Chapter 1

Introduction

My earlier book on philosophy of probability, [Burdzy (2009)], was focused on finding the best scientific formulation of the foundations of probability. This book presents much of the same program but it also tries to determine the main sources of the success of probability. This analysis will take me well beyond the narrow topic of philosophy of probability, to the frightening depths of epistemology.

1.1 Knowledge

"There is no truth" — this claim, in different forms, was made by a number of philosophers. What is surprising to me is that philosophers as different as Karl Popper and Thomas Kuhn, arguably the best known 20th century philosophers of science, shied away from the clear declaration that science constantly brings us closer to the truth about the objectively existing universe. I cannot prove that the objective universe exists or that we can find the truth about it. Nobody can. I am a 100% skeptic. But skepticism is a dead end in philosophy. The interesting direction in philosophy is to describe how we arrive at statements that we consider true. Once we understand the process, or rather many different processes, we can join other people in pursuing the truth in one of the established ways or we can seek our own alternative way. Scientists arrive at the truth in their own way. The prevalence of religion proves that there is not even a slightest chance, in the present society, for a consensus on how to find the truth. Despite the obvious lack of consensus on the truth and ways of attaining it, I believe that there is an important and universal element of knowledge acquisition (I will call it "resonance") that received too little attention from philosophers. I will argue that resonance is the missing link in the known philosophical theories of probability. If resonance proves to be a viable concept outside philosophy of probability, I will consider this a welcome bonus.

1.2 Probability

Two and two makes four. Imagine a mathematical theory which says that it makes no sense to talk about the result of addition of two and two. Imagine another mathematical theory which says that the result of addition of two and two is whatever you think it is. Would you consider any of these theories a reasonable foundation of science? Would you think that they are relevant to ordinary life?

If you toss a coin, the probability of heads is 1/2. According to the frequency philosophy of probability, it makes no sense to talk about the probability of heads on a single toss of a coin. According to the subjective philosophy of probability, the probability of heads is whatever you think it is. Would you consider any of these theories to be a reasonable foundation of science? Would you think that they are relevant to ordinary life?

The frequency philosophy of probability is usually considered to be the basis of the "frequency" statistics and the subjective philosophy of probability is often regarded as the basis of the "Bayesian" statistics (readers unfamiliar with these terms should consult Chapter 18). According to the frequency philosophy of probability, the concept of probability is limited to long runs of identical experiments or observations, and the probability of an event is the relative frequency of the event in the long sequence. The subjective philosophy claims that there is no objective probability and so probabilities are subjective views; they are rational and useful only if they are "consistent," that is, if they satisfy the usual mathematical probability formulas.

Von Mises, who created the frequency philosophy, claimed that ([von Mises (1957), p. 11]),

> We can say nothing about the probability of death of an individual [within a year] even if we know his condition of life and health in detail.

De Finetti, who proposed the subjective philosophy, asserted that ([de Finetti (1974), p. x]),

> Probability does not exist.

The standard education in probability and statistics is a process of indoctrination in which students are taught, explicitly or implicitly, that individual events have probabilities, and some methods of computing probabilities are scientific and rational. An alien visiting our planet from a different galaxy would have never guessed from our textbooks on probability and statistics that the two main branches of statistics are related to the philosophical claims cited above. I believe that the two cited philosophical claims are incomprehensible to all statisticians except for a handful of aficionados of philosophy. I will try to explain their meaning and context in this book. I will also argue that the quoted claims are not mere footnotes but they constitute the essence of the two failed philosophical theories.

1.3 Summary of the Main Claims

1.3.1 Resonance

The acquisition of information and creating knowledge (this includes both facts and theories) can be divided into two steps. I will call the first step "resonance" for reasons explained later in the book. This process is very fast in most cases, subconscious and very reliable in a great variety of situations. My guess is that resonance is not based on logic in any reasonable sense of the word "logic." Resonance is a crude but reasonably reliable filter of information arriving at our senses.

Resonance is fallible in many situations recognized as significant to individual lives, society and science. Logic, probability, induction and all other named and unnamed ingredients of science provide the second filter, much more refined and reliable than resonance. Traditionally, philosophy was focused on the second filter because we have almost no access to resonance via our consciousness. This situation created various misconceptions concerning the sources of reliable truth. One of these is a tendency to ignore resonance despite the fact that resonance is at least as important to science and general knowledge as logic. Another common misconception is that resonance (under the name of "subjectivity") is unreliable or not needed. Some other myths go in the opposite direction and invest intuition, subjective opinions and mystical experiences in powers that these sources of opinion and information do not have.

