Introduction

1. Intellectual Origins of the Essays Assembled Here

In this introduction we shall array a family of fundamental questions pertaining to probability, especially as it has been judged to bear upon the guidance of life. Applications and uses of probability theory need either to address some or all of these questions, or to tell us why they don't. The essays assembled in this volume bring integrative perspectives on this family of questions. We asked the authors to describe in their own voices the intellectual histories of their contributions, so as to shed further light upon the philosophical interest of their projects, and their particular integrative approaches, within the broader context we have sketched. The authors' comments precede the essays.

2. Fallibility Rediscovered

Centuries ago David Hume struck a deep chord that has been resonating ever since. While philosophical predecessors and contemporaries enjoyed confidence in the human capacity to know our environment, whether by means of the senses or by means of reason, Hume insisted instead upon a philosophical account of how we pull this off—how we succeed in getting beyond the most superficial of sensory information, by means of a combination of the evidence of the senses and the power of reason, and (though aware of his continental contemporaries' rationalist solutions to the question) he contended that no such account has ever been produced. "Nature," he wrote, "has kept us at a great distance from all her secrets, and has afforded us only the [direct] knowledge of a few superficial qualities of objects; while she conceals from us those powers and
principles on which the influence of those objects entirely depends" (Enquiry Concerning Human Understanding).\(^1\)

Hume's contemporaries were very much in agreement with one another on the proposition that nature's secrets can be unlocked, although they might have disagreed on the key to that achievement. Thinkers on the continent (like Descartes and Leibniz) held that the key—the foundation of knowledge—lies fundamentally in the power of reason, while many of Hume's own compatriots held that the key lies instead in the evidence of the senses. Their confidence in human access to nature's secrets served as foundation for unexamined faith in the principle that the future shall resemble the past—the principle on which Hume's arguments cast a long shadow of doubt. Against his contemporaries and the preponderance of his predecessors, Hume resolutely maintained that confidence in our opinions vis-à-vis nature's secrets is thoroughly unfounded—while nonetheless affirming that confounding human nature foists this confidence upon us, and indeed that this (baseless) confidence serves our mundane interests surprisingly well. It is not without a strong measure of irony then that he wrote: "only a fool or madman will ever pretend to dispute the authority of experience, or to reject that great guide of human life."\(^2\) And indeed he seemed to favor the proposal that experience makes many of those opinions to which we cling with undeserved confidence, "merely probable."

Hume's legacy is in part the lesson that the proposition to the effect that the future shall resemble the past is probable at best, never certain. It is a legacy that faces us to the fallibility of human understanding, under the very best of circumstances. To be sure, Aristotle before him would doubtless have affirmed this thesis of fallibility. But Hume had to rediscover it. For intervening between Aristotle and Hume was a Roman empire, which was followed by an epoch in which the Christian Church controlled dissemination of intellectual ideas, and then by an age of Reason Alone ushered in by dissolution of Church rule over Reason. Hume's legacy to us is, in part, a rediscovered fallibility, a rediscovered empiricism, which is by its nature more vigorous than the original. It is more vigorous in that it urgently presses the question of the sources of fallibility. Indeed it amounts to discovery of the qualifier "probable" as an entire field for philosophical—as against purely mathematical—inquiry. Before Hume it was a term for which there existed only the most negligible philosophical market.

One thing is unshakably certain: we cannot have unshakable certainty. So perhaps belief in some proposition pertaining to the future is a form of wagering, on something that is not certain but probable. This leads to substantive philosophical questions about the meaning of probability statements, and their applications to actual day-to-day instances. And it leads too to questions as to the relationship between knowledge and wagering.

