Chapter 1

Introduction

"The subject of probability theory is the foundation upon which all of statistics is built, providing a means for modeling populations, experiments, or almost anything else that could be considered a random phenomenon. Through these models, statisticians are able to draw inferences about populations, inferences based on examination of only a part of the whole". - Statistical Inference, Casella and Berger (2002)

1.1 Why Probability Theory for Statistics?

There are four (4) things to note from the quote above:

1. **Foundation**

   As a branch of Science, Statistics is never exempted by the requirement for its core concepts to be *founded in mathematical rigor* . This is to ensure that, as a body of knowledge, every process involved and results derived thereof is logically consistent, i.e., one could never arrive to contradictory (i.e. nonsense) outcomes.

2. **Modelling**

   A *model* in Science is a simplified representation of something in reality called *phenomenon* that is often very complex to observe directly but are useful to be ignored. The simplification may remove some subtle details from what it actually represents. However, it allows scientists to draw a partial yet much quicker and useful explanation of the said complex phenomenon.

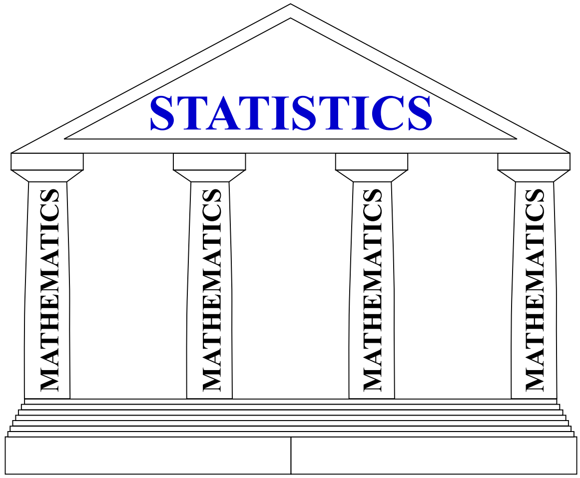

Figure 1.1: Mathematics being a foundation of Statistics

Consider the phenomenon of gravitation between two celestial bodies in space (see Figure 1.2. To explain this force, two well-known mathematical models are proposed: *Newton's law of gravitation* and *Einstein's general relativity*. Both models attempt to give an explanation of this phenomenon but each differs (1) what things are assumed and (2) consequently what variables to include.

3. **Random Phenomenon**

If there is one phenomenon that is prevalent in reality, that should be *random phenomenon* . In simple terms, a random phenomenon is any event or experience whose actual occurrence is *uncertain* . The fluctuation of stock prices, your grade on this subject, the exact position and velocity of a subatomic particle at a given time, the number of babies born around the world every second, the weather tomorrow, and the survival of the cancer patient tomorrow are just some of the many examples of random phenomena. The information that is systematically measured, collected and stored of the actual *realizations* (i.e. outcomes) of such uncertain events comprises to that of what we come to call as *data* . The uncertainty inherent in such source of important information can be quantified using the concept of *probability* . The sub-field of Mathematics upon which the concepts of probability are rigorously formalized is what we call *Probability Theory*.

A classic example of a random phonemona is the result of a coin toss (see Figure 1.3). Unless every possible variables (e.g. environmental factors) are

