## Contents

Foreword

1 Sets and Counting
   1.1 Sets ........................................................................... 1
   1.1.1 Set Algebra .............................................................. 2
   1.1.2 Words of Caution ...................................................... 3
   1.2 Fundamental Principle of Counting ................................. 4
   1.3 Venn Diagrams and Counting ........................................... 6
   1.4 The Multiplication Principle ........................................... 9
   1.5 Permutations and Combinations ..................................... 11
   1.5.1 Permutations .......................................................... 12
   1.5.2 Combinations ......................................................... 13

2 Probability
   2.1 Sample Spaces and Events ............................................. 20
   2.2 Assigning and Calculating Probabilities ............................ 23
   2.3 Conditional Probability ............................................... 29
   2.4 Independent Events ................................................... 32
   2.5 Bayes’ Theorem ......................................................... 34
   2.6 Binomial Trials ......................................................... 38
   2.7 Expected Value ......................................................... 39

3 Matrices
   3.1 Systems of Linear Equations ....................................... 41
   3.2 Arithmetic with Matrices ............................................. 46
Foreword

These notes are mostly\(^1\) a transcription of an edition Diane Hampshire’s lecture notes for Finite Mathematics taught in the Spring of 2017 at Indiana University Kokomo. The current author assumes all responsibility of any mistakes introduced in the transcription.

Of particular note, the section 1.1.2 is included only for those students with a curiosity exceeding the course’s expectations.

\(^1\)some personal flair and other embellishments added
1 Sets and Counting

1.1 Sets

Definition 1. A set is a collection of objects. The objects of a set are called the elements.

Note. Sometimes we call an element of a set a member of the set.

We can define a set by listing all of its elements between curly braces; e.g.

\{Sun., Mon., Tues., Wed., Thurs., Fri., Sat.\}

We can also define a set by giving a description of its elements using set builder notation; e.g.

\{x : x is a day of the week\}

Notation. We write \(x \in A\) to mean that “\(x\) is an element of the set \(A\)” and \(x \notin B\) to mean that “\(x\) is not an element of the set \(B\)”.

Example 1. \(a \in \{a, b, c\}\)

Example 2. \(d \notin \{a, b, c\}\)

Definition 2. The number of elements in a set is called the cardinality, denoted by \(n(A)\), where \(A\) is a set.

Example 3. \(n(\{x : x is a day of the week\}) = 7\)

Example 4. \(n(\{a, b, c\}) = 3\)

Definition 3. If two sets \(A\) and \(B\) have the same elements, then \(A\) equals \(B\).

Example 5. If

\[D = \{\text{Sun., Mon., Tues., Wed., Thurs., Fri., Sat.}\}\]

and

\[W = \{x : x is a day of the week\},\]

then \(D = W\).

Definition 4. A set \(A\) is a subset of another set \(B\), denoted by \(A \subseteq B\), if every element of \(A\) is also an element of \(B\).

Example 6. Let \(A = \{a, b, c\}, B = \{a, b, c, d, e, f, g\}\), and \(C = \{a, z\}\). Then

- \(A \subseteq B\)
- \(C \not\subseteq B\)

Definition 5. A set \(A\) is a proper subset of \(B\), denoted \(A \subset B\), if a set \(A\) is a subset of \(B\) but \(A \neq B\).

Example 7. Let \(A = \{e, a, r\}\) and \(B = \{y, e, a, r\}\). Then \(A \subset B\).

Example 8. Let \(X = \{e, a, t\}\) and \(Y = \{a, t, e\}\). Then \(X \subseteq Y\).
Note. \( A \subseteq A \) for any set \( A \).

**Definition 6.** The **empty set** (or **null set**) is the set that contains no elements, denoted \( \emptyset \) or \( \{ \} \).

Note. \( n(\emptyset) = 0 \).

Note. \( \emptyset \subseteq A \) for any set \( A \).

**Example 9.** List the subsets of \( \{a, b, c\} \).

\[
\emptyset, \\
\{a\}, \{b\}, \{c\}, \\
\{a, b\}, \{a, c\}, \{b, c\}, \\
\{a, b, c\}
\]

### 1.1.1 Set Algebra

**Definition 7.** If \( A \) and \( B \) are sets, the **intersection** of \( A \) and \( B \), denoted by \( A \cap B \), is the set of elements that belong to both \( A \) and \( B \).

**Example 10.** Let \( A = \{a, b, c, d\} \) and \( B = \{f, a, c, e\} \). Then \( A \cap B = \{a, c\} \).

Note. Intersection \( \cap \) like an “N”: \( A \) and \( B \)

**Definition 8.** If \( A \) and \( B \) are sets, the **union** of \( A \) and \( B \), denoted by \( A \cup B \), is the set of all elements that belong to either \( A \) or \( B \) (or both).

**Example 11.** Let \( A = \{a, b, c, d\} \) and \( B = \{f, a, c, e\} \). Then \( A \cup B = \{a, b, c, d, e, f\} \).

Note. Union \( \cup \) like a “U” for union.

**Example 12.** Let \( A = \{a, b, c, d\} \) and \( B = \{e, f, g\} \). Then \( A \cap B = \emptyset \) (\( A \) and \( B \) are **disjoint**). Also, \( A \cup B = \{a, b, c, d, e, f, g\} \).

**Definition 9.** In any particular problem or scenario, there is usually an understood set \( U \), which we call the **universal set**, consisting of all things being considered. Then any other set we discuss will be a subset of \( U \).

**Definition 10.** If \( A \) is a set, the **complement** of \( A \), denoted by \( A' \), is the set of elements in the universal set \( U \) that are not in \( A \).

**Example 13.** Let \( U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \), \( A = \{0, 2, 4, 6, 8\} \), \( B = \{3, 6, 9\} \), and \( C = \{5, 7, 9\} \). Then

- \( A' = \{1, 3, 5, 7, 9\} \)
- \( B' = \{0, 1, 2, 4, 5, 7, 8\} \)
- \( A' \cup B' = \{0, 1, 2, 3, 4, 5, 7, 8, 9\} \)
- \( B \cup C = \{3, 5, 6, 7, 9\} \)
• \(A' \cap (B \cup C) = \{3, 5, 7, 9\}\)
• \(A \cap B = \{6\}\)
• \((A \cap B)' = \{0, 1, 2, 3, 4, 5, 7, 8, 9\}\) (notice a similarity to \(A' \cup B'\)?)
• \(A \cap B \cap C = \emptyset\)

**Note.** For any set \(A\),
• \(A \cap A' = \emptyset\)
• \(A \cup A' = U\)

**Note.** Think of
• \(\cup\) the cup that holds everything
• \(\cap\) cap the common

### 1.1.2 Words of Caution

One may be tempted to think that any description of things yields a set. For example, one can define \(G\) to be the set of all geese that have existed, exist, and will ever exist. Unfortunately, we have to exercise some amount of care when defining sets if we want to avoid contradictions.

**The Barber Paradox.** In the small town of Russellville, there is a barber who shaves the beards all of those, and only those, who do not shave themselves\(^2\). We can express this using sets with

\[S = \{P : P \text{ does not shave himself} \}\].

Let \(U\) be the set of all Russellville residents that shave their beards. Can the barber shave himself?

**Russell’s Paradox.** This is similar to the Barber Paradox but much more formal. Consider a potential set definition

\[R = \{x : x \not\in x\},\]

the set of all sets which don’t contain themselves.

As an example to see that there might be something to this kind of property, consider

\[A = \{x : x \text{ is not an orange}\}\].

The collection \(A\) surely isn’t an orange so \(A \in A\).

Now we return to \(R\). If \(R\) is itself a set, then we have
• \(R \in R \implies R \not\in R\)
• \(R \not\in R \implies R \in R\)

In either scenario, we have a contradiction. Therefore, \(R\) cannot be a set.

\(^2\)so everyone's beard gets shaved
1.2 Fundamental Principle of Counting

**Inclusion-Exclusion Principle.** If \( A \) and \( B \) are sets, then

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

Equivalently,

\[
n(A \cup B) + n(A \cap B) = n(A) + n(B).
\]

If we assume, in addition, that \( A \) and \( B \) are disjoint (i.e., \( A \cap B = \emptyset \)), then \( n(A \cup B) = n(A) + n(B) \) since \( n(A \cap B) = 0 \).

**Example 14.** Let \( A = \{a, b, c, d, e\} \) and \( B = \{a, b, f, g\} \). Then

- \( n(A \cup B) = 7 \) \( \leftarrow \) \( A \cup B = \{a, b, c, d, e, f, g\} \)
- \( n(A \cap B) = 2 \) \( \leftarrow \) \( A \cap B = \{a, b\} \)
- \( n(A) = 5 \)
- \( n(B) = 4 \)

\[
\begin{align*}
n(A \cup B) &= n(A) + n(B) - n(A \cap B) \\
7 &= 5 + 4 - 2
\end{align*}
\]

**Example 15.** Given \( n(A \cup B) = 40 \), \( n(A) = 28 \), and \( n(A \cap B) = 10 \), find \( n(B) \).

From the equation \( n(B) = n(A \cup B) + n(A \cap B) - n(A) \) we compute that

\[
n(B) = 40 + 10 - 28 = 22.
\]

**Example 16.** If 80 pigs fly, 60 pigs talk, and 52 pigs fly and talk, determine

(a) how many pigs either fly or talk (or do both) and

(b) how many pigs fly but don’t talk.

Let \( F \) be the set of pigs that fly and \( T \) be the set of pigs that talk. We want to find \( n(F \cup T) \) and we know that

- \( n(F) = 80 \)
- \( n(T) = 60 \)
- \( n(F \cap T) = 52 \).

So,

\[
n(F \cup T) = n(F) + n(T) - n(F \cap T) \\
= 80 + 60 - 52 \\
= 88.
\]

Though there are 60 pigs that talk, only 52 of those talking pigs also fly. To find how many pigs fly but don’t talk, we need only count the number of pigs that fly and, of those, take away the pigs that also talk. In sets, we are interested in \( F \cap T' \) and we compute

\[
n(F \cap T') = n(F) - n(F \cap T) = 80 - 52 = 28.
\]
**Venn diagrams** are often used to help visualize sets. For example, we can draw

\[ A \cup A' \]

which graphically represents the universal set \( U \), the set \( A \), and the set \( A' \).

