Introduction

Humans have been thinking in probabilistic terms since antiquity. They have been thinking systematically and philosophizing about probability since the seventeenth century. And they have been formalizing probability since the end of the nineteenth century. The twentieth century saw intense philosophical work done on *interpreting* probability, in a sort of attempt to find out its essence. The twenty-first century, I argue, will bring a focus on more practical endeavours, concerning mainly the methodologies of data analysis and statistical modelling. The essence of probability, it turns out, lies in the diversity of its uses. So, the methodological study of the use of probability is what brings humans closer to a comprehensive understanding of its nature.

These and other ideas expounded in this Element developed out of a Marie Curie project on probability and propensities that I carried out at the Institute of Philosophy of the School of Advanced Study at London University during 2013–15. I came out of that project with the distinct impression that the study of practice was of primary importance; and that much philosophy of probability is still to come to terms with it. This Element is my first attempt at the bare bones of a new research programme into the methodology of statistical modelling. Most of the Element is devoted to justifying this methodology – on the grounds of practical involvement with the scientific modelling practice but also, I argue, on account of the limitations of the traditional interpretative approaches to the topic.

Thus, the first half of the Element (Sections 1–7) is entirely a state-of-the-art review of the historiography of probability and its ensuing impact upon the interpretative endeavour. This is fitting for a Cambridge *Elements* volume, which allows for a profuse setting of the stage. And it is anyway needed in order to understand why nothing other than a study of the practice of statistical model building will do for a full understanding of objective probability. I first explore (in Section 1) the dual character of the notion of probability from its inception – the subjective and objective aspects of probability that are essential to any understanding the concept. The twentieth century brought in several *interpretations* of probability. But one way or another, they all aim to reduce probability to either subjective or objective elements, thus doing away with the duality; and one way or another they all fail, precisely because they do away with the duality. In the remaining sections in this half of the Element, I analyse in detail the many objections against both the main subjective interpretations (the logical and personalist or Bayesian interpretations), and the main objective interpretations (the frequency and propensity interpretations). To make most of these interpretations work, and overcome the objections, demands some

acknowledgment of the complex duality of probability. This is by now widely accepted, and the Element first reviews the roots and consequences of pluralism about objective probability.

The second half of the Element (Sections 8–13) then centres upon the objective aspects of probability, but now without any pretence of a reduction of the whole concept. The discussion is focused entirely on objective probability, and it contains most of the original material. I advance a number of novel theses, which I defend in various original ways as well as proposing a number of new avenues for research. The starting point is pluralist, and it accepts the duality insofar as it argues that there are important matters of judgement in the selection of crucial aspects of the application of objective probability in practice. Here, the critical distinction, advanced in Sections 8 and 9, is between the traditional project to merely *interpret* probability and a distinct project to study the *application* of probability. On the other hand, I go considerably beyond the pluralism defended in the first half of the Element and, in Section 10, I embrace novel forms of pluralism and pragmatism regarding objective probability.

The central idea of the second half, which also informs the Element as a whole and looms large through most of its discussions, is what I have elsewhere called the 'tripartite conception' of objective probability (Suárez, 2017a). This is the idea that the failure to reduce chance to either propensity or frequency ought to lead to the acceptance of all three concepts as distinct, insufficient yet necessary, parts of the larger notion of 'objective probability'. This tripartite conception is introduced in Section 10, which also assesses the role of judgement and various subjective components. Sections 11, 12, and 13 are then devoted to modelling methodology, and the application of the tripartite conception in statistical modelling practice in particular, in what I call the 'complex nexus of chance' (CNC). The thought running through these sections is new and radical: objective probability is constituted by a thick array of interlinked practices in its application; these are practices that essentially involve the three distinct notions pointed to above; and since none of these notions is theoretically reducible to any combination or set of the other two, this means that the overall methodology remains unavoidably 'complex'. There is no philosophical theory that may explicate fully the concept of objective probability, or chance, by reducing this complexity, and this already sheds light on the limitations of the interpretations reviewed in the Element's first half.

What's more, the second half of the Element also continues to illustrate the fundamental duality of probability unearthed in the historiographical material reviewed in the first seven sections. It does so in three different yet interrelated ways. First of all, it leaves open that subjective elements may come into the nature of the single-case chances that make up the tripartite conception.

