UNIT 1

SCIENTIFIC METHODOLOGY

The scientific method

The practice of attempting to approach the objective truth as closely as possible is known as the scientific method. It is a set of procedures that individuals may use to learn more about the world they live in, advance their understanding of it, and make an effort to explain why and/or how things happen. With this approach, observations are made, questions are formulated, hypotheses are formed, an experiment is conducted, data is analyzed, and a conclusion is drawn. But part of the process is to keep searching for the universe's laws, keep refining your findings, and ask new questions.

A proposed explanation of the scientific process is the **hypothetico-deductive model** or technique. It states that the process of scientific investigation begins with the formulation of a hypothesis in a form that may be verified or refuted by an experiment on observable data with an unknown consequence.

- A test result that could have, or really does, defy the hypothesis's predictions is seen as proof that the hypothesis is false.
- A test outcome that could have, but does not run contrary to the hypothesis corroborates the theory.

The results are then compared with other competing hypotheses to determine (stringently) the validity of the proposed theory.

Hypothesis

A hypothesis is an assumption, an idea that is proposed for the sake of argument so that it can be tested to see if it might be true. For a hypothesis to be a **scientific hypothesis**, the scientific method requires that one can test it. Scientific hypotheses are often based on past findings that the current body of knowledge is unable to adequately explain.

Hypothesis vs Theory

A theory is a system of explanations that ties together a whole bunch of facts. It not only explains those facts, but predicts what you ought to find from other observations and experiments.

A theory, is a principle that has been formed as an attempt to explain things that have already been substantiated by data. Because of the rigors of experimentation and control, it is understood that theory to be more likely to be true than a hypothesis.

In non-scientific use, however, hypothesis and theory are often used interchangeably to mean simply an idea, speculation, or hunch, with theory being the more common choice.

Some Famous Theories

In no particular order, below are some of the well known theories that have stood against the face of time.

The Big Bang Theory

The Heliocentric Theory

The Theory of General Relativity

The Theory of Evolution by Natural Selection

Experiment

In science, an experiment is simply a test of a hypothesis in the scientific method. It is a controlled examination of cause and effect.

The two key parts of an experiment are the **independent** and **dependent** variables. The independent variable is the one factor that you control or change in an experiment. The dependent variable is the factor that you measure that responds to the independent variable. In a science experiment, a variable is any factor, attribute, or value that describes an object or situation and is subject to change.

NOTE: There is another type of variable called **confounding** variable. A confounding variable is a variable that has a hidden effect on the results. Sometimes, once you identify a confounding variable, you can turn it into a controlled variable in a later experiment.

Some Famous Experiments

Galileo Galilei and the Leaning Tower of Pisa Experiment

Aristotle had proposed that objects fell at different rates because gravity would act more strongly on heavier objects, but it turns out that the feather falls slower only because of air resistance. If you could perform the same experiment in a vacuum, the feather and ball will hit the ground at exactly the same time. It is difficult to separate fact from legend, but the story goes that Aristotle's theory of gravity went unchallenged until Italian polymath Galileo Galilei disproved it.

Mendel's peas

Augustinian friar Gregor Johann Mendel crossed-bred peas with varying traits to assess the inheritance patterns of various traits in their progeny. His research concentrated on pea plants and their seven distinguishable characteristics: plant height; flower location; seed, pod, and bloom color; and pod and seed morphologies.

He watched about $28,000$ pea plants throughout the course of the eight-year investigation. Mendel discovered that many plant generations displayed varying ratios of green to yellow peas, with yellow being the predominant color, while examining the color of the peas that were generated. He found that genes are paired, and that the dominant and recessive expression of those genes is determined by the mathematical pattern that is observed throughout generations.

Rutherford strikes gold

Ernest Rutherford — Hans Geiger and Ernest Marsden performed a series of experiments between 1908–1913 to prove Rutherford's theory of an atomic model, which resembled planets orbiting the Sun.

The physicists used a radioactive substance to bombard a thin piece of gold foil with positively charged alpha particles. The majority of particles passed through the foil without any deflection, suggesting that atoms had a great deal of open space. some were deflected from the gold foil at different angles, which meant that those particular particles had hit something with the same charge.

This meant that rather than a positive charge engulfing electrons, a smaller positive charge was held in the dense middle, thus heralding the discovery of the atomic nucleus.