Resonance is a necessary first filter because it is impossible to process all information available to us in a logical way in a timely manner. This should not be interpreted as a claim that resonance was created deliberately by humans. Quite the opposite, resonance is the result of the blind evolution process selecting the fittest individuals.

I will illustrate the role of resonance in the knowledge acquisition process by analyzing four classical philosophical problems: induction, consciousness, intelligence and free will.

I will argue that the classical problem of induction is ill-posed. Our "knowledge" of facts is based on the same ontological and epistemological assumptions as our predictions of "unknown" events. The reliability of induction is a law of nature or rather the confluence of several laws of nature, including evolution. These laws of nature are specific to our universe so no general logical justification of induction can exist.

Consciousness is (among other things) an ability to observe, memorize and analyze one's own information processing. Resonance is inaccessible to consciousness so it may appear to be irrational. This is a misleading impression. A process that does not follow the classical logic and is inaccessible to conscious analysis does not have to be arbitrary, subjective or unreliable.

Intelligence is the highest form of resonance. It does not have roots in observations of repetitive phenomena. Its essence is the ability to select facts or highly probable theories that are relevant to the current interests of the individual or society from the practically infinite amount of information and potential explanations of observations.

The analysis of free will cannot profit from relating free will to the deterministic or stochastic nature of our universe. Free will is an impression of one individual about another one due to the inability of a highly complex mind to create a model of another equally complex mind that would generate reliable predictions.

1.3.2 Critique of frequency and subjective philosophies of probability

In a nutshell, each of the two most popular philosophies of probability, frequency and subjective, failed in two distinct ways. First, both theories are very weak. The frequency philosophy of von Mises, developed in the first half of the 20th century, provides analysis of long sequences of independent and identical events only. The subjective philosophy of de Finetti (developed in parallel to that of von Mises, more or less) offers an argument in support of the mathematical rules of probability, with no hint on how the rules can be matched with the real world. Second, each of the two philosophical theories failed in a "technical" sense. The frequency theory is based on "collectives," a notion that was completely abandoned by the scientific community long time ago. The subjective theory is based on an argument which fails to give any justification for the use of the Bayes theorem. Even one of the two types failures would be sufficient to disqualify these theories. The double failure makes each of the theories an embarrassment for the scientific community.

The philosophical contents of the theories of von Mises and de Finetti may be split into (i) positive philosophical ideas, (ii) negative philosophical ideas, and (iii) innovative technical ideas. There is nothing new about the positive philosophical ideas in either theory. The negative philosophical ideas are pure fantasy. The technical ideas proved to be completely useless. I will now discuss these elements of the two theories in more detail.

1.3.2.1 *Positive philosophical ideas*

The central idea in the frequentist view of the world is that probability and (relative) frequency can be identified, at least approximately, and at least in propitious circumstances. It is inevitable that, at least at the subconscious level, von Mises is credited with the discovery of the close relationship between probability and frequency. Nothing can be further from the truth. At the empirical level, one could claim that a relationship between probability and frequency is known even to animals, and was certainly known to ancient people. The mythical beginning of the modern probability theory was an exchange of ideas between Chevalier de Mere, a gambler, Pierre de Fermat and Blaise Pascal, two mathematicians, in 1654. It is clear from the context that Chevalier de Mere identified probabilities with frequencies and the two mathematicians developed algebraic formulas. On the theoretical side, the approximate equality of relative frequency and probability of an event is known as the Law of Large Numbers. An early version of this mathematical theorem was proved by Jacob Bernoulli in 1713.

The main philosophical and scientific ideas associated with subjectivism and Bayesian statistics are, obviously, the Bayes theorem and the claim that probability is a personal opinion. Once again, one can subconsciously give credit to de Finetti for discovering the Bayes theorem or for inventing the idea that probability is a subjective opinion. The Bayes theorem was proved by Thomas Bayes, of course, and published in 1763 (although it appears that the theorem was known before Bayes). De Finetti was not the first person to suggest that the Bayes theorem should be used in science and other avenues of life, such as the justice system. In fact, this approach was well known and quite popular in the 19th century.