Secondly, confirmation theory comes into the assessment of evidence for and against different models. And, finally, there are irreducible subjective judgements involved in the pragmatist methodology advocated in the later sections. For instance, in Section 11 I argue that choosing the appropriate parametrization of the phenomenon to be modelled is a critical part; and there is no algorithm or automatic procedure to do this – the choice of free parameters is subject to some fundamentally 'subjective' estimate of what is most appropriate in the context for the purposes of the model at hand. Once again, the 'subjective' and the 'objective' aspects of probability meet in fundamental ways (see Gelman and Hennig (2017) as well as my response Suárez (2017b) for an account of such a merge in practice). Another related sense of subjectivism in statistical modelling is sometimes referred to as the 'art of statistical modelling' and concerns the choice of a correlative outcome or attribute space. There is nothing arbitrary about this 'subjectivity' though, since it answers precisely to specific pragmatic constraints: it is a highly contextual and purpose-driven judgement.

On my view, each of the parametrizations of a phenomenon involves a description of its propensities, dispositions, or causal powers. What is relevant about propensities is that they do not fall in the domain of the chance functions that they generate (Suárez, 2018). Rather a propensity is related to a chance function in the way that possibilities are related to probabilities: the propensity sets the range of possible outcomes, the full description of the outcome space, while the chance function defined over this space then determines the precise single-case chance ascribed to each of these outcomes. A different parametrization would involve a different description of the system's propensities, perhaps at a different level of generality or abstraction (and no parametrization is infinitely precise); and focusing on a different set of propensities may well issue in a different set of possible outcomes, hence a different outcome space, over which a different chance function shall lay out its probabilities. Since the parametrizations obey pragmatic constraints that require appropriate judgements within the context of application, it follows that the outcome spaces will correspondingly depend on such judgements. In other words, a chance function is not just a description of objective probabilities for objectively possible outcomes; it is one amongst many such descriptions for a particular system, made relevant by appropriate judgements of salience, always within a particular context of inquiry. Here, again, the 'subjective' and the 'objective' aspects of probability merge.

1 The Archaeology of Probability

The philosophy of probability is a well-established field within the philosophy of science, which focuses upon questions regarding the nature and interpretation

of the notion of probability, the connections between probability and metaphysical chance, and the role that the notion of probability plays in statistical modelling practice across the sciences. Philosophical reflection upon probability is as old as the concept of probability itself, which historians tend to place originally in the late seventeenth century. As the concept developed, it also acquired increasing formal precision, culminating in the so-called Kolmogorov axioms first formulated in 1933. Ever since, philosophical discussions regarding the interpretation of probability have often been restricted to the interpretation of this formal mathematical concept, yet the history of the concept of probability is enormously rich and varied. I thus begin with a review of some of the relevant history, heavily indebted to Ian Hacking's (Hacking, 1975, 1990) and Lorraine Daston's (Daston, 1988) accounts. Throughout this historical review I emphasize the non-eliminability of objective chance. I then turn to a detailed description of the different views on the nature of probability, beginning with the classical interpretation (often ascribed to Laplace, and anticipated by Leibniz), and then moving on to the logical interpretation (Keynes), the subjective interpretation (Ramsey, De Finetti), the frequency interpretation (Mises, Reichenbach), and ending in a detailed analysis of the propensity interpretation in many of its variants (including the views of Peirce, Popper, Mellor, Gillies, and my own contributions). The discussion is driven by the 'doctrine of chances' and the recognition that objective chance is an ineliminable and essential dimension of our contemporary concept of probability. In particular I argue that the logical and subjective interpretations require for their intelligibility a notion of objective chance and that the frequency interpretation is motivated by a form of empiricism that is in tension with an honest and literal realism about objective chance.

Hacking's archaeology of probability revealed unsuspected layers of meaning in the term 'probability', unearthed a fundamental duality in the concept, and revealed that, although the concept itself in its modern guise only fully appears around 1660 (most notably in the Pascal–Fermat correspondence), the imprint of the antecedent marks (i.e. of the 'prehistory' of probability) are even to this day considerable. The legacy of Hacking's inquiries into probability is an increased understanding of the transformative processes that turned the Renaissance's concept of 'probability' into our contemporary concept of probability. The new concept finally comes through strongly in the writings of the Jansenist members of Port Royal (mainly Arnauld and Pascal), but it has both antecedents and contemporaries in some of leading thinkers on signs, chance, and evidence, including Paracelsus, Fracastoro, Galileo, Gassendi, and most notably the contemporaneous Leibniz and Huygens.

