1 Probability

1.1 Sample Spaces and Events

Probability allows us to quantify the variability in the outcome of any experiment whose exact outcome cannot be predicted with certainty. (For example rolling a die.) However, before we can introduce probability, it is necessary to specify the space of outcomes and the events on which it will be defined.

- An **experiment** is any action or process whose **outcome** is subject to uncertainty (in a very broad sense).
- The **sample space** S of a random phenomenon is the set of all possible outcomes.
- An **event** is an outcome or a set of outcomes of a random phenomenon. That is, an event is a subset of the sample space.

A **discrete sample space** is an itemized list of outcomes, where each outcome has an associated probability.

Example 1.1.1 Rolling a fair 6-sided die (an experiment).

A **continuous sample space** is an interval of outcomes.

Example 1.1.2 Random Number Generation (an experiment).

Suppose that we want to choose a number at random between 0 and 1. We could use a random number generator (RStudio, Python, etc). How could we describe the sample space S?

Example 1.1.3 Family Planning.

A couple wants to have three children. Assume that the probabilities of a newborn being male or female are the same and that the sex of one child does not influence the sex of another child. There are eight possible arrangements of males and females.

(a) What is the sample space for having three children (sex of the first, second, and third child)? All eight arrangements are (approximately) equally likely.

(b) The future parents wonder how many males they might get if they have three children. Give the sample space for the number of males.

Events: Union, Intersection and Compliment

If A and B are any two events (sets) in a sample space S,

- their **union** $A \cup B$ is the subset of S that contains all elements that are either in A, in B, or in both;
- their **intersection** $A \cap B$ is the subset of S that contains all elements that are in both A and B;
- and the **complement** $\overline{A}$ of A is the subset of S that contains all the elements of S that are not in A.
- When A and B have no outcomes in common, they are said to be **disjoint** or **mutually exclusive** events. As a result $A \cap B = \emptyset$ where $\emptyset$ denotes the "null" or "empty" set.

Venn Diagrams

Sample spaces and events are often visualized using Venn diagrams. Typically, the sample space is represented by a rectangle, and events are represented by regions within the rectangle (i.e. by circles or parts of circles).

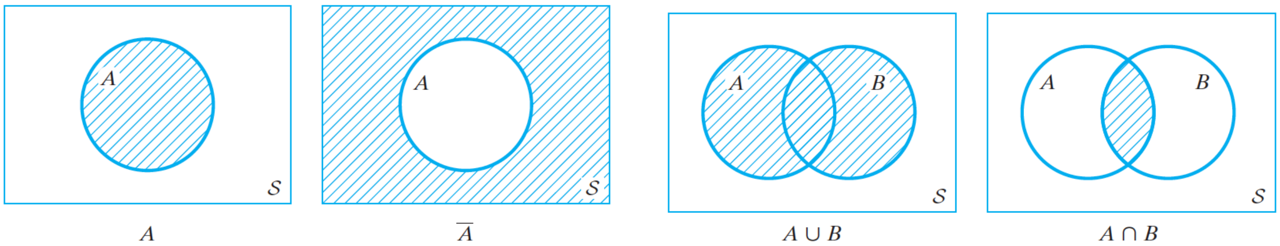
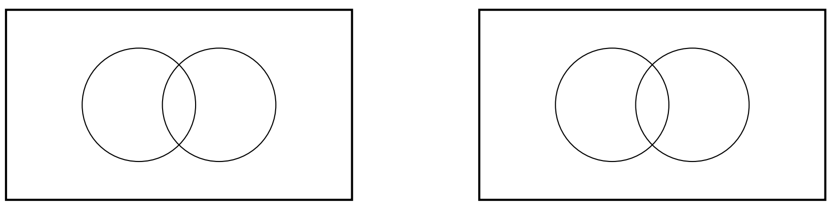


Figure 3.2 Johnson, Miller & Freund’s Probability and Statistics for Engineers, 9E, © Pearson

Example 1.1.4
$\overline{A \cup B} = \overline{A} \cap \overline{B}$



Example 1.1.5
Suppose that vehicles taking a particular freeway exit can turn right (R), turn left (L), or go straight (S). Consider observing the direction for each of three successive vehicles.

(a) List all outcomes in the event A that all three vehicles go in the same direction.

