1 The Scientific Method

1.1 What the Brain Does

From the moment a child is born, her brain is flooded with signals generated by her senses, including smells, sounds, images, touches, and tastes, as well as various other messages produced by her internal organs. Her brain readily engages in processes of pattern recognition and classification to "organize" all these data. This pattern recognition forms the basic elements of perception or views of reality the child will gather from this moment on. Slowly but surely, her brain will start to organize the data into mental models of the world that surrounds her: *If I cry, somebody comes, feeds me, holds me, and it feels good*.

At first, her perceptions are crude and simple, but with time, they progressively become more and more sophisticated. And soon, with the emergence of language abilities, her modeling of the world acquires tags and labels. She communicates the things she likes and dislikes, her wants, and what she doesn't want. And, as her senses and motor skills develop, she becomes an avid explorer of her immediate environment, discovering and adding to the categories, that is, the models already partially in place. The process is slow, sometimes laborious, but ineluctably, her brain engages in increasingly sophisticated pattern recognition and progressively forms complex and subtle representations of the world. This process eventually explodes in a rich and intense search for meaning, through her adolescence and early adulthood, as her brain nears full maturity. In time, her brain becomes a sophisticated pattern recognition engine that excels at detecting, sometimes inventing, patterns of all kinds, and making connections between the various entities that inhabit her world. Quite naturally, and with perhaps relatively limited awareness of her mental processes, she becomes an intuitive model builder, arguably some form of prescientific state of mind endowed with curiosity, a vast capacity for inquiry, and extensive intellectual resources that enable her not only to witness the world, but also become one of its actors, to experiment, and even to shape reality.

Though the storyline of the development of a human from infant to adulthood is fascinating, the reader might ask why it is relevant as an introduction to a book on data analysis techniques?

The answer is rather simple: much like a growing child, modeling the world and experimenting is in fact what scientists do. Indeed, the raison d'être of scientists is to discover, observe, and formulate models of the world and reality. Could it be then that we humans are scientists by our very nature, that is, by simple virtue of the inner workings of our minds? Surely, the evolution of a child's brain and mental processes just briefly depicted seems a rather universal process, and this does suggest that all humans are endowed with a natural and innate ability to become scientists. The fact that only relatively few among us end up making science their primary occupation and profession may not diminish this capacity in the least, but we need to acknowledge that as each of us emerges into adulthood, we develop varied interests and skill sets and engage in diverse activities, each accompanied by distinct and at times seemingly incompatible forms of discourse. Our capacity to detect or make up patterns does not vanish, however, and we continue through adulthood to seek understanding and meaning in the people and events that surround us. Evidently, given that our models of the world are based on our individual experiences and circumstances (including formal and informal education), we collectively end up having a plurality of views and interpretations of reality, some of which may clash drastically, or even violently.

Can all these views and forms of discourse be true simultaneously, or are there models that constitute a closer representation of reality? And if so, is there a privileged form of discourse and method that can enable us to reach, progressively perhaps, a more robust and truthful model of reality? Is there, in fact, such a thing as reality? These are both powerful and complex questions that many great thinkers have reflected upon through time. Indeed, ideas on such matters abound and many philosophers, through history, have claimed ownership of the truth. Amid all these ideas, one particular form of discourse and inquiry has risen and developed. It is both humble and powerful: it admits its innate incapacity to reach a perfect truth, but provides the means to progressively and systematically identify better views of reality. We call it the **scientific method**.

1.2 Critics of the Scientific Method

Science, as an empirical form of inquiry, finds its roots in the work of Copernicus, Galileo, and Newton. In Europe, before and around their times, the dominant philosophical view was that human perceptions and reasonings are intrinsically fallible and cannot be trusted, and that we should consequently rely on the word of God embodied in the Bible to guide our views and interpretation of reality. Copernicus, who had an interest in the motion of heavenly bodies, most particularly the planets, came to the realization that the Ptolemaic tables that had been used to describe the motion of the planets for several centuries were quite inaccurate and developed the notion that better observations of their positions through time would reveal the geocentric Ptolemaic model is wrong and that the planets, Earth included, all revolve around the Sun. Galileo and Newton, much like Copernicus, would champion the notion that the heavens are not perfect, and that careful observations can reveal much about the nature of things.

Arguably, it was Galileo who made the greatest breakthrough. Equipped with the telescopes he had built, he proceeded to discover mountains on the Moon, spots on the surface of the Sun, the Sun's rotation, phases of Venus, and satellites orbiting Jupiter. He would then conclude that the heavens are not immutable or perfect, and demonstrably show that empirical observations are not only possible but also powerful in their capacity to reveal new phenomena and new worlds.