The fundamental change traced by Hacking concerns the notion of evidence which, in its contemporary sense, also emerges at around the same time. In the old order, the justification of 'probable' claims was thought to be provided by the testimony of authority (usually religious authority). But the Renaissance brings along a reading of natural and, in particular, medical and physiological phenomena where certain 'signs' are taken to impart a corresponding testimony, under the authority of the book of nature. 'Probable' is then whatever is warranted by the relevant authority in the interpretation of the 'signs' of nature. But what to do in cases of conflict of authorities? Hacking (1975, ch. 5) chronicles the fascinating dispute between the Jesuit casuistry tradition – which considers the consequences of each authority and chooses accordingly – and the protesting Jansenists' novel emphasis on locating the one true testimony – typically the testimony provided by nature herself. The transformation of the testimony of earthily authority into the evidence of nature thus configures the background to the emergence of probability. Hacking's careful 'archaeology' then reveals that the most striking imprint of the old order upon the new is the dualistic or Janus-faced character of probability. Our modern concept of probability is born around 1660 and characteristically exhibits both epistemological and ontological aspects. It inherits the dualism from the medieval and Renaissance conceptual schemes which, however otherwise fundamentally different, also exhibited a similar duality. Thus, in the old order and parlance, 'probable' stood roughly for both the opinion of the authority and the evidence of nature's signs, while in the parlance of the new order, 'probable' stands both for logical or subjective degree of belief and for objective chance, tendency or disposition.

My aim in the first half of the Element is to review the present state of the philosophy of probability with an eye on this fundamental duality or pluralism. I shall emphasize how an appropriate articulation of subjective probability is facilitated by a proper regard for the objective dimension of probability. And conversely, a fair theory of objective chance needs to make room and accommodate subjective elements. First, in Section 2 I continue the historical review by introducing the notion of equipossibility in Leibniz and Laplace. I then move on in Section 3 to the logical interpretation and the principle of indifference as they appear mainly in the work of John Maynard Keynes. In both cases I aim to show the role of objective notions of probability in the background of the argument and development of the logical interpretation of probability. In Section 4 I follow a similar strategy with the subjective interpretation of Ramsey and De Finetti, in an attempt to display the ways in which the interpretation ultimately calls for objective notions in order to overcome its difficulties. Section 5 retakes the historical account in order to review the history of metaphysical chance and its ultimate vindication in the late nineteenth century,

particularly in relation to the work of the American pragmatist philosopher, Charles Peirce. In Section 6 I introduce and review different versions of the frequency interpretation of probability (finite frequentism and hypothetical frequentism). I show that subjective notions appear in the formulation of these theories, or at any rate in those formulations that manage to overcome the objections. Finally, in Section 7, I review in detail some of the main propensity accounts of probability, pointing out some of their resorts to subjective notions.

2 The Classical Interpretation: Equipossibility

The classical interpretation of probability is supposed to be first enunciated in the works of Pierre Simon Laplace, in particular in his influential *Essai Philosophique sur les Probabilités* (1814). But there are important antecedents to both classical probability and the notion of equipossibility that ground it in the writings of many of the seventeenth-century probabilists,[1] particularly Leibniz's and Bernouilli's about a century earlier. Ian Hacking (1975) chronicles the appearance of the notion of equipossibility in the metaphysical writings of Leibniz, and the connection is apposite since it is an essentially modal notion that nowadays can best be understood by means of possible world semantics. I first review the historical developments that give rise to the Laplacean definition, and only then address some of the difficulties with the classical view in more contemporary terms.

Leibniz seems to have developed his views on probability against the background of an antecedent distinction between two types of possibility, which roughly coincide with our present-day notions of *de re* and *de dicto* possibility (Hacking, 1975, p. 124). In English we mark the distinction between epistemological and physical possibilities by means of different prepositions on the word 'possible'. There is first a 'possible that' epistemological modality: 'It is possible that Laplace just adopted Leibniz's distinction' expresses an epistemological possibility; for all we know it remains possible that Laplace did in fact copy Leibniz's distinction. The statement is in the present because it reflects our own lack of knowledge now. Contrast it with the following 'possible for' statement: 'It was possible for Laplace to adopt Leibniz's distinction' expresses a physical possibility at Laplace's time, namely that Laplace had the resources at his disposal, and sufficient access to Leibniz's work, and was not in any other way physically impeded from reproducing the distinction in his own work. More prosaic examples abound: 'It is possible that my child rode his bicycle' is

[1] Gigerenzer et al. (1989, ch. 1.9) even argue that by the time of Poisson´s subsequent writings circa 1837, the classical interpretation was already in decline!

epistemological, while 'it is possible for my child to ride his bicycle' is physical, and so on.