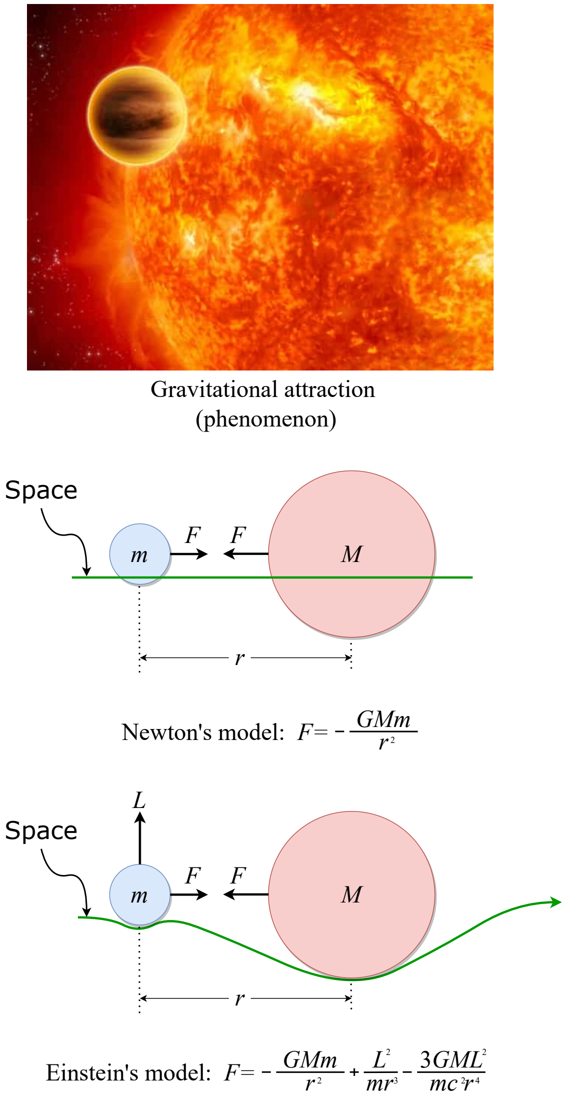

Figure 1.2: Two models for the gravitational attraction between two mass bodies.

perfectly quantified, no one can predict with 100% certainty the result of a coin toss.

Figure 1.3: An actual demonstration of a coin toss

4. **Draw Inference**

   With models serving the purpose of allowing scientists to deduce a simplified explanations about the phenomenon it represents, Probability Theory provides statisticians and data analysts alike the modelling tools to represent random phenomena. The process of drawing explanations about a random phenomena based on a statistical model is what we call *inference* .

   Consider again the coin toss scenario introduced above. Suppose we denote X as the outcome of the toss such that $X = 1$ if the result is Head and zero if the outcome is Tail. A theoretical model that quantifies the degree of certainty (i.e. probability) what the next outcome X is is given by the following equation:

$$f(x|p) = p^x(1-p)^{1-x}, \ x = 0, 1 \tag{1.1}$$

for some $p \in (0, 1)$[1]. Suppose we toss the coin in the scenario above $N = 100$ times resulting to sequence of 0's and 1's:

$$\mathsf{Data} = \{0, 0, 0, 1, 1, 0, \dots, 1, 1, 1, 0, 1, 1\}$$

The goal of inference is to *estimate* the value p for a given the Data. An estimate of p is

$$\widehat{p} = \frac{\sum\limits_{x \in \mathsf{Data}} x}{N}.$$

For example, if $\widehat{p} = 0.5$ then a quantified answer to the question

> "*What is the chance the next toss results to Head?*"

is $f(x = 1|p = 0.5) = 0.5^1(1-0.5)^{1-1} = 0.5$. One interpretation for this result is

> "There's a 50% chance that the next toss results to Head."

Equation (1.1) is an example of what we call *probability distribution* .

We will discuss more about these in the succeeding chapters.

[1] Please don't get intimidated at this point with this equation. Everything will be clearer in the succeeding chapters.

ASSIGNMENT

- **Problem.** Think of a profession. Discuss in one paragraph (between 3 to 10 sentences)a specific scenario in their field of expertise that involves uncertain events. Enumerate at least two possible realizations for such scenario.

- **Examples.** The following are example discussions of a scenario:

  Example 1. A medical doctor is attending to a Leukemia patient. The patient's parents are asking the doctor for an advice on deciding whether their child should undergo chemotherapy or not. The uncertain event here is final decision of the doctor that has two possible outcomes: "undergo" or "not undergo".

  Example 2. A meteriologist held a press conference discussing the issues of drought on a certain province. One of the reporters from certain TV network asked her what would be the expected temperature next week on the said province. The uncertain event in this scenario would be the estimated temperature announced by the meteorologist which the values could range from negative to positive real numbers of a specific unit (e.g. C°, F°, etc).

- **Submission.** Instructions for submission are as follows:

  1. First, log-in to your my.iit accounts in `my.iit.edu.ph`.
  2. Write your answers on a Word document (e.g. MS Word or Google Docs) with your `g.msuiit.edu.ph` email address written on the upper write corner like this example.
  3. Save your document with the naming format `surname.firstname-stt131-assignment-01` like this example.
  4. Submit your answers to this drive folder. This drive folder will be closed 6 hours after the class and will no longer accept any other submissions.