Newton would later use precise observations of Mars, in particular, to demonstrate that the planets follow elliptical orbits with the Sun at one foci, and formulate a theory postulating the existence of a force acting at a distance between the Sun and the planets: gravity.

Together, the work of Galileo and of Newton would exemplify the notion that empirical knowledge is possible, particularly when the senses are enhanced (e.g., with a telescope) and reasonings framed into a powerful mathematical language (e.g., calculus). Together, these works would provide the impetus for a new view of the world and a new approach to scholarly works: empirical observations coupled with detailed mathematical representations of world entities and their relations can lead to great advances in our ability to understand and shape the world. Science, or the Scientific Method, as we now think of it, was born.

In spite of these great early successes, several philosophers would argue that while empirical methods have merits and do enable the formulation of models or theories of the world that work, such models could never be proven absolutely correct. David Hume (1711–1776), in particular, argued that it is factually impossible to deduce universal generalizations (i.e., models that always apply) from series of finitely many observations, and consequently, that inductive reasoning, and therefore causality, cannot, ultimately, be justified rationally. The notion has obvious merits. For instance, to use a rather trivial example, consider whether the fact that all zebras *you* might have so far seen in your life featured a black and white striped mane implies that all zebras are necessarily striped black and white? Obviously not: you have not seen all zebras in existence on our planet and thus cannot conclude all zebras feature black and white stripes. In fact, golden zebras, featuring a pigmentation abnormality characterized by the lack of melanin color pigments, do exist.

A less trivial example of Hume's point involves Newton's three laws of mechanics and his Law of Universal Gravitation. Following the initial successes of the theory in explaining the observations of Mars' orbit, Newton's laws of mechanics and gravity were tested repeatedly throughout the eighteenth and nineteenth centuries and found to be exquisitely accurate. Universal Gravitation also featured great predictive power best exemplified by the discovery of the planet Neptune based on calculations by the French astronomer Urbain Le Verrier, that accounted for observed anomalies in the orbit of Uranus. It seemed fitting, indeed, to qualify the law as universal. Yet, neither the three laws of mechanics nor the law of gravity would indefinitely survive the test of time. Indeed, with the publication of a paper, in 1907, by Albert Einstein, it emerged that the laws of Newton were in fact incorrect, or as many physicists prefer to say, incomplete. Evidently, demonstrating the inadequacy of Newton's laws (or more properly stated, the underlying principle of Galilean transformations) would require observations involving light and objects moving at large velocities. A few years later, in 1915, with the publication of a paper on general relativity, Einstein would also put into question Newton's amazingly successful theory of gravity. General relativity would eventually find confirmation in observations of the precession of the orbit of Mercury and the deflection of light by the Sun's gravity measured by Sir Arthur Eddington in 1919 [78]. No matter how many observations could be successfully explained by Newton's Law of Universal Gravitation, its failure to successfully explain the magnitude of Mercury's precession and the proper deflection of light by the Sun would provide tangible evidence of its inaccuracy as a model of reality. Hume was certainly right. Yet, he also obviously overstated the case. Though not perfect, Newton's laws proved to be immensely useful in explaining the world as well as building devices and artifacts that would ease humans' lives. Indeed, effective empirical knowledge is possible, even though it is bound to forever be tentative.

1.3 Falsifiability and Predictive Power

The notion of empirical knowledge as tentative was properly clarified by Sir Karl Popper, who argued that although a scientific theory cannot be proven correct by any finite number of experiments, it can be **falsified**; that is, it may be proven false (wrong) in the appropriate context. Indeed, it suffices to observe a single non-black and white striped zebra to conclude that the theory that all zebras are black and white striped is incorrect. Likewise, measurements of the precession of Mercury and the deflection of light would demonstrate the inadequacy of Newton's Law of Universal Gravitation. The theory could not be proven right, but it could be proven wrong. Limited empirical knowledge is thus possible, insofar as it can be falsified, that is, demonstrated false in an appropriate context. We will see in Chapter 6 that **falsifiability** has led to the notion of **null hypothesis** and (scientific) **hypothesis testing**. A null hypothesis typically champions the accepted theory (e.g., Law of Universal Gravitation). Experiments are then conducted to test the null hypothesis and the theory remains unchallenged as long as it is deemed acceptable based on measured data. An accumulation of observations that support the null hypothesis increases its plausibility (degree of belief) without ever proving it is universally correct. However, if (reliably) measured data are observed to deviate sizably from the null hypothesis (i.e., data values predicted based on the null hypothesis), the null hypothesis is considered rejected (false) and the theory challenged. And if the number of observations that challenge the null hypotheses becomes large, or the observed deviations can be considered irreconcilably large, the theory is eventually abandoned.