Now, epistemological possibility is typically *de dicto* (it pertains to what we know or state), while ontological possibility is *de re* (it pertains to how things are in the world independently of what we say or state about it). So, the 'possible that' phrase tends to express a *de dicto* possibility, while the 'possible for' phrase expresses *de re* possibilities. The two are obviously related – for one physical possibility may be thought to be a precondition for epistemological possibility since there is no *de dicto* without *de re*. For Leibniz the connection was, if anything, stricter – they were two sides of the same concept of possibility. And in building his notion of probability out of possibility, Leibniz transferred this dualism onto the very concept of probability: 'Quod facile est in re, id probabile est in mente' (quoted in Hacking, 1975, p. 128). The link expresses Leibniz's belief that the dual physical and epistemological aspects of probability track the duality of *de re* and *de dicto* possibility.

This tight conceptual connection is also the source of Leibniz's emphasis on equipossibility as the grounds for the allocation of equal probabilities, and it in turn underwrites Bernouilli's and Laplace's similar uses of the notion. Leibniz employs two separate arguments for the equiprobability of equipossible events: the first derives from the principle of sufficient reason and is essentially epistemological; the other one derives from physical causality and is essentially ontological or physical (Hacking, 1975, p. 127). According to the first, if we cannot find any reason for one outcome to be any more 'possible' than another, we judge them epistemically equiprobable. According to the latter, if none of the outcomes is in fact more 'facile' than any other, they are physically equiprobable.

The duality of probability (and its grounding in the similar duality of possibility) becomes gradually lost in the advent of the classical interpretation of probability, which is often presented in a purely epistemic fashion, as asserting that probabilities represent merely our lack of knowledge. The eighteenth century brought an increasing emphasis on the underlying determinism of random looking phenomena, in the wake of Newtonian dynamics, and probability in such a deterministic universe can only signal the imperfection of our knowledge. By the time of the publication of the treatise that established the classical interpretation (i.e. Laplace's *Essai sur les Probabilités*) in 1814, the deterministic paradigm had become so imperious, and the demise of probability to the strict confines of the epistemology so marked, that Laplace could confidently assert that a superior omniscient intelligence would have no time or purpose for probability. If so, the fact that ordinary agents have use for non-trivial (i.e. other than 0 or 1) probabilities comes to show our cognitive

limitations and entails that probability is essentially an epistemic consequence of our ignorance. The connection is at the foundation of subjective views on probability and is nowadays embodied in what is known as *Laplace's demon*: 'an intelligence which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intelligence were also vast enough to submit these data to analysis, she would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intelligence nothing would be uncertain and the future just like the past would be present before her eyes' (Laplace, 1814, p. 4, my own translation).

Yet, the Laplacean formal definition of probability as the ratio of favourable to possible cases, of course, only makes sense against the background of equipossible events, as: $P(a) = \frac{\#(a)}{\#(t)}$ where #(a) is the number of positive cases of a, and #(t) is the number of total cases. Thus, in the case of an unbiased coin, the probability of the coin landing heads if tossed is given by the ratio of the cases in which it lands heads divided by the total number of cases (i.e. either outcome). But this, of course, assumes that each case is equipossible – that is, that the tosses are independent in the strong sense of there being no causal influences that determine different degrees of possibility for the different outcomes. If, for instance, landing heads on the first trial made it more likely for the coin to land heads in the second trial, the probability of heads in the second or any other trial in the series would not be given by the ratio. Laplace himself was acutely aware of the issue. As he writes: 'The preceding notion of probability supposes that, in increasing in the same ratio the number of favourable cases and that of all the cases possible, the probability remains the same' (Laplace, 1814, ch. 6).

Commentators through the years have pointed out repeatedly how any purely epistemic reading of the condition of equipossibility would render Laplace's definition of probability hopelessly circular: it defines the notion of probability back in terms of the equivalent notion of equal possibility – the very grounds for epistemic equiprobability. Hence, we find Hans Reichenbach (1935/1949, p. 353) stating as part of his critique of epistemological theories: 'Cases that satisfy the principle of "no reason to the contrary" are said to be equipossible and therefore equiprobable. This addition certainly does not improve the argument, even if it originates with a mathematician as eminent as Laplace, since it obviously represents a vicious circle. Equipossible is equivalent to equiprobable.' However, the realization that Leibniz and Bernouilli in fact entertained mixed notions of probability and possibility, incorporating both epistemic and ontological dimensions, allows for a distinct resolution of this issue. If the equipossibility is ontological, if, for example, it is physically there in nature,

then the assumption of equal probabilities follows without any appeal to sufficient reason. There seems to be no circularity involved here as long as physical possibility may be independently understood.