If we have two sets \( A \) and \( B \), we can use shading to represent particular regions:

- \( A \cap B' \) shaded
- \( A' \cap B' \) shaded
- \( (A \cup B)' \) shaded

**Example 17.** Draw Venn diagrams with two sets \( A \) and \( B \) and shade the regions

- \( A \cap B' \)
- \( A' \cap B' \)
- \( (A \cup B)' \)
De Morgan’s Laws. For sets $A$ and $B$,

- $(A \cup B)' = A' \cap B'$
- $(A \cap B)' = A' \cup B'$

1.3 Venn Diagrams and Counting

Example 18. Let $U$ be the collection of all 120 people who participated in a survey. In this survey,

- 75 people said they liked apples (set $A$),
- 60 said they liked bananas (set $B$), and
- 42 said they liked both.

Draw a Venn diagram describing the survey and determine

(a) how many only like apples,
(b) how many only like bananas,
(c) how many like apples or bananas, and
(d) how many don’t like apples or bananas.
(a) To find the number of people that only liked apples, we take the number of people that liked apples and subtract the number of people that also like bananas:

\[ n(A) - n(A \cap B) = 75 - 42 = 33. \]

(b) Similarly, to find the number of people that only liked bananas, we take the number of people that liked bananas and subtract the number of people that also like apples:

\[ n(B) - n(A \cap B) = 60 - 42 = 18. \]

(c) Those that liked apples or bananas are represented by the set \( A \cup B \). So here we could apply the inclusion-exclusion principle:

\[ n(A \cup B) = 75 + 60 - 42 = 93. \]

Alternatively, we could sum up the number of people who only like apples, the number of people who only like bananas, and those that like both:

\[ 33 + 18 + 42 = 93. \]

(d) To count those who don’t like either, we look at everyone and subtract those that like apples or bananas: \( 120 - 93 = 27 \).

**Example 19.** In a survey of 80 gym-goers,

<table>
<thead>
<tr>
<th>Activity</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jog (set ( J ))</td>
<td>38</td>
</tr>
<tr>
<td>Swim (set ( S ))</td>
<td>46</td>
</tr>
<tr>
<td>Ride bike (set ( B ))</td>
<td>28</td>
</tr>
<tr>
<td>Jog and swim</td>
<td>24</td>
</tr>
<tr>
<td>Jog and bike</td>
<td>18</td>
</tr>
<tr>
<td>Swim and bike</td>
<td>16</td>
</tr>
<tr>
<td>Do all three</td>
<td>11</td>
</tr>
</tbody>
</table>

Draw a Venn diagram to describe the information and determine

(a) How many do none of these?
(b) How many do exactly two of these activities?
(c) How many do at least two of these activities?
(d) How many do at most two of these activities?

• To make sure we keep of the regions distinct, we start by writing 11 in the intersection of $J$, $S$, and $B$, the number of people that do all three activities.

• To fill in the other intersections (e.g. those in $J \cap S$ but that aren’t in $J \cap S \cap B$), we count the number in the intersection and take away those already accounted for that do all three. For example, the number of people that job and swim but don’t bike is $24 - 11 = 13$.

• To find the remaining portion of each single activity (e.g. those who jog), we look at the number of people who do the activity and subtract those already accounted for in the subregions. For example, for the remaining people that job, we calculate

$$38 - (13 + 11 + 7) = 7.$$ 

(a) To find the number that do none of these activities, we take the total number of gym-goers, 80, and subtract those that partake in activities:

$$80 - (17 + 7 + 5 + 13 + 5 + 7 + 11) = 15.$$ 

(b) Those that do exactly two are colored in blue so we simply add them up:

$$13 + 5 + 7 = 25.$$ 

(c) Those that do at least two activities either do two activities or three activities. So we can add the number of those that do three to those that do exactly two:

$$25 + 11 = 36.$$
To count the number of people that do at most two activities, we need to account for those that

- do no activities,
- do only one activity, and
- do exactly two activities.

So we calculate

\[
15 + (17 + 7 + 5) + 25 = 69
\]

Alternatively, we can count all gym-goers and take away those that do three activities:

\[
80 - 11 = 69.
\]

### 1.4 The Multiplication Principle

**Example 20.** Fresh Appeal offers polo shirts in sizes S, M, L, and XL and in colors white, green, and black. How many choices of polo are there?

The following **tree diagram** is helpful:

```
size →
S  M  L  XL

color → W  G  B  W  G  B  W  G  B
```

In the end, we see that there are

\[
4 \cdot 3 = 12
\]

different choices:

<table>
<thead>
<tr>
<th></th>
<th>white</th>
<th>green</th>
<th>black</th>
</tr>
</thead>
<tbody>
<tr>
<td>small</td>
<td>small white</td>
<td>small green</td>
<td>small black</td>
</tr>
<tr>
<td>medium</td>
<td>medium white</td>
<td>medium green</td>
<td>medium black</td>
</tr>
<tr>
<td>large</td>
<td>large white</td>
<td>large green</td>
<td>large black</td>
</tr>
<tr>
<td>extra large</td>
<td>extra large white</td>
<td>extra large green</td>
<td>extra large black</td>
</tr>
</tbody>
</table>

**The Multiplication Principle.** Suppose a task consists of two choices. If choice 1 can be made in \( m \) ways and choice 2 can be made in \( n \) ways, then the complete task can be completed in \( m \cdot n \) ways.
The Generalized Multiplication Principle. Suppose a task consists of \( k \) choices. If choice \( j \) can be made in \( m_j \) ways, then the complete task can be completed in
\[
m_1 \cdot m_2 \cdots m_k
\]
ways.

Example 21. At Party Bonanza, balloons come in 4 colors, 2 sizes, and 7 shapes. How many choices for a balloon are there?

There are
\[
4 \cdot 2 \cdot 7 = 56
\]
different options.

Definition 11. By a word we mean any string of letters; e.g., \( abc \) is a word of length 3.

Example 22. How many four-letter words are there?

There are
\[
26 \cdot 26 \cdot 26 \cdot 26 = 26^4 = 456,976
\]
different words of length 4.

Example 23. How many four-letter words can be made without repeating any letters?

There are
\[
26 \cdot 25 \cdot 24 \cdot 23 = 358,800
\]
four-letter words without any repeated letters.

Example 24. How many four-letter words are there that start with a “k” or “w”?

There are
\[
2 \cdot 26 \cdot 26 \cdot 26 = 2 \cdot 26^3 = 35,152
\]
different words that start with a “k” or “w”.

Example 25. How many four-letter words are there that start with a “k” or “w” and don’t allow for repeated letters?

There are
\[
2 \cdot 25 \cdot 24 \cdot 23 = 27,600
\]
different words that start with a “k” or “w” and that don’t repeat any letters.

Example 26. How many four-letter words are there that don’t have repeated letters and begin with a “w” or end with a “z”?

Let \( A \) be the set of four-letter words without repetition of letters that start with “w” and \( B \) be the set of four-letter words without repetition of letters that end with “z”. Then we can use the inclusion-exclusion principle
\[
n(A \cup B) = n(A) + n(B) - n(A \cap B)
\]
to count what we want. Note that
\[ n(A) = 1 \cdot 25 \cdot 24 \cdot 23 = 13800, \]
\[ n(B) = 25 \cdot 24 \cdot 23 \cdot 1 = 13800, \]
and
\[ n(A \cap B) = 1 \cdot 24 \cdot 23 \cdot 1 = 552. \]
Finally, we obtain
\[ n(A \cup B) = 13800 + 13800 - 552 = 27048. \]

**Example 27.** How many four-letter words have repeated letters?

We know from Example 22 that there are 456,976 four-letter words and we know from Example 23 that there are 358,800 four-letter words without any repeated letters. So there must be
\[ 456976 - 358800 = 98176 \]
four-letter words with a repeated letter.

**Example 28.** In how many ways can a 5-question True/False exam be completed?

Each of the five questions has two possible responses so there are
\[ 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5 = 32 \]
ways to complete the exam.

**Example 29.** If questions can be left blank, in how many ways can a 5-question True/False exam be completed?

Now each of the questions has 3 possible responses: True, False, or blank. So there are
\[ 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 = 3^5 = 243 \]
ways to complete the exam.

**Example 30.** Three couples go on a movie date. In how many ways can they be seated in a row of six so that each couple is seated together?

There are
\[ 6 \cdot 1 \cdot 4 \cdot 1 \cdot 2 \cdot 1 = 48 \]
ways for them to be arranged in the desired manner.

### 1.5 Permutations and Combinations

In this section, the factorial function will be critical. Recall that, for a positive integer \( n \),
\[ n! = 1 \cdot 2 \cdot 3 \cdots (n - 1) \cdot n, \]
the product of all positive integers \( \leq n \). The quantity \( n! \) is referred to as “\( n \) factorial”. 
Example 31. There are four members of a relay team. How many ways can the runners be lined up?

There are

$$4 \cdot 3 \cdot 2 \cdot 1 = 4! = 24$$

ways to line up the four runners.

1.5.1 Permutations

Definition 12. An ordered arrangement where no item is used more than once is called a permutation.

Remark. In the definition of permutations, order and no repetition are vital.

Example 32. Eight suspects are to be lined up by the police. In how many ways can they line up all 8 suspects?

There are

$$8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 8! = 40,320$$

ways to line up the 8 suspects.

In general, given $n$ items, there are $n!$ ways to order those items. From this, we see the convention that $0! = 1$ since there is exactly one way to line up no items.

Example 33. There are 10 members on a track team. In how many ways can we make a 4-person relay team?

Here, since we only need four people, we count

$$\frac{10 \cdot 9 \cdot 8 \cdot 7}{1!} = 5,040$$

ways to line them up as a relay team.

**Permutations.** The number of permutations of $n$ objects taken $r$ at a time is

$$P(n, r) := \frac{n!}{(n-r)!}.$$  

Example 34. How many three-letter words without repeated letters can be formed from the set $\{a, b, c, d\}$?

There are

$$P(4, 3) = \frac{4!}{1!} = 24$$

three-letter words that can be made.
1.5.2 Combinations

Example 35. How many 3-person committees can be formed from a set of four people \{a, b, c, d\}?

Consider the list of 24 three-letter words that can be made:

\[
\begin{array}{cccc}
abc & abd & acd & bcd \\
acb & adb & adc & bdc \\
bac & bad & cad & cdb \\
aca & bda & cda & cdb \\
cab & dab & dac & dbc \\
\end{array}
\]

Since order doesn’t matter in the making of a committee, there are

\[
\frac{24}{6} = \frac{P(4, 3)}{3!} = 4
\]

possible three-person committees.