Alas, nothing is that simple. An experiment can go wrong and produce results that improperly reject the null hypothesis. The rejection is then considered an **error of the first kind (type I)**. Alternatively, experimental results can also falsely support the null hypothesis, and the unwarranted acceptance of the hypothesis is known as an **error of the second kind (type II)**. Regrettably, pushing this argument to its absurd limit, postmodern philosophers have formulated the notion that even falsification is impossible. Is the golden zebra actually a zebra or something else? Can one be sure of anything at all; can one reliably say anything at all? This absurd line of argument has led to the postmodern notion that all theories are equally valid; that all forms of discourse and inquiries are equally valid; and science, most particularly the scientific method, does not constitute a privileged vehicle to acquire and validate (or falsify) knowledge. This view, however, is considered far too extreme by most modern scientists given that it blindly neglects the tangible and significant advances made by science in the last three centuries. The post-postmodernistic view, as one might call it, accepted by most modern scientists, is that Popper was in fact essentially correct but the notion of **predictive power** must augment basic falsification. In other words, scientific models must be capable of making predictions that can be tested against careful and reproducible observations or measurements.

Indeed, much like a child's brain evolves to become a model builder and learn to learn, philosophers and scientists have, over time, examined and pondered how to learn from Nature and eventually formulated what is now known as the **scientific method**. This said, the scientific method is not fundamentally different today than it was, say, fifty or one hundred years ago. Although certainly more articulated, debated, and written about, it remains rather simple at its very core: observe a phenomenon of interest, formulate a mathematical model, verify how the model accounts for past observations, and use it to predict variants of the phenomenon and future observations. And, following Popper, models shall be readily abandoned if falsified by observed data. However, falsifiability is often largely insufficient: models based on distinct and perhaps incompatible assumptions may often be concomitantly supported by a specific dataset. It is thus necessary to identify extensions to existing measurements and examine where conflicting models may deviate appreciably from each other, and thereby provide grounds for additional testing and falsification.

1.4 A Flawed but Effective Process!

At this point, it is important to stress that perceptions and experimental measurements do not in fact need to be perfect for science to progress. As we will discuss at length in Chapter 12, no perception or measurement is in fact ever perfect: there are invariably resolution effects that smear the values of measured (physical) quantities; there are signal recognition or reconstruction issues that lead to signal losses and biases; and there may be background signals that may interfere with a measurement. But it is in the very nature of the scientific method to base measurements on techniques and processes that are well established and can be modeled with a high degree of reliability, and as such, can be corrected for experimental or instrumental effects. Indeed, a detailed understanding of the measurement process enables the formulation of (measurement) models that enable precise and reliable corrections of measured data. Although our sensory perceptions may not provide an objective view of reality, our technology-based measurements can. And if and when corrections applied to data are deemed questionable, the correction procedures can usually be studied and improved through better design of the measurement apparatus and protocol.

1.5 Science as an Empirical Study of Nature

The first model assumption to be made is perhaps that a particular phenomenon might be interesting or useful to study. Although such a statement may seem trivial, particularly in modern scientific cultures, it amounts to a relatively new idea, which dates back to the times of Copernicus and Galileo. Before the Copernican revolution,[1] the prevalent

---

[1] And sadly still today in some conservative religious cultures.

attitude, among learned individuals and scholars, was that wisdom cannot be gained by observations of the world but must rather be based on sacred texts, such as the Bible, and the writings of ancient Greek philosophers, most particularly those of Aristotle and Plato. But the tremendous discoveries Galileo achieved with his small telescope made it abundantly clear that new knowledge and wisdom can in fact be gained by observing Nature with our own senses, or with "machines" that enhance them.

Early scientific observations mostly involved direct sensory inputs. But in time it became clear that our senses suffer from several limitations. Indeed, there is great variance as to what might be considered a loud sound or a bright light intensity by different individuals. There is also great variability or lack of reproducibility in the observations of a single observer. Our eyes and ears, in particular, dynamically adapt to the environment and thus do not provide reliable measurements of luminosity and loudness, among several other observables. Our senses are also limited in their sensitivity. For instance, our eyes and ears cannot perceive very weak light and sound signals, and they may easily be damaged by excessively bright or loud sources. They also have rather crude abilities to detect the difference between two distinct sensory inputs (e.g., just noticeable loudness difference, pitch difference, etc.). In stark contrast, technologies based on previously acquired scientific knowledge alleviate most, if not all, of our senses' shortcomings.