Between Newton and Einstein, the unquestioned scientific view of the world was that of a clockwise mechanism. There was nothing random about the physical processes. Einstein himself was reluctant to accept the fact that quantum mechanics was inseparable from randomness. Hence, before the 20th century, probability was necessarily an expression of limited human knowledge of reality. Many details of de Finetti's theory of subjective probability were definitely new but the general idea that probability was a personal opinion was anything but new.

1.3.2.2 *Negative philosophical ideas*

Both von Mises and de Finetti took, as a starting point, a very reasonable observation that not all everyday uses of the concept of probability deserve to be elevated to the status of science. A good example to have in mind is the concept of "work" which is very useful in everyday life but had to be considerably modified to be equally useful in physics.

One of the greatest challenges for a philosopher of probability is the question of how to measure the probability of a given event. Common sense suggests observing the frequency of the event in a sequence of similar experiments, or under similar circumstances. It is disappointing that quite often there is no obvious choice of "similar" observations, for example, if we want to find the probability that a given presidential candidate will win the elections. Even when we can easily generate a sequence of identical experiments, all we can get is the relative frequency which characterizes the whole sequence, not any particular event. The observed frequency is not necessarily equal to the true probability (if there is such a thing), according to the mathematical theory of probability. The observed frequency is highly probable to be close to the true probability, but applying this argument seems to be circular — we are using the concept of probability ("highly probable") before we determined that the concept is meaningful.

Von Mises and de Finetti considered philosophical difficulties posed by the measurement of probability of an event and concluded that a single event does not have a probability. This intellectual decision was similar to that of a philosopher coming to the conclusion that God does not exist because the concept of God is mired in logical paradoxes. The atheist philosophical option has a number of intellectual advantages — one does not have to think about whether God can make a stone so heavy that He cannot lift it himself. More significantly, one does not have to resolve the apparent contradiction between God's omnipotence and omnibenevolence on one hand, and all the evil in the world on the other. Likewise, von Mises and de Finetti do not have to explain how one can measure the probability of a single event.

While the philosophical position of von Mises and de Finetti is very convenient, it also makes their philosophies totally alienated from science and other branches of life. In practical life, all people have to assign probabilities to single events and they have to follow rules worked out by probabilists, statisticians and other scientists. Declaring that a single event does not have a probability has as much practical significance as declaring that complex numbers do not exist.

The claim that "God does not exist" is a reasonable philosophical option. The claim that "religion does not exist" is nonsensical. The greatest philosophical challenge in the area of probability is a probabilistic counterpart of the question "What does a particular religion say?" This challenge is deceptively simple — philosophers found it very hard to pinpoint what the basic rules for assigning probabilities are. This is exemplified by some outright silly proposals by the "logical" school of probability. While other philosophers tried to extend the list of basic rules of probability, von Mises and de Finetti removed some items from the list, most notably symmetry.

The fundamental philosophical claim of von Mises and de Finetti, that events do not have probabilities, was like a straitjacket that tied their hands and forced them to develop very distinct but equally bizarre theories. Their fundamental claim cannot be softened or circumvented. For a philosopher, it is impossible to be an atheist and believe in God just a little bit. Creating a philosophical theory of God that exists just a little bit is not any easier than creating a theory of God that fully exists. Similarly, creating a philosophy of probability which includes some events with a somewhat objective probability is as hard as inventing a philosophy claiming that all events have fully objective probability.

The two philosophies can be considered normative. Then their failure manifests itself in the fact that they are totally ignored. If the two theories are regarded as descriptive then they are complete failures because the two philosophers proved unable to make simple observations.