Definition 13. An un-ordered arrangement where no item is used more than once is called a combination.

<table>
<thead>
<tr>
<th>Combinations. The number of combinations of (n) objects taken (r) at a time is</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C(n, r) := \frac{P(n, r)}{r!} = \frac{n!}{(n - r)!r!} ).</td>
</tr>
</tbody>
</table>

Example 36. How many ways can 5 people be assigned to seats in a 12-seat row?

We can approach this problem in two ways.

1. We can list the people and choose their seats which gives

\[
\text{# of seats: } 12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95,040
\]

Notice that this is \(P(12, 5)\).

2. We can also split it up into two tasks:

- choose 5 seats and then
- arrange the 5 people.

There are \(C(12, 5) = 792\) ways to choose 5 seats of the 12. Then there are \(5!\) many ways to arrange the 5 people. So there should be

\[792 \cdot 5! = 95,040\]

ways to seat these 5 people in a 12-seat row.
Example 37. A hand in the game of Poker consists of 5 cards dealt from a deck of 52 cards. How many possible hands are there?

There are
\[ C(52, 5) = 2\,598\,960 \]
possible hands.

Example 38. Suppose 5 men and 7 women have submitted applications to 4 different positions at The Academy of Swole. In how many ways can The Academy choose

(a) 2 men and 2 women?
(b) 4 of them without regard to gender?
(c) at least 3 females?

- We can choose 2 men in \( C(5, 2) = 10 \) ways and we can choose 2 women in \( C(7, 2) = 21 \) ways. So we can choose 2 men and 2 women in

\[ 10 \cdot 21 = 210 \]

ways.

- If we ignore the genders, there are \( C(12, 4) = 495 \) ways to select 4 candidates.

- If we wish to pick at least 3 women, there are two possibilities:
  
  \[ \begin{align*}
  &\Diamond \text{ 3 women and one man} \\
  &\Diamond \text{ 4 women and no men}
  \end{align*} \]

There are
\[ C(7, 3) \cdot C(5, 1) = 35 \cdot 5 = 175 \]
ways to pick 3 women and a man. There are
\[ C(7, 4) \cdot C(5, 0) = 35 \cdot 1 = 35 \]
ways to pick 4 women and no men.

Now, these aren’t two steps in one process but two separate processes. So the number of ways to pick

\[ \begin{align*}
  &\Diamond \text{ 3 women and a man or} \\
  &\Diamond \text{ 4 women and no men}
  \end{align*} \]

is \( 175 + 35 = 210 \).

Example 39. You are to receive a package of 50 light bulbs, 6 of which are defective. You will randomly pick four light bulbs.
(a) How many samples are possible?
(b) How many samples have 2 defective bulbs?
(c) How many samples consist of only good bulbs?
(d) How many samples have at least one defective bulb?

- There are \( \binom{50}{4} = 230300 \) samples.
- Since there are 6 defective, there are 44 good bulbs. There are \( \binom{44}{2} = 946 \) ways to get 2 good bulbs and \( \binom{6}{2} = 15 \) ways to get 2 defective bulbs. Then there are \( 946 \cdot 15 = 14190 \) ways to get a sample of 4 where 2 bulbs are defective.
- There are \( \binom{44}{4} = 135751 \) samples without any defective bulbs.
- There are
  - \( \binom{6}{1} \cdot \binom{44}{3} = 6 \cdot 13244 = 79464 \) ways to get one defective bulb and 3 good bulbs
  - \( \binom{6}{2} \cdot \binom{44}{2} = 15 \cdot 946 = 14190 \) ways to get two defective bulbs and 2 good bulbs
  - \( \binom{6}{3} \cdot \binom{44}{1} = 20 \cdot 44 = 880 \) ways to get three defective bulbs and 1 good bulb
  - \( \binom{6}{4} \cdot \binom{44}{0} = 15 \cdot 1 = 15 \) ways to get four defective bulbs and no good bulbs
so there are \( 79464 + 14190 + 880 + 15 = 94549 \) samples with at least one defective bulb.

Alternatively, we know that there are 230300 total samples and 135751 with only good bulbs. So there are \( 230300 - 135751 = 94549 \) samples with at least one defective bulb.

Example 40. Toss a fair coin 10 times and note the sequence of heads and tails.
(a) How many outcomes are possible?
(b) How many outcomes have exactly two heads?
(c) How many outcomes have at most two heads?
(d) How many outcomes have at least two heads?

- Consider the 10 blank slots:

```
1  2  3  4  5  6  7  8  9  10
```

For each slot, it can either be a heads or tails; e.g.

```
H H H T H T T H T
```

So there are \(2^{10} = 1024\) different possible sequences.

- The outcomes with exactly 2 heads can be determined by their place in the sequence. Here, order doesn’t matter since the two heads occurring are indistinguishable. For example, consider the sequences with colored heads:

```
T T H T H H T T T
```

and

```
T T H T H H T T T
```

Without coloration, the two sequences are not distinguishable from each other. There are \(C(10, 2) = 45\) ways to select two spots to place heads in from the 10 available.

- To get at most 2 heads, we can get
  - no heads,
  - 1 head, or
  - 2 heads.

Hence, we count that there are

\[
C(10, 0) + C(10, 1) + C(10, 2) = 1 + 10 + 45 = 56
\]

ways to get at most two heads.

- With similar reasoning, we see that there should be

\[
C(10, 2) + C(10, 3) + C(10, 4) + C(10, 5) + \cdots + C(10, 9) + C(10, 10)
\]

ways to get at least two heads. But as before, we can see that there are

\[
\frac{1024}{\text{total}} - \frac{C(10, 0) + C(10, 1)}{\text{at most 1H}} = 1013
\]

ways to get at least 2 heads.
Remark. Again we see that it is sometimes easier to count the elements in the complement of a set and then subtract from the total number of possibilities.

Example 41. At Glacial Surprise, you can get an ice cream bowl in one of 4 flavors and choose from any of the 12 toppings.

(a) How many possible ice cream bowls are there?

(b) How many possible ice cream bowls are there with exactly 2 toppings?

• We can build an ice cream bowl by first choosing the flavor and then going through each option of topping and deciding whether we want that topping or not. So, for each topping, the possibilities can be represented by a “Yes” or “No”. That gives us 2 choices for each topping. Therefore, there are

\[ 4 \cdot 2^{12} = 16384 \]

different ice cream bowls possible.

• If we wish to choose exactly 2 flavors, there are

\[ 4 \cdot C(12, 2) = 264 \]

options.

Example 42. In general, a set with \( n \) elements has \( 2^n \) subsets. This can be seen by listing out the \( n \) elements in the set and noting that a subset is essentially assigning a “Yes” or “No” to each element; Y to those in the set and N to those not.

Example 43. Consider the following arrangement of city blocks.

Ernest Hemingway is at Fun Palace at point A but needs to get to his flat at point B. How many ways can he get from point A to point B if he only moves to the South or to the East at each block?

Observe that Ernest must go 5 Easts and 4 Souths. He can choose these in any order he desires and he will have to make 9 decisions. Of those 9 decisions, his path is determined by the placement of the choices of South so he can make it from A to B in

\[ C(9, 4) = 126 \]

ways.
**Example 44.** Ernest just remembered that he has to pick up Gertrude Stein on his way back home. Gertrude is at point C in the city block arrangement that follows.

If they can only travel to the South or to the East at each block, how many ways can Ernest make it from point A to point B passing through C?

In a similar way to the above, we reason that Ernest has

\[ C(5, 2) = 10 \]

ways to make it from A to C and then

\[ C(4, 2) = 6 \]

ways to make it from C to B. Since the trip must first reach C and then reach B, there are

\[ 10 \cdot 6 = 60 \]

ways to make the trip in the desired way.

**Example 45.** In this scenario, Agnes Denes is at Pyramid Point at point A and needs to get back to her studio at point B.

If Agnes travels South twice in a row, the Earth evaporates. In how many ways can she make it from point A to point B only going South or East at each block without ever going South twice in a row?

Agnes needs to travel East 5 times and South 4 times. Consider the diagram:
Of the 6 spots available, Agnes needs to choose 4 of them to go South. So there are

\[ C(6, 4) = 15 \]

ways for her to get from A to B in the desired way.
2 Probability

2.1 Sample Spaces and Events

Definition 14. A **probability** expresses the long-term likelihood of an event occurring and is a number between 0 and 1, inclusive.

Remark. 0 corresponds to 0% and 1 corresponds to 100%.

Example 46. Something that has a probability of \(0.9 = 90\% = \frac{9}{10}\) is very likely to happen.

Something that has a probability of \(0.5 = 50\% = \frac{5}{10} = \frac{1}{2}\) is just as likely to occur as not to occur.

Definition 15. We offer the following terms.

- **Experiment** — an activity with an observable result
- **Trial** — each repetition of an experiment
- **Outcome** — each possible result of an experiment
- **Sample Space** — the set of all possible outcomes of an experiment
- **Event** — a subset of the sample space (i.e., a collection of outcomes)

Example 47. An experiment consists of rolling a 6-sided die. The sample space is \(\{1, 2, 3, 4, 5, 6\}\). The event of rolling an odd number can be expressed in set notation: \(\{1, 3, 5\}\). Notice that \(\{1, 3, 5\} \subseteq \{1, 2, 3, 4, 5, 6\}\).

Example 48. An experiment consists of picking 3 balls from a bowl containing red, green, blue, and yellow balls where order of the selection doesn’t matter.

(a) Find the cardinality of the sample space \(S\).

(b) Let \(E\) be the event of picking at least one red ball. Find the cardinality of \(E\).

(c) Let \(F\) be the event of picking at least one purple. Find the cardinality of \(F\).