In the course of time, scientists and engineers have learned to design and build devices that vastly surpass human sensory capabilities in sensitivity, precision, and accuracy, as well as dynamic range. Consider that Galileo's first telescope, with an aperture of 26 mm, could collect roughly four times the light his unaided eyes could and provided modest magnification (14×). So equipped, Galileo was able to "see" phenomena and features of Nature that were otherwise impossible to detect with human eyes alone, such as the phases of Venus, mountains on the Moon, and so forth. But modern telescopes have far surpassed the reach and prowess of Galileo's first telescope. The Hubble Space Telescope (HST) in Earth's orbit and the Keck telescopes atop Mauna Kea, Hawaii, have effective apertures 90,000 and 1,600,000 times larger than those of the human eye, respectively, thereby enabling astronomers to "see" objects at distances that Galileo himself could perhaps not even comprehend. Other orbital telescopes, such as the Wilkinson Microwave Anisotropy Probe, the Spitzer Space Telescope, the Chandra X-ray Observatory, and the Fermi Gamma-ray Space Telescope have extended the range of the human eye so it is now essentially possible to exploit the entire electromagnetic spectrum in our study of the Universe. Recent technological advances have also made it possible to detect and study gravitational waves produced by large objects in rapid motion, thereby also extending humans' very crude and primitive ability to sense gravity.

Technology not only amplifies, extends, or improves human perceptions; it also enables scientists to select, prepare, and **repeat** the conditions of particular observations. It is then possible to bring specific phenomena into focus while eliminating others, or at the very least suppress uninteresting and spurious effects. This is perhaps best epitomized by experiments at the Large Hadron Collider (LHC), such as ATLAS, CMS, and ALICE, that study collisions of specific beams of well-defined energy with vast arrays of high-granularity and high-sensitivity sensors that enable precise detection of particles produced by collisions. Much like the Keck and HST telescopes, the LHC detectors provide observational capabilities that far surpass any individual human ability, and thereby enable detailed exploration of the structure of elementary particles and the forces that govern them.

Obtaining such fantastic capabilities evidently involves many challenges. The cost of these facilities is extremely large, and their complexity is commensurate with their cost. The design, construction, and operation of these very complex machines require large international collaborations with scientists and engineers of varied and advanced skill sets. Complexity also brings challenges in the areas of student training, detector maintenance and operation, as well as data analysis, and thus often necessitates narrow training and specialization. It also brings about the need for elaborate detector calibration and data correction procedures. Fortunately, these large experimental facilities stand on the shoulders of prior facilities and experiments. They were indeed not designed totally from scratch and scientists involved in these organizations have inherited and perfected clear and precise protocols to handle all matters of data calibration and reconstruction, some of which are briefly discussed in Chapters 12 and 14.

1.6 Defining Scientific Models

While the notion that humans are natural-born modelers is enticing, it tells us very little about the requirements for scientific modeling, that is, the elaboration of scientific models or theories of the world that are falsifiable and endowed with predictive power. Surely, the plurality of religions and philosophies humans adopt or inherit from their parents should be a clear sign that reaching an objective view of reality is anything but a simple process. Yet, the fact that we all share a common predicament is a good indicator that we partake in the same reality and that although it may be difficult to reach an objective and comprehensive view of this reality, the task nonetheless remains feasible, if only by small increments. What then should be the defining elements of a scientific model of reality?

Broadly speaking, a model may consist of any constructs used to represent, describe, or predict observations. In this context, quantum mechanics, the theory of evolution, and creationism may then qualify as models of reality, at least to the extent that they provide a means to represent and interpret reality. These models, however, differ greatly in their capacity for falsification as well as in their predictive power. To be scientific, and thus used reliably or in a credible authoritative fashion, a model must satisfy a few minimal requirements that, alas, are not met by all models elaborated, shared, or inherited by humans in their daily lives. To be scientific, a model must reasonably circumscribe its range or scope of applicability; it must identify and clearly define observables or quantities of interest to the model; it must be internally consistent and logical; and it must be capable, following Popper's argument, of falsification as well as genuine and meaningful predictive power.

