcome of an experiment occurs with equal probability. For example, if you toss a **fair**, six-sided die, each face (1, 2, 3, 4, 5, or 6) is as likely to occur as any other face. If you toss a fair coin, a Head (H) and a Tail (T) are equally likely to occur. If you randomly guess the answer to a true/false question on an exam, you are equally likely to select a correct answer or an incorrect answer.

To calculate the probability of an event *A* when all outcomes in the sample space are equally likely, count the number of outcomes for event A and divide by the total number of outcomes in the sample space. For example, if you toss a fair dime and a fair nickel, the sample space is $\{HH, TH, HT, TT\}$ where T = tails and H = heads. The sample space has four outcomes. A = getting one head. There are two outcomes that meet this condition $\{HT, TH\}$, so $P(A) = \frac{2}{4} = 0.5$.

Suppose you roll one fair six-sided die, with the numbers {1, 2, 3, 4, 5, 6} on its faces. Let event E = rolling a number that is at least five. There are two outcomes {5, 6}. $P(E) = \frac{2}{6}$. If you were to roll the die only a few times, you would not be surprised if your observed results did not match the probability. If you were to roll the die a very large number of times, you would expect that, overall, $\frac{2}{6}$ of the rolls would result in an outcome of "at least five". You would not expect exactly $\frac{2}{6}$. The long-term relative frequency of obtaining this result would approach the theoretical probability of $\frac{2}{6}$ as the number of repetitions grows larger and larger.

This important characteristic of probability experiments is known as the **law of large numbers** which states that as the number of repetitions of an experiment is increased, the relative frequency obtained in the experiment tends to become closer and closer to the theoretical probability. Even though the outcomes do not happen according to any set pattern or order, overall, the long-term observed relative frequency will approach the theoretical probability. (The word **empirical** is often used instead of the word observed.)

It is important to realize that in many situations, the outcomes are not equally likely. A coin or die may be **unfair**, or **biased**. Two math professors in Europe had their statistics students test the Belgian one Euro coin and discovered that in 250 trials, a head was obtained 56% of the time and a tail was obtained 44% of the time. The data seem to show that the coin is not a fair coin; more repetitions would be helpful to draw a more accurate conclusion about such bias. Some dice may be biased. Look at the dice in a game you have at home; the spots on each face are usually small holes carved out and then painted to make the spots visible. Your dice may or may not be biased; it is possible that the outcomes may be affected by the slight weight differences due to the different numbers of holes in the faces. Gambling casinos make a lot of money depending on outcomes from rolling dice, so casino dice are made differently to eliminate bias. Casino dice have flat faces; the holes are completely filled with paint having the same density as the material that the dice are made out of so that each face is equally likely to occur. Later we will learn techniques to use to work with probabilities for events that are not equally likely.

" ∪ " Event: The Union

An outcome is in the event $A \cup B$ if the outcome is in A or is in B or is in both A and B. For example, let A = {1, 2, 3, 4, 5} and B = {4, 5, 6, 7, 8}. $A \cup B$ = {1, 2, 3, 4, 5, 6, 7, 8}. Notice that 4 and 5 are NOT listed twice.

" ∩ " Event: The Intersection

An outcome is in the event $A \cap B$ if the outcome is in both A and B at the same time. For example, let A and B be {1, 2, 3, 4, 5} and {4, 5, 6, 7, 8}, respectively. Then $A \cap B$ = {4, 5}.

The **complement** of event A is denoted A' (read "A prime"). A' consists of all outcomes that are **NOT** in A. Notice that $P(A) + P(A') = 1$. For example, let S = {1, 2, 3, 4, 5, 6} and let A = {1, 2, 3, 4}. Then, A' = {5, 6}. $P(A) = \frac{4}{6}$, $P(A') = \frac{2}{6}$, and $P(A) + P(A') = \frac{4}{6} + \frac{2}{6} = 1$

The **conditional probability** of A given B is written $P(A|B)$. $P(A|B)$ is the probability that event A will occur given that the event B has already occurred. **A conditional reduces the sample space**. We calculate the probability of A from the reduced sample space B. The formula to calculate $P(A|B)$ is $P(A|B) = \frac{P(A \cap B)}{P(B)}$ where $P(B)$ is greater than zero.