- Notice that there are
  - \(C(4, 1) = 4\) ways to choose one color (a one-color selection of 3 balls)
  - \(C(4, 2) = 6\) ways to pick two colors and 2 ways to choose which color appears twice in the selection. So there are \(6 \cdot 2 = 12\) ways in which we can select two colors, one of which appears twice in the selection. Alternatively, there are \(C(4, 1) = 4\) ways to choose the color which appears twice and \(C(3, 1) = 3\) ways to choose the other color. So there are \(4 \cdot 3 = 12\) ways to make this selection of one color appearing twice and another color appearing once.
  - \(C(4, 3) = 4\) ways to pick out three colors (a three-color selection of 3 balls)
So \( n(S) = 20 \). We can also visualize the sample space as follows:

\[
S = \{ \text{RRR, GGG, BBB, YYY,} \\
\text{RRG, RRB, RRY,} \\
\text{GGR, GGB, GGY,} \\
\text{BBR, BBG, BBY,} \\
\text{YYR, YYG, YYB,} \\
\text{RGB, RGY, RBY, GBY} \}
\]

- To pick no red balls, notice that there are
  - \( C(3, 1) = 3 \) ways to choose one color which is not red
  - \( C(3, 2) = 3 \) ways to pick two colors that are not red and 2 ways to choose which color appears twice in the selection. So there are \( 3 \cdot 2 = 6 \) ways in which we can select two non-red colors, one of which appears twice in the selection.
    Alternatively, there are \( C(3, 1) = 3 \) ways to choose a non-red color which appears twice. Then there are \( C(2, 1) = 2 \) ways to choose another non-red color to appear once. This leaves us with \( 3 \cdot 2 = 6 \) possible ways to assign a color to appear twice and another color to appear once, neither of which is red.
  - \( C(3, 3) = 1 \) way to pick out three non-red colors

So there are \( 3 + 6 + 1 = 10 \) ways to pick no red balls. Hence, there are \( 20 - 10 = 10 \) ways to pick at least one red ball. We can also refer to the sample space and pick out the outcomes where a red appears:

\[
E = \{ \text{RRR, RRG, RRB, RRY, GGR, BBR, YYR, RGB, RGY, RBY} \}.
\]

- Since there are no purple balls, there are 0 ways to pick at least one purple ball. More explicitly, \( n(F) = 0 \).

**Definition 16.** We say that \( \emptyset \) is an **impossible event** since it never occurs. We say that the whole sample space \( S \) is a **certain event** because any trial of the experiment must product an outcome in the sample space.

**Example 49.** In Example 48 (c) the event \( F \) is an impossible event.

**Remark.** Let \( E \) and \( F \) be events.

- The event where either \( E \) or \( F \) occurs is \( E \cup F \).
- The event where both \( E \) and \( F \) occur is \( E \cap F \).
- The event where \( E \) does not occur is \( E' \).

**Example 50.** An experiment consists of rolling a 6-sided die twice.

(a) Describe the sample space.
(b) Write the event \( A = \) “the sum is greater than 9” in set notation.

(c) Write the event \( B = \) “the numbers on both rolls are equal” in set notation.

(d) Write the event \( C = \) “the sum is 7” in set notation.

(e) List the elements of \( A \cap B \) and \( A \cup B \).

(f) Find \( n((A \cup B)') \).

(g) Investigate \( A \cap C \) and \( B \cap C \).

The sample space is represented with

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

Then observe that

- \( A = \{ (6,4), (5,5), (4,6), (6,5), (5,6), (6,6) \} \)
- \( B = \{ (1,1), (2,2), (3,3), (4,4), (5,5), (6,6) \} \)
- \( C = \{ (6,1), (5,2), (4,3), (3,4), (2,5), (1,6) \} \)
- Notice that \( A \cap B \cap C = \emptyset \).

Then observe that

- First, note that \( n(S) = 36 \) and \( n(A \cup B) = 10 \). Then \( n((A \cup B)') = 36 - 10 = 26 \).
- Since \( A \) and \( C \) have no elements in common, we have that \( A \cap C = \emptyset \). Similarly, \( B \cap C = \emptyset \).

**Definition 17.** Let \( E \) and \( F \) be events in a sample space. Then \( E \) and \( F \) are said to be **mutually exclusive** if their intersection is empty; i.e., \( E \cap F = \emptyset \).

**Remark.** Mutually exclusive events can’t occur at the same time.

**Example 51.** Three balls are selected from a box containing red, green, blue, and yellow balls. Let \( A \) be the event of selecting balls of the same color and \( B \) be the event of selecting balls, all of which have a different color. Show that the events \( A \) and \( B \) are mutually exclusive.

Note that

\[
A = \{ \text{RRR, GGG, BBB, YYY} \} \\
B = \{ \text{RGB, RGY, RBY, GBY} \}
\]

which shows that \( A \cap B = \emptyset \).
Note. A standard deck of cards contains 52 cards. There are

• 13 ranks, in ascending order: 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, A
• and 4 suits: ♦, ♠, ♥, ♦

The cards in the ranks J, Q, and K are called face cards. Each suit contains the 13 ranks. The suits clubs ♠ and spades ♦ are colored black. The suits hearts ♠ and diamonds ♦ are colored red.

Example 52. Select one card from a deck of 52. Let

• A be the event that the card drawn is a K,
• B be the event that the card drawn is a Q,
• C be the event of drawing a heart,
• D be the event of drawing a diamond, and
• E be the event of drawing a face card.

Show that

(a) A and B are mutually exclusive,
(b) A and C are not mutually exclusive,
(c) D and E are not mutually exclusive, and
(d) C and D are mutually exclusive.

• K and Q are different cards so $A \cap B = \emptyset$.
• There is a K of hearts so $A \cap C \neq \emptyset$.
• There are three face cards in the suit of diamonds so $D \cap E \neq \emptyset$.
• The diamond and heart suits are different suits so $C \cap D = \emptyset$.

2.2 Assigning and Calculating Probabilities

Definition 18. A probability distribution for a sample space $S = \{a_1, a_2, \ldots, a_n\}$ assigns a probability to each outcome so that the sum of the probabilities equals 1. If the probability distribution assigns the same probability to each outcome, then we say that all outcomes are equally probable or equally likely.

Theorem 19. If $S$ is a sample space where all outcomes are equally likely, then for any event $E$,

$$\Pr(E) = \frac{n(E)}{n(S)} = \frac{\text{number of outcomes in the event } E}{\text{number of possible outcomes}}.$$
Example 53. Toss a fair coin. Then

<table>
<thead>
<tr>
<th>outcome</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>1/2</td>
</tr>
<tr>
<td>T</td>
<td>1/2</td>
</tr>
</tbody>
</table>

is a table expressing the probability distribution. Notice that both outcomes are equally likely.

Definition 20. An experimental probability is calculated after conducting many trials and examining the data. The probability calculated is sometimes called the relative frequency.

Example 54. Three hundred freshmen were asked how many colleges they had applied to and the data is summarized below:

<table>
<thead>
<tr>
<th># of colleges</th>
<th>frequency</th>
<th>probability (decimal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51</td>
<td>51/300 = 0.17</td>
</tr>
<tr>
<td>2</td>
<td>36</td>
<td>36/300 = 0.12</td>
</tr>
<tr>
<td>3</td>
<td>42</td>
<td>42/300 = 0.14</td>
</tr>
<tr>
<td>4</td>
<td>48</td>
<td>48/300 = 0.16</td>
</tr>
<tr>
<td>5+</td>
<td>123</td>
<td>123/300 = 0.41</td>
</tr>
</tbody>
</table>

Fundamental Properties of Probability. For a sample space \( \{a_1, a_2, \ldots, a_n\} \), suppose the probability distribution is \( \{p_1, p_2, \ldots, p_n\} \) where \( p_j \) is the probability of \( a_j \).

(P1) The probability of any individual outcome is between 0 and 1, inclusive; i.e., \( 0 \leq p_j \leq 1 \).

(P2) \( p_1 + p_2 + \cdots + p_n = 1 \)

(P3) The probability of an event \( E \), denoted \( \Pr(E) \), is equal to the sum of the probabilities of the outcomes in \( E \).

Remark. Property (P3) is known as the **Addition Principle** for probability.

Example 55. An experiment consists of rolling two 6-sided dice. Find the probability that the sum of the two rolls is 7.

In Example 50 (d), we showed that there are 6 outcomes so that the sum of the two rolls is 7. Therefore, the probability that the sum of the two rolls is 7 is \( \frac{6}{36} = \frac{1}{6} \).

Inclusion-Exclusion Principle. For any two events \( E \) and \( F \),

\[
\Pr(E \cup F) = \Pr(E) + \Pr(F) - \Pr(E \cap F).
\]

Equivalently,

\[
\Pr(E \cup F) + \Pr(E \cap F) = \Pr(E) + \Pr(F).
\]

If \( E \) and \( F \) are mutually exclusive events, then \( \Pr(E \cup F) = \Pr(E) + \Pr(F) \) since \( E \cap F = \emptyset \) and \( \Pr(\emptyset) = 0 \).
**Example 56.** If a random card is drawn from a standard deck of cards, find the probability that the card drawn is a Queen or a red card.

Note that there are
- 52 cards, total
- 4 Queens
- 26 red cards
- 2 red Queens.

So,

\[
\Pr(Q \text{ or red}) = \Pr(Q) + \Pr(\text{red}) - \Pr(Q \text{ and red})
\]

\[
= \frac{4}{52} + \frac{26}{52} - \frac{2}{52}
\]

\[
= \frac{7}{13}.
\]

**Example 57.** Suppose \( \Pr(E) = 0.5, \Pr(F) = 0.6, \) and \( \Pr(E \cup F) = 0.9. \) Find \( \Pr(E \cap F). \)

By the Inclusion-Exclusion Principle,

\[
\frac{\Pr(E \cup F) + \Pr(E \cap F)}{0.9 + \Pr(E \cap F)} = \frac{0.5 + 0.6}{1.1} = 0.9
\]

\[
\Pr(E \cap F) = 0.2
\]

**Example 58.** A fair coin is tossed 4 times and the number of heads is noted. Describe the sample space and find the probability of getting at most one head.

There are \( 2^4 = 16 \) different sequences. We list them below:

\[
\text{TTTT, HTTT, THTT, TTHT,}
\]

\[
\text{TTTH, HHTT, HTHT, HTTH,}
\]

\[
\text{THHT, THHT, TTHH, HHHT,}
\]

\[
\text{HHTH, HTHH, THHH, HHHH.}
\]

There are \( C(4, 0) = 1 \) ways to get no heads and \( C(4, 1) = 1 \) ways to get exactly one head. Then there are \( 1 + 4 = 5 \) may ways to get at most one head. Hence, the probability of getting at most one head is

\[
\Pr(\text{at most 1H}) = \frac{5}{16}.
\]

We could also see this by counting the sequences with at most one head appearing in the list of all possible sequences above.

**Example 59.** An urn contains 6 white balls and 5 red balls. A sample of four balls is taken at random. Calculate

(a) the probability of getting only white balls,
(b) the probability of getting 2 white and 2 red balls,

(c) the probability of getting at least one red ball.