To be clear, Popper's notion of models as tentative implies models are not reality but at best fragmentary representations or images of reality. But three hundred years of modern science have taught us that although all models remain tentative, one can nonetheless have very high expectations from scientific models, and that in fact it is possible to reach increasingly sophisticated and powerful insights into the inner workings of our universe.

Parenthetically, it should be stressed that a person or organization presenting a model or set of ideas as the ultimate and final view of reality (i.e., reality itself) should most likely not be trusted...

Scientific models are built on the basis of well-defined entities and designed to describe these and the relations between them. Although a mathematical formulation of these relations is not absolutely required, it is typically useful because it enables, in most cases, a clear path to falsifiability and model predictability.

Roadmaps constitute an interesting basic example of a model of reality. Typically presented on a flat surface, they depict the position of entities of interest (e.g., roads, landmarks, and various artifacts) relative to a reference position, and are designed to provide guidance to travelers. Representation as flat surfaces is of course not meant to imply the world is flat but merely a necessity imposed by the media used to present maps. Roads, landmarks, and artifacts are labeled and coded to help users find their location and a path to their destination. Although somewhat simplistic as a model of reality, a map provides a good example of a basic scientific model. It is based on well-defined entities, has a clear purpose (representation of the human environment), provides predictive power (i.e., how to reach one's destination), and is falsifiable. The reliability of a map is particularly important. If roads are improperly represented or labeled, users of the map may experience delays and frustrations in reaching their destination. The map is thus easily falsified: if you reach an intersection and the road names posted at the section do not match those shown on the map, suggests that the map may not be accurate. There might be a typo on the map. It could be dated and not representative of recent changes in the road structure, and so forth. But it is also conceivable that postings are missing or improperly placed. The user could also be confused about his or her actual location. In essence, both the data (observed street names) and the model (road names shown on the map) can be wrong. The map remains nonetheless useful insofar as the number of typos or mistakes is relatively modest. A practical user would thus not dispose of the map simply because it features a few mistakes or inaccuracies. In fact, a savvy road traveler would figure out how to use her location to update and correct the map. The model could then be salvaged and thus reusable in the future without a repeat of the same frustrations. But then again, if whole regions of a city have been remodeled, an old map has lost its utility and should be disposed of, unless perhaps its uniqueness justifies keeping it as a museum piece.

While somewhat trivial, the roadmap example illustrates many of the facets and properties of a scientific model. It models well-defined entities of the real world, it indicates their relations (e.g., the Empire State building is south of the Chrysler building), it is falsifiable, and has some predictive power (i.e., how to get to one's destination). But as a model of reality, it is rather primitive and limited. It does not tell us why the streets were built the way they were, who named them, why the given names were chosen, and so forth. It is descriptive and useful but features no dynamics, evolution, or causal relationships.

Scientific models formulated in the physical and biological sciences typically seek to provide not only a descriptive account of reality but also the dynamics, that is, the causal and evolutionary interrelationships connecting the entities of a model. Classical mechanics (Newton's laws) is a prime example of a scientific model featuring descriptive components (kinematics: representation of motion) and causal components (dynamics: forces and causes of the motion). Likewise, the theory of evolution involves a descriptive component, the taxonomy of species, and a dynamic component that describes how species are connected through time, how they evolve. In both cases, it is the dynamical component of the theory that is of greatest interest because it tells us how systems change, and indeed how they evolve. Dynamical models feature great predictive power and capacity for falsifiability. For instance, not only does classical mechanics provide for a description of the motion of objects, called **kinematics**; it also features **dynamics**, which enables predictions of where objects will be in the future based on their current location and models of the forces through which they interact. If these force models are wrong, so will be the predictions. Likewise, the theory of evolution, first formulated by Darwin, empowers us to understand the relations and connections between species and how environmental constraints shaped them over long periods of time to become what they are today.

