

This is “Sample Spaces, Events, and Their Probabilities”, section 3.1 from the book [Beginning Statistics](#) (v. 1.0). For details on it (including licensing), [click here](#).

For more information on the source of this book, or why it is available for free, please see [the project's home page](#). You can browse or download additional books there.

Has this book helped you? Consider passing it on:



Help Creative Commons

Creative Commons supports free culture from music to education. Their licenses helped make this book available to you.



Help a Public School

DonorsChoose.org helps people like you help teachers fund their classroom projects, from art supplies to books to calculators.



[Previous Section](#)



[Table of Contents](#)



[Next Section](#)

3.1 Sample Spaces, Events, and Their Probabilities

LEARNING OBJECTIVES

1. To learn the concept of the sample space associated with a random experiment.
2. To learn the concept of an event associated with a random experiment.
3. To learn the concept of the probability of an event.

Sample Spaces and Events

Rolling an ordinary six-sided die is a familiar example of a *random experiment*, an action for which all possible outcomes can be listed, but for which the actual outcome on any given trial of the experiment cannot be predicted with certainty. In such a situation we wish to assign to each outcome, such as rolling a two, a number, called the *probability* of the outcome, that indicates how likely it is that

the outcome will occur. Similarly, we would like to assign a probability to any *event*, or collection of outcomes, such as rolling an even number, which indicates how likely it is that the event will occur if the experiment is performed. This section provides a framework for discussing probability problems, using the terms just mentioned.

Definition

A **random experiment** is a mechanism that produces a definite outcome that cannot be predicted with certainty. The **sample space** associated with a random experiment is the set of all possible outcomes. An **event** is a subset of the sample space.

Definition

An event E is said to **occur** on a particular trial of the experiment if the outcome observed is an element of the set E .

EXAMPLE 1

Construct a sample space for the experiment that consists of tossing a single coin.

Solution:

The outcomes could be labeled h for heads and t for tails. Then the sample space is the set $S = \{h, t\}$.

EXAMPLE 2

Construct a sample space for the experiment that consists of rolling a single die. Find the events that correspond to the phrases “an even number is rolled” and “a number greater than two is rolled.”

Solution:

The outcomes could be labeled according to the number of dots on the top face of the die. Then the sample space is the set $S = \{1,2,3,4,5,6\}$.

The outcomes that are even are 2, 4, and 6, so the event that corresponds to the phrase “an even number is rolled” is the set $\{2,4,6\}$, which it is natural to denote by the letter E . We write $E = \{2,4,6\}$.

Similarly the event that corresponds to the phrase “a number greater than two is rolled” is the set $T = \{3,4,5,6\}$, which we have denoted T .

A graphical representation of a sample space and events is a **Venn diagram**, as shown in [Figure 3.1 "Venn Diagrams for Two Sample Spaces"](#) for [Note 3.6 "Example 1"](#) and [Note 3.7 "Example 2"](#). In general the sample space S is represented by a rectangle, outcomes by points within the rectangle, and events by ovals that enclose the outcomes that compose them.

Figure 3.1 Venn Diagrams for Two Sample Spaces



(a): One Coin Toss

(b): One Die Roll

EXAMPLE 3

A random experiment consists of tossing two coins.

- a. Construct a sample space for the situation that the coins are indistinguishable, such as two brand new pennies.
- b. Construct a sample space for the situation that the coins are distinguishable, such as one a penny and the other a nickel.

Solution:

- a. After the coins are tossed one sees either two heads, which could be labeled $2h$, two tails, which could be labeled $2t$, or coins that differ, which could be labeled d . Thus a sample space is $S = \{2h, 2t, d\}$.
- b. Since we can tell the coins apart, there are now two ways for the coins to differ: the penny heads and the nickel tails, or the penny tails and the nickel heads. We can label each outcome as a pair of letters, the first of which indicates how the penny landed and the second of which indicates how the nickel landed. A sample space is then $S' = \{hh, ht, th, tt\}$.

A device that can be helpful in identifying all possible outcomes of a random experiment, particularly one that can be viewed as proceeding in stages, is what is called a **tree diagram**. It is described in the following example.

EXAMPLE 4

Construct a sample space that describes all three-child families according to the genders of the children with respect to birth order.