Eddington and the eclipse

The year 1919 had a total solar eclipse, which gave Eddington a rare chance to see the night sky during the day. Eddington studied star positions at night and again during the false night of an eclipse after sailing to Príncipe Island to witness the finest possible solar eclipse and verify Einstein's hypothesis. This implied that he could watch to see if the Sun's gravity had changed the stars' apparent locations, which it had. This demonstrated that Einstein was right—light had been bent throughout its travel to Earth due to the force of the Sun.

Examples of what are NOT experiments

- Making observations does not constitute an experiment. Initial observations often lead to an experiment, but are not a substitute for one.
- Making a model or a poster is not an experiment.
- Just trying something to see what happens is not an experiment. You need a hypothesis or prediction about the outcome.
- Changing a lot of things at once isn't an experiment.

Error Analysis

Science is all about building knowledge based on reliable evidence. But no measurement is perfect; errors always creep in, no matter how carefully one performs the experiment. Error analysis is a crucial part of scientific inquiry that helps us understand how much trust we can place in our results.

Errors are deviations between the measured value and the actual value. **Uncertainty** on the other hand gives us the range of possible values within which the true value *likely* lies.

Error analysis in science is the process of evaluating the uncertainties associated with measurements and experimental results. It's not about achieving perfect measurements (which are impossible), but understanding how close your results are likely to be to the true value and how much confidence you can have in them. Error analysis involves two main types of errors:

NOTE: As an exercise try to find out all the possible sources of errors while performing an experiment.

Random Errors Fluctuations happening by chance, causing measurements to scatter. Think of throwing darts randomly around the bullseye. These can be minimized by taking multiple measurements and averaging the results.

Systematic Errors Consistent biases that push measurements in one direction (overestimating or underestimating). Imagine a tilted dartboard where darts consistently land off-center. These require careful analysis of the experiment's setup and instruments.

By analyzing errors, we can:

- Estimate the uncertainty in our measurements.
- Evaluate the reliability of our results.
- Draw valid conclusions from the experiment, considering the limitations of the measurements.

What Error Analysis Doesn't Do: It doesn't eliminate errors entirely. Errors are inevitable in any measurement or experiment. However, error analysis helps us account for these uncertainties and build trust in our findings.

Accuracy and Precision

Accuracy quantifies how closely a measured value aligns with the actual or true value of the quantity being measured. It reflects the "bullseye" of scientific measurement, striving to hit the mark. For example, a thermometer that consistently reads your body temperature at 98.6°F is considered accurate, assuming that's your true body temperature. A single measurement can, in theory, be accurate. If a surveyor's instrument yields a building height of exactly 100 meters, and that's the true height, then the measurement is accurate.

Precision describes the closeness of multiple measurements of the same quantity to each other. It reflects the reproducibility or repeatability of a measurement process, like clustering your darts tightly on the dartboard. Precision is assessed by analyzing the spread or variability across a series of measurements. A set of measurements can be precise, exhibiting minimal variation between them, yet inaccurate if they all deviate from the true value.

Feature	Accuracy	Precision
Definition	How close a measurement is to the true or actual value	How close repeated measurements are to each other
Analogy (Think dartboard)	How close a throw lands to the bullseye	How close multiple throws are grouped together
Example	A scale that measures your weight exactly at 150 lbs (assuming that's your true weight) is accurate.	Throwing darts that all land within a 1-inch radius of each other is precise (even if they're not on the bullseye).
Dependence	Depends on the true value	Independent of the true value
Multiple measurements needed	Not necessarily. A single measurement can be accurate by chance.	Yes. Precision refers to consistency across multiple measurements.
Impact of random errors	Random errors can cause inaccuracy by scattering measurements away from the true value.	Random errors can affect precision by causing throws to be spread out more.
Impact of systematic errors	Systematic errors can cause inaccuracy by consistently pushing measurements in one direction (over or underestimating).	Systematic errors won't affect precision (throws will still be grouped together) but will affect their location relative to the bullseye (accuracy).
Importance	Crucial for ensuring measurements reflect reality. Inaccurate data can lead to misleading conclusions.	Important for ensuring consistent and repeatable results.

Quantifying errors

Significant Figures and Round off

The precision of an experimental result is reflected not only in the specific value reported, but also in the way it is written. To convey this precision, scientists use the concept of significant figures. The number of significant figures in a result signifies the digits that are considered reliable and contribute to the measured value. Here are the rules of how to determine significant figures:

- Most Significant Digit: The leftmost non-zero digit is always significant.
- Least Significant Digit (No Decimal): If there's no decimal point, the rightmost non-zero digit is the least significant.
- Least Significant Digit (With Decimal): When a decimal is present, all digits to the right of the decimal are significant, even trailing zeros.