1.2 Some Preliminary Mathematics

As with every standard undergraduate textbook in Probability theory, introducing the basics of sets and counting techniques is inevitable. The notions of sets enables us to introduce necessary notations while counting techniques give us a precursor to introduce simple probability distributions like Equation (1.1).

1.2.1 Sets

Sets are often used in solving probability problems. The same with Probability Theory, a sub-field of mathematics dedicated to formalizing the concepts of sets is called *Set Theory*. We provide a brief review of set theory here.

Important: The following discussion assumes that you have a basic working knowledge of intervals and the Cartesian coordinate system.

Definition 1.1. A ***set*** is a collection of objects (elements).

Remark. Upper-case letters are typically used to denote sets, for example, A, B. The notion of a set is a very general one, as will be seen in the example below.

Example 1.1. Consider the following four sets.

$$\begin{aligned}
A &= \{1, 2, \dots, 100\} \\
B &= \{x : x \text{ is a positive integer less than } 101.\} \\
C &= \{\text{MSU-IIT, MSU-Marawi}\} \\
D &= \{(x, y) \,|\, 0 < x < 1,\ 0 < y < 2\} \\
E &= \{\omega \in \Omega \,|\, X(\omega) = 1\} \\
F &= (-1, 7]
\end{aligned}$$

Respectively, the above sets are read as follows:

- A is the set containing counting numbers between 1 to 100.
- B is the set of x's such that x is a positive integer less than 101.
- C is the set containing the names MSU-IIT and MSU-Marawi.
- D is the set of all (Cartesian coordinate) points (x, y) such that x is between 0 and 1 while y ranges between 0 to 2.
- E is the set of all omega (ω) belonging to some Ω such that if we plug it in to the function X, the output is 1.
- F is the set of all real numbers between -1 and 7 including 7.

Remark. Notice that we read both the colon ":" and vertical slash "|" as "*such that*".

Remark. Set F is what we call an ***interval***. If we denote $\mathbb{R}$ as the set of all real numbers, then

$$F = (-1, 7] = \{x \in \mathbb{R} : -1 < x \leq 7\}.$$

Remark. It is possible for sets to have elements that are also sets. To differentiate this from the common sets, we denote it with script letters $\mathscr{A}, \mathscr{B}, \mathscr{S}$. The following are example of such sets:

$$\mathscr{A} = \{\{1\}, \{1, 2\}, \{1, 2, 3\}\}$$
$$\mathscr{B} = \left\{\left(-\frac{1}{n}, 1 - \frac{1}{n}\right] : n = 1, 2, 3, \ldots\right\}$$

Set $\mathscr{A}$ is a set cotaining the sets $\{1\}$, $\{1, 2\}$, and $\{1, 2, 3\}$ while set $\mathscr{B}$ is a collection of intervals of the form $(-\frac{1}{n}, 1 - \frac{1}{n}]$ for $n = 1, 2, 3, \ldots$. Examples of such intervals are $(-1, 0]$, $(-\frac{1}{2}, \frac{1}{2}]$, and $(-\frac{1}{3}, \frac{2}{3}]$.

Remark. Oftentimes it is more convenient to write sets using directly the rules defining the elements belonging to it specially when some conditions are already known. Consider set E in the example above. Suppose it is already known that every set in a given context, including E, contains only ω's as elements. Then, in place of $E = \{\omega \in \Omega \,|\, X(\omega) = 1\}$, we can simply drop the ω and write using only the condition as $\{X = 1\}$. More examples are as follows:

Original	Simplified (assuming ω in context)
$\{\omega : 0 \leq X(\omega) \leq 10\}$	$\{1 \leq X \leq 10\}$
$\{\omega : X(\omega) > 12\}$	$\{X > 12\}$
$\{\omega : 0 \leq X(\omega) < 1,\ -2 \leq Y(\omega) \leq 2\}$	$\{0 \leq X < 1,\ -2 \leq Y \leq 2\}$
$\{\omega : X(\omega) \in (0, 1]\}$	$\{X \in (0, 1]\}$

This notation will be used in later chapters.

Definition 1.2. If an object belongs to a set, it is said to be an ***element of*** the set. The notation $\in$ is used to denote membership in a set.

