Section 7.1 Basic Concepts of Probability

Question 1 – What is an event?

Question 2 – What is probability?

Question 3 - How is probability assigned?

Question 1 – What is an event?

Key Terms

Probability	Outcomes	Event	Experiment
Sample space			

Summary

Probability is used to measure the likelihood of something happening. Implicit in the idea of likelihood is chance. We are uncertain what will happen. In this section, we'll learn how to assign a number from 0 to 1 that reflects how likely an event will occur. A probability of 0 means the event will not occur. A probability of 1 means the event will occur. A probability between 0 and 1 reflects varying degrees to which the event might occur. If the chance of rain is 0.1 (often read as a 10% chance of rain), it probably won't rain. However, a probability of 0.9 (a 90% chance of rain) indicates that it probably will rain. A 50% chance of rain (probability equal to 0.5) means is just as likely to rain as not rain.

An **experiment** is a process that generates uncertain occurrences. These occurrences are called the **outcomes** of the experiment.

The collection of all possible outcomes of an experiment is called the **sample space**. An **event** is any collection of outcomes in the sample space.

For instance, suppose a manufacturer is producing batteries that are sold in a two pack. If a package of batteries is selected from the production line, the batteries in the package may be examined to determine whether they work or are defective. The process of examining whether the batteries in the package are defective is an experiment. The outcome of the experiment may be listed by indicating whether each battery is working (W) or defective (D).

First Battery	Second Battery
W	W
W	D
D	W
D	D

We can specify the first outcome of the experiment as (*W*, *W*). Other outcomes can be written in a similar manner. Written this way, this first letter indicates whether the first battery in the package is working or defective. The second letter indicates whether the second battery in the package is working or defective. We can refer to these outcomes collectively as

$$S = \{(W,W),(W,D),(D,W),(D,D)\}$$

The experiment is carried out many times with each outcome being uncertain. These repetitions of the experiment are called trials.

Notes

Guided Example 1	Practice
A marketing specialist administers a three-question test. Each question is answered yes or no. a. Find the sample space if we are interested in knowing how many questions were answered yes. **Solution** The sample space consists of all of the possible ways to answer yes. Since there are three questions, it is possible to answer no questions, 1 question, 2 questions, or three questions yes. This makes the sample space, $S = \{0, 1, 2, 3\}$	A marketing specialist administers a four-question test. Each question is answered yes or no. a. Find the sample space if we are interested in knowing how many questions were answered yes.
b. Find the sample space if we are interested in knowing how the tests were answered. **Solution** Think of each survey as consisting a sequence of three *Y*'s or *N*'s. For instance, *NYY* would correspond to the outcome where the answer to the first question is no and the answer to the last two questions is yes. The sample space is $S = \{YYY, YNN, NYN, NNY, NYY, YNY, YYN, NNN\}$	b. Find the sample space if we are interested in knowing how the tests were answered.
c. For the sample space in part b, list the outcomes in the event "more questions answered yes than no". **Solution** The event "more questions answered yes than no" consists of events in the sample space with more Y's than N's. Examining the sample space S in part b, the event is $\{YYY, YYN, NYY, YNY\}$	c. For the sample space in part b, list the outcomes in the event "more questions answered yes than no".

Question 2 – What is probability?

Key Terms

Probability

Summary

Probability is defined in terms of the outcomes in the sample space of an experiment. Suppose we have an experiment with a finite number of outcomes in the sample space. Let's represent the outcomes with the letter e followed by a subscript. If there are n outcomes from the experiment, then the sample space S is

$$S = \{e_1, e_2, \ldots, e_n\}$$

The probability of each outcome is symbolized by writing $P(e_1)$, $P(e_2)$, …, $P(e_n)$. We can assign a probability to each outcome as long as the probability satisfies certain requirements. Each outcome of an experiment must meet two requirements.

1. The probability of each outcome is a number from 0 to 1,

$$0 \le P(e_1) \le 1, \quad 0 \le P(e_2) \le 1, \ \ldots, \quad 0 \le P(e_n) \le 1$$

2. The sum of the probabilities of all outcomes is equal to 1,

$$P(e_1) + P(e_2) + \cdots + P(e_n) = 1$$

Probability is a number that indicates the likelihood of an occurrence happening. This number may be as low as 0 indicating the occurrence will not happen. If the probability is equal to 1, the occurrence is certain to happen. Probabilities between 0 and 1 indicate the varying levels of uncertainty about the occurrence.

