Section 8.1 – Properties of Probability

Section 8.1
Properties of Probability

A probability is a function that assigns a value between 0 and 1 to an event, describing the likelihood of that event happening. Probability retains a strong intuitive appeal: an impossible event is assigned a 0 probability of happening, and probabilities close to zero are “unlikely” events. As the probability values increase, the events are correspondingly more likely to happen, so that probabilities close to 1 are considered “likely” events, with absolutely certain events being assigned a probability of 1 (most people would say “100%”). An event that is equally likely to be true as not be true is assigned a probability of 0.5 (again, most people would say “50%”).

The formal study of probability assumes a strong knowledge of sets (Preliminary Section 8.0) and their various concepts and operations. In this section it is assumed you have this knowledge of the concepts and of the terminology.

Intuitive Examples and Definitions

Consider the following example, whose outcome should be obvious:

Example 1. A fair coin is flipped once. What is the probability it lands heads?

Discussion. The “obvious” answer seems to be 50%. Note however the requirement for the coin to be fair, meaning it is not weighted or suspect in any fashion. It is true that under this condition, the probability of the coin landing heads is 50%. ■ ■

The following argument supports this claim: a coin can land two possible ways: heads or tails. Assuming this is a fair coin, the two sides are (intuitively) equally likely. We want the specific outcome of “heads”, which can happen just 1 way out of the 2 possible ways. Hence, the reasonable answer of 50%. This suggests a formal way to define a probability:

An experiment generates a set of (equally likely) outcomes called the sample space $S$. Any subset of the sample space is called an event space (or event for short), usually represented by $E$. Therefore, the probability that event $E$ occurs is the number of elements in $E$ divided by the number of elements in $S$:

$$P(E) = \frac{n(E)}{n(S)}$$

where $n(E)$ represents the number of elements in $E$, and $n(S)$ represents the number of elements in $S$. Since $0 \leq n(E) \leq n(S)$, then $0 \leq P(E) \leq 1$ always.

In example 1 above, the argument can be made even more formally: $S$ is the sample space (the universe) of outcomes of flipping a coin: $S = \{t, h\}$, where $h = \text{heads}$ and $t = \text{tails}$. The event we are interested in is “heads”, or formally, a subset of the sample space $E = \{h\}$. Since there is just one element in the event space and two elements in the sample space, it follows that the probability of the coin landing heads on one flip is $\frac{1}{2}$, or 0.5, or 50%, all of which are equivalent.
Example 2: A fair die is rolled once. What is the probability is comes up “2”?

Discussion. The sample space (universe) is \( S = \{1, 2, 3, 4, 5, 6\} \) and the event space (subset) is \( E = \{2\} \). Thus, the probability of a “2” is \( P(“2”) = \frac{n(E)}{n(S)} = \frac{1}{6} \), which agrees with intuition.

Example 3: A roulette wheel has 38 slots, numbered 1-36 and two numbered 0 and 00. What is the probability any particular slot comes up as the winner?

Discussion. It’s not a requirement to list every element in the sample space when we can write \( n(S) = 38 \) and \( n(E) = 1 \). Thus, the probability for any particular number coming up, including your lucky number, is \( P(E) = \frac{n(E)}{n(S)} = \frac{1}{38} \).

Comment: The size of the sample space can become quite large, making it impractical to list every element. However, it is true that we need to know the size of the sample and event spaces exactly, and to be able to generate elements if need be, in order to properly determine probabilities. This is illustrated in the following examples and discussed further in the next section where methods of enumeration are presented to expedite these calculations.

Example 4: Two fair dice are rolled once. What is the probability the sum of the two dice is 12? 7?

Discussion. It is imperative to identify the correct sample space in which each element is equally likely to occur. Thus, instead of just looking at the sums, we look at every possible roll:

\[
S = \begin{pmatrix}
(1,1) & (2,1) & (3,1) & (4,1) & (5,1) & (6,1) \\
(1,2) & (2,2) & (3,2) & (4,2) & (5,2) & (6,2) \\
(1,3) & (2,3) & (3,3) & (4,3) & (5,3) & (6,3) \\
(1,4) & (2,4) & (3,4) & (4,4) & (5,4) & (6,4) \\
(1,5) & (2,5) & (3,5) & (4,5) & (5,5) & (6,5) \\
(1,6) & (2,6) & (3,6) & (4,6) & (5,6) & (6,6)
\end{pmatrix}
\]

There are 36 possible outcomes. In the case of rolling a 12, there is just one way to do that, so \( E = \{(6,6)\} \) and \( n(E) = 1 \), thus \( P(sum = 12) = \frac{n(E)}{n(S)} = \frac{1}{36} \).