Example 1.2. Using the sets defined in Example 1.1,

$$17 \in A, \quad 99 \in B, \quad (\tfrac{2}{3}, 1) \in D, \quad \text{UP-Diliman} \notin C.$$

Definition 1.3. If every element of a set A is also an element of a set B, then we say that A is ***subset of*** of B and write $A \subseteq B$.

Definition 1.4. If $A \subseteq B$ but $B \subseteq A$, then we say that A is ***equal to*** of B and write $A = B$.

Definition 1.5. If $A \subseteq B$ but $B \nsubseteq A$, then we say that A is ***proper subset of*** of B and write $A \subset B$.

Example 1.3. Let $\mathbb{N}$, $\mathbb{W}$, $\mathbb{Z}$, $\mathbb{Z}^+$, $\mathbb{Z}^{\geq 0}$, $\mathbb{Q}$, $\mathbb{Q}^c$, and $\mathbb{R}$ be defined as follows:

$$\begin{aligned}
\mathbb{N} &= \{1, 2, 3, 4, \dots, n, n+1, \dots\} \\
\mathbb{W} &= \{0, 1, 2, 3, 4, \dots, n, n+1, \dots\} \\
\mathbb{Z} &= \{\dots, -3, -2, -1, 0, 1, 2, 3, 4, \dots\} \\
\mathbb{Z}^+ &= \{1, 2, 3, 4, \dots\} \\
\mathbb{Z}^{\geq 0} &= \{0, 1, 2, 3, 4, \dots\} \\
\mathbb{Q} &= \{\tfrac{m}{n} : m \in \mathbb{Z},\, n \in \mathbb{N}\} \\
\mathbb{Q}^c &= \{x : x \text{ has a non-terminating} \\
&\qquad\quad \text{non-repeating decimal representation} \\
&\qquad\quad \text{(e.g. } \pi = 3.1415\dots)\} \\
\mathbb{R} &= \{x : x \in \mathbb{Q} \text{ or } x \in \mathbb{Q}^c\}
\end{aligned}$$

Then

$$\mathbb{N} \subseteq \mathbb{Z}^+ \subset \mathbb{Z}^{\geq 0} \subseteq \mathbb{W} \subset \mathbb{Q} \subset \mathbb{R}$$

Example 1.4. Consider the simplified set notation with the context that the elements all belong to a universal set Ω, i.e., $\omega \in \Omega$, and X is real-valued function on Ω, i.e., $X : \Omega \to \mathbb{R}$. Then

$$\begin{aligned}
\{X \in (0, 1]\} &= \{0 < X \leq 1\} \\
\{X \leq r\} &\subset \{X \leq s\}, \text{ given } s > r
\end{aligned}$$

Definition 1.6. The set containing no elements is what call as the ***null*** or ***empty*** set and is denoted by $\emptyset$.

Definition 1.7. For a given $A \subseteq B$, the ***complement of*** A ***with respect to*** B is defined as the set of all elements in B that are not in A and we write either as

$$B \setminus A = \{x \in B : x \notin A\}$$

Remark. When B is already assumed in context, we simply say "*A complement*" and write A^c instead of $B \setminus A$. We also read $B \setminus A$ as "*the set difference of A from B (or from B to A)*" .

Example 1.5. Using the sets defined in Example 1.3:

$$\mathbb{Q}^c = \mathbb{R} \setminus \mathbb{Q}, \quad \{0\} = \mathbb{W} \setminus \mathbb{N}, \quad \mathbb{R} \setminus (5, +\infty) = (-\infty, 5]$$

Example 1.6. Suppose every set we are going to talk about are assumed to be subset of $\mathbb{R}$. Then $(-\infty, 5]^c = (5, +\infty)$.

Definition 1.8 (Union). $A \cup B = \{x : x \in A \text{ or } x \in B\}$.

Definition 1.9 (Intersection). $A \cap B = \{x : x \in A \text{ and } x \in B\}$.

Example 1.7.

$$\mathbb{R} = \mathbb{Q} \cup \mathbb{Q}^c, \quad \emptyset = \mathbb{Q} \cap \mathbb{Q}^c, \quad \mathbb{R} \setminus \{\tfrac{1}{2}\} = (-\infty, \tfrac{1}{2}) \cup (\tfrac{1}{2}, +\infty)$$

Example 1.8.

$$\mathbb{R} \setminus \{\tfrac{1}{2}, 0, 1\} = (-\infty, \tfrac{1}{2}) \cup (\tfrac{1}{2}, 0) \cup (0, 1) \cup (1, +\infty)$$

The following example is common in simplified notations of sets where $\omega \in \Omega$ is assumed.