(b) List all outcomes in the event B that all three vehicles take different directions.

(c) List all outcomes in the event C that exactly two of the three vehicles turn right.

(d) List all outcomes in the event D that exactly two vehicles go in the same direction.

(e) List outcomes in $\overline{D}$, $C \cup D$, and $C \cap D$.

1.2 Counting

At times it can be quite difficult, or at least tedious, to determine the number of elements in a finite sample space by direct enumeration. To illustrate, suppose all newer used cars in a large city can be classified as low, medium, or high current mileage; moderate or high priced; and be inexpensive, average, or expensive to operate. In how many ways can a used car be categorized? Each car (event) will have three elements: mileage, price, operation. To handle this kind of problem systematically, it helps to draw a **tree diagram**.

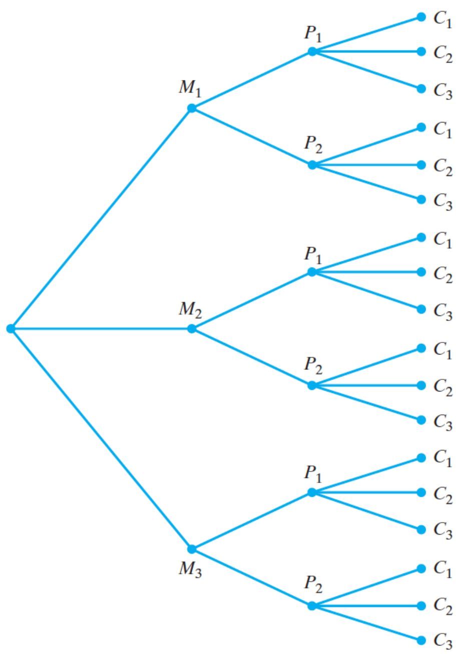


Figure 3.5 Johnson, Miller & Freund’s Probability and Statistics for Engineers, 9E, © Pearson

Theorem 1.2.1
If sets $A_1, A_2, \cdots, A_k$ contain, respectively, $n_1, n_2, \cdots, n_k$ elements, there are $n_1 \cdot n_2 \cdots n_k$ ways of choosing first an element of A_1, then an element of $A_2, \cdots$, and finally an element of A_k.

Example 1.2.1

If a test consists of 12 true-false questions, in how many different ways can a student mark the test paper with one answer to each question?

Example 1.2.2

In how many different ways can the phi club (with a membership of 25) choose a vice president and a president?

Factorial

For $n \in \mathbb{N}$ $\boxed{n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1}$ NOTE: $0! = 1$

Permutations

When describing sample spaces and events, we are often concerned with the number of different ways that we can choose and order subsets (events) from the sample space.

Theorem 1.2.2

The number of **permutations** of r objects selected from a set of n distinct objects is

$$_nP_r = n(n-1)\cdots(n-r+1) = \frac{n!}{(n-r)!}$$

*This is the number of different ways that we can choose and order the r objects - the order matters!

Example 1.2.3

From an alphabet consisting of 10 digits, 26 lower-case and 26 capital letters, how many different 8-character passwords can one create? (Assume no repeated characters.)

At a speed of 1 million passwords per second, it will take a spy program almost 7 years to try all of them. Thus, on the average, it will guess your password in about 3.5 years.

Combinations

There are many problems in which we must find the number of ways in which r objects can be selected from a set of n objects, but we do not care about the order in which the selection is made.

Theorem 1.2.3

The number of ways in which r objects can be selected from a set of n distinct objects is

$$\binom{n}{r} = \frac{n(n-1)\cdots(n-r+1)}{r!} = \frac{n!}{r!(n-r)!}$$

*This is the number of different ways that we can choose the r objects - but the order does not matter!

Example 1.2.4 OLG Lotto Max Winners

The top prize for the weekly Lotto Max (Up to $70 million) is shared between all tickets that match all 7 of the randomly drawn numbers between 1 and 49 (i.e. there are 49 distinct numbers, and 7 numbers drawn **without replacement**). How many different combinations of winning numbers are possible?

1.3 Probability

So far we have considered the possible outcomes and events in a given situation. Next we consider what is probable (and what is improbable).