First, note that there are
\[ C(11, 4) = 330 \]
different possible samples. Of these, there are
\[ C(6, 4) = 15 \]
that consist only of white balls. So the probability of getting only white balls is
\[ \Pr(\text{only W}) = \frac{15}{330} = \frac{1}{22}. \]

Now, there are
\[ C(6, 2) \cdot C(5, 2) = 15 \cdot 10 = 150 \]
samples with exactly two white and two red balls. So the probability of getting a sample with two white and two red balls is
\[ \Pr(2W \text{ and } 2R) = \frac{150}{330} = \frac{5}{11}. \]

To calculate the probability of getting at least one red ball, we can consider the failures to at least one red. That is, a sample fails to have at least one red when it contains only white balls. Of the 330 total samples, we found 15 to consist only of white balls. So there are
\[ 330 - 15 = 315 \]
samples with at least one red ball. Thus, the probability of getting a sample with at least one red ball is
\[ \Pr(\text{at least 1R}) = \frac{315}{330} = \frac{21}{22}. \]

Note. For any event \( E \), \( \Pr(E) + \Pr(E') = 1 \) and, equivalently, \( \Pr(E) = 1 - \Pr(E') \). So
\[ \Pr(\text{at least one}) = 1 - \Pr(\text{none}). \]

Example 60. A bag contains 9 tomatoes, one of which is rotten. A sample of three is chosen. What is the probability the sample contains the rotten tomato?

First, note that there are \( C(9, 3) = 84 \) total possible samples. There is \( C(1, 1) = 1 \) way to choose the rotten tomato and \( C(8, 2) = 28 \) ways to choose two more tomatoes from the remaining 8. So the probability that a random sample of three tomatoes contains the rotten tomato is
\[ \Pr(\text{contains rotten}) = \frac{28}{84} = \frac{1}{3}. \]

Example 61. A coin is tossed 4 times and the sequence of results is recorded. What is the probability that there are more heads than tails?
To ensure more heads than tails, we need either 3 heads or 4 heads in the sequence. There are \( C(4, 3) = 4 \) ways to get 3 heads and \( C(4, 4) = 1 \) way to get 4 heads. So there are \( 4 + 1 = 5 \) ways that result in more heads than tails. Since there are \( 2^4 = 16 \) total possible sequences,

\[
\Pr(\text{more H than T}) = \frac{5}{16}.
\]

**Example 62.** A 6-sided die is rolled 4 times and the sequence of results is recorded.

(a) What is the probability that 4 different numbers are rolled?

(b) What is the probability that exactly two 3’s are rolled?

(c) What is the probability that no 6’s appear?

(d) What is the probability that all the numbers appearing are odd?

First, note that there are

\[ 6^4 = 1296 \]

different possible sequences of rolls. There are

\[ P(6, 4) = 6 \cdot 5 \cdot 4 \cdot 3 = 360 \]

ways to get four different numbers rolled in sequence. Hence,

\[
\Pr(\text{all different}) = \frac{360}{1296} = \frac{5}{18}.
\]

There are \( C(4, 2) = 6 \) ways for exactly two 3’s to appear. Once those 3’s are placed, there are \( 5 \cdot 5 = 25 \) ways to assign values other than 3 to the remaining two spots. So there are \( 6 \cdot 25 = 150 \) ways to get exactly two 3’s which means

\[
\Pr(\text{exactly two 3’s}) = \frac{150}{1296} = \frac{25}{216}.
\]

There are

\[ 5^4 = 625 \]

sequences with no 6’s appearing so

\[
\Pr(\text{no 6’s}) = \frac{625}{1296}.
\]

There are 3 odd numbers between 1 and 6: 1, 3, and 5. So there are

\[ 3^4 = 81 \]

sequences with only odd numbers appearing. Therefore,

\[
\Pr(\text{all odd}) = \frac{81}{1296} = \frac{1}{16}.
\]
Example 63. Suppose 40% of pigs can talk, 52% of pigs can fly, and 29% of pigs can both fly and talk. Draw a Venn diagram to describe the information, label each region, and determine the probability that a randomly selected pig

(a) can fly or talk,
(b) can fly but not talk, and
(c) can neither fly nor talk.

From this we see that

- \( \Pr(T \cup F) = 0.11 + 0.29 + 0.23 = 0.63, \)
- \( \Pr(F \cap T') = 0.23, \)
- \( \Pr((T \cup F)') = 0.37. \)

Example 64. A coin is tossed 12 times. Find

(a) the probability there are exactly 3 heads,
(b) the probability there are at least 2 heads, and
(c) the probability there are 4 heads or 7 heads.

First, observe that there are \( 2^{12} = 4096 \) different possibilities. Of those, \( C(12, 3) = 220 \) have exactly 3 heads. So

\[ \Pr(3H) = \frac{220}{4096} = \frac{55}{1024}. \]

To compute the number of sequences where at least 2 heads appear, we note that

- there are \( C(12, 0) = 1 \) ways for no heads to appear and
- there are \( C(12, 1) = 12 \) ways for one head to appear.

Then, we see that there are \( 4096 - (1 + 12) = 4083 \) sequences with at least two heads. Thus,

\[ \Pr(\text{at least 2H}) = \frac{4083}{4096}. \]

Lastly, we note that there are
• $C(12, 4) = 495$ ways to get exactly 4 heads and
• $C(12, 7) = 792$ ways to get exactly 7 heads.

Hence,

$$\Pr(4H \text{ or } 7H) = \frac{495 + 792}{4096} = \frac{1287}{4096}.$$ 

**Example 65.** Find the probability that, in a class of 35 students, at least two students share the same birthday.

For simplicity’s sake, we will not count February 29th. Now, imagine every student’s name listed in a roster along with the day and month of their birthday. There are $365^{35}$ different possible sequences of birthdays.

Since we are interested in the event of at least two students having the same birthday, it may be easier to find out how many ways all students can have different birthdays. Indeed, there are $P(365, 35)$ ways in which all 35 students have a different birthday. What we get in the end is that the probability that at least two students share a birthday is

$$\Pr(\text{at least 2 sharing}) = 1 - \frac{P(365, 35)}{365^{35}} \approx 81.438\%.$$ 

**Example 66.** A True/False test has 10 questions. What is the probability that a student who randomly guesses each problem gets 7 or more questions correct?

There are

• $C(10, 7) = 120$ ways to get exactly 7 right,
• $C(10, 8) = 45$ ways to get exactly 8 right,
• $C(10, 9) = 10$ ways to get exactly 9 right, and
• $C(10, 10) = 1$ ways to get them all right.

There are $2^{10} = 1024$ different ways to answer all of the questions on the test. Therefore, the probability this guessing student gets at least 7 questions correct is

$$\frac{120 + 45 + 10 + 1}{1024} = \frac{11}{64}.$$ 

### 2.3 Conditional Probability

A conditional probability is essentially one relating two events $E$ and $F$ by finding the probability that $E$ occurs given that $F$ has already occurred. By assuming that $F$ has already occurred, we can think of limiting the sample space to compute the newly framed probability of $E$.

**Notation.** We write $\Pr(E|F)$ to denote the probability of $E$ given $F$.

**Example 67.** Find the probability of randomly drawing a Jack given that you have drawn a face card.
There are 12 face cards and 4 Jacks. Let $J$ be the event of drawing a Jack and $F$ be the event of drawing a face card. All Jacks are face cards so $J \cap F = J$. Treating $F$ as the restricted sample space, that is, as we’re computing the probability of drawing a Jack given that we’ve drawn a face card, we see that

$$\Pr(J|F) = \frac{4}{12} = \frac{1}{3}.$$  

**Definition 21.** In general, the *conditional probability* of $E$ given $F$, where $\Pr(F) \neq 0$, is defined to be

$$\Pr(E|F) = \frac{\Pr(E \cap F)}{\Pr(F)}.$$  

Referring back to Example 67, we see that

$$\Pr(J|F) = \frac{\Pr(J \cap F)}{\Pr(F)} = \frac{4}{52} \cdot \frac{52}{12} = \frac{1}{3},$$

consistent with our informal calculations.

**Example 68.** Of the 26 students in a speech class at Academico, 12 are Canadian, 17 are freshmen, and 5 are neither Canadian nor freshmen. Use a Venn diagram for cardinalities to find

(a) the probability a randomly selected student is Canadian and a freshman

(b) the probability a randomly selected student is Canadian given that they are a freshman

(c) the probability a randomly selected student is a freshman given that they are Canadian

First, notice that $n(C \cup F) = 26 - 5 = 21$. Then

$$\frac{n(C \cup F) + n(C \cap F)}{21 + n(C \cap F)} = \frac{n(C) + n(F)}{12 + 17}$$

which allows us to fill in the Venn diagram:
So the probability that a randomly selected student is a Canadian and a freshmen is
\[
Pr(C \cap F) = \frac{8}{26} = \frac{4}{13}.
\]
By the definition of conditional probability, the probability that a randomly selected student is Canadian given that they are a freshman is
\[
Pr(C|F) = \frac{Pr(C \cap F)}{Pr(F)} = \frac{4}{13} \cdot \frac{26}{17} = \frac{8}{17}.
\]
Similarly, the probability that a randomly selected student is a freshman given that they are Canadian is
\[
Pr(F|C) = \frac{Pr(F \cap C)}{Pr(C)} = \frac{4}{13} \cdot \frac{26}{12} = \frac{2}{3}.
\]
**Remark.** In general, if \( S \) is a sample space where all outcomes are equally likely and \( n(F) \neq 0 \),
\[
Pr(E|F) = \frac{Pr(E \cap F)}{Pr(F)} = \frac{n(E \cap F)}{n(F)} \cdot \frac{n(S)}{n(S)} = \frac{n(E \cap F)}{n(F)}.
\]

**Example 69.** We roll a 6-sided die twice.

(a) What is the probability the sum of the rolls is 7 given that one of the rolls is a 3?

(b) What is the probability that one of the rolls was a 3 given that the sum of the rolls is 7?

Let \( A \) be the event that the sum of the rolls is 7 and \( B \) be the event that one of the rolls is a 3. Then we list

- \( A = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\} \)
- \( B = \{(1, 3), (2, 3), (3, 3), (4, 3), (5, 3), (6, 3), (3, 1), (3, 2), (3, 4), (3, 5), (3, 6)\} \)
- \( A \cap B = \{(3, 4), (4, 3)\} \)

Since all outcomes, of the 36 possible outcomes, are equally likely, we see that

- \( Pr(A|B) = \frac{2}{11} \)
- \( Pr(B|A) = \frac{2}{6} = \frac{1}{3} \)
Example 70. Two cards are drawn at random, in sequence, from a standard deck of 52 cards. What is the probability that first card is an Ace and the second card is a face card?