Example 1.9. The following are sets defined under the context of Example 1.4.

$$\begin{aligned}\{X < 1, -1 \le X \le 4\} &= \{X < 1\} \cap \{-1 \le X \le 4\} \\ &= \{-1 \le X < 1\} \\ &= \{X \in [-1, 1)\}\end{aligned}$$

Remark. Notice that the "*comma*" is interpreted as an *intersection*. We will continue to do so for simplified notations in the succeeding chapters.

Proposition 1.1 (DeMorgan's Laws). *Let $A, B \subset S$.*

$$(A \cup B)^c = A^c \cap B^c, \quad (A \cap B)^c = A^c \cup B^c.$$

Definition 1.10 (Generalized Union and Intersection). Let $I \subseteq \mathbb{W}$. Then, the ***generalized union*** and ***generalized intersection*** of given sets A_i, $i \in I$ are respectively denoted by

$$\bigcup_{i \in I} A_i \qquad \text{and} \qquad \bigcap_{i \in I} A_i.$$

Remark. Generalized union and intersection can be interpreted as follows:

- $x \in \bigcup_{i \in I} A_i$ means "*x belongs to **at least one** of the A_i*".
- $x \in \bigcap_{i \in I} A_i$ means "*x **must** belong to **every** A_i*".

Example 1.10. Let $I = \{1, 2, 3, 4, 5\}$, and $A_n = \{n - 1\}$, for $n \in I$. Then

$$\bigcup_{i \in I} A_i = \{0\} \cup \{1\} \cup \{2\} \cup \{3\} \cup \{4\} = \{0, 1, 2, 3, 4\}.$$

What do you think is $\bigcup_{i \in I} A_i$ when $I = \mathbb{N}$?

Remark. We can simplify the notations of both generalized union and intersection under the following special cases:

- When $I = \{1, 2, 3, \dots, n\}$:

$$\bigcup_{i\in I} A_i = \bigcup_{i=1}^{n} A_i \quad \text{and} \quad \bigcap_{i\in I} A_i = \bigcap_{i=1}^{n} A_i$$

- When $I = \mathbb{N}$:

$$\bigcup_{i\in I} A_i = \bigcup_{i=1}^{\infty} A_i \quad \text{and} \quad \bigcap_{i\in I} A_i = \bigcap_{i=1}^{\infty} A_i$$

- When $I = \mathbb{Z}$:

$$\bigcup_{i\in I} A_i = \bigcup_{i=-\infty}^{\infty} A_i \quad \text{and} \quad \bigcap_{i\in I} A_i = \bigcap_{i=-\infty}^{\infty} A_i$$

Example 1.11. Define the sets A_i and B_i, for $i \in \mathbb{N}$ as $A_i = \left(-\frac{1}{i}, 1-\frac{1}{i}\right]$ and $B_i = \left\{0, \frac{1}{i}, \frac{2}{i}, \dots, \frac{i-1}{i}, 1\right\}$. Then

$$\bigcup_{i\in I} A_i = \bigcup_{i=1}^{\infty} \left(-\frac{1}{i}, 1-\frac{1}{i}\right] = (-1, 1)$$

$$\bigcap_{i\in I} A_i = \bigcap_{i=1}^{\infty} \left(-\frac{1}{i}, 1-\frac{1}{i}\right] = \{0\}$$

$$\bigcup_{i\in I} B_i = \bigcup_{i=1}^{\infty} \left\{0, \frac{1}{i}, \frac{2}{i}, \dots, \frac{i-1}{i}, 1\right\} = \mathbb{Q} \cap [0, 1]$$

$$\bigcap_{i\in I} B_i = \bigcap_{i=1}^{\infty} \left\{0, \frac{1}{i}, \frac{2}{i}, \dots, \frac{i-1}{i}, 1\right\} = \{0, 1\}$$

Definition 1.11 (Generalized DeMorgan Law's)**.** Let $I \subseteq \mathbb{W}$. Then, the ***generalized DeMorgan Law's*** of given sets A_i, $i \in I$ are as follows:

$$\left(\bigcup_{i\in I} A_i\right)^c = \bigcap_{i\in I} A_i^c \quad \text{and} \quad \left(\bigcap_{i\in I} A_i\right)^c = \bigcup_{i\in I} A_i^c$$

Definition 1.12. Two sets A and B are said to be ***disjoint*** (or ***mutually exclusive***) if $A \cap B = \emptyset$.