The event space for rolling a 7 is \( E = \{(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)\} \). Thus, the probability of rolling a 7 is \( P(sum = 7) = \frac{n(E)}{n(S)} = \frac{6}{36} = \frac{1}{6} \).

Comment: You may see a short-cut to identify the size of \( S \): there are six ways to roll the first die and six ways to roll the second die, hence there are \( 6 \times 6 = 36 \) ways to roll two dice simultaneously. This is true and uses the Multiplication Principal (section 8.2) to determine the quantity.
Section 8.1 – Properties of Probability

Comment: Note that “rolling two dice simultaneously” is the same as “rolling one die twice”. They produce the exact same sample space so for all intents and purposes are the exact same experiment. This helps explain why the outcomes (5,2) and (2,5) in example 4 are treated as different outcome elements.

Basic Properties of Probability

Many probability questions can be solved using concepts from sets. Many are quite intuitive and you may find yourself using these concepts without being aware. The next example illustrates the use of complementary sets in calculating probabilities:

Example 5: Two fair dice are rolled, and the sample space is the same as that in example 4. What is the probability the sum is not 12? Is not 7?

Discussion. Since there is just one way to roll a 12, there must be 35 ways to not roll a 12. Thus the probability of not rolling a 12 is \( P(sum \neq 12) = \frac{n(E^c)}{n(S)} = \frac{35}{36} \).

Similarly, since there are 6 ways to roll a 7, there must be 30 ways to not roll a 7, and the probability of not rolling a 7 is \( P(sum \neq 7) = \frac{n(E^c)}{n(S)} = \frac{30}{36} = \frac{5}{6} \).

If \( E \) is an event, then \( E^c \) is the complement of \( E \), that is, all the elements in the sample space \( S \) that are not in \( E \). This suggests a useful property involving an event \( E \) and its complement \( E^c \):

\[
\text{Complementary probabilities: } P(E) + P(E^c) = 1
\]

Notice that in example 5, there was no need to manually count out every single way to not roll a 12. It’s faster to count the ways to roll a 12, and subtract from the known quantity in the sample space.

Example 6: What is the probability that a randomly chosen person has a birthday different from yours, ignoring leap years and assuming all birthdays are equally likely (uniformly distributed)?

Discussion. The sample space is the days in a year, thus \( n(S) = 365 \). The event space is your birthday itself which is just one day, so \( n(E) = 1 \). It would be time consuming to list every single day that is not your birthday. Just note that your birthday is one day in the year, so there must be 364 other days available; that is, \( n(E^c) = 364 \). Thus, the answer is \( \frac{364}{365} \), which is intuitive what you should expect.

Example 7: If the probability of \( A \) is \( P(A) = \frac{3}{11} \), what is the probability of “not \( A \)”?

Discussion: Since \( P(A) + P(A^c) = 1 \), the probability of “not \( A \)” must be \( P(A^c) = 1 - \frac{3}{11} = \frac{8}{11} \).
Example 8: A football team has a probability of 0.52 of winning any given game, and a probability of 0.06 of tying any given game. What is the probability this football team loses any given game? Assume these three options make up the universe of all possible outcomes for a game.

Discussion. Since this football team must win, lose or tie a game, we must have that \( P(\text{win}) + P(\text{lose}) + P(\text{tie}) = 1 \), so \( P(\text{lose}) = 1 - P(\text{win}) - P(\text{tie}) = 1 - 0.52 - 0.06 = 0.42 = 42\% \). The event \{lose\} is the complement of the event space \{win, tie\}.

### Probability of Unions

An experiment is performed and a sample space (universe) of outcomes is generated. Consider two different event spaces (subsets) for the experiment; call them \( A \) and \( B \). It is then natural to ask, “what is the probability of \( A \) or \( B \) occurring?” The key word in this sentence is “or”, and from sets this suggests we are looking at the union of the two events \( A \) and \( B \). We need to be careful to not “double-count” some of the elements. The following example is a heuristic approach to developing a formula for this situation.

Example 9: Two fair dice are rolled (see the sample space in example 4). Let \( A = \) “the two dice show the same numbers” and \( B = \) “at least one of the dice is a 1”. What is \( P(A \cup B) \)? In other words, what is the probability that the two dice show the same number or at least one of the dice is a 1?