By stark contrast, creationism, as well as intelligent design, are models of reality that are totally devoid of content, falsifiability, and predictive power. To be sure, one is obviously at a liberty to posit that God created the Universe on a Sunday, exactly 6,000 years ago. That includes, of course, photons traveling through space as if they came from galaxies located hundreds of millions of light years away. And given there are billions of galaxies in the visible universe (e.g., a simple extrapolation from the Ultra Deep Field survey completed by the Hubble Space Telescope), that makes God an amazing being indeed. But why are some stars red and others blue? Why are some galaxies shaped like spirals and others like footballs? Why does Earth have an atmosphere rich in oxygen and capable of sustaining life, while the other planets do not? Because God made it that way? But why did God make these things that way? The faithful respond: Don't ask. It is the mystery of the creation. But how does this inform us about the world we live in? Are you sure the Universe was created on a Monday, not a Tuesday, or was it a Saturday? Did the wise men who wrote the Bible know about other galaxies, dinosaurs, and stellar nucleo-synthesis? Does the Bible provide a path to such discoveries and for a falsifiable representation of reality? Sadly, it does not. For sure, it tells an evocative story. But the story has no reliable markers, no real capacity for cross-checks and thus no falsifiability. And more importantly, it has no factual predictive power. It is thus of little use as a basis to model reality. Change the religion, change the book (e.g., the Koran, the Torah, the Bhagavad Gita, the Vedas, the Avestan, etc.), and the actors change names, the narrative changes, the commandments also change a little, but the conclusion remains the same: sadly, as models of reality these books have no trustworthy content, no falsifiability, no predictive power, and thus no real usefulness.

No reliable model of reality means no trustworthy path to knowledge, no tangible and reliable source of meaning and ethics. It means chaos. And chaos it is across our beautiful blue planet. Witness cultural and religious factions claiming they own their lands as well as the truth, and worst, readily conducting genocides, in the name of God, to eliminate whatever groups disagree with them and stand in their path. But it does not have to be that way. The scientific method does work. It is slow but robust and the models (knowledge) of reality it provides, while innately tentative and incomplete, are steadily bringing our scientific civilization to a greater and clearer vision of our Universe, our origins, and our nature. To quip, I would suggest that those who wish to get a true spiritual experience should pick up a physics or biology textbook, because there is sure no better way to embrace reality than science. But brace yourself, it takes work!

1.7 Challenges of the Scientific Method: Paradigm Shifts and Occam's Razor

Though science works, scientific modeling of reality is not without challenges of its own. One such challenge was clearly put to light by Thomas Kuhn in his book, *The Structure of Scientific Revolutions* [132]. To understand the issue, let us briefly consider the transition between the geocentric and heliocentric views of the world that occurred at the eve of the scientific revolution following the works and writings of Nicolaus Copernicus.

Born in 1473 in the town of Torun, Poland, Copernicus was orphaned at an early age and taken under the tutelage of his maternal uncle, Lucas Watzenrode the Younger (1447–1512), a very influential bishop of Poland, who provided for his education. Copernicus studied law and medicine, but his true passion was astronomy. Noticing that the positions of the planets were considerably off compared to predictions provided by the Alfonsine tables,[2] he became convinced that Ptolemy's geocentric view of the world was incorrect and he proposed a heliocentric model in which all the planets revolve around the Sun, except the Moon, which revolves around the Earth. Much time would pass before Copernicus's heliocentric model became widely accepted, but it eventually did, thanks in part to the work of Kepler, Galileo, and Newton. In time, astronomers would successively also dethrone the Sun and our galaxy as the center of the Universe.

The central point of this story is the geocentric model with its deferents and epicycles. The fact of the matter is that an arbitrary number of nested epicycles could be added to the geocentric model to fix it and provide a very accurate model of the apparent motion of the planets. With sufficiently many epicycles, the model could be made reliable for several decades, perhaps centuries. Kuhn's point is that based on observations of the apparent motion of the planets alone, it would not have been possible to readily falsify the geocentric model augmented with an arbitrary number of nested epicycles. Thankfully, several other observations, including the fact that Venus has phases incompatible with Ptolemy's geocentric model, the aberration of light, and Foucault's pendulum, would provide incontrovertible falsification power to reject the geocentric model in favor of the heliocentric model.

The capacity to mathematically represent the apparent motion of the planets with the wrong model, however, remains a serious issue. Thomas Kuhn realized the same type of issue could arise within models discussing various other aspects of reality. In essence, Kuhn understood that mathematical models describing a portion of reality (e.g., geocentric motion of the planets, classical description of particle motion, etc.) may be artificially

---

[2] Astronomical tables prepared on the request of thirteenth century King Alfonso X of Castile, based on Claudius Ptolemy's geocentric model of the motion of planets.

augmented with fixes and artifices until drastically different types of observations render a model untenable and requires what he called a **paradigm shift**.