Notes

Guided Example 2	Practice
A company monitors the snack boxes of fig newtons coming off a production line. They measure the number of fig newtons in each package and assign the probabilities below. Determine if the probability assignment is valid.	A company monitors the graphics chips coming off a production line. They measure the number of defective chips and assign the probabilities below. Determine if the probability assignment is valid.
a. $P(0)=10, P(1)=10, P(2)=80, P(3)=20$ **Solution** To be a valid assignment, each outcome must be assigned a probability from 0 to 1. The outcomes in this assignment are greater than 1 so this is not a valid probability assignment.	a. $P(0)=1, P(2)=99, P(3)=5$
b. $P(0)=0.01, P(1)=0.01, P(2)=0.95,$ $P(3)=0.03$ **Solution** Each probability is in the interval [0, 1] so the first conditional on probability assignments is met. The second condition indicates that the sum of the probabilities must be 1. Adding the four probabilities gives $$0.01+0.01+0.95+0.03=1$$ Since both conditions are satisfied, this is a valid probability assignment.	b. $P(0)=0.98, P(2)=0.01, P(3)=0.01$
c. $P(0)=0.02, P(1)=0.01, P(2)=0.92,$ $P(3)=0.01$ **Solution** All probabilities are in the interval [0, 1]. Checking the sum of the probabilities, we see that $$0.02+0.01+0.92+0.01=0.98$$ The sum of the probabilities is not 1 so this is not a valid probability assignment.	c. $P(0)=0.95, P(1)=0.02, P(2)=0.01$

Question 3 – How is probability assigned?

Key Terms

Equally likely Empirical probability

Summary

There are several ways to assign probability to outcomes in an experiment. The simplest method is to assume that each outcome in the sample space is **equally likely**. In this case, the probability of each outcome in the sample space is the same as any other outcome in the sample space.

Probability of Equally Likely Outcomes

Suppose the outcomes from an experiment are equally likely. If the sample space for the experiment contains *n* outcomes,

$$S = \{e_1, e_2, \ldots, e_n\}$$

then the probabilities of the outcomes are

$$P(e_1) = P(e_2) = \cdots = P(e_n) = \frac{1}{n}$$

The assumption that the outcomes are equally likely is a powerful assumption. It allows us to roll a fair die with six sides and compute the probability of getting a six as $\frac{1}{6}$. We can also use this assumption to compute the probability of selecting the king of clubs from a 52-card deck as $\frac{1}{52}$. However, this assumption may lead to probabilities that are not realistic.

Suppose a factory worker tests randomly selected items from a production line to determine whether they are defective or not defective. If these two outcomes are assumed to be equally likely,

$$P(\text{defective}) = \frac{1}{2}, \quad P(\text{nondefective}) = \frac{1}{2}$$

This factory has a serious problem with quality control! The worker knows from experience that he is much more likely to find that the item is not defective. The equally likely assumption must not be valid.

To get an idea of how likely it is to test an item and find whether it is defective or not defective, the factory worker repeats the testing experiment many times. Out of 500 items, he finds 10 defective products and 490 not defective products. Based on these results, he calculates the probabilities

$$P(\text{defective}) = \frac{10}{500} = 0.02, \quad P(\text{nondefective}) = \frac{490}{500} = 0.98$$

These numbers are the relative frequencies of each outcome in the sample space. We can estimate probabilities of outcomes by repeating an experiment many times and calculating the relative frequency of each outcome. Probability estimated from relative frequencies in a sample of trials from an experiment is called **empirical probability**.

Empirical Probability

If an experiment is performed many times, the probability of an outcome to the experiment is

$$P(e_i) \approx \frac{\text{Number of times } e_i \text{ occurs}}{\text{Total number of trials}}$$

where e_i is any outcome in the sample space of the experiment.

Notes

Guided Example 3

In an analysis of airplane crashes, a researcher notes the primary causes of a crash.

Cause of Crash
Mechanical failure
Weather
Terrorism
Collision with another object
Pilot error

If each of the outcomes are assumed to be equally likely, what is the probability that a crash is caused by terrorism.

Solution Since the each of the outcomes are equally likely, the probability of all outcomes are the same and each is equal to $\frac{1}{n}$ where n is the number of outcomes in the sample space. Since there are 5 outcomes in the sample space, each has a probability of $\frac{1}{5}$. In particular,

$$P(\text{terrorism}) = \frac{1}{5}$$

Practice

A meteorologist notes that on a certain date with certain weather data, the following weather conditions are observed at noon.

Conditions
Sunny
Partly cloudy
Mostly cloudy
Rain

If each of the outcomes are assumed to be equally likely, what is the probability of rain on this date with the weather data observed.

.

Guided Example 4

In an analysis of airplane crashes, a researcher notes the primary cause of a crash and the corresponding frequency of those causes.

Cause of Crash	Frequency
Mechanical failure	123
Weather	52
Terrorism	1
Collision with another object	3
Pilot error	21

Use the frequencies to find the probability that the primary cause of an airplane crash is weather.

Solution To compute the relative frequency of a crash caused by weather we need to know the

Practice

A meteorologist notes that on a certain date with certain weather data, the following weather conditions are observed at noon.

Conditions	Frequency
Sunny	76
Partly cloudy	23
Mostly cloudy	5
Rain	4

Use the frequencies to find the probability of a sunny day on this date with the weather data observed.

total number of trials from the table. The sum of the frequencies is 200. The probability that a crash is caused by weather is $P(\text{weather}) = \dfrac{\text{Number of weather crashes}}{\text{Total number of trials}}$ $= \dfrac{52}{200}$ $= 0.26$ Based on the data in the table, the likelihood of a crash caused by weather is 26%.	