The Classical Probability Concept

If there are m equally likely possibilities, of which one must occur and s are satisfy a condition ("success"), then the probability of a "success" is given by $\dfrac{s}{m}$.

Example 1.3.1

What is the probability of drawing an ace from a well-shuffled deck of 52 playing cards?

Example 1.3.2 Return to Lotto Max

The top prize for the weekly Lotto Max (Up to $70 million) is shared between all tickets that match all 7 of the randomly drawn numbers between 1 and 49 (i.e. there are 49 distinct numbers, and 7 numbers drawn **without replacement**).

(a) Find the probability that a randomly chosen Lotto Max entry wins the Jackpot.

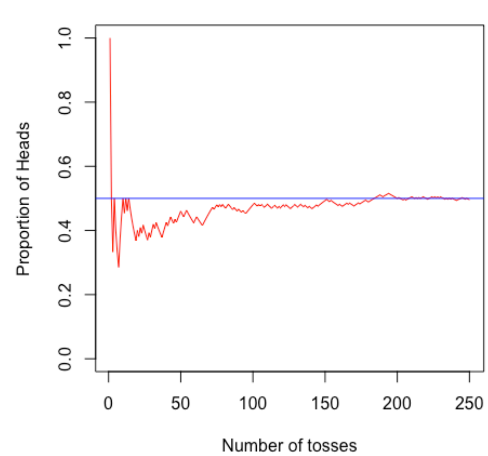
(b) A much smaller prize (say $5000) is awarded to all entries with 6 out of 7 matching numbers. What is the probability of this occurring, for a single play (one set of randomly selected Lotto Max numbers)?

The Frequency Interpretation of Probability

The **probability** of an event (or outcome) is the proportion of times the event will occur in a long run of repeated experiments.

Example 1.3.3

Consider tossing a coin 250 times in a row and calculate, for example, the proportion of tosses that gave a head.



Computing probabilities based on what happens in the long run is called a **frequentist** approach to defining probabilities, because we rely on the relative frequency (proportion) of one particular outcome among very many observations of the random phenomenon.

Example 1.3.4

Probability is a measure of how likely an event is to occur. Match the probabilities that follow with each statement of likelihood given. (The probability value is usually a more exact measure of likelihood than is the verbal statement.)

0	0.01	0.6	0.99	1

(a) This event is impossible. It can never occur.

(b) This event is certain. It will occur on every trial.

(c) This event is very unlikely, but it will occur once in a while in a long sequence of trials.

(d) This event will occur more often than not.

1.4 Probability Rules

Probability Axioms (Rules)

Rules for assigning probabilities to any sample space (discrete or continuous):

Axiom 1. The probability $P(A)$ for any event A satisfies $0 \leq P(A) \leq 1$.

Axiom 2. If S is the sample space in a probability model, then $P(S) = 1$.

Axiom 3. Recall: Sets A and B are **disjoint (mutually exclusive)** if $A \cap B = \emptyset$. If A and B are disjoint, then

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

This is the **addition rule for disjoint events**.

For any event A,

$$P(A \text{ does not occur }) = 1 - P(A).$$

The event "A does not occur" is called the **compliment** of event A, denoted $\overline{A}$.

Theorem 1.4.1 Generalization of the Third Rule

If $A_1, A_2, \ldots, A_n$ are mutually exclusive events in a sample space S, then

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$$

Example 1.4.1 Rabies in Florida

Rabies is a viral disease of mammals transmitted through the bite of a rabid animal. The virus infects the central nervous system, causing encephalopathy and ultimately death. The Florida Department of Health reports the distribution of documented cases of rabies for all of 2016.

Species	Raccoon	Bat	Fox	Other
Probability	0.53	0.22	0.10	?

(a) What probability should replace "?" in the distribution?

(b) What is the probability that a reported case of rabies is not a raccoon?

(c) What is the probability that a reported case of rabies is either a bat or a fox?

Theorem 1.4.2

If A is an event in the finite sample space S, the $P(A)$ equals the sum of the probabilities of the individual outcomes comprising A.

Example 1.4.2 Soda consumption

A national survey asked a random sample of Canadian adults about their soda consumption. Let's call X the number of glasses of soda consumed on a typical day. The survey found the following probabilities for X.