For example, suppose we toss one fair, six-sided die. The sample space S = {1, 2, 3, 4, 5, 6}. Let A = face is 2 or 3 and B =

face is even (2, 4, 6). To calculate $P(A \mid B)$, we count the number of outcomes 2 or 3 in the sample space $B = \{2, 4, 6\}$. Then we divide that by the number of outcomes B (rather than S).

We get the same result by using the formula. Remember that S has six outcomes.

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{(\text{the number of outcomes that are 2 or 3 and even in} S)}{6}}{\frac{(\text{the number of outcomes that are even in} S)}{6}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

Odds

The odds of an event presents the probability as a ratio of success to failure. This is common in various gambling formats. Mathematically, the odds of an event can be defined as:

$$\frac{P(A)}{1 - P(A)}$$

where $P(A)$ is the probability of success and of course $1 - P(A)$ is the probability of failure. Odds are always quoted as "numerator to denominator," e.g. 2 to 1. Here the probability of winning is twice that of losing; thus, the probability of winning is 0.66. A probability of winning of 0.60 would generate odds in favor of winning of 3 to 2. While the calculation of odds can be useful in gambling venues in determining payoff amounts, it is not helpful for understanding probability or statistical theory.

Understanding Terminology and Symbols

It is important to read each problem carefully to think about and understand what the events are. Understanding the wording is the first very important step in solving probability problems. Reread the problem several times if necessary. Clearly identify the event of interest. Determine whether there is a condition stated in the wording that would indicate that the probability is conditional; carefully identify the condition, if any.

Example 3.1

The sample space S is the whole numbers starting at one and less than 20.

a. S = _______________________________
Let event A = the even numbers and event B = numbers greater than 13.

b. A = ______________________, B = ______________________

c. $P(A)$ = ______________, $P(B)$ = _________________

d. $A \cap B$ = _____________________, A OR B = _________________

e. $P(A \cap B)$ = __________, $P(A \cup B)$ = ______________

f. A' = ______________, $P(A')$ = ______________

g. $P(A) + P(A')$ = _____________

h. $P(A \mid B)$ = ____________, $P(B \mid A)$ = ______________; are the probabilities equal?

Solution 3.1

a. $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19\}$

b. $A = \{2, 4, 6, 8, 10, 12, 14, 16, 18\}$, $B = \{14, 15, 16, 17, 18, 19\}$

c. $P(A) = \frac{9}{19}$, $P(B) = \frac{6}{19}$

d. $A \cap B = \{14,16,18\}$, A OR $B = \{2, 4, 6, 8, 10, 12, 14, 15, 16, 17, 18, 19\}$

e. $P(A \cap B) = \frac{3}{19}$, $P(A \cup B) = \frac{12}{19}$

f. A' = 1, 3, 5, 7, 9, 11, 13, 15, 17, 19; $P(A') = \frac{10}{19}$

g. $P(A) + P(A') = 1$ ($\frac{9}{19} + \frac{10}{19} = 1$)

h. $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{3}{6}$, $P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{3}{9}$, No

Try It Σ

3.1 The sample space S is all the ordered pairs of two whole numbers, the first from one to three and the second from one to four (Example: (1, 4)).

a. S = ______________________________

Let event A = the sum is even and event B = the first number is prime.

b. A = ______________________, B = ______________________

c. $P(A)$ = ______________, $P(B)$ = ________________

d. $A \cap B$ = ____________________, $A \cup B$ = ________________

e. $P(A \cap B)$ = _________, $P(A \cup B)$ = _____________

f. B' = ______________, $P(B')$ = _____________

g. $P(A) + P(A')$ = ____________

h. $P(A \mid B)$ = ___________, $P(B \mid A)$ = _____________; are the probabilities equal?