1.3.2.3 *Innovative technical ideas*

Von Mises came to the conclusion that the only scientific application of probability was in the context of long sequences of identical experiments or observations. Nowadays, such sequences are modeled mathematically by “i.i.d.” random variables (i.i.d. is an acronym for “independent identically distributed”). Since individual events do not have probabilities in the von Mises’ view of the world, one cannot decide in any way whether two given elements of the sequence are independent, or have identical distribution. Hence, von Mises invented a notion of a “collective,” a mathematical formalization of the same class of real sequences. Collectives are sequences in which the same stable frequencies of an event hold for all subsequences chosen without prophetic powers. Collectives have been abandoned by scientists long time ago. One of the basic theorems about i.i.d. sequences that scientists like to use is the Central Limit Theorem. I do not know whether this theorem was proved for collectives and I do not think that there is a single scientist who would like to know whether it was. De Finetti proposed to consider probability as a purely mathematical technique that can be used to coordinate families of decisions, or to make them “consistent.” This idea may be interpreted in a more generous or less generous way. The more generous way is to say that de Finetti had nothing to say about the real practical choices between innumerable consistent decision strategies. The less generous way is to say that he claimed that all consistent probability assignments were equally good. In practice, taking the last claim seriously would lead to chaos. The second significant failure of de Finetti’s idea is that in a typical statistical situation, there are no multiple decisions to be coordinated. And finally and crucially, I will show that de Finetti’s theory cannot justify the Bayes theorem — the essence of Bayesian statistics. De Finetti’s theory applies only to a handful of artificial textbook examples, and only those where no data are collected.

1.3.3 Scientific laws of probability

I will argue that the following laws are the de facto standard of applications of probability in all sciences.

(L1) Probabilities are numbers between 0 and 1 (including 0 and 1), assigned to events whose outcome may be unknown.

(L2) If events A and B cannot happen at the same time then the probability that A or B will occur is the sum of the probabilities of the individual events, that is, $P(A \text{ or } B) = P(A) + P(B)$.

(L3) If events A and B are physically independent then they are independent in the mathematical sense, that is, $P(A \text{ and } B) = P(A)P(B)$.

(L4) If events A and B are symmetric then the two events have equal probabilities, that is, $P(A) = P(B)$.

(L5) When an event A is observed then the probability of B changes from $P(B)$ to $P(A \text{ and } B)/P(A)$.

(L6) An event has probability 0 if and only if it cannot occur. An event has probability 1 if and only if it must occur.

The shocking aspect of the above laws is the same as in "the Emperor has no clothes." There is nothing new about the laws — they are implicit in all textbooks. The laws (L1)–(L6) provide a codification of the science of probability at the same level as laws known in some fields of physics, such as thermodynamics or electromagnetism. People familiar with the probability theory at the college level will notice that (L1)–(L6) are a concise summary of the first few chapters of any standard undergraduate probability textbook. It is surprising that probabilists and statisticians, as a community, cling to odd philosophical theories incompatible with (L1)–(L6), and at the same time they teach (L1)–(L6) implicitly, using examples. I will argue that both frequency statistics and Bayesian statistics fit quite well within the framework of (L1)–(L6).

The laws (L1)–(L6) include ideas from the "classical" philosophy of probability and Popper's "falsifiability" approach to science in the probabilistic context. Hence, the laws can hardly be called new. However, I am not aware of any published system of probability laws that are equally simple and match the contents of current textbooks equally well.

1.3.4 Statistics and philosophy

I will argue that frequency statistics has nothing (essential) in common with the frequency philosophy of probability and Bayesian statistics has nothing (essential) in common with the subjective philosophy of probability. The two branches of statistics and the two corresponding philosophical theories have roots in the same intuitive ideas based on everyday observations. However, the intellectual goals of science and philosophy pulled the developing theories apart. The basic intuition behind the frequency statistics and the frequency philosophy of probability derives from the fact that frequencies of some events appear to be stable over long periods of time. For example, stable frequencies have been observed by gamblers playing with dice. Stable frequencies are commonly observed in biology, for example, the percentage of individuals with a particular trait is often stable within a population. The frequency philosophy of probability formalizes the notion of stable frequency but it does not stop here. It makes an extra claim that the concept of probability does not apply to individual events. This claim is hardly needed or noticed by frequency statisticians. They need the concept of frequency to justify their computations performed under the assumption of a "fixed but unknown" parameter (implicitly, a physical quantity). Hence, frequency statisticians turned von Mises' philosophy on its head. Von Mises' philosophy can be summarized by saying that "If you have an observable sequence, you can apply probability theory." Frequency statisticians transformed this claim into "If you have a probability statement, you can interpret it using long run frequency."