### 3. Possibility and Probability

The theory of probability grew up in gaming rooms. One of the first serious studies of probability was performed in the mid-1500s by an Italian physicist and mathematician named Girolamo Cardano, who was perhaps the first to think about chance events in a systematic way. He discovered that the likelihood of getting a particular sum when rolling two dice (assumed fair and distinct from one another) exactly equals the number of ways of obtaining that sum, divided by the total number of possible ways that the pair can fall. Thus he could say that the likelihood of securing a 7, on any given roll of fair dice, is exactly 6/36, that it is the largest likelihood for any sum on a roll of dice, and that this likelihood is exactly 1.5 times as large as the likelihood of turning up a 9 or a 5. These facts had been known, no doubt, for thousands of years; but only as lore, never as founded upon principles.

It would be another century before the next conceptual landmark in the field of probability would emerge: the notion of expected outcome. In 1657 the Dutch physicist Christiaan Huygens proposed a precise formula for computing the expected outcome of a game. It would be the sum of all possible outcomes, weighted by the relative chance of each occurring. So for example, in a lottery with two equally likely monetary payoffs, a and b, the expected outcome (in monetary units) of a ticket is \(1/2 (a) + 1/2 (b)\). In a lottery that is equally likely to pay one of either 0 or $1000, the expected outcome is $500. And the expected outcome, said Huygens, is a "fair" asking price: a ticket in such a lottery is worth $500, on the open market. Should one always be prepared to pay the fair asking price? What kind of sense does it make, to ask what price one should pay to undertake a risky venture? Is there a price on risk itself? To answer these and related philosophical questions, we require a theory of decision under risk. Such a theory would take nearly three centuries to become fully articulated, as it did with publication in the 1940s of John von Neumann and Oskar Morgenstern's Theory of Games and Economic Behavior.\(^3\)

With the rise of mercantilism, a special form of gambling became important: insurance. Insurance rates were not based on the counting of possible cases, à la Cardano, but on the relative frequencies known through experience. Uncertainty became a practical matter, not just for gamers, whose probabilities could be computed on the basis of numbers of possible cases, but also in business. And through this route it made its way into science itself. The mathematics of probability came to be developed for use in a number of scientific areas, from the treatment of errors of measurement, to biology and astronomy. The great mathematician and astronomer Laplace recognized the essential circularity of defining probability as a ratio of possible cases—it only works if the cases are equally probable. He therefore
proposed a principle of indifference, according to which two symmetrical alternatives would be assigned the same probability.

Although there was some discussion of the meaning of the term "probability" before the twentieth century, the three standard views were articulated mainly after 1928. Richard von Mises in 1928 identified probability with limiting frequency: the probability of the kind of event $K$, within the collective $C$, is $p$, just in case the limiting relative frequency of events of kind $K$ among the individuals of an initial segment of $C$, approaches $p$ as longer and longer initial segments of $C$ are considered. This interpretation of probability reflects the concern with relative frequency. Probability statements, on this view, are empirical statements, concerning frequencies or limits of frequencies in the real world. They are true or false, absolutely and without qualification, and independent of every corpus of propositions, believed or entertained.

At the opposite end of the spectrum of definitions we find those of Carnap and his followers. According to this approach, probability represents a kind of partial entailment. This is a view not unlike Laplace's, except that here inconsistency is avoided through requiring that the cases denote the most detailed possible stipulations of ways the world might be. These stipulated ways the world might be were called state descriptions. So a state description specifies, for every individual and every property term of the language in use, whether or not the property term applies to that individual; and for every pair of individuals, and every two-place relation, whether or not those individuals stand in that relation; and so on.

On the Carnapian approach each state description is assigned a measure—for example, each state description in a finite language could be assigned the same measure. Every statement of the language is equivalent to a disjunction of state descriptions. Thus every statement has a measure determined by the sum of the measures of the state descriptions whose disjunction it is equivalent to. Then, for an agent with background knowledge $K$, the probability of a statement $S$, $\text{prob}(S|K)$ is the ratio of the measure of $S$ and $K$ to the measure of $K$. This is the degree of rational belief the agent with background $K$ ought to have in $S$.