Solution:

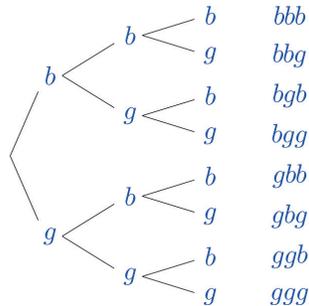
Two of the outcomes are "two boys then a girl," which we might denote bbg , and "a girl then two boys," which we would denote gbb . Clearly there are many outcomes, and when we try to list all of them it could be difficult to be sure that we have found them all unless we proceed systematically. The tree diagram shown in [Figure 3.2 "Tree Diagram For Three-Child](#)

Families", gives a systematic approach.

Figure 3.2

Tree Diagram For

Three-Child Families



The diagram was constructed as follows. There are two possibilities for the first child, boy or girl, so we draw two line segments coming out of a starting point, one ending in a b for “boy” and the other ending in a g for “girl.” For each of these two possibilities for the first child there are two possibilities for the second child, “boy” or “girl,” so from each of the b and g we draw two line segments, one segment ending in a b and one in a g . For each of the four ending points now in the diagram there are two possibilities for the third child, so we repeat the process once more.

The line segments are called **branches** of the tree. The right ending point of each branch is called a **node**. The nodes on the extreme right are the **final nodes**; to each one there corresponds an outcome, as shown in the figure.

From the tree it is easy to read off the eight outcomes of the experiment, so the sample space is, reading from the top to the bottom of the final nodes in the tree,

$$S = \{bbb, bbg, bgb, bgg, gbb, gbg, ggb, ggg\}$$

Probability

Definition

The **probability of an outcome** e in a sample space S is a number p between 0 and 1 that measures the likelihood that e will occur on a single trial of the corresponding random experiment. The value $p = 0$ corresponds to the outcome e being impossible and the value $p = 1$ corresponds to the outcome e being certain.

Definition

The **probability of an event** A is the sum of the probabilities of the individual outcomes of which it is composed. It is denoted $P(A)$.

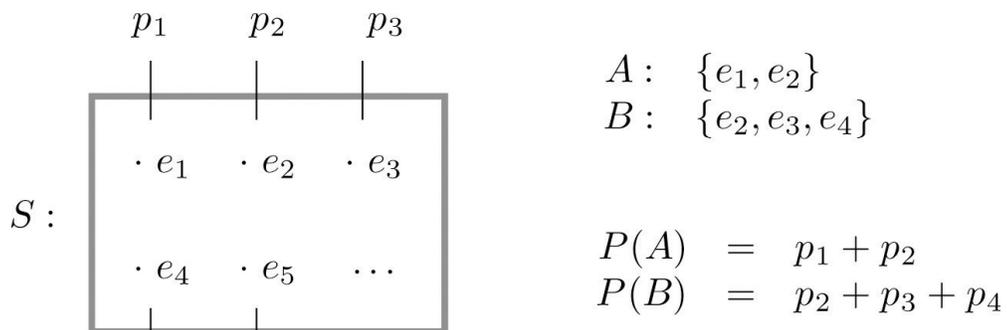
The following formula expresses the content of the definition of the probability of an event:

If an event E is $E = \{e_1, e_2, \dots, e_k\}$, then

$$P(E) = P(e_1) + P(e_2) + \dots + P(e_k)$$

Figure 3.3 "Sample Spaces and Probability" graphically illustrates the definitions.

Figure 3.3 Sample Spaces and Probability



$$\begin{array}{c|c} | & | \\ \hline p_4 & p_5 \end{array}$$

Since the whole sample space S is an event that is certain to occur, the sum of the probabilities of all the outcomes must be the number 1.

In ordinary language probabilities are frequently expressed as percentages. For example, we would say that there is a 70% chance of rain tomorrow, meaning that the probability of rain is 0.70. We will use this practice here, but in all the computational formulas that follow we will use the form 0.70 and not 70%.

EXAMPLE 5

A coin is called “balanced” or “fair” if each side is equally likely to land up. Assign a probability to each outcome in the sample space for the experiment that consists of tossing a single fair coin.

Solution:

With the outcomes labeled h for heads and t for tails, the sample space is the set $S = \{h, t\}$. Since the outcomes have the same probabilities, which must add up to 1, each outcome is assigned probability $1/2$.

EXAMPLE 6

A die is called “balanced” or “fair” if each side is equally likely to land on top. Assign a probability to each outcome in the sample space for the experiment that consists of tossing a single fair die. Find the probabilities of the events E : “an even number is rolled” and T : “a number greater than two is rolled.”

Solution:

With outcomes labeled according to the number of dots on the top face of the die, the

sample space is the set $S = \{1,2,3,4,5,6\}$. Since there are six equally likely outcomes, which must add up to 1, each is assigned probability $1/6$.

Since $E = \{2,4,6\}$, $P(E) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2$.