X	0	1	2	3	4+
Probability	0.52	0.28	0.09	0.04	0.07

Consider the events

$$A = \{\text{number of glasses of soda is 1 or greater}\}$$
$$B = \{\text{number of glasses of soda is 2 or less}\}$$

(a) What outcomes make up the event A? What is $P(A)$?

(b) What outcomes make up the event B? What is $P(B)$?

(c) What outcomes make up the event "$A \cup B$"? What is $P(A \cup B)$? Why is this probability not equal to $P(A) + P(B)$?

General Addition Rule for Any Two Events

For two disjoint events A and B, $P(A \cup B) = P(A) + P(B)$

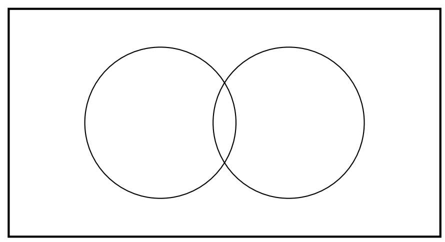
For any two events A and B,

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

(subtract the overlap that's counted twice)

Example 1.4.2 Soda consumption continued

(c) What is $P(A \cup B)$



Example 1.4.2 Income and Stress

The following data is taken from a study on the link between stress levels and income levels using 5000 individuals. Here an individual is identified as "stressed" if they reported feeling extremely or quite stressed on most days, and "not stressed" if they felt a bit stressed, not very stressed, or not at all stressed on most days.

	Income Level			
	Low	Medium	High	Totals
Stressed	230	321	144	695
Not Stressed	1920	1025	1360	4305
Totals	2150	1346	1504	5000

Let L, M, H represent the events that a randomly selected individual has low, medium or high income respectively. Let S and N denote the events that a randomly selected individual is "stressed" and "not stressed", respectively. Find the following probabilities:

(a) $P(H)$

(b) $P(\overline{L} \cap S)$

(c) $P(L \cup M)$

(d) $P(L \cup S)$

1.5 Conditional Probability and Independence (Text: 3.6)

Suppose a coin is flipped twice. If the first flip is a head, does that mean the probability of flipping a head on the second flip will be less than 0.5?

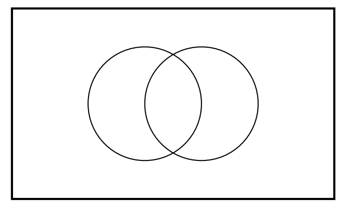
Be careful! This is different from the probability of flipping two heads in a row. In this case, we haven't flipped a coin yet. In the original question, a coin has already been flipped and has come up heads.

Let B be the event that the first coin flip is a head. Let A be the event that the second coin flip is a head.

Conditional Probability

If A and B are any events in S and $P(B) \neq 0$, the **conditional probability** of A given B is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$



Independence

Events A and B are **independent** if occurrence of B does not affect the probability of A, i.e.,

$$P(A|B) = P(A)$$

and thus, we get the following Theorem.

Theorem 1.5.1 Multiplication Rule for Independent Events
If A and B are independent,

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

Conversely, if this condition is not satisfied, then events A and B are **dependent**.

Example 1.5.1 Flights.

Ninety percent of flights depart on time. Eighty percent of flights arrive on time. Seventy-five percent of flights depart on time and arrive on time.

(a) You are meeting a flight that departed on time. What is the probability that it will arrive on time?

(b) You have met a flight, and it arrived on time. What is the probability that it departed on time?

(c) Are the events, departing on time and arriving on time, independent?

Disjoint versus Independent

- Be careful not to confuse ”disjoint” with ”independent”.
- If A and B are disjoint, then the fact that A occurs tells us that B cannot occur.
- Thus, disjoint events are not independent; disjoint events are dependent.

Theorem 1.5.2 General Multiplication Rule for Two Events

For <u>any</u> two events A and B,

$$P(A \cap B) = P(A|B) \cdot P(B) \qquad \text{if } P(B) \neq 0$$
$$P(A \cap B) = P(B|A) \cdot P(A) \qquad \text{if } P(A) \neq 0$$

Example 1.5.2 Cards

The colours of successive cards dealt from the same deck are not independent. Knowing the outcome of the first card dealt changes the probability for the second. A standard 52-card deck contains 26 red cards and 26 black cards. What is the probability that first two cards dealt from the top are both red?