Example 3.2

A fair, six-sided die is rolled. Describe the sample space S, identify each of the following events with a subset of S and compute its probability (an outcome is the number of dots that show up).

a. Event T = the outcome is two.

b. Event A = the outcome is an even number.

c. Event B = the outcome is less than four.

d. The complement of A.

e. $A \mid B$

f. $B \mid A$

g. $A \cap B$

h. $A \cup B$

i. $A \cup B'$

j. Event N = the outcome is a prime number.

k. Event I = the outcome is seven.

Solution 3.2

a. $T = \{2\}$, $P(T) = \frac{1}{6}$

b. $A = \{2, 4, 6\}$, $P(A) = \frac{1}{2}$

c. $B = \{1, 2, 3\}$, $P(B) = \frac{1}{2}$

d. $A' = \{1, 3, 5\}$, $P(A') = \frac{1}{2}$

e. $A | B = \{2\}$, $P(A | B) = \frac{1}{3}$

f. $B | A = \{2\}$, $P(B | A) = \frac{1}{3}$

g. $A \cap B = \{2\}$, $P(A \cap B) = \frac{1}{6}$

h. $A \cup B = \{1, 2, 3, 4, 6\}$, $P(A \cup B) = \frac{5}{6}$

i. $A \cup B' = \{2, 4, 5, 6\}$, $P(A \cup B') = \frac{2}{3}$

j. $N = \{2, 3, 5\}$, $P(N) = \frac{1}{2}$

k. A six-sided die does not have seven dots. $P(7) = 0$.

Example 3.3

Table 3.1 describes the distribution of a random sample S of 100 individuals, organized by gender and whether they are right- or left-handed.

	Right-handed	**Left-handed**
Males	43	9
Females	44	4

Table 3.1

Let's denote the events M = the subject is male, F = the subject is female, R = the subject is right-handed, L = the subject is left-handed. Compute the following probabilities:

a. $P(M)$

b. $P(F)$

c. $P(R)$

d. $P(L)$

e. $P(M \cap R)$

f. $P(F \cap L)$

g. $P(M \cup F)$

h. $P(M \cup R)$

i. $P(F \cup L)$

j. $P(M')$

k. $P(R \mid M)$
l. $P(F \mid L)$
m. $P(L \mid F)$

Solution 3.3

a. $P(M) = 0.52$
b. $P(F) = 0.48$
c. $P(R) = 0.87$
d. $P(L) = 0.13$
e. $P(M \cap R) = 0.43$
f. $P(F \cap L) = 0.04$
g. $P(M \cup F) = 1$
h. $P(M \cup R) = 0.96$
i. $P(F \cup L) = 0.57$
j. $P(M') = 0.48$
k. $P(R \mid M) = 0.8269$ (rounded to four decimal places)
l. $P(F \mid L) = 0.3077$ (rounded to four decimal places)
m. $P(L \mid F) = 0.0833$

3.2 | Independent and Mutually Exclusive Events

Independent and mutually exclusive do **not** mean the same thing.

Independent Events

Two events are independent if one of the following are true:

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

Two events *A* and *B* are **independent** if the knowledge that one occurred does not affect the chance the other occurs. For example, the outcomes of two roles of a fair die are independent events. The outcome of the first roll does not change the probability for the outcome of the second roll. To show two events are independent, you must show **only one** of the above conditions. If two events are NOT independent, then we say that they are **dependent**.

Sampling may be done **with replacement** or **without replacement**.

- **With replacement**: If each member of a population is replaced after it is picked, then that member has the possibility of being chosen more than once. When sampling is done with replacement, then events are considered to be independent, meaning the result of the first pick will not change the probabilities for the second pick.
- **Without replacement**: When sampling is done without replacement, each member of a population may be chosen only once. In this case, the probabilities for the second pick are affected by the result of the first pick. The events are considered to be dependent or not independent.

If it is not known whether *A* and *B* are independent or dependent, **assume they are dependent until you can show otherwise**.