Since $T = \{3,4,5,6\}$, $P(T) = 4/6 = 2/3$.

EXAMPLE 7

Two fair coins are tossed. Find the probability that the coins match, i.e., either both land heads or both land tails.

Solution:

In [Note 3.8 "Example 3"](#) we constructed the sample space $S = \{2h, 2t, d\}$ for the situation in which the coins are identical and the sample space $S' = \{hh, ht, th, tt\}$ for the situation in which the two coins can be told apart.

The theory of probability does not tell us *how* to assign probabilities to the outcomes, only what to do with them once they are assigned. Specifically, using sample space S , matching coins is the event $M = \{2h, 2t\}$, which has probability $P(2h) + P(2t)$. Using sample space S' , matching coins is the event $M' = \{hh, tt\}$, which has probability $P(hh) + P(tt)$. In the physical world it should make no difference whether the coins are identical or not, and so we would like to assign probabilities to the outcomes so that the numbers $P(M)$ and $P(M')$ are the same and best match what we observe when actual physical experiments are performed with coins that seem to be fair. Actual experience suggests that the outcomes in S' are equally likely, so we assign to each probability $1/4$, and then

$$P(M') = P(hh) + P(tt) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Similarly, from experience appropriate choices for the outcomes in S are:

$$P(2h) = \frac{1}{4} \quad P(2t) = \frac{1}{4} \quad P(d) = \frac{1}{2}$$

which give the same final answer

$$P(M) = P(2h) + P(2t) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

The previous three examples illustrate how probabilities can be computed simply by counting when the sample space consists of a finite number of equally likely outcomes. In some situations the individual outcomes of any sample space that represents the experiment are unavoidably unequally likely, in which case probabilities cannot be computed merely by counting, but the computational formula given in the definition of the probability of an event must be used.

EXAMPLE 8

The breakdown of the student body in a local high school according to race and ethnicity is 51% white, 27% black, 11% Hispanic, 6% Asian, and 5% for all others. A student is randomly selected from this high school. (To select “randomly” means that every student has the same chance of being selected.) Find the probabilities of the following events:

- a. B : the student is black,
- b. M : the student is minority (that is, not white),
- c. N : the student is not black.

Solution:

The experiment is the action of randomly selecting a student from the student population of the high school. An obvious sample space is $S = \{w, b, h, a, o\}$. Since 51% of the students are white and all students have the same chance of being selected, $P(w) = 0.51$, and similarly for the other outcomes. This information is summarized in the following table:

Outcome	w	b	h	a	o
Probability	0.51	0.27	0.11	0.06	0.05

- a. Since $B = \{b\}$, $P(B) = P(b) = 0.27$.

b. Since $M = \{b, h, a, o\}$,

$$P(M) = P(b) + P(h) + P(a) + P(o) = 0.27 + 0.11 + 0.06 + 0.05 = 0.49$$

c. Since $N = \{w, h, a, o\}$,

$$P(N) = P(w) + P(h) + P(a) + P(o) = 0.51 + 0.11 + 0.06 + 0.05 = 0.73$$

EXAMPLE 9

The student body in the high school considered in [Note 3.18 "Example 8"](#) may be broken down into ten categories as follows: 25% white male, 26% white female, 12% black male, 15% black female, 6% Hispanic male, 5% Hispanic female, 3% Asian male, 3% Asian female, 1% male of other minorities combined, and 4% female of other minorities combined. A student is randomly selected from this high school. Find the probabilities of the following events:

- B : the student is black,
- MF : the student is minority female,
- FN : the student is female and is not black.

Solution:

Now the sample space is $S = \{wm, bm, hm, am, om, wf, bf, hf, af, of\}$. The information given in the example can be summarized in the following table, called a *two-way contingency table*:

Gender	Race / Ethnicity				
	White	Black	Hispanic	Asian	Others
Male	0.25	0.12	0.06	0.03	0.01
Female	0.26	0.15	0.05	0.03	0.04

- Since $B = \{bm, bf\}$, $P(B) = P(bm) + P(bf) = 0.12 + 0.15 = 0.27$.
- Since $MF = \{bf, hf, af, of\}$,
 $P(M) = P(bf) + P(hf) + P(af) + P(of) = 0.15 + 0.05 + 0.03 + 0.04 = 0.27$
- Since $FN = \{wf, hf, af, of\}$,
 $P(FN) = P(wf) + P(hf) + P(af) + P(of) = 0.26 + 0.05 + 0.03 + 0.04 = 0.38$

KEY TAKEAWAYS

- The sample space of a random experiment is the collection of all possible outcomes.
- An event associated with a random experiment is a subset of the sample space.