Example 1.5.3 Hardware Problems

Suppose that after 10 years of service, 40% of computers have problems with motherboards (MB), 30% have problems with hard drives (HD), and 15% have problems with both MB and HD. What is the probability that a 10-year old computer still has fully functioning MB and HD? Are the events of have MB problems and having HD problems independent?

1.6 Bayes' Theorem (Text: 3.7)

The general multiplication rules are useful in solving many problems in which the ultimate outcome of an experiment depends on the outcomes of various intermediate stages.

Example 1.6.1 Defective Tablet Screens

A manufacturer of tablets receives its LED screens from three different suppliers, 60% from supplier B_1, 30% from supplier B_2, and 10% from supplier B_3. Also suppose that 95% of the LED screens from B_1, 80% of those from B_2, and 65% of those from B_3 perform according to specifications. We would like to know the probability that any one LED screen received by the plant will perform according to specifications. Let A be the event that a LED screen received by the plant performs according to specifications, and B_1, B_2, and B_3 be the events that it comes from the respective suppliers.

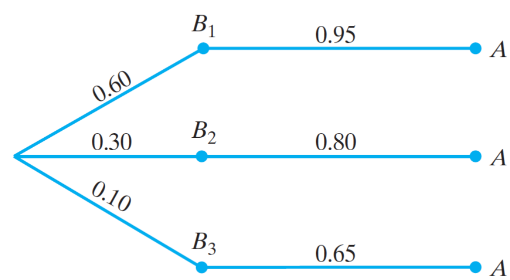


Figure 3.15 Johnson, 9E, © Pearson

This example had only three possibilities at the intermediate stage (i.e. three possible suppliers, B_1, B_2, B_3), but we can extend this idea of summing the product of the branches when there are n mutually exclusive possibilities. Thus we get the following **rule of total probability**.

Theorem 1.6.1 Rule of Total Probability

If $B_1, B_2, \dots, B_n$ are mutually exclusive events of which one must occur, then

$$P(A) = \sum_{i=1}^{n} P(B_i)P(A|B_i)$$

Example 1.6.2 Defective Tablet Screens, cont.

Suppose we want to know the probability that a particular LED screen, which is known to perform according to specifications, came from supplier B_3. Symbolically, we want to know the value of $P(B_3|A)$.

Theorem 1.6.2 Bayes' Theorem
If $B_1, B_2, \ldots, B_n$ are mutually exclusive events of which one must occur, then

$$P(B_r|A) = \frac{P(B_r \cap A)}{P(A)} = \frac{P(B_r)P(A|B_r)}{\sum_{i=1}^{n} P(B_i)P(A|B_i)}$$

Example 1.6.3 Identifying Spam

A first step towards identifying spam is to create a list of words that are more likely to appear in spam than in normal messages. For instance, words like "buy" or the brand name of an enhancement drug are more likely to occur in spam messages than in normal messages. Suppose a specified list of words is available and that your data base of 5000 messages contains 1700 that are spam. Among the spam messages, 1343 contain words in the list. Of the 3300 normal messages, only 297 contain words in the list. Obtain the probability that a message is spam given that the message contains words in the list.

Define $A =$ [message contains words in list] (i.e. message is identified as spam),

$B_1 =$ [message is spam], and $B_2 =$ [message is normal].

Example 1.6.4 Reliability of a Test

There exists a test for a certain viral infection (including a virus attack on a computer network). It is 95% reliable for infected patients and 99% reliable for the healthy ones. That is, if a patient has the virus (event V), the test shows that (event S) with probability $P(S|V) = 0.95$, and if the patient does not have the virus, the test shows that with probability $P(\overline{S}|\overline{V}) = 0.99$.

Consider a patient whose test result is positive (i.e., the test shows that the patient has the virus). Knowing that sometimes the test is wrong, naturally, the patient is eager to know the probability that he or she indeed has the virus. I.e. We want $P(V|S)$.

Suppose that 4% of all the patients are infected with the virus, $P(V) = 0.04$. Find $P(V|S)$

Recommended Homework: Chapter 3 (Johnson, 9th ed.) Exercises 3.81 to 3.99 (odd numbered problems)