Figuratively speaking, the need for the addition of epicycles may in part arise because of the limited quality of the data or poorly understood features of the measurement process. It could also amount to an attempt to temporarily fix the model while waiting for additional and better quality data, or a new model capable of providing a more encompassing view of the phenomenon or system of interest, and so on. One could obviously expound further on this topic, but the main point of interest, as far as this book is concerned, is that mathematical models used for the representation of natural phenomena can always be made more complicated to account for unusual features of the data. How then can scientists judge whether a scientific model and its mathematical realization provide a proper and sufficient representation of a phenomenon of interest (i.e., reality)? Why is Newton's model of orbital motion better than Ptolemy's? Why was it necessary to invent quantum mechanics?

An answer to such questions is often provided in the form of Occam's razor,[3] a principle stating that among competing hypotheses, the simplest, that is, the one with the fewest assumptions, should be selected. In mathematical terms, this translates into selecting the (fit) model with the least number of free parameters that is consistent with the data. Indeed, why use a cubic or quartic polynomial to fit a set of data if the precision of the data does not warrant it? A straight line or parabola might be sufficient, unless perhaps one has prior reasons to believe a higher degree polynomial must be used. An important aspect of data analysis then involves the evaluation of errors and of the techniques to assess or test whether a model or hypothesis constitutes an appropriate and sufficient representation of the data.

Understanding experimental errors and how they propagate to model properties derived from the data is thus a central aspect of the scientific method. We thus devote several sections, throughout the book, on this very important topic. An intuitive notion of error and techniques of error propagation are first discussed in Chapter 2 after the introduction of the concept of probability. A more precise definition of the concept of errors is introduced in Chapter 4 on the basis of estimators and statistics. A full characterization of errors, however, requires the notions of confidence level and confidence intervals discussed in Chapter 6. The notion of confidence interval is slightly modified in Bayesian inference and renamed credible interval in Chapter 7. Equipped with the notions of confidence intervals and data probability models, it then becomes possible, in Chapter 6, to fully address the notion of (scientific) hypothesis tests. A mathematical implementation of Occam's razor, based on the Bayesian interpretation is finally discussed in Chapter 7.

Occam's razor alone, however, is usually insufficient to abandon a particular model. A more convincing line of arguments is generally needed. Indeed, and to get back to our example, it is not Occam's razor that made scientists reject Ptolemy's model, but the observation of phenomena completely inconsistent with a geocentric universe, including the existence of Venus phases, the aberration of light, and Foucault's pendulum, and just as importantly the immense predictive power of Newton's mechanics and Law of Universal Gravitation. Newton's model is not better than Ptolemy's merely because it has fewer

---

[3] A problem-solving principle attributed to William of Ockham (c. 1287–1347).

parameters (in fact, it has quite a few as well, the mass of the planets, the size and eccentricity of their orbits, etc.) but because its higher level of abstraction provides a unifying principle (e.g., the force of gravity) that enables falsifiable predictions of the motion of objects on Earth as well as in the heavens. Likewise, the nonclassical concepts of wave function and quantization enabled accurate quantitative descriptions of large classes of phenomena that were otherwise intractable within the classical physics paradigm.

Kuhn argued that competing paradigms are frequently incommensurable, as competing and irreconcilable accounts of reality, and that scientists cannot rely on objectivity alone. Accordingly, the necessity for paradigm shifts may then involve a certain degree of subjectivity. It remains nonetheless that it would have been completely impossible to realize a flyby of Pluto (NASA Horizon spacecraft, summer 2015) based on Ptolemy's geocentric model but the exploit was readily achievable based on Newton mechanics. Likewise, designing microtransistors and computer chips would be inconceivable within the framework of classical mechanics but became a tremendously successful outcome of the development of quantum mechanics. Similar conclusions may also be stated for the tremendous progress achieved in biology, most particularly in genetics. It is rather clear that such advances could not have been possible without the guiding principles of Darwin's theory of evolution. There is little or no subjectivity associated with these facts. Falsifiability may be temporarily compromised by the addition of epicycles but the overarching predictive power and technical prowess of scientific theories developed after successive paradigm shifts are objective indications that although our models of reality remain tentative, science as a whole has made tremendous progress in formulating meaningful and powerful theories of reality. There is indeed little doubt that science is closer to the true nature of reality today than it was 50, 100, or 2,000 years ago.