The third main interpretation of probability is the subjective or Bayesian interpretation, developed by the philosopher Frank Ramsey and independently also by the probabilist, Bruno de Finetti. This view is similar to Carnap's logical view, but without the Carnapian proviso to the effect that state descriptions are assigned measures on logical grounds. Probabilities here are tied directly to individual choices or predispositions to choose: the probability of the statement $S$, for John, is the amount he would pay for a ticket that would return $1.00$ in case $S$ turned out to be true. Thus the probability of securing a one on the next roll of a die, for John, is $1/3$ just in case he would pay $0.33$ for a ticket that would return him $1$ if that roll did in fact result in a one. Bayesians insist that the probabilities of a rational individual, over any finite algebra of statements, should satisfy the axioms of the probability calculus, either as a matter of consistency or as a matter simply of immunity to manipulation by others.

As a guide to life, each of these three interpretations of probability suffers shortcomings. The probability of a certain kind of event, on the frequency interpretation, depends on the particular collectivity, class or sequence in which it is regarded as embedded. The probability that the insurance applicant Fred survives for ten years depends on whether he is considered a forty-year-old smoker, a forty-year-old College Professor, a skier, . . . There are many collectivities and classes to which he belongs, each with a different frequency of ten-year survival. By contrast, subjective probability is a direct reflection on the subject whose probability it is. According to the subjective-Bayesian account of probability, any distribution of probabilities will do, provided only that it satisfy the axioms of the calculus and refrains from assigning the extreme values of 0 and 1 to uncertain propositions. There is no requirement here that a probability assignment be constrained by any facts of the world. And the logical interpretation of probability—the third proposal canvassed above and developed by Carnap and his followers—never became settled. Optimism amongst logical theorists gradually faded: they never identified grounds ex ante on which to found probability assignments to state descriptions in a useful language.

Thus despite the optimism mid twentieth century, there was found no simple solution to the problem of applying probability as an integral part of a recipe for guiding life. The frequency interpretation countenanced multiple values of probability, pending identification of the relevant reference class; the subjective interpretation could accommodate any consistent values whatever, but did not dispel discomforts expressed by detractors vis-à-vis its deference to whim and vagary; and the question amongst logical theorists went in the end unanswered. Where there is advisement at all (amongst the frequency theorists), there is altogether too much—so much as to undo consistency. And where consistency is possible (among the subjectivists), there is too much freedom, and the choice amongst possibilities left up to groundless, arbitrary whim.

4. Probability and Belief-as-Decision

At about the time of Huygens, the French mathematician Blaise Pascal appealed to the notion of expected value to commend belief in God, as a form of decision-making. He therefore pioneered the idea that believing is a form of wagering. Assuming that the value of a wager is equal to the sum
of the stakes involved in each of winning and losing (normally with opposite affective signs), weighted by the relative likelihoods of each, Pascal argued that since the stakes if God exist are infinite (on the one side there is eternal blessing minus small inconveniences of attending church and the small cost of candles, on the other side there is eternal damnation outweighing the pleasures of a finite earthly existence passed one's own way), then it is definitely worth believing even if the likelihood of God existing is itself very small. Besides taking for granted that one can think of believing as a form of choosing, this early argument assumes that the value to someone, of a wager, is the same as the fair asking price, itself equated (just as Huygens equated it) with the expected outcome. But there is cause here to tread with some caution.

Daniel Bernoulli, writing in 1738, argued that it is wrong to equate expected outcome with the value of a wager to the risk-taker. The example he used to drive the message home was introduced twenty-five years earlier by his brother Nicolas, and is now known as the St. Petersburg Paradox (with capital letters serving to caution that one must exercise one's very best judgment when accepting any particular solution to it). Peter promises Paul that he will toss a fair coin as many times as it takes for it to land heads, and will pay Paul one dollar if it lands heads on the first trial, two if it lands heads on the second, four if it lands heads on the third, 8 if it lands heads on the fourth, and so on (that is, $2^n$ dollars, where $n$ is the trial number on which it finally turns up heads). How much should Paul be willing to pay for an opportunity to play this game? Bernoulli, for his part, asserted that no reasonable person would pay more than $20 to play the game, although its expected outcome is infinite ($\frac{1}{2} \times 1 + \frac{1}{4} \times 2 + \frac{1}{8} \times 4 + \frac{1}{16} \times 8 + \ldots = \infty$). And he explained this empirical fact by arguing—and defending it as rational—that people do not determine how much to pay for an opportunity on the basis of expected outcome, but rather on what he referred to as expected utility.