There are several intuitive sources of Bayesian statistics and the subjective philosophy of probability. People often feel that some events are likely and other events are not likely to occur. People have to make decisions in uncertain situations and they believe that despite the lack of deterministic predictions, some decision strategies are better than others. People "learn" when they make new observations, in the sense that they change their assessment of the likelihood of future events. The subjective philosophy of probability formalizes all these intuitive ideas and observable facts but it also makes an extra assertion that there is no objective probability. The last claim is clearly an embarrassment for Bayesian statisticians so they rarely mention it. Their scientific method is based on a mathematical result called the Bayes theorem. The Bayes theorem and Bayesian statistics are hardly related to the subjective philosophy. Just like frequency statisticians, Bayesian statisticians turned a philosophy on its head. A brief summary of de Finetti's philosophy is "No matter how much information you have, there is no scientific method to assign a probability to an event." Bayesian statisticians transformed this claim into "No matter how little information you have, you can assign a probability to an event in a scientifically acceptable way." Some Bayesian statisticians feel that they need the last claim to justify their use of prior distributions.

I do not see anything absurd in using the frequency and subjective interpretations of probability as mental devices that help people do abstract research and apply probability in real life. Frequency statisticians use probability outside the context of long runs of experiments or observations, but they may imagine long runs of experiments or observations, and doing this may help them conduct research. In this sense, the frequency theory is a purely philosophical theory — some people regard long run frequency as the true essence of probability and this conviction may help them apply probability even in those situations where no real long runs of experiments exist.

Some Bayesian statisticians consider probability to be a tool used for coordination of decisions in a rational way, in agreement with the philosophical theory of de Finetti. All Bayesian statisticians apply probability irrespective of whether there is a need to coordinate any decisions. Bayesian statisticians may believe that coordination of decisions is the essence of probability and this purely philosophical belief may help them conduct research.

1.4 Historical and Social Context

In order to avoid unnecessary controversy and misunderstanding, it is important for me to say what claims I do *not* make. The controversy surrounding probability has at least two axes, a scientific axis and a philosophical axis. The two controversies were often identified in the past, sometimes for good reasons. I will *not* discuss the scientific controversy, that is, I will not take any position in support of one of the branches of the science of statistics, frequency or Bayesian; this is a job for statisticians and other scientists using statistics. I will limit myself to the following remarks. Both frequency statistics and Bayesian statistics are excellent scientific theories. This is not a judgment of any particular method proposed by any of these sciences in a specific situation — all sciences are more successful in some circumstances than others, and the two branches of statistics are not necessarily equally successful in all cases. My judgment is based on the overall assessment of the role of statistics in our civilization, and the perception of its value among its users.

A reader not familiar with the history of statistics may be astounded by the audacity of my criticism of the frequency and subjective philosophical theories of probability. In fact, there is nothing new about it, except that some of my predecessors were not so bold in their choice of language. Countless arguments against the frequency and subjective philosophies were advanced in the past and much of the material in this book consists of a new presentation of known ideas.

I will be mostly concerned with the substance of philosophical claims and their relationship with statistics. One is tempted, though, to ask why it is that thousands of statisticians seem to be blind to apparently evident truth. Why did philosophical and scientific theories, rooted in the same elementary observations, develop in directions that are totally incompatible? Although these questions are only weakly related to the main philosophical arguments in this book, I will now attempt to provide a brief diagnosis.

Statisticians have been engaged for a long time in a healthy, legitimate and quite animated scientific dispute concerning the best methods to analyze data. Currently, the competition is viewed as a rivalry between "frequency" and "Bayesian" statistics but this scientific controversy precedes the crystallization of these two branches of statistics into well defined scientific theories in the second half of the 20th century. An excellent book [Howie (2002)] is devoted to the dispute between Fisher and Jeffreys, representing competing statistical views, at the beginning of the 20th century. The scientific dispute within statistics was always tainted by philosophical controversy. Some statisticians considered understanding philosophical aspects of probability to be vitally important to scientific success of the field. My impression, though, is that philosophy was and is treated in a purely instrumental way by many, perhaps most, statisticians. They are hardly interested in philosophical questions such as whether probability is an objective quantity. They treat ideology as a weapon in scientific discussions, just like many politicians treat religion as a weapon during a war. Most statisticians find little time to read and think about philosophy of probability and they find it convenient to maintain superficial loyalty to the same philosophy of probability that other statisticians in the same branch of statistics profess. Moreover, many statisticians feel that they have no real choice. They may feel that their own philosophy of probability might be imperfect but they do not find any alternative philosophy more enticing.