Let $F$ be the event that the first card drawn is a face card and $A$ be the event that the second card drawn is an Ace. Notice that we can calculate $\Pr(F) = \frac{12}{52} = \frac{3}{13}$ and $\Pr(A|F) = \frac{4}{51}$.

We want to find $\Pr(A \cap F)$ and, by definition of conditional probability,

\[
\Pr(A|F) = \frac{\Pr(A \cap F)}{\Pr(F)}.
\]

Multiplying by $\Pr(F)$ yields

\[
\Pr(A \cap F) = \Pr(F) \cdot \Pr(A|F) = \frac{3}{13} \cdot \frac{4}{51} = \frac{4}{221}.
\]

**Product Rule.** If $\Pr(F) \neq 0$, then

\[
\Pr(E \cap F) = \Pr(F) \cdot \Pr(E|F).
\]

Example 71. WoodyMammoth has recently outsourced some of their guitar manufacturing to Nevada and 30% of all WoodyMammoth guitars are made at the Nevada branch. The quality control team has determined that 5% of guitars made in the Nevada branch are defective. What is the probability that a randomly purchased WoodyMammoth guitar was made in Nevada and is defective?

Let $N$ be the event that a guitar is made in Nevada and $D$ be the event that a guitar is defective. We want to find $\Pr(N \cap D)$. The information given is $\Pr(N) = .3$ and $\Pr(D|N) = .05$. By the Product Rule,

\[
\Pr(N \cap D) = \Pr(N) \cdot \Pr(D|N) = (0.3)(0.05) = 0.015
\]

so the probability that a randomly purchased WoodyMammoth guitar was made in Nevada and is defective is 1.5%.

### 2.4 Independent Events

We introduce the concept of independence with an example.

**Example 72.** Flip a fair coin twice. Let $A$ be the event that the first toss results in a heads and $B$ be the event that the second toss results in a heads. Find $\Pr(A)$, $\Pr(A|B)$, $\Pr(B)$, and $\Pr(B|A)$.

We can do this by listing out the sample space:

\{TT, HT, TH, HH\}.
Note that \( \Pr(A) = \frac{3}{4} = \frac{1}{2} \) and \( \Pr(B) = \frac{2}{4} = \frac{1}{2} \).

Now, since \( \Pr(A \cap B) = \frac{1}{4} \), we calculate
\[
\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{1}{4} \cdot \frac{2}{1} = \frac{1}{2}
\]
and
\[
\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)} = \frac{1}{2}.
\]

What we see here is that \( \Pr(A) = \Pr(A|B) \) and \( \Pr(B) = \Pr(B|A) \).

In this example, we conceptually see that \( A \) is independent of \( B \) in the sense that the occurrence of \( B \) does not affect the probability of \( A \); i.e., \( \Pr(A) = \Pr(A|B) \). In general, if \( A \) and \( B \) are events so that \( \Pr(B) \neq 0 \) and \( \Pr(A) = \Pr(A|B) \), we have
\[
\Pr(A) = \Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} \implies \Pr(A \cap B) = \Pr(A) \cdot \Pr(B).
\]

This is what drives the following definition.

**Definition 22.** Two events \( E \) and \( F \) are said to be independent events if
\[
\Pr(E \cap F) = \Pr(E) \cdot \Pr(F).
\]

**Example 73.** Suppose there are two entrance exams that need to be taken when applying to Mathemagix University. To be accepted, you need only pass one of the two exams. There is a 40% chance you pass Exam I, a 30% chance you pass Exam II, and both events are independent. What is the probability you get accepted?

The outcomes we are interested in are
- pass I and fail II
- fail I and pass II
- pass I and pass II

so we calculate the probability:
\[
(0.4)(0.7) + (0.6)(0.3) + (0.4)(0.3) = 0.58
\]

Hence, there is a 58% you get accepted.

**Example 74.** Roll a 6-sided die twice and let \( E \) be the event that the first roll is a 6 and \( F \) be the event that the second roll is a 4. Show that \( E \) and \( F \) are independent events.

We need to show that \( \Pr(E \cap F) = \Pr(E) \cdot \Pr(F) \). Firstly,
\[
\Pr(E \cap F) = \frac{1}{36}
\]
since the only outcome in \( E \cap F \) is \( (6, 4) \).
Since
\[ E = \{(6, 1), (6, 2), (6, 3), (6, 4), (6, 5), (6, 6)\} \]
and
\[ F = \{(1, 4), (2, 4), (3, 4), (4, 4), (5, 4), (6, 4)\} \]
we see that \( \Pr(E) = \frac{1}{6} \) and \( \Pr(F) = \frac{1}{6} \).

Thus,
\[ \Pr(E) \cdot \Pr(F) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} = \Pr(E \cap F) \]
which establishes that \( E \) and \( F \) are independent.

**Example 75.** Two tests are used to determine whether or not someone has been contaminated with nanomachines. Test I is correct 70% of the time and Test II is correct 80% of the time. At least one of them is correct 99% of the time. Determine whether or not the correct identification by these tests are independent.

By the Inclusion-Exclusion Principle,
\[ \Pr(I \cup II) = \Pr(I) + \Pr(II) - \Pr(I \cap II) \]
\[ 0.99 = 0.7 + 0.8 - \Pr(I \cap II) \]
which provides
\[ \Pr(I \cap II) = 0.51 \]

Now, notice that
\[ \Pr(I) \cdot \Pr(II) = (0.7)(0.8) = 0.56 \neq 0.51 = \Pr(I \cap II) \]
so the correct identification by these tests are not independent; i.e., they are dependent.

**Example 76.** Suppose \( E \) and \( F \) are mutually exclusive with \( \Pr(E) \neq 0 \) and \( \Pr(F) \neq 0 \). Determine whether or not \( E \) and \( F \) are independent.

Since \( E \) and \( F \) are mutually exclusive,
\[ \Pr(E \cap F) = \Pr(\emptyset) = 0 \neq \Pr(E) \cdot \Pr(F). \]
That is, \( E \) and \( F \) are dependent.

### 2.5 Bayes’ Theorem

We’ll build up to Bayes’ Theorem by first introducing tree diagrams and using them to evaluate certain probabilities.

**Example 77.** Two stones are drawn from a cup containing 2 white and 8 black stones.

(a) What is the probability of drawing a white stone on the second draw?

(b) What is the probability of drawing a black stone on the second draw?
We are interested in the second stone being white but this depends on the first stone drawn. There are two possibilities for the first stone: it can be white or black. Let $W_2$ be the event that the second stone drawn is white, $W_1$ be the event that the first stone drawn is white, and $B$ be the event that the first stone drawn is black. Let’s investigate both cases.

- Assuming the first stone drawn is white, we compute
  $$\Pr(W_1 \cap W_2) = \Pr(W_1) \cdot \Pr(W_2|W_1) = \frac{2}{10} \cdot \frac{1}{9} = \frac{1}{45}.$$ 

- Now, assuming the first stone drawn is black, we compute
  $$\Pr(B \cap W_2) = \Pr(B) \cdot \Pr(W_2|B) = \frac{8}{10} \cdot \frac{2}{9} = \frac{8}{45}.$$ 

Since these two cases are separate and capture all possibilities where the second draw results in a white stone, we see

$$\Pr(W_2) = \frac{1}{45} + \frac{8}{45} = \frac{1}{5}.$$ 

Now, we can capture this process using trees.

![Tree Diagram](https://via.placeholder.com/150)

From the tree, we can compute the probability that the second stone drawn is black:

$$\frac{8}{45} + \frac{28}{45} = \frac{4}{5}.$$ 

**Example 78.** Two stones are drawn from a cup containing 2 white and 8 black stones. What is the probability that the first stone was black given that the second stone is white?

Note that

$$\Pr(1^\text{st} \, B|2^\text{nd} \, W) = \frac{\Pr(1^\text{st} \, B \cap 2^\text{nd} \, W)}{\Pr(2^\text{nd} \, W)}$$

which we compute by referring to the tree in Example 77:

$$\Pr(1^\text{st} \, B|2^\text{nd} \, W) = \frac{\Pr(1^\text{st} \, B \cap 2^\text{nd} \, W)}{\Pr(2^\text{nd} \, W)} = \frac{\frac{8}{45}}{\frac{1}{5} + \frac{8}{45}} = \frac{8}{9}.$$ 

In general, we have
**Theorem 23** (Bayes’ Theorem). Given a sample space $S$, mutually exclusive events $F_1, F_2, \ldots, F_n$ so that $S = F_1 \cup F_2 \cup \cdots \cup F_n$, and an event $E$ with $\Pr(E) \neq 0$,

$$\Pr(F_j|E) = \frac{\Pr(F_j) \cdot \Pr(E|F_j)}{\Pr(F_1) \cdot \Pr(E|F_1) + \Pr(F_2) \cdot \Pr(E|F_2) + \cdots + \Pr(F_n) \cdot \Pr(E|F_n)}.$$ 

**Example 79.** We have discovered that aliens have implanted nanomachines in 3% of the human population and a new test has an accuracy rate of 98%. That is, it correctly identifies a person with nanomachines 98% of the time and it correctly identifies someone without nanomachines 98% of the time. If you have tested positive for nanomachines, find the probability that you actually have nanomachines.

Using a tree diagram:

```
  +-----------------+-----------------+
  |                 |                 |
  | 98%             | 2%              |
  |                 |                 |
  +-----------------+-----------------+
  |                 |                 |
  | 3%              | 97%             |
  |                 |                 |
  +-----------------+-----------------+
  |                 |                 |
  | have            | don’t have      |
  |                 |                 |
  +-----------------+-----------------+
  |                 |                 |
  | 0.03 \cdot 0.98 = 0.0294 |
  | 0.03 \cdot 0.02 = 0.0006 |
  |                 |                 |
  +-----------------+-----------------+
  |                 |                 |
  | 0.97 \cdot 0.02 = 0.0194 |
  | 0.97 \cdot 0.98 = 0.9506 |
  |                 |                 |
  +-----------------+-----------------+
```

From here, we see that $0.0294 + 0.0194 = 0.0488$ tested positive, $0.0294$ of those are correctly identified so

$$\frac{0.0294}{0.0488} \approx 0.6025$$

would actually have nanomachines given that they tested positive for them.