Definition 1.13. Let $\mathscr{A} = \{A_i : i \in I\}$ be collection of sets. If for all $i, j \in I$ with $i \neq j$, $A_i \cap A_j = \emptyset$ we say that $\mathscr{A}$ is a ***pairwise disjoint*** (or ***pairwise mutually exclusive***) collection of sets.

Definition 1.14. Let S be a non-empty set and $\mathscr{A} = \{A_i : i \in I\}$ be collection of sets. If the following two conditions are satisfied:

a. $\mathscr{A}$ is a pairwise disjoint collection and
b. $S = \bigcup_{i \in I} A_i$

we say that $\mathscr{A}$ forms a ***partition*** of the set S.

Proposition 1.2. *Let S be a non-empty set and $\mathscr{A} = \{A_i : i \in \mathbb{N}\}$ be a collection of sets such that $S = \bigcup_{i \in I} A_i$. Define the collection $\mathscr{B} = \{B_i : i \in \mathbb{N}\}$ such that:*

$$B_i = \begin{cases} A_1 & \textit{if } i = 1 \\ A_i \setminus \left(\bigcup_{k=1}^{i-1} A_k\right) & \textit{if } i \geq 2. \end{cases}$$

Then, $\mathscr{B}$ is a partition of S.

Proposition 1.3. *Let $A, B, C \subseteq S$ and $A_i \subseteq S$ for all $i \in I$. Moreover, denote $\sigma(I)$ as a rearrangment of the elements of I, i.e., $I = \sigma(I)$ but the elements are written in different order of enumeration. The following properties hold:*

1. *Distributive laws:*
   a. $C \cap (A \cup B) = (C \cap A) \cup (C \cap B)$
   b. $C \cup (A \cap B) = (C \cup A) \cap (C \cup B)$
   c. $C \cap \left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} (C \cap A_i)$
   d. $C \cup \left(\bigcap_{i \in I} A_i\right) = \bigcap_{i \in I} (C \cup A_i)$
2. *Commutative laws:*
   a. $A \cup B = B \cup A$
   b. $A \cap B = B \cap A$
   c. $\bigcup_{i \in I} A_i = \bigcup_{i \in \sigma(I)} A_i$
   d. $\bigcap_{i \in I} A_i = \bigcap_{i \in \sigma(I)} A_i$
3. *Set difference:*
   a. $A \setminus B = A \cap B^c$
   b. $A \setminus \left(\bigcup_{i \in I} A_i\right) = \bigcap_{i \in I} (A \setminus A_i)$
   c. $A \setminus \left(\bigcap_{i \in I} A_i\right) = \bigcup_{i \in I} (A \setminus A_i)$
4. *Miscellaneous properties: Let $A, B, C \subset S$ and $\mathscr{A} = \{A_i : i \in I\}$ be a collection of subsets of S.*
   a. $A \cup A^c = S$ *and* $A \cap A^c = \emptyset$

b. $A \cap B \subseteq A \subseteq A \cup B$
c. $A \cap B \subseteq B \subseteq A \cup B$
d. $\bigcap_{i \in I} A_i \subseteq A_k \subseteq \bigcup_{i \in I} A_i$ *for all* $k \in I$.
e. If $\mathscr{A}$ *is a partition of* S, *then*

$$B = \bigcup_{i \in I} (B \cap A_i).$$

1.2.2 Countable and Uncountable Sets

Definition 1.15. A set A is said to be *countable* if there exist a bijective function $f : K \to A$ for some $K \subseteq \mathbb{N}$. Otherwise, we say that A is *uncountable.*

Proposition 1.4. *Every finite set is a countable set.*

Proposition 1.5. *Let* A *and* B *be countable sets. Then,*

a. $A \cap B$ *and* $A \cup B$ *are countable sets;*
b. if C *is a set such that* $C \subset A$, *then* C *is countable.*

Corollary 1.1. *The empty set is a countable set.*

Definition 1.16. A set that is not finite but is countable is called a *countably infinite* set.