Philosophers and statisticians try to understand the same simple observations, such as more or less stable frequency of girls among babies, or people's beliefs about the stock market direction. Philosophy and science differ not only in that they use different methods but they also have their own intellectual goals. Statisticians are primarily interested in understanding complex situations involving data and uncertainty. Philosophers are trying to determine the nature of the phenomenon of probability and they are content with deep analysis of simple examples. It is a historical accident that frequency statistics and the frequency philosophy of probability developed at about the same time and they both involved some frequency ideas. These philosophical and scientific theories diverged because they had different goals and there was no sufficient interest in coordinating the two sides of the frequency analysis — it was much easier for statisticians to ignore the inconvenient claims of the frequency philosophy. The same can be said, more or less, about Bayesian statistics. The roots of Bayesian statistics go back to Thomas Bayes in the 18th century but its modern revival coincides, roughly, with the creation of the subjective philosophy of probability. The needs of philosophy and science pushed the two intellectual currents in incompatible directions but scientists preferred to keep their eyes shut rather than to admit that Bayesian statistics had nothing in common with the subjective philosophy.

One of my main theses is that the original theories of von Mises and de Finetti are completely unrelated to statistics and totally unrealistic. So, why bother to discuss them? It is because they are the only fully developed and mostly logically consistent intellectual structures, one based on the idea that probabilities are frequencies, and the other one based on the idea that probabilities are subjective opinions. Both assert that individual events do not have probabilities. Some later variants of these theories were less extreme in their assertions and hence more palatable. But none of these variants achieved the fame of the original theories, and for a good reason. The alternative versions of the original theories are often focused on arcane philosophical points and muddle the controversial but reasonably clear original ideas.

1.5 Disclaimers

I will cite many sources in this book but I am not able to trace every one of my philosophical claims to an earlier philosophy or scientific theory. I believe that the concept of "resonance" is my original contribution but my "resonance" theory is clearly an amalgam of various known philosophical and scientific ideas. So, the value of the resonance theory, if any, lies in bringing various elements together. I had doubts about the value of and need for my philosophical project. But I found so much nonsense and hypocrisy in various supposedly respectable philosophical writings that I came to the conclusion that a new dose of common sense in the philosophical literature is needed, even if it is partly repetitive, to buttress the camp of reason.

The philosophical material to which I refer is easily accessible and well organized in books and articles. Both de Finetti and von Mises wrote major books with detailed expositions of their theories. These were followed by many commentaries. I felt that these writings were often contradictory and confusing but I had enough material to form my own understanding of the frequency and subjective philosophical theories. Needless to say, this does not necessarily imply that my understanding is correct and my very low opinion about the two theories is justified. If any of my claims are factually incorrect, I have nobody but myself to blame.

When it comes to statistics, the situation is much different. On the purely mathematical side, both frequency and Bayesian statistics are very clear. However, the philosophical views of professional statisticians span a whole spectrum of opinions, from complete indifference to philosophical issues to fanatical support for the extreme interpretation of one of the two popular philosophies. For this reason, whenever I write about statisticians' views or practices, I necessarily have to choose positions that I consider typical. I regret any misrepresentation of statisticians' philosophical positions, overt or implied.

I feel that I have to make another explicit disclaimer, so that I am not considered ignorant and rude (at least not for the wrong reasons). Both von Mises and de Finetti were not only philosophers but also scientists. My claim that their ideas are complete intellectual failures refers only to their philosophical theories. Their scientific contributions are quite solid. For example, de Finetti's representation of exchangeable sequences as mixtures of i.i.d. sequences is one of the most beautiful and significant theorems in the mathematical theory of probability.

I end the introduction with an explanation of the usage of a few terms, because readers who are not familiar with probability and statistics might be confused when I refer to "philosophy of probability" as a foundation for statistics rather than probability. I am a "probabilist." Among my colleagues, this word refers to a mathematician whose focus is a field of mathematics called "probability." The probability theory is applied in all natural sciences, social sciences, business, politics, etc., but there is only one field of natural science (as opposed to the deductive science of mathematics) where probability is the central object of study and not just a tool — this field is called "statistics." For historical reasons, the phrase "philosophy of probability" often refers to the philosophical and scientific foundations of statistics.