1.8 To Err Is Human, to Control Errors, Science!

One of the very first things science students learn in the laboratory is that making mistakes in the execution of an experimental procedure and the acquisition of data is terribly easy. Committed students usually figure out how to improve their lab work and reduce or even eliminate the number of mistakes made in their execution of experimental protocols. But students must also learn that although the capacity to conduct reliable experiments is a skill one can hone through training and repetition, the risk of errors, and the need to understand them, never goes away. There are many types and sources of errors, of course, but at the end of the day, even if an experimental procedure is executed flawlessly, there remain irreducible sources of errors and uncertainties determined by the measurement process itself, known as **process noise**, and the instrument read-out, known as **measurement noise**. And although improvements in the design of an experimental apparatus or procedure may reduce these errors, they can never eliminate them completely.

But, as we have argued earlier, the true power of the scientific method resides in its capacity to falsify models and its ability to make accurate predictions. This implies it is absolutely necessary to understand the errors of a measurement, identify their sources and types, and make the best efforts to reduce their amplitude. Indeed, the capacity to reject a model (falsification of the null hypothesis), and adopt another one (adoption of an alternative hypothesis), based on a specific measurement is chiefly determined by the measurement's precision.

Evidently, having an estimate of the error involved in a measurement does not mean one knows the value of the error (unless perhaps, the value being measured is already known). If one did, it would suffice to subtract the error from the measured value and one would achieve an error-free measurement. Having an estimate of the error of measurement instead means that one can assess the likelihood of deviations of any given size from the true value of the observable. In this context, measured values are viewed as random variables, that is, observables that may deviate uncontrollably from their true value due either to their intrinsic nature (e.g., the exact decay time of a radioactive nuclei cannot be predicted, only the average lifetime is known) or variability associated with a macroscopic process involving a large number of elementary subprocesses (e.g., fluctuations of the number of collisions experienced by a high-energy particle traversing a large chunk of material). One is thus constrained to model measurements according to the language of probability and statistics. Assessing experimental errors thus requires a probabilistic model of the measurement process, and extraction of meaningful scientific results necessitates statistical analysis and inference. This is largely what this book is about.

1.9 Goals, Structure, and Layout of This Book

Traditionally, scientists, most particularly physicists, have made use of probability and statistical techniques rooted in the so-called frequentist interpretation of probabilities, which basically regards probabilities as relative frequencies observed in the limit of an infinite number of trials or measurements. However, a growing number of scientists, including physicists, now make use of techniques based on the Bayesian interpretation of probabilities, which assigns hypotheses a certain degree of belief or plausibility. This rapid increase in the adoption of Baysian probabilities and inference techniques stems in part from the relatively recent developments of the theory of probability as logic by Jaynes and others, the elegance and power of the interpretation, as well as the development and articulation of numerous tools to deal with practical scientific problems. The majority of works and scientific publications, however, are still based on the traditional methods of the frequentist interpretation. This book thus attempts to cover data analysis techniques commonly used in astro-, nuclear, and particle physics based on both approaches.

This book targets graduate students and young scientists. It is not intended as a text for mathematicians or statisticians. So, while we spend a fair amount of time introducing the foundations of the theory of probability, and discuss, among other topics, properties of estimators and statistical inference in great detail, we leave out detailed proofs of many of the theorems and results presented in the book, and rather concentrate on the interpretation and applications of the concepts. Note that at variance with more elementary books on probability and statistics, we assume the reader to be reasonably proficient in elementary mathematical and computational techniques, and thus also leave out many of the detailed numerical calculations and derivations found in such texts.

The book is divided into three parts. The first part provides a foundation in probability and statistics beginning with a formal definition of the concepts of probability and detailed discussions of notions of statistics, estimators, fitting methods, confidence intervals, and statistical tests. These concepts are first explored from a classical or frequentist perspective in Chapters 4–6, but also discussed within the Bayesian paradigm in Chapter 7.

The second part of the book introduces a variety of measurements techniques commonly utilized in nuclear and particle physics. It begins, in Chapter 8, with the introduction of basic observables of interest in particle/nuclear physics and exemplars of basic measurement methods. Selected advanced topics are presented in Chapters 9–11. Chapter 9 introduces the notion of event reconstruction, and presents examples of track reconstruction and fitting, as well as primary and secondary vertex reconstruction techniques, while Chapters 10 and 11 present extensive descriptions of correlation function measurement methods. These topics then form grounds for a detailed discussion of error assessment and data correction methods, including efficiency corrections, spectrum unfolding techniques, and various other techniques involved in the correction of correlation functions and flow measurements, presented in Chapter 12.

Part III introduces basic Monte Carlo techniques, in Chapter 13, and elementary examples of their use for simulations of real-world phenomena, as well as for detailed evaluations of the performance and response of complex detection apparatuses, in Chapter 14.

There is much to discuss. Let us begin!