Aiming to explain also why people of good sense refuse fair bets (such as, for example, an equal chance of winning or losing some large sum of money $S$), instead of being simply indifferent between accepting and refusing them, Bernoulli originated the doctrine we now refer to familiarly as the diminishing marginal utility of money. This doctrine states that the value, to someone, of one more monetary unit will be strictly smaller than the value, to that same person, of one he currently holds. The effect is that my millionth dollar is worth much less to me than my tenth, or my hundredth, or even my thousandth (possibly because when one has only ten, or one hundred, dollars, one has much less in the way of a discretionary budget than if one had one million dollars).

The idea that it is rational to pay more or less than the “fair” or “market” price (understood as the expected outcome), establishes that value and good are distinct things. Value is set by us, under the prevailing conditions, upon those goods, commodities, and services that are on offer. Goods (and services too) are created and destroyed, where by contrast value is set and withdrawn. This is a metaphysical distinction. It implies that goods and values belong to different metaphysical realms or categories. This separation of good and value does not presuppose monism (the doctrine that there is one species of intrinsic goods—goods sought or worth seeking for their own sake—of which one can have more or less). It presupposes, instead, that value is conferred and not found in the objects valued. And it leaves open entirely the question of what makes a good what it is—what makes a good a good. The doctrine that value is conferred rather than discovered also descends from Hume, and is taken on by the (utilitarian) architects of decision theory who come after him. The cornerstone of this edifice, whose foundation was laid by the British utilitarians (for example, Jeremy Bentham and John Stuart Mill) and which was brought to completion by von Neumann and Morgenstern, is the doctrine of expected utility—to the effect that the supremely rational decision maker maximizes expected utility, not expected value.

Now if believing can be a form of wagering, as Pascal suggested, then it is a form of decision making, and hence subject to the doctrine of expected utility. Belief formation becomes a specimen of practical reasoning, and the distinction between theoretical reasoning and practical reasoning as nonoverlapping forms of reason—a distinction that was insisted upon with nary a second thought before the nineteenth century—disappears. This is the Bayesian turn, which was in many ways anticipated by Karl Popper.

Bayesianism is an uneasy contemporary marriage between eighteenth-century empiricism, on the one side, and modern decision theory, on the other side. Bayesianism, like many of its rivals, divides reasoning ab initio into two kinds: theoretical (or disciplinary) reasoning and practical reasoning. Practical reasoning has the function of controlling action or decision: its point is to figure out what to do. This is contrasted with disciplinary reasoning, whose aim is to figure out how things stand in the world, rather than what to do about them. Bayesians moreover maintain that disciplinary and practical reasoning operate independently—each without consultation with the other. This (if true) preserves the impartiality of scientific disciplines, by way of ensuring that the opinions we (as scientists) hold as to how things stand in the world, are not influenced by how we might wish things stood (our utilities). And this, of course, guarantees that science and other theoretical achievements committed within a disciplinary setting are never tainted by wishful thinking.

Now Bayesians also believe that all reasoning, even disciplinary reasoning, is at bottom a form of decision making, to which a cost-benefit calculus applies at the ground floor. Thus, while the Bayesian marriage
between theoretical reason and practical reason is supposed to be a marriage of equals, it really is not. And so the contemporary Bayesian, against the likes of Hume, Descartes, Carnap, and Kant, maintains the idea that the fundamental imperatives for all of human life are practical, and that disciplinary reasoning ultimately resolves into practical reasoning. Everything in the cognitive cornucopia, according to the Bayesian, is poured out, in the first instance, as a guide to life.