- The probability of any outcome is a number between 0 and 1. The probabilities of all the outcomes add up to 1.
- The probability of any event A is the sum of the probabilities of the outcomes in A .

EXERCISES

BASIC

1. A box contains 10 white and 10 black marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, two marbles in succession and noting the color each time. (To draw “with replacement” means that the first marble is put back before the second marble is drawn.)
2. A box contains 16 white and 16 black marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, three marbles in succession and noting the color each time. (To draw “with replacement” means that each marble is put back before the next marble is drawn.)
3. A box contains 8 red, 8 yellow, and 8 green marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, two marbles in succession and noting the color each time.
4. A box contains 6 red, 6 yellow, and 6 green marbles. Construct a sample space for the experiment of randomly drawing out, with replacement, three marbles in succession and noting the color each time.
5. In the situation of Exercise 1, list the outcomes that comprise each of the following events.
 - a. At least one marble of each color is drawn.
 - b. No white marble is drawn.
6. In the situation of Exercise 2, list the outcomes that comprise each of the following events.
 - a. At least one marble of each color is drawn.
 - b. No white marble is drawn.
 - c. More black than white marbles are drawn.

7. In the situation of Exercise 3, list the outcomes that comprise each of the following events.
- No yellow marble is drawn.
 - The two marbles drawn have the same color.
 - At least one marble of each color is drawn.

8. In the situation of Exercise 4, list the outcomes that comprise each of the following events.
- No yellow marble is drawn.
 - The three marbles drawn have the same color.
 - At least one marble of each color is drawn.

9. Assuming that each outcome is equally likely, find the probability of each event in Exercise 5.

10. Assuming that each outcome is equally likely, find the probability of each event in Exercise 6.

11. Assuming that each outcome is equally likely, find the probability of each event in Exercise 7.

12. Assuming that each outcome is equally likely, find the probability of each event in Exercise 8.

13. A sample space is $S = \{a, b, c, d, e\}$. Identify two events as $U = \{a, b, d\}$ and $V = \{b, c, d\}$.

Suppose $P(a)$ and $P(b)$ are each 0.2 and $P(c)$ and $P(d)$ are each 0.1.

- Determine what $P(e)$ must be.
 - Find $P(U)$.
 - Find $P(V)$.
14. A sample space is $S = \{u, v, w, x\}$. Identify two events as $A = \{v, w\}$ and $B = \{u, w, x\}$.

Suppose $P(u) = 0.22$, $P(w) = 0.36$, and $P(x) = 0.27$.

- Determine what $P(v)$ must be.
 - Find $P(A)$.
 - Find $P(B)$.
15. A sample space is $S = \{m, n, q, r, s\}$. Identify two events as $U = \{m, q, s\}$ and $V = \{n, q, r\}$. The probabilities of some of the outcomes are given by the following table:

Outcome	m	n	q	r	s
Probability	0.18	0.16	0.24	0.21	

- Determine what $P(q)$ must be.
- Find $P(U)$.

- c. Find $P(V)$.
16. A sample space is $S = \{d, e, f, g, h\}$. Identify two events as $M = \{e, f, g, h\}$ and $N = \{d, g\}$.
The probabilities of some of the outcomes are given by the following table:

Outcome	d	e	f	g	h
Probability	0.22	0.13	0.27		0.19

- a. Determine what $P(g)$ must be.
- b. Find $P(M)$.
- c. Find $P(N)$.
- APPLICATIONS**
17. The sample space that describes all three-child families according to the genders of the children with respect to birth order was constructed in [Note 3.9 "Example 4"](#). Identify the outcomes that comprise each of the following events in the experiment of selecting a three-child family at random.
- At least one child is a girl.
 - At most one child is a girl.
 - All of the children are girls.
 - Exactly two of the children are girls.
 - The first born is a girl.
18. The sample space that describes three tosses of a coin is the same as the one constructed in [Note 3.9 "Example 4"](#) with "boy" replaced by "heads" and "girl" replaced by "tails." Identify the outcomes that comprise each of the following events in the experiment of tossing a coin three times.
- The coin lands heads more often than tails.
 - The coin lands heads the same number of times as it lands tails.
 - The coin lands heads at least twice.
 - The coin lands heads on the last toss.
19. Assuming that the outcomes are equally likely, find the probability of each event in Exercise 17.
20. Assuming that the outcomes are equally likely, find the probability of each event in Exercise 18.