Our standard contemporary understanding of modality is in terms of possible world semantics. A statement of possibility is understood as a statement about what is the case in some possible world, which may but need not be the actual world. Equipossibility is trickier since it involves comparisons across possible worlds, and these are notoriously hard to pin down quantitatively. Measures of similarity are sometimes used. For two statements of possibility to be quantitatively equivalent it needs to be the case, for example, that the number of possible worlds that make them true be the same, or that the 'distance' of such worlds from the actual world be the same, or that the similarity of those worlds to the actual world be quantitatively identical. Whichever measure is adopted, it does seem to follow that some objective relation across worlds warrants a claim as to identical probability. The quantitative measures of equipossibility are not necessarily probability measures – but they can be seen 'to inject' a probability measure at least with respect to the equally possible alternatives. It is at least intuitive that physical equipossibility may give rise to equiprobability without circularity. The upshot is that what looks like an eminently reasonable purely epistemological definition of probability as the ratio of favourable to possible cases in fact presupposes a fair amount of ontology – and a concomitantly robust and unusually finely graded notion of objective physical possibility.

3 The Logical Interpretation: Indifference

There are two schools of thought that assume that probability is not objective or ontological – not a matter of what the facts of the world are, but rather a matter of the mind – one of our understanding or knowledge of the world. These accounts follow the main lines of the most common interpretation of the classical theory. According to the logical interpretation, probability is a matter of the logical relations between propositions – a question thus regarding the relational properties of propositions. According to the subjective interpretation, by contrast, probability is a matter of our degrees of belief – a question that regards therefore our mental states, and in particular our belief states. These interpretations developed particularly during the twentieth century. The logical interpretation was championed by John Maynard Keynes, Harold Jeffreys, and Rudolf Carnap (for what Carnap called probability$_1$ statements, which he distinguished from objective probability$_2$ statements); while the subjective interpretation was defended by Frank Ramsey, Bruno de Finetti, and Leonard

Savage. In this section I review the logical interpretation, mainly as espoused by Keynes, and in Section 4 I look at the subjective interpretation, particularly in Ramsey's version.

Keynes argues that probability is a logical relation between propositions akin to logical entailment but weaker – whereby two propositions A and B are related by means of logical entailment if and only if A cannot be true and B fail to be so; while A and B are more weakly related by partial degree of entailment if and only if A cannot be true and B fail to have some probability, however short of certainty, or probability one. So, the first caveat that must be introduced at this point is the fact that for Keynes probability is not in fact subjective but objective. However, we must be careful with our use of language here – 'objective' for Keynes does not stand for 'ontological' but for non-arbitrary or relative to known fact. More particularly, Keynes held that the probability of a proposition is always the relation of partial degree of entailment of that proposition by some background body of knowledge. That is, given some background knowledge, a proposition is entailed to a certain degree. As he writes (Keynes, 1921, p. 4): 'In the sense important to logic, probability is not subjective. It is not, that is to say, subject to human caprice. A proposition is not probable because we think it so. Once the facts are given that determine our knowledge, what is probable or improbable in these circumstances has been fixed objectively, and it is independent of our opinion.'

The fundamental insight here is the thought that probability is a logical relation amongst propositions. Thus, if I claim now that 'the probability that it will rain tomorrow is 50 per cent', I am making a claim about how probable this proposition is on account of the knowledge I now have of any facts relative to it – weather patterns, dynamical laws, the present isobaric facts, and so on. If and when my information changes, so does my probability estimate. But this is perfectly consistent with the relational character of probability: it is always a property of a proposition relative to background knowledge, which will naturally vary with time, as new information accrues. Therefore, the probability of the proposition in question becomes zero or one not at the time the event comes to be – or fails to be – but rather at the time we as agents gain the relevant background information. Yet, there is a normative dimension to probability according to Keynes, as we saw in the quote above. What this means is that there is some background information that is objectively relevant at each time for each proposition. The rational agent is normatively constrained by it in the sense that, were the agent to be aware of all the relevant facts, she would ascribe the corresponding probability. We can thus say that objectively the probability of the proposition is given by its relation to the background facts that are relevant to our knowledge *regardless of whether anyone is in fact aware or not of those*