Example 1.12. The sets $\mathbb{N}$, $\mathbb{W}$, $\mathbb{Z}$, and $\mathbb{Q}$ are countably infinite sets.

Example 1.13. Intervals are typical examples of uncountable sets: (a, b), $(a, b]$, $[a, b)$, $[a, b]$, $[a, +\infty)$, $(-\infty, b]$, $(a, +\infty)$, $(-\infty, b)$ and $\mathbb{R} = (-\infty, +\infty)$.

1.3 Counting

When you study probability and statistics, there are problems that require ***enumerating the possible outcomes*** of a random phenomenon. Questions like the "*What is the probability that tossing a coin 15 times results to a total of 10 Heads?*" sometimes requires counting all the possible number of outcomes in 15 tosses. If we simplify this experiment into a total of 3, the possible outcomes are enumerated as follows:

$$\begin{array}{cc} HHH & TTT \\ HHT & TTH \\ HTH & THT \\ THH & HTT \end{array}$$

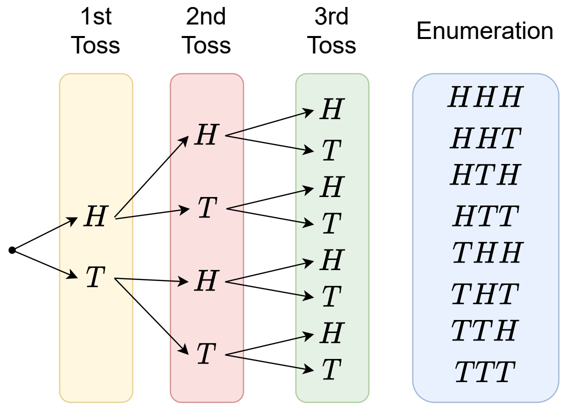

Figure 1.4: Tree diagram enumeration of outcomes when a coin is tossed three times.

Using tree diagram it would look something like the one shown in Figure 1.4.

Thus, there are exactly 8 possible outcomes when you toss a coin three times. But how about when you toss it 15 times. Would the enumeration or tree still be practical to use just to determine the total number of possible outcomes? That answer would surely be no. Therefore, we need some mathematical tools to solve this in a more elegant way. This leads us to the subfield of mathematics called ***Combinatorics*** which provides the tools dedicated for such problems. This tools are what we commonly call as ***counting techniques***.

In this section, we will discuss the following counting techniques and their variations:

1. The multiplication rule
2. Permutations[2]
3. Combinations

1.3.1 The Multiplication Rule

Theorem 1.1 (Multiplication Rule)**.** *Assume that there are r decisions to be made. If there are n_1 ways to make decision 1, n_2 ways to make decision 2, ..., n_r ways to make decision r, then the total number of ways to make all decisions is*

$$n_1 \times n_2 \times n_3 \times \ldots \times n_r.$$

[2]This is a special case of multiplication rule

Example 1.14. Let's go to our scenario of tossing a coin 15 times. Since for every toss (decision), we have two outcomes (ways) which is either Head or Tail, then by the multiplication rule, then the total number of outcomes (total ways to make all decisions) is

$$\underbrace{2 \times 2 \times \cdots \times 2}_{\text{15-times}} = 2^{15} = 32,768.$$

1.3.2 Permutations

Definition 1.17. A ***permutation*** is an ordered arrangment of r objectss selected from a set of n distinct objects without replacement.

Example 1.15. List the permutations of the set $\{a, b, c, d\}$ selected 2 at a time.

By Definition 1.17 with $n = 4$ and $r = 2$ there are 12 odered pairs (permutations):

$$\begin{array}{cc} (a,b) & (b,a) \\ (a,c) & (c,a) \\ (a,d) & (d,a) \\ (b,c) & (c,b) \\ (b,d) & (d,b) \\ (c,d) & (d,c) \end{array}$$

Notice that unlike sets, the order of elements in permutations matter.