Now, to help make this clearer, imagine that the population is 10 000. Then 300 have nanomachines. Of those 300 with nanomachines, 98% of them will test positively; i.e., 294 of them will be correctly identified. Of those 9700 without nanomachines, 98% of them will test negatively; i.e., 9506 of them will be correctly identified. Let’s organize this in a table:

<table>
<thead>
<tr>
<th>actually positive</th>
<th>tested positive</th>
<th>tested negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>actually negative</td>
<td>294</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>194</td>
<td>9506</td>
</tr>
</tbody>
</table>

So there are 488 people which test positive, 294 of those actually have it. Hence, the probability that you actually have nanomachines given you tested positive for them is $294/488 \approx 0.6025$.

**Example 80.** An alien sensing device has 6 different components. Failure rates and proportions are given below.
If any one of the components fails\(^3\), the whole device fails.

(a) What is the probability that the device doesn’t fail?

(b) Assuming the device has failed, what is the probability that it was caused by component 1?

(c) Assuming the device has failed, which component most likely caused the failure?

To find the probability that the device doesn’t fail, we first calculate the probability that the device fails:

\[
\Pr(\text{failure}) = (0.35)(0.0007) + (0.25)(0.0013) + (0.15)(0.0002) \\
+ (0.12)(0.05) + (0.1)(0.01) + (0.03)(0.009) \\
= 0.00787
\]

So the probability that the device doesn’t fail is

\[
1 - 0.00787 = 0.99213
\]

To find the probability that component 1 failed given that the device has failed, we let \( A \) be the event that component 1 fails and \( B \) be the event that the device fails. Above, we found \( \Pr(B) = 0.00787 \). Since \( A \cap B = A \), we see that

\[
\Pr(A \cap B) = \Pr(A) = (0.35)(0.0007) = 0.000245
\]

Therefore,

\[
\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{0.000245}{0.00787} \approx 3.113\%.
\]

Consider the table of probabilities:

<table>
<thead>
<tr>
<th>Component</th>
<th>Proportion</th>
<th>Failure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.35</td>
<td>0.0007</td>
</tr>
<tr>
<td>2</td>
<td>0.25</td>
<td>0.0013</td>
</tr>
<tr>
<td>3</td>
<td>0.15</td>
<td>0.0002</td>
</tr>
<tr>
<td>4</td>
<td>0.12</td>
<td>0.05</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>0.01</td>
</tr>
<tr>
<td>6</td>
<td>0.03</td>
<td>0.009</td>
</tr>
</tbody>
</table>

\(^3\)and only one component can fail for any given device
2.6 Binomial Trials

Definition 24. A random variable is any rule $X$ which assigns numerical values to each outcome of a sample space $S$.

Example 81. Toss a coin 15 times and let $X$ be the number of heads appearing. Then $X$ is a random variable.

Definition 25. A binomial experiment has the following properties:

- The number of trials in the experiment is a fixed number $n$.
- There are two outcomes: Success or Failure.
- The probability of success in each trial is the same, $p$. Since there are only two outcomes, the probability of failure is $q = 1 - p$.
- The trials are independent of each other.

Remark. For a binomial experiment, we usually let $X$ be the random variable giving the number of successes.

Example 82. Toss a coin 12 times and consider heads the “success” case. Then let $X$ be the random variable which counts the number of successes. We write $X = 4$ to mean four heads appeared.

Notation. For a binomial experiment, we write $\Pr(X = k)$ to denote the probability that $X = k$; i.e., that there are $k$ successes.

Formula. Given a binomial experiment with $n$ trials where $p$ is the probability of success, $q = 1 - p$ is the probability of failure, and $X$ is the number of successes, the probability of $k$ successes is

$$\Pr(X = k) = \binom{n}{k} p^k q^{n-k}.$$ 

Example 83. When a pair of 6-sided dice is rolled, the probability of obtains a sum of 7 is $\frac{1}{6}$. Suppose a pair of dice is rolled 25 times. Find the probability that the sum is 7 eight times.

This is a binomial experiment where

- $n = 25$,
- success is the sum being 7,
- $p = \frac{1}{6}$,
- $q = \frac{5}{6}$, and
- $k = 8$.

By the formula,

$$\Pr(X = 8) = \binom{25}{8} \left(\frac{1}{6}\right)^8 \left(\frac{5}{6}\right)^{17} \approx 2.902\%.$$
Note. Notice that, doing Example 83 combinatorially, we would count that there are

- \( C(25, 8) \) to choose the rolls which sum to 7,
- 6 ways that a pair of dice add up to 7,
- 6^8 ways that eight pairs add up to 7,
- 36 - 6 = 30 ways that a pair of dice don’t add up to 7,
- 30^{17} ways that 17 pairs don’t add up to 7.

Then we would calculate the probability that eight pairs add up to 7 to be

\[
\frac{C(25, 8) \cdot 6^8 \cdot 30^{17}}{36^{25}} \approx 2.902\%.
\]

Example 84. A coin is weighted so that when it is flipped, there is a 70% chance it ends up tails. If we toss this coin 15 times, find

(a) the probability that we end up with exactly 10 tails,

(b) the probability that we end up with at most 10 tails, and

(c) the probability that we end up with at least 11 tails.

2.7 Expected Value

Conceptually, the expected value of a random value is the value which we can expect on average. In general, the expected value doesn’t have to be a value that the random variable can possibly assume.

Definition 26. For a random variable \( X \), the expected value of \( X \) is defined to be

\[
E(X) = x_1p_1 + x_2p_2 + \cdots + x_np_n
\]

where \( x_1, x_2, \ldots, x_n \) are the values of \( X \) and \( p_j = \Pr(X = x_j) \) for \( 1 \leq j \leq n \).

Example 85. A 6-sided die is rolled. Find the expected value.

Check the table:

<table>
<thead>
<tr>
<th>outcome, ( x )</th>
<th>( \Pr(X = x) )</th>
<th>( x \cdot \Pr(X = x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( \frac{1}{6} )</td>
<td>( \frac{1}{6} )</td>
</tr>
<tr>
<td>2</td>
<td>( \frac{1}{6} )</td>
<td>( \frac{2}{6} )</td>
</tr>
<tr>
<td>3</td>
<td>( \frac{1}{6} )</td>
<td>( \frac{3}{6} )</td>
</tr>
<tr>
<td>4</td>
<td>( \frac{1}{6} )</td>
<td>( \frac{4}{6} )</td>
</tr>
<tr>
<td>5</td>
<td>( \frac{1}{6} )</td>
<td>( \frac{5}{6} )</td>
</tr>
<tr>
<td>6</td>
<td>( \frac{1}{6} )</td>
<td>( \frac{6}{6} )</td>
</tr>
</tbody>
</table>
which shows that

\[ E(X) = \frac{21}{6} = 3.5 \]

Of course, we can never roll a 3.5 but expected value doesn’t have to be possible outcomes.

**Example 86.** Roll a 6-sided die twice and let \( X \) be the sum of the rolls. Find the expected value of \( X \).

Recall that

<table>
<thead>
<tr>
<th>( x )</th>
<th>( f(x) )</th>
<th>( x )</th>
<th>( f(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1/36</td>
<td>8</td>
<td>5/36</td>
</tr>
<tr>
<td>3</td>
<td>2/36</td>
<td>9</td>
<td>4/36</td>
</tr>
<tr>
<td>4</td>
<td>3/36</td>
<td>10</td>
<td>3/36</td>
</tr>
<tr>
<td>5</td>
<td>4/36</td>
<td>11</td>
<td>2/36</td>
</tr>
<tr>
<td>6</td>
<td>5/36</td>
<td>12</td>
<td>1/36</td>
</tr>
<tr>
<td>7</td>
<td>6/36</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

which shows that \( E(X) = 7 \). Coincidentally, in this case, this is also the most likely outcome.

**Example 87.** You pay $1 to draw a card from a standard deck of 52 cards. If you draw

- an Ace, you win $9,
- a face card, you win $2, and
- otherwise, you win nothing.

If \( X \) is your net earning, find the expected value of \( X \). Is this game worth playing?

Incorporating the cost, notice that

\[ E(X) = 8 \cdot \frac{4}{52} + 1 \cdot \frac{12}{52} + (-1) \cdot \frac{36}{52} = \frac{2}{13} \approx \$0.15. \]

So, on average, we should expect to win 15 cents. As this is a positive value, this is a good game to play.
3 Matrices

3.1 Systems of Linear Equations

Notice that, for constant values \( a, b, \) and \( c \), \( ax + by = c \), as long as \( b \neq 0 \), can be written as

\[
y = \frac{-a}{b} \cdot x + \frac{c}{b},
\]

the familiar form of a line. For this reason, we say that the equation \( ax + by = c \) is linear. For three variables, a linear equation will be one of the form

\[
ax + by + cz = d.
\]

Definition 27. A system of linear equations is any collection of linear equations sharing the same variables.

Example 88. The following is a system of equations:

\[
\begin{cases}
5x - 3y - 6z = -4 \\
3x + y - 5z = 5 \\
4x - 2y + z = -13
\end{cases}
\]

We’ll be interested in finding solution sets to these equations. First, let’s consider some 2-dimensional examples by graphing.

Example 89. The system

\[
\begin{cases}
-6x + 3y = 1 \\
2x - y = 2
\end{cases}
\]

has no solutions. This is seen by graphing the two lines and noting that they are parallel and not equal.

Example 90. The system

\[
\begin{cases}
6x + y = 1 \\
2x - 3y = 7
\end{cases}
\]

has exactly one solution. This is seen by graphing the two lines and noting that they intersect at one point.

Example 91. The system

\[
\begin{cases}
-6x + 3y = -9 \\
2x - y = 3
\end{cases}
\]

has infinitely many solutions. This is seen by graphing and noting that they are the same line.

When solving systems of linear equations, we want to get them into diagonal form which is

\[
\begin{cases}
x = a \\
y = b \\
z = c
\end{cases}
\]

We can do this using the Gauss-Jordan elimination method which will involve matrices and a combination of elementary row operations and pivots.
Elementary Row Operations. For systems of equations, the elementary row operations are:

1. Interchange any two equations.
2. Multiply an equation by a non-zero number.
3. Change an equation by adding to it a multiple of another equation.

Remark. Row operation 2 is used to get the one and row operation 3 is to get the zeroes. We will discuss this more when we get to pivots.

Example 92. Here, we will give examples of each elementary row operation.