Discussion: The two event spaces are viewed separately at first:

\[
A = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\} \Rightarrow n(A) = 6 \Rightarrow P(A) = \frac{6}{36} = \frac{1}{6}
\]

\[
B = \{(1,1), (2,1), (3,1), (4,1), (5,1), (6,1), (1,2), (1,3), (1,4), (1,5), (1,6)\} \Rightarrow n(B) = 11 \Rightarrow P(B) = \frac{11}{36}
\]

Note the outcome (1,1) appears in both event sets \( A \) and \( B \), the intersection \( A \cap B \). It gets counted twice, which is not allowed. It should only be counted once, so we subtract its probability of \( \frac{1}{36} \) out once to balance the calculation:

\[
P(A \cup B) = \frac{6}{36} + \frac{11}{36} - \frac{1}{36} = \frac{16}{36}
\]

This suggests a formula for the probability of a union:

\[
\text{Probability of the Union: } P(A \cup B) = P(A) + P(B) - P(A \cap B)
\]

It is possible that events \( A \) and \( B \) have no common elements; the events \( A \) and \( B \) are mutually exclusive. The intersection \( A \cap B \) will be empty, and the probability \( P(A \cap B) = 0 \). In other words, it is impossible for events \( A \) and \( B \) to occur simultaneously. The probability of the union in this case reduces to \( P(A \cup B) = P(A) + P(B) \). Consider the following example (next page):
Example 10: What is the probability that when two dice are rolled, the sum is exactly 12 or the sum is exactly 7?

Discussion. Let \( A = \) “sum is exactly 12” and \( B = \) “sum is exactly 7”. It should be clear that it is impossible (probability = 0) for these two events to happen simultaneously. Therefore, their intersection is empty, and the probability of “rolling exactly a 12 or exactly a 7” is

\[
P(A \cup B) = \frac{1}{36} + \frac{3}{36} - 0 = \frac{7}{36}.
\]

Comment: Be especially aware of the words “and” and “or” in probability questions. The word “and” requires that the two events \( A \) and \( B \) occur simultaneously and thus corresponds to the intersection of the two event spaces. The word “or” requires that at least one of the events occurs and corresponds to the union of the two event spaces.

Example 11: A card is drawn from a deck of 52 well-shuffled cards. Let \( A = \) “the card is a King” and \( B = \) “the card is a Diamond”. What is the probability (a) the card drawn is a King and a Diamond? What is the probability the card drawn is a King or a Diamond?

Discussion. In (a), the “and” suggest we look for the card(s) that are simultaneously a King and a Diamond card. There is just one such card: the King of Diamonds card. Thus, the probability the card drawn is a King and a Diamond is \( P(A \cap B) = \frac{1}{52} \).

In (b) the “or” suggests the card must be one or the other, possibly both. There are four Kings in the deck, so \( P(A) = \frac{4}{52} \), and there are thirteen Diamonds in the deck, so \( P(B) = \frac{13}{52} \). However, the King of Diamond card has been counted twice, once as a King and once as a Diamond. It must be subtracted out once to regain the balance. Thus,

\[
P(A \cup B) = P(A) + P(B) - P(A \cap B)
\]

\[
\Rightarrow P(A \cup B) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} \approx 0.307
\]

Example 12: A corner shopping center has two main stores, Bob’s Cleaners and Jim’s Bait Shop, plus some other stores. A study of traffic entering onto the parking lot shows that 45% of those who pull in shop at Bob’s while 37% shop at Jim’s. Furthermore, 15% shop at both (this is included in the individual figures for Bob’s and Jim’s). Based on this information, (a) what percentage shop just at Bob’s? (b) What percentage shop at Bob’s or Jim’s? And (c) what percentage shop at neither store?

Discussion. A Venn Diagram is useful to parse the data into strict subdivisions (next page):
Section 8.1 – Properties of Probability

The Venn Diagram is filled in first at the intersection of Bob’s and Jim’s, representing 15% of the shoppers. Working outward, we know that 45% overall shop at Bob’s, including the 15% who shop at both stores, so evidently 30% shop just at Bob’s, answering part (a) of the question.

For part (b), this is the union of those who shop at the two stores. With the Venn Diagram, the union is just the sum of the probabilities within the Bob’s and Jim’s circles. Thus, $30\% + 15\% + 22\% = 67\%$ shop at Bob’s or Jim’s.

This can also be solved using the probability of the union formula:

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

$$\Rightarrow P(B \cup J) = 0.45 + 0.37 - 0.15 = 0.67$$

For part (c), the Venn Diagram shows the result to be 33%, the complement of all those who shop at one store or the other. ■ ■

Empirical Probabilities

Empirical probabilities are calculated by direct experimentation as opposed to the theoretical, mathematical approach used in the preceding (and forthcoming) sections. The probability of some events can be found easily by a purely mathematical approach as well as by direct experimentation. For example, if flipping a fair coin, determining that heads comes up with a probability of 0.5 is easily “proven” to be accurate by actually flipping a coin numerous times and tracking the outcomes.