- Digits Between: All digits between the most and least significant digits are considered significant.

Following these rules, several examples illustrate numbers with four significant figures: $1,234$; $123,400$; 123.4; $1,001$; 10.10 and 0.0001010. However, ambiguity arises when dealing with trailing zeros in numbers without a decimal. For instance, the number $1,010$ might have a physically significant last digit, but by convention, it's interpreted as having only three significant figures. To avoid this ambiguity, it is better to supply decimal points or write such numbers in exponent form as an argument in decimal notation times the appropriate power of 10. Thus, our example of $1,010$ would be written as $1,010.$ or 1.010×10^3 if all four digits are significant.

Rules of error propagation

Error propagation refers to how uncertainties in individual measurements translate into the uncertainty of a final result obtained through calculations. Here's a concise overview with key equations:

- Addition/Subtraction: The uncertainty of the sum/difference equals the square root of the sum of the squared uncertainties of individual measurements.

$$\Delta z = \sqrt{(\Delta x^2 + \Delta y^2 + ...)} \tag{1.1}$$

(where Δz is the combined uncertainty, Δx, Δy, $\ldots$ are individual uncertainties)

- Multiplication/Division: The relative uncertainty of the product/quotient is the sum of the relative uncertainties of individual measurements.

$$(\Delta z/z) = (\Delta x/x) + (\Delta y/y) + ... \tag{1.2}$$

(where z is the final result, x and y are individual measurements)

Relation between Z and (A, B)	**Relation between errors Δz and (ΔA, ΔB)**
$Z = A + B$	$(\Delta Z)^2 = (\Delta A)^2 + (\Delta B)^2$
$Z = A - B$	$(\Delta Z)^2 = (\Delta A)^2 + (\Delta B)^2$
$Z = A \cdot B$	$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2$
$Z = A/B$	$\left(\frac{\Delta Z}{Z}\right)^2 = \left(\frac{\Delta A}{A}\right)^2 + \left(\frac{\Delta B}{B}\right)^2$
$Z = A^n$	$\frac{\Delta Z}{Z} = n\left(\frac{\Delta A}{A}\right)$
$Z = \ln A$	$\Delta Z = \frac{\Delta A}{A}$
$Z = e^A$	$\frac{\Delta Z}{Z} = \Delta A$

Mean, Variance and Standard deviation

Mean

The mean provides a measure of central tendency, summarizing the dataset with a single value can give an idea of what is to be expected when the experiment is performed again (sometimes can be referred to as the expected value of the experiment).

It allows researchers to compare different groups or conditions in an experiment by simplifying a large dataset to an easily represented and understandable value. This is because most of the results of the experiment follow the Gaussian curve, however one must note that the mean is extremely sensitive to the extreme values.

Mathematically, the mean (or average) is the sum of all data points in a dataset divided by the number of data points given by:

$$\text{Mean} = \frac{\sum X}{N}$$

Where:

$\sum X$ is the sum of all the observations (data points).

N is the total number of observations.

Variance

Variance tells us how far the data points are from the mean, it allows for comparison between different datasets in terms of their spread. A smaller variance indicates consistent results since the data points are closer to the mean, while a larger variance indicates greater variability since the data points are more spread out.

Mathematically, it is given by:

$$\text{Variance}(\sigma^2) = \frac{\sum(X-\mu)^2}{N}$$

Where:

X represents each data point.

μ is the mean of the dataset.

N is the total number of observations.

Standard Deviation

Standard deviation is the square root of the variance and is a measure of the amount of variation or dispersion in a dataset. Unlike variance, which is in squared units, standard deviation is in the same units as the data, making it easier to interpret the data. It is useful for identifying outliers, as data points more than two or three standard deviations away from the mean are often considered unusual.

In a normally distributed dataset, approximately 68% of data points fall within one standard deviation of the mean, and 95% fall within two standard deviations.

It is mathematically given by,

$$\text{Standard Deviation}(\sigma) = \sqrt{\frac{\sum(X-\mu)^2}{N}} \implies \sigma = \sqrt{\frac{\sum X^2}{N} - \mu^2}$$

Where:

σ is the standard deviation.

X represents each data point.

μ is the mean of the dataset.

N is the total number of observations.