1. This is an example of interchanging two equations along with the appropriate notation:

\[
\begin{align*}
5x - 3y - 6z &= -4 \\
3x + y - 5z &= 5 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

\[
\begin{align*}
3x + y - 5z &= 5 \\
5x - 3y - 6z &= -4 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

- \( R_{1} \leftrightarrow R_{2} \rightarrow \)

2. This is an example of multiplying an equation by a non-zero number with the appropriate notation:

\[
\begin{align*}
5x - 3y - 6z &= -4 \\
3x + y - 5z &= 5 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

\[
\begin{align*}
2R_{3} \rightarrow \quad 5x - 3y - 6z &= -4 \\
3x + y - 5z &= 5 \\
8x - 4y + 2z &= -26 \\
\end{align*}
\]

3. This is an example of changing an equation by adding a multiple of another row to it with the appropriate notation:

\[
\begin{align*}
5x - 3y - 6z &= -4 \\
3x + y - 5z &= 5 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

\[
\begin{align*}
R_{2} - 2R_{3} \rightarrow \quad 5x - 3y - 6z &= -4 \\
-5x + 5y - 7z &= 31 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

Example 93. We will often be performing elementary row operations multiple times. For example,

\[
\begin{align*}
5x - 3y - 6z &= -4 \\
3x + y - 5z &= 5 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

\[
\begin{align*}
R_{2} - 2R_{3} \rightarrow \quad 5x - 3y - 6z &= -4 \\
-5x + 5y - 7z &= 31 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

\[
\begin{align*}
R_{1} + R_{2} \rightarrow \quad -5x + 5y - 7z &= 31 \\
5x - 3y - 6z &= -4 \\
4x - 2y + z &= -13 \\
\end{align*}
\]

\[
\begin{align*}
R_{2} + R_{1} \rightarrow \quad -5x + 5y - 7z &= 31 \\
2y - 13z &= 27 \\
4x - 2y + z &= -13 \\
\end{align*}
\]
Definition 28. For a system of equations
\[
\begin{align*}
  a_1x + b_1y + c_1z &= d_1 \\
  a_2x + b_2y + c_2z &= d_2 \\
  a_3x + b_3y + c_3z &= d_3
\end{align*}
\]
the corresponding **augmented matrix** is
\[
\begin{bmatrix}
  a_1 & b_1 & c_1 & d_1 \\
  a_2 & b_2 & c_2 & d_2 \\
  a_3 & b_3 & c_3 & d_3
\end{bmatrix}
\]
Column 1 consists of the coefficients for \(x\), column 2 the coefficients for \(y\), and column 3 the coefficients for \(z\). The vertical bar is used as a reminder that column 4 consists of the constant values.

Using the augmented matrix, we can symbolically simplify the elementary row operations.

**Elementary Row Operations.** For matrices, the elementary row operations are:

1. Interchange any two rows.
2. Multiply a row by a non-zero number.
3. Change a row by adding to it a multiple of another row.

**Example 94.** Repeat Example 93 using the augmented matrices.
\[
\begin{bmatrix}
  5 & -3 & -6 & -4 \\
  3 & 1 & -5 & 5 \\
  4 & -2 & 1 & -13
\end{bmatrix}
\]
\[
\overset{R_2-2R_3}{\longrightarrow}
\begin{bmatrix}
  5 & -3 & -6 & -4 \\
  -5 & 5 & -7 & 31 \\
  4 & -2 & 1 & -13
\end{bmatrix}
\]
\[
\overset{R_1\leftrightarrow R_2}{\longrightarrow}
\begin{bmatrix}
  -5 & 5 & -7 & 31 \\
  5 & -3 & -6 & -4 \\
  4 & -2 & 1 & -13
\end{bmatrix}
\]
\[
\overset{R_2+R_1}{\longrightarrow}
\begin{bmatrix}
  -5 & 5 & -7 & 31 \\
  0 & 2 & -13 & 27 \\
  4 & -2 & 1 & -13
\end{bmatrix}
\]

**Pivot Method.** To pivot a matrix about a given non-zero entry,

1. Multiply the row by the reciprocal of the entry.
2. Transform all other entries of the same column to zero (by using the one in the pivot entry).

**Example 95.** Pivot around the circled entry in
\[
\begin{align*}
  4x - 2y + 3z &= 4 \\
  8x - 3y + 5z &= 7 \\
  7x - 2y + 4z &= 6
\end{align*}
\]
Observe:

\[
\begin{cases}
4x - 2y + 3z = 4 \\
8x - 3y + 5z = 7 \\
7x - 2y + 4z = 6
\end{cases}
\begin{array}{c}
\xrightarrow{R_1} \quad \begin{cases}
x - \frac{1}{2} y + \frac{3}{4} z = 1 \\
8x - 3y + 5z = 7 \\
7x - 2y + 4z = 6
\end{cases}
\end{array}
\begin{array}{c}
\xrightarrow{R_2 - 8R_1} \quad \begin{cases}
x - \frac{1}{2} y + \frac{3}{4} z = 1 \\
y - z = -1 \\
7x - 2y + 4z = 6
\end{cases}
\end{array}
\begin{array}{c}
\xrightarrow{R_3 - 7R_1} \quad \begin{cases}
x - \frac{1}{2} y + \frac{3}{4} z = 1 \\
y - z = -1 \\
\frac{3}{2} y - \frac{5}{4} z = -1
\end{cases}
\end{array}
\]

We will see how this method of pivoting looks in the context of augmented matrices when we learn about the Gauss-Jordan elimination method.

Remark. Generally, the Gauss-Jordan elimination method on an augmented matrix

\[
\begin{bmatrix}
a_1 & b_1 & c_1 & d_1 \\
a_2 & b_2 & c_2 & d_2 \\
a_3 & b_3 & c_3 & d_3
\end{bmatrix}
\]

follows a particular recipe:

1. If \(a_1 = 0\), swap \(R_1\) with another row where \(a_j \neq 0\).

2. With \(a_1 \neq 0\), perhaps after relabeling, pivot the matrix around \(a_1\). This results in a new augmented matrix of the form

\[
\begin{bmatrix}
1 & b_1^* & c_1^* & d_1^* \\
0 & b_2^* & c_2^* & d_2^* \\
0 & b_3^* & c_3^* & d_3^*
\end{bmatrix}
\]

3. As long as \(b_2^* \neq 0\), now pivot the matrix about \(b_2^*\) which results in a matrix of the form

\[
\begin{bmatrix}
1 & 0 & c_1^{**} & d_1^{**} \\
0 & 1 & c_2^{**} & d_2^{**} \\
0 & 0 & c_3^{**} & d_3^{**}
\end{bmatrix}
\]

4. As long as \(c_3^{**} \neq 0\), pivot the matrix about \(c_3^{**}\) which results in a matrix of the form

\[
\begin{bmatrix}
1 & 0 & 0 & d_1^{***} \\
0 & 1 & 0 & d_2^{***} \\
0 & 0 & 1 & d_3^{***}
\end{bmatrix}
\]
Note. Generally, if we end up with a matrix that looks like
\[ \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ 0 & 0 & 0 & 0 \end{bmatrix} \]
during the course of applying the Gauss-Jordan elimination method, then the original system has **infinitely many solutions**.

Note. Generally, if we end up with a matrix that looks like
\[ \begin{bmatrix} ? & ? & ? & ? \\ ? & ? & ? & ? \\ 0 & 0 & 0 & t \end{bmatrix} \]
where \( t \neq 0 \) during the course of applying the Gauss-Jordan elimination method, then the original system has **no solutions**.

---

**Example 96.** Solve the system
\[
\begin{align*}
4x - 2y + 3z &= 4 \\
8x - 3y + 5z &= 7 \\
7x - 2y + 4z &= 6
\end{align*}
\]
using the Gauss-Jordan elimination method and augmented matrices.

Recall that the goal of the Gauss-Jordan elimination method is to apply elementary row operations in a formulaic way to arrive at a diagonal matrix. Let’s see how it’s done:

\[
\begin{bmatrix}
\begin{array}{cccc}
\frac{1}{4} & -2 & 3 & 4 \\
8 & -3 & 5 & 7 \\
7 & -2 & 4 & 6
\end{array}
\end{bmatrix}
\xrightarrow{\frac{1}{4}R_1}
\begin{bmatrix}
\begin{array}{cccc}
1 & -1/2 & 3/4 & 1 \\
8 & -3 & 5 & 7 \\
7 & -2 & 4 & 6
\end{array}
\end{bmatrix}
\xrightarrow{R_2-8R_1}
\begin{bmatrix}
\begin{array}{cccc}
1 & -1/2 & 3/4 & 1 \\
0 & 1 & -1 & -1 \\
7 & -2 & 4 & 6
\end{array}
\end{bmatrix}
\xrightarrow{R_3-7R_1}
\begin{bmatrix}
\begin{array}{cccc}
1 & -1/2 & 3/4 & 1 \\
0 & 1 & -1 & -1 \\
0 & 3/2 & -5/4 & -1
\end{array}
\end{bmatrix}
\xrightarrow{R_1+\frac{1}{4}R_2}
\begin{bmatrix}
\begin{array}{cccc}
1 & 0 & 1/4 & 1/2 \\
0 & 1 & -1 & -1 \\
0 & 3/2 & -5/4 & -1
\end{array}
\end{bmatrix}
\xrightarrow{R_3-\frac{3}{4}R_2}
\begin{bmatrix}
\begin{array}{cccc}
1 & 0 & 1/4 & 1/2 \\
0 & 1 & -1 & -1 \\
0 & 0 & 1/4 & 1/2
\end{array}
\end{bmatrix}
\]
This completes the Gauss-Jordan elimination method. If we rewrite the final augmented matrix as a system of equations, we have

\[
\begin{cases}
x &= 0 \\
y &= 1 \\
z &= 2
\end{cases}
\]

and we see that our solution is \( x = 0, y = 1, z = 2 \).

**Example 97.** Solve

\[
\begin{cases}
x - 3y + z &= 5 \\
-2x + 7y - 6z &= -9 \\
x - 2y - 3z &= 6
\end{cases}
\]

**Example 98.** Solve

\[
\begin{cases}
x + 2y + 3z &= 1 \\
4x + 5y + 6z &= -9 \\
x + 2y + 3z &= 6
\end{cases}
\]

**Example 99.** Solve

\[
\begin{cases}
3y + 3z &= 6 \\
x - 5y + z &= 1 \\
3z &= 9
\end{cases}
\]

### 3.2 Arithmetic with Matrices