The probability of other scenarios can only be determined either through a mathematical approach, or by direct experimentation. In the dice game Yahtzee, a player throws five dice at once. A “yahtzee” itself is all five dice showing the same value. The probability of such an event is found mathematically to be $\frac{1}{1,296}$. Very few people would have the patience to demonstrate this empirically as it would require many thousands of trials, so the mathematical figure is accepted as fact. On the other hand, a basketball player’s shot accuracy (shots made divided by attempts) is essentially a probability of success. This figure can only be determined by past performance; there is no purely mathematical way to determine this figure.
Section 8.1 – Properties of Probability

A key concept in comparing empirical probabilities to theoretical probabilities is a notion known as the *Law of Large Numbers*. It simply means that the empirical approach must be performed many many times—as many times as possible—before the probabilities derived from the experiment can be accepted as trustworthy and in agreement with the theoretical figures.

**Example 13:** A coin is flipped twice, and comes up heads both times. Is this proof that the coin is not fair?

**Discussion:** Two flips is hardly enough experimentation to determine the nature of the coin. The answer would be no. ■ ■

There is no hard and fast rule as to how many trials must be performed, or when the figures “agree” or “disagree” or are “proven”. The only requirement of the Law of Large Numbers is to perform the experiment many times. If the math and the experiment are done properly and with care, the two should agree after many trials (and stay in agreement after many trials).

An interesting avenue of discussion appeals to a person’s subjective sense of how much experimentation is enough, and whether the figures are in agreement with the theoretical. Some questions arise: Should you perform more trials? Perhaps the theoretical figure is wrong? Suppose the coin is flipped 10 times and comes up heads 3 times, tails 7 times. Most people probably would not be surprised at this, figuring it is within the vagaries of chance. However, after 100 flips of a coin, a 70-30 split may be suspicious (or maybe not). Everyone has a different sense as to when to feel suspicious about the coin itself, or whether to just flip the coin some more.

**Example 14:** A die is rolled, and the theoretical probability of it showing a “2” is \( \frac{1}{6} \) (rounded, see example 2 in this section). How does the actual throwing of a die many times compare to this theoretical figure?

**Discussion:** This experiment was performed by students in three sections of this course at Arizona State University during Spring term, 2008. Each student rolled a die 25 times and tallied the number of times each face came up. The figures were collected (in no particular order) and tallied on a spreadsheet. A total of 82 students participated, with a total of 2,050 rolls (trials).

The first sheet showed that “2” had come up twice in 25 rolls, a \( \frac{2}{25} = 0.08 \) figure. However, two sheets later the student had rolled a “2” seven times in 25 rolls. In a way, these two people balanced one another out. The running tally figures were:

<table>
<thead>
<tr>
<th>Number of rolls:</th>
<th>100</th>
<th>200</th>
<th>300</th>
<th>400</th>
<th>500</th>
<th>1000</th>
<th>1500</th>
<th>2000</th>
<th>2050</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of “2”:</td>
<td>17</td>
<td>38</td>
<td>52</td>
<td>68</td>
<td>83</td>
<td>166</td>
<td>268</td>
<td>338</td>
<td>347</td>
</tr>
<tr>
<td>Empirical Probability:</td>
<td>0.17</td>
<td>0.19</td>
<td>0.173</td>
<td>0.17</td>
<td>0.166</td>
<td>0.166</td>
<td>0.179</td>
<td>0.169</td>
<td>0.169</td>
</tr>
</tbody>
</table>

After over 2,000 rolls, the empirical figures were in strong alignment with the theoretical figure of \( \frac{1}{6} = 0.167 \). Figures for the other values were in strong alignment as well. The expected number of “2”’s (or any other number) would be 341-342 (2,050 divided by 6). Given we were off by 5 or 6 still seems very reasonable. ■ ■
Section 8.1 – Properties of Probability

Summary

• An experiment generates a sample space (universe) $S$ of outcomes.

• An event $E$ is any subset of $S$.

• The probability of an event is

$$P(E) = \frac{n(E)}{n(S)}$$

• The probability of an impossible event is 0.

• The probability of a certain event is 1.

• Since $0 \leq n(E) \leq n(S)$, then $0 \leq P(E) \leq 1$.

• Given an event $E$, we can consider its complement $E'$, and their probabilities are related by the complementary probabilities equation

$$P(E) + P(E') = 1$$

• The probability of the union of two events $A$ and $B$ is

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

• If $A$ and $B$ are mutually exclusive events, then these probabilities are true:

$$P(A \cap B) = 0$$

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

• An empirical probability is one that is found by direct experimentation.

• The Law of Large Numbers suggests that an experiment must be performed many times before the empirical probabilities become in strong agreement with the theoretical probabilities.