Question 3: How is probability assigned?

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.

In part c of Example 5, the nanobrewery conducts an experiment to determine what beer is ordered by customers. The sample space for this experiment has three outcomes, brown ale (b), pale ale (p), and lager (l). When the outcomes are equally likely, the probability of each outcome is $\frac{1}{3}$. This means a customer is just as likely to order a brown ale as a pale ale or lager.

If the outcomes for an experiment are equally likely, it is easy to determine the probability of each outcome.

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

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

As long as the number of outcomes in the sample space is finite, we can use this relationship to assign probabilities.

Example 6 Find the Probability of an Outcome

A mining company collects samples from along the Hassayampa River in Central Arizona. They examine the levels of gold present in the sample. The samples are labeled

Label	Outcome
No detectable gold present	e_1
Gold present at a level of 0.07 ounces per ton or less	e_2
Gold present at levels greater than 0.07 ounces per ton, but less than or equal to 0.14 ounces per ton	e_3
Gold present at levels greater than 0.14 ounces per ton	e_4

If the outcomes are assumed to be equally likely, find the probability that there will be no detectable gold in the sample.

Solution There are four equally likely outcomes to this experiment. The probability of any outcome is $\frac{1}{4}$. Specifically, the probability of no detectable gold present is

$$P(e_1) = \frac{1}{4}$$

■

Example 7 Find the Probability of an Outcome

In Example 2, we found the sample space for an experiment where a marketing company administers a three question survey where each question is answered yes (Y) or no (N).

If each outcome to this experiment is assumed equally likely, what is the probability that each question is answered no?

Solution In Example 2, a tree diagram was used to determine that the sample space for the experiment is

$$S=\{(Y,Y,Y),(Y,Y,N),(Y,N,Y),(Y,N,N),(N,Y,Y),(N,Y,N),(N,N,Y),(N,N,N)\}$$

The sample space contains eight outcomes. Since the outcomes are equally likely, the probability of any of these outcomes is $\frac{1}{8}$. In particular,

$$P((N,N,N))=\frac{1}{8}$$

■

In the example above, we did not attempt correspond the outcome (N,N,N) to an outcome like e_8. It was simpler to indicate the outcome by writing the actual outcome instead of the correspondence. Technically, we should indicate that this outcome is part of a collection like the sample space and use braces around it,

$$P(\{(N,N,N)\})=\frac{1}{8}$$

Since this complicates the notation, we will typically leave out the braces unless they are needed to clarify the probability being calculated.

Assuming 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.

Example 8   Find the Probability of Outcomes

A nanobrewery records the type of beers it sells over several days to customers. The frequencies are recorded in the table below.

Type	**Frequency**
Brown Ale (b)	90
Pale Ale (p)	150
Lager (l)	160

Use relative frequencies to find the probabilities of each outcome.

Solution The total number of beers sold is $90+150+160$ or 400. The probabilities are estimated by dividing each frequency by the total,

$$P(b)=\frac{90}{400}=0.225$$
$$P(p)=\frac{150}{400}=0.375$$
$$P(l)=\frac{160}{400}=0.4$$

■

Relative frequencies computed from a sample result in probabilities that are estimates of the actual probabilities. If we were to record the number of beers sold over a different set of days or even a few hours, the relative frequencies we would calculate would not be the same as the ones above. If one sample is from a larger proportion of women, we might also expect the relative frequencies to be different.

This may be because the trials are not all comparable. If the nanobrewery carries out the trials during the winter versus during the summer, the brown ale might be more popular. Such beers are typically heartier and traditionally consumed during the winter. Light beers such as a lager are more popular during the summer because of their thirst quench properties. For trials to be comparable, we want to make sure all factors are taken into consideration so that the time of year or day, the gender of the drinker, or other factors do not skew the relative frequencies.

If the number of trials is very small, the probability assessment may be inaccurate. For instance, suppose the nanobrewery calculates the probability of selling lager based on the fact that the last five customers have purchased lagers. According to relative frequencies,

$$P(l)=\frac{10}{10}=1$$

Using this probability, we would assume that there is no possibility of selling pale or brown ale. There is certainly some chance that a pale or brown ale will be sold. Basing the probability on only ten sales leads to a probability assessment that is unreliable.

To ensure the probability assessment is reliable, the number of trials in the experiment must be large and all trails are comparable. When this is done, the relative frequencies for different samples will be approximately the same.

Assuming that outcomes are equally likely or using past history to assign probabilities may not be appropriate. For instance, how likely will it be that the United States economy will suffer a recession next year? Assessing this probability is best done using the experience and judgment. An economist may take all factors into account such as the economies of major trading partners, energy prices, unemployment levels, or debt levels and predict that the likelihood is 10%. Essentially, she is saying there is a 1 in 10 chance of a recession next year. This estimate is based on the knowledge and skill of the economist, not any relative frequencies or logic. A probability assessment like this is a subjective probability assessment. Subjective probability assessments are made when no data exists to calculate relative frequencies or the assumption that outcomes are equally likely is not valid.

Example 9 Find Probability Assessments Subjectively

An economist estimates that the US economy is four times as likely to not suffer a recession as suffer a recession next year. Using this information, find the probability that the US economy will suffer a recession next year.

Solution There are two outcomes to this experiment,

$$e_1 = \text{"the US economy will suffer a recession next year"}$$
$$e_2 = \text{"the US economy will not suffer a recession next year"}$$

Using the information in the problem, the probability of each outcome is written as

$$P(e_1) = x$$
$$P(e_2) = 4x$$

Since the sum of the probabilities of the outcomes is 1, we can solve for x,

$$\begin{aligned} x + 4x &= 1 \\ 5x &= 1 \\ x &= \frac{1}{5} \end{aligned}$$

The probability of each outcomes is

$$P(e_1) = \frac{1}{5} = 0.2$$
$$P(e_2) = 4\left(\frac{1}{5}\right) = 0.8$$

These values match the statement, “the US economy is four times as likely to not suffer a recession as suffer a recession”. The likelihood that the US economy will suffer a recession next year is 0.20. This number might be documented by saying there is a 20% chance that the US economy will suffer a recession next year.

■