5. Probability and Causality

Hume defined the notion of cause twice over, within the space of paragraphs: once in terms of necessary connections or secret powers connecting events together in time, and a second time in terms of how human impressions of objects in the world are related. In the first instance we have a cause as “an object, followed by another, and where all objects similar to the first are followed by objects similar to the second” (Enquiry Concerning Human Understanding). In the second, a cause is “an object followed by another, and whose appearance always conveys the thought to that other.” This definition twice over could be interpreted (as many have suggested) as a reductionist strategy, in which the notion of cause is reduced to that of a chance- or probability-raiser. For we can harmonize Hume’s two definitions through the proposal that a cause (by definition) functions to render its effect more likely or more probable. So that (once again) probability should serve as a guide to life, by serving as a guide to making the shortest, most cost-effective, or in some other way best, beeline to our (human) objectives.

Hume’s remarks about causation, besides creating a crisis in the theory of knowledge, created a stir also in the discipline of metaphysics, by drawing attention to the fact that too little attention had been given to discussion of the metaphysical referent of the term “cause,” understood in Aristotle’s sense of “efficient cause.” To be sure, Malebranche and Leibniz treated the question of how it is possible to have world of causes alongside an omnipotent and omniscient God, but they did not elaborate on what the term “cause” refers to. The debate Hume sparked on this point rages today, if anything more than ever before, with the Humean occupying a distinctive position in the spectrum.

The Humean reduction falls nowadays under a larger class of proposals that can be termed reductionistic. The reductionist seeks to identify the core of the concept of cause, with a central function of the causal concept in the economy of human life. The most resilient reductionist proposal extant today is traceable to Hume himself. It claims for its foundation the idea that causal facts are nothing more than configurations of probability relations—specifically, facts about certain factors improving the chances of others. This is the proposal that lies behind all probabilistic accounts of causality. Probabilistic accounts draw heavily upon highly sophisticated statistical analyses. They confer upon the notion of chance—or at any rate that of probability—a fundamental role to play in the universe with which humans must cope, both practically and scientifically. And so they both service and cultivate a class of coping sciences—the engineering sciences, under which fall all the medical and clinical disciplines.

But inference from statistical data to causal hypothesis is treacherous, fraught with danger and drama. One of the most well known problems is Simpson’s Paradox, and is famously illustrated by a certain constellation of incidents at Berkeley. The graduate school at Berkeley was once accused of discriminating against women in admissions, so that the question was raised, “Does being a woman cause one to be rejected at Berkeley?” The accusation rested on the numbers: women seeking graduate admission at Berkeley were rejected at a much higher rate than men. But when the data were more closely scrutinized, it was found that in a majority of the 85 graduate departments, the probability of admission for women was just about the same as for men, and in some even higher for women than for men. The reason for the overall rate of rejection of women being higher, is simply that women tended to apply to departments with higher rejection rates overall than men did. This illustrates the reality that inference from numbers to causes, even if possible in principle, is very difficult.

6. The Larger Uses of Probability

In life, the more probable routinely happens more often than the less probable—if there’s anything in a name. But this only underscores the fact that sometimes the less probable happens. Does this require explanation? Does nature owe us an explanation when the less likely happens in preference to the more likely? This question raises, albeit in a narrower context than usual, the question of what explanation consists in, and the role of probability in explanation. And it raises in an oblique fashion the question of the relationship between giving explanation and attending to relations of cause and effect. At the center of the debate over explanation are those who equate explanation-giving with the illumination of causes. Naturally, their adversaries separate the terms of this equation. However this debate plays itself out, the question will still remain as to how knowledge of probabilities should bear upon issues of moral and legal responsibility. Are we entitled to use statistical knowledge in assessing guilt? How shall we use probabilities responsibly, on the way to framing scientific explanations? And how shall we use them responsibly in the courtroom?
NOTES

4. It isn’t true, therefore, that the Humean is not in a position to make a distinction between wanting and valuing, as Gary Watson suggests in the much-cited article “Free Agency,” *Journal of Philosophy* 72 (1975): 205–20. In fact, the Humean must make that distinction—and does—for the sake of her utility theory.

PART I

Probability, Frequency, and Inference