Definition 1.18. ***Factorial*** of $n \in \mathbb{N}$ is denoted by $n!$ and defined as

$$n! = \begin{cases} n(n-1) \cdot \ldots \cdot 2 \cdots 1 & \text{if } n \geq 1 \\ 1 & \text{if } n = 0 \end{cases}$$

Example 1.16. $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120.$

Theorem 1.2. *The number of permutations of n distinct objects selected r at a time without replacement is*

$$n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-r+1) = \frac{n!}{(n-r)!}$$

Example 1.17. Applying Theorem 1.2 to the scenario of Example 1.15 with $n = 4$ and $r = 2$, the total number permutation of the letters is

$$\frac{4!}{(4-2)!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{2 \cdot 1} = 4 \cdot 3 = 12.$$

Theorem 1.3 (Circular Permutation). *The number of permutations of n distinct objects arranged in a circle is $(n-1)!$*

Theorem 1.4 (Nondistinct Permutation). *The number of nondistinct permutations of n objects of which n_1 are of the first type, n_2 are of the second type, ..., n_r are of the r-th type, is*

$$\frac{n!}{n_1!n_2!\dots n_r!}$$

where $n_1 + n_2 + \cdots + n_r = n$.

Example 1.18. How many ways are there to arrange the letters in the word "door"?

There are $n = 4$ objects (the letters) and $r = 3$ of them are distinct. Thus, there are

$$\frac{4!}{1!2!1!} = \frac{24}{2} = 12$$

different arrangements. The 2! in the denominator accounts for swapping the two indistinguishable "o" letters.

Example 1.19. If we toss a coin 3 times, we all know that the number of possible outcomes is $2^3 = 8$. However, there is another way of arriving to the same answer by partitioning the counts in terms of counts in the number of heads present as follows:

- There are no heads: $\{TTT\}$
- There is one head: $\{HTT, THT, TTH\}$
- There are two heads: $\{THH, HTH, HHT\}$
- There are three heads: $\{HHH\}$

Let x be the number of heads. Then each of the above partition can be thought of arranging x number indistinguishable letter H and $3-x$ indistinguishable letter T. Thus, using 1.4, a part of the partition with x heads contains exactly

$$\frac{3!}{x!(3-x)!}$$

number of elements. That is,

Description	No. of Elements in the Partition	Elements in the Partition
There are no heads	$\frac{3!}{0!(3-0)!} = 1$	$\{TTT\}$

Description	No. of Elements in the Partition	Elements in the Partition
There is one head	$\frac{3!}{1!(3-1)!} = 3$	$\{HTT, THT, TTH\}$
There are two heads	$\frac{3!}{2!(3-2)!} = 3$	$\{THH, HTH, HHT\}$
There are three heads	$\frac{3!}{3!(3-3)!} = 1$	$\{HHH\}$

Observe that,

$$2^3 = 8 = 1 + 3 + 3 + 1$$

In general, if we a toss coin n times and let x be the number heads, then the number possible outcomes is

$$2^n = \sum_{x=0}^{n} \frac{n!}{x!(n-x)!}$$

In the next section, the summand above is denoted by a special notation called the combination notation.

1.3.3 Combinations

Definition 1.19. A set of r objects taken from a set of n distinct objects without replacement is a ***combination***.

Example 1.20. List the combinations of 2 elements taken from $\{a, b, c, d\}$.

Since the order is not relevant, there will be fewer combinations than permutations (Examle 1.15 listed 12 permutations of 2 elements taken from the 4 elements):

$$\{a, b\},\ \{a, c\},\ \{a, d\},\ \{b, c\},\ \{b, d\},\ \{c, d\}$$

Theorem 1.5. *The number of combinations of r objects taken without replacement from n distinct objects is*

$$\binom{n}{r} = \frac{n!}{(n-r)!r!}$$

Example 1.21. How many ways are ther to pick a *committee* of three people from seven "volunteers"?

$$\binom{7}{3} = \frac{7!}{4!3!} = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1}$$

Example 1.22. Recall the last example from the previous section. The total number of possible outcomes when tossing a coin n times can be written as

$$2^n = \sum_{x=1}^{n} \binom{n}{x}$$

where x denotes the number of heads.

Theorem 1.6. *The number of ways of partitioning a set nondistinct permutations of n distinct objects into k subsets with n_1 in the first subset, n_2 into the second subset, ..., n_k are in the k-th subset, is*

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1! n_2! \ldots n_k!}$$

where $n_1 + n_2 + \cdots + n_= n$.

Example 1.23. The Glen family consists of 9 people. How many arrangements are there for them to watch the nightly news seated on four sofas: one that seats three and the others seat two?

$$\binom{9}{3, 2, 2, 2} = \frac{9!}{3!2!2!2!} = 7560.$$