ADDITIONAL EXERCISES

21. The following two-way contingency table gives the breakdown of the population in a particular locale according to age and tobacco usage:

Age	Tobacco Use	
	Smoker	Non-smoker
Under 30	0.05	0.20
Over 30	0.20	0.55

A person is selected at random. Find the probability of each of the following events.

- The person is a smoker.
- The person is under 30.
- The person is a smoker who is under 30.

22. The following two-way contingency table gives the breakdown of the population in a particular locale according to party affiliation (A, B, C, or *None*) and opinion on a bond issue:

Affiliation	Opinion		
	Favors	Opposes	Undecided
A	0.12	0.09	0.07
B	0.16	0.12	0.14
C	0.04	0.03	0.06
<i>None</i>	0.08	0.06	0.03

A person is selected at random. Find the probability of each of the following events.

- The person is affiliated with party B.
- The person is affiliated with some party.
- The person is in favor of the bond issue.
- The person has no party affiliation and is undecided about the bond issue.

23. The following two-way contingency table gives the breakdown of the population of married or

previously married women beyond child-bearing age in a particular locale according to age at first marriage and number of children:

Age	Number of Children		
	0	1 or 2	3 or More
Under 20	0.02	0.14	0.08
20-29	0.07	0.37	0.11
30 and above	0.10	0.10	0.01

A woman is selected at random. Find the probability of each of the following events.

- The woman was in her twenties at her first marriage.
- The woman was 20 or older at her first marriage.
- The woman had no children.
- The woman was in her twenties at her first marriage and had at least three children.

24. The following two-way contingency table gives the breakdown of the population of adults in a particular locale according to highest level of education and whether or not the individual regularly takes dietary supplements:

Education	Use of Supplements	
	Takes	Does Not Take
No High School Diploma	0.04	0.06
High School Diploma	0.06	0.44
Undergraduate Degree	0.09	0.28
Graduate Degree	0.01	0.02

An adult is selected at random. Find the probability of each of the following events.

- The person has a high school diploma and takes dietary supplements regularly.
- The person has an undergraduate degree and takes dietary supplements regularly.
- The person takes dietary supplements regularly.
- The person does not take dietary supplements regularly.

LARGE DATA SET EXERCISES

25. Large Data Sets 4 and 4A record the results of 500 tosses of a coin. Find the relative frequency of each outcome 1, 2, 3, 4, 5, and 6. Does the coin appear to be “balanced” or “fair”?

<http://www.flatworldknowledge.com/sites/all/files/data4.xls>

<http://www.flatworldknowledge.com/sites/all/files/data4A.xls>

26. Large Data Sets 6, 6A, and 6B record results of a random survey of 200 voters in each of two regions, in which they were asked to express whether they prefer Candidate A for a U.S. Senate seat or prefer some other candidate.

- Find the probability that a randomly selected voter among these 400 prefers Candidate A.
- Find the probability that a randomly selected voter among the 200 who live in Region 1 prefers Candidate A (separately recorded in Large Data Set 6A).
- Find the probability that a randomly selected voter among the 200 who live in Region 2 prefers Candidate A (separately recorded in Large Data Set 6B).

<http://www.flatworldknowledge.com/sites/all/files/data6.xls>

<http://www.flatworldknowledge.com/sites/all/files/data6A.xls>

<http://www.flatworldknowledge.com/sites/all/files/data6B.xls>

ANSWERS

- $S = \{bb, bw, wb, ww\}$
- $S = \{rr, ry, rg, yr, yy, yg, gr, gy, gg\}$
- $\{bw, wb\}$
 - $\{bb\}$
- $\{rr, rg, gr, gg\}$
 - $\{rr, yy, gg\}$
 - \emptyset

9. a. $2/4$
b. $1/4$
11. a. $4/9$
b. $3/9$
c. 0
13. a. 0.4
b. 0.5
c. 0.4
15. a. 0.21
b. 0.6
c. 0.61
17. a. $\{bbg, bgb, bgg, gbb, bgg, ggb, ggg\}$
b. $\{bbb, bbg, bgb, gbb\}$
c. $\{ggg\}$
d. $\{bgg, gbg, ggb\}$
e. $\{gbb, bgg, ggb, ggg\}$
19. a. $7/8$
b. $4/8$
c. $1/8$
d. $3/8$
e. $4/8$
21. a. 0.25
b. 0.25
c. 0.05
23. a. 0.55
b. 0.76
c. 0.19
d. 0.11
25. The relative frequencies for 1 through 6 are 0.16, 0.194, 0.162, 0.164, 0.154 and 0.166. It would

appear that the die is not balanced.



[Previous Section](#)



[Table of Contents](#)

[Next Section](#)

