80-600
Minds, Machines and Knowledge
w/ Horacio Arlo-Costa

Keith Douglas
kdouglas@andrew

A Survey and Review - What is Probability?

Introduction

In this paper I survey and examine notions of probability in philosophical (particularly, metaphysical and epistemological) contexts and scientific/technological contexts. The aim is to raise awareness about many different understandings of probability and their conflations and confusions. Specifically, I shall pay close attention to the "same" probability function being defined over heterogeneous sets of objects and to possibly spurious arguments for use of probability in various domains as well as the arguments for and against certain applications and interpretations.

The philosophical and scientific viewpoints surveyed in this paper are not meant to be exhaustive of different nuances of viewpoint in the literature. The sample is selected from those often associated with "probabilism" or criticism of certain aspects of it. It is perhaps more historical than contemporary as many current viewpoints consider the issues I investigate to have been settled. I do not, however, stray much beyond the early 20th century for manageability reasons. As noted above, my aim is primarily "consciousness raising" and as such could easily be extended with many more case studies.

This paper consists of four sections. Generally, the first section shall be mainly concerned with those who (at least in the works under consideration) uncritically accept the view that probability can be attributed to propositions, sentences or other (pseudo)linguistic items. This preference is a bias on my part; it is this viewpoint (and that of "frequentism") for which much foundation work has been done and tacitly accepted. It behooves an investigation, then, to see whether this work does what it is claimed to do.

The second investigates the viewpoints of several influential works which attempt to develop an account of one or more of these notions. The third is a brief survey of scientific and technological uses of the concepts of probability in attempt to see whether there is any consensus on the actual use of probability (rather than "merely" by philosophers analyzing it). A primary goal here is to provide data on said actual uses for any future work in probability by philosophers and foundation workers. The scientific and technological uses thus should constrain and shape any possible philosophical views on the subject. (This is not to say they rigidly rule out certain interpretations, but they at the very least rule some in.) The fourth section draws some relevant lessons from these uses and arguments met in the second and third sections.

It must be noted also that this survey is not meant to be exhaustive, even with a restriction to the twentieth (and twenty first!) century alone. No doubt, countless other views exist out there; my job in this paper is simply to raise awareness about the various different notions and the arguments raised for and against each.

I close this introductory section by noting that "probability theory" as often understood is a branch of pure mathematics[1] . I am only concerned with applications of this mathematics. I shall take it for granted that the mathematics by itself does not specify an, interpretation or use of the theory. Though, as we shall see, some have thought formal considerations can rule out some uses.

Section 1 - Some taking of notions of probability for granted

[1] For instance, "probability measure" in purely formal terms is defined in Weisstein (2001) as:

> "Consider a probability space specified by the triple (S, S, P), where (S, S) is a measurable space, with S the domain and S is its measurable subsets, and P is a measure on S with P(S) =1. Then the measure P is said to be a probability measure. Equivalently, P is said to be normalized."

Along these lines, it is important to note that throughout I shall be inquiring into how given thinkers know that their functions have range [0,1] and the like. Obviously, if the formal characteristics of the function are of a probability measure in the sense above, this will hold true. It should be obvious then that I am asking how they know this "maps onto the world" appropriately (or in whatever idiom one likes should one not be a realist).

This section shows that there is some need to clarify which notions of probability have been used in various contexts. Examples such as these can be multiplied; here I focus only on four.

I shall start by providing an example from a well-known philosopher. Peter Gärdenfors writes as follows (1988, pp. 105, italics in original):

> "Bayesianism comes in two parts. The first part of the doctrine is that epistemic states can be represented by *probability functions* defined over sentences of an appropriate language. This part is usually defended by the Dutch book theorem or some related coherence argument. In the main part of this chapter, I assume this representation of states of belief."

Gärdenfors' book thus makes use of probability functions with domain "sentences." Beliefs are said to be something sentential, and degrees of belief are the thus interpretation of the probability attached to each sentence. His justification for this (the Dutch book theorem) is a typical viewpoint shared by many others and thus bears some analysis. But, what does the Dutch book theorem show? It assumes that one can place a probability measure on sentences corresponding to credences. Since Gärdenfors' objective in this book is to develop a system for understanding belief dynamics, we thus have one area where the notion of probability needs clarification.

Another example of a work where some notions of probability are taken for granted is in Levi's *Hard Choices* (1986). Here, Levi assumes justification of his use of probability functions of domain propositions. Levi references his earlier work on the subject. Levi's main task in *Hard Choices* is about decision making, and so this is another area where one needs an elucidation of probability. It might appear at first glance that Levi is merely not dealing with everything of importance in one place, and so this remark about *Hard Choices* is unjustified. I claim that this rejoinder misses that many of the book's central issues need it to be properly grounded. For instance, §7.4 (on expected value) includes passages such as the following (pp. 113):

"Suppose Smith assigns h_1 the credal probability 0.1 ..."

Thus, to evaluate whether Levi's decision theory is a good one requires us to understand his use of probability.

Probability has also been said to be useful in elucidating the notion of truth. Popper, for instance, has attempted to do so. Whether this is justified or not is another story: the interpretation of probability here is critical. In some (perhaps all) interpretations of probability theory, the notions of logic are prior, making Popper's goal (apparently) impossible. But, perhaps an a-logical interpretation of probability could be developed, thus making this task possible.

Related to the above, Reichenbach (1949, ch. 10) has tried use the probability calculus not to elucidate truth, but instead replace the two truth values of classical logic with a continuum of truth values. This requires that the domain of a factual probability function to be the same as the domain of the (factual) truth valuation function.

Many other possible areas of uses of probability are possible. We now turn to a discussion of several representative authors on our theme.

Section 2

One person whose views on probability have been very influential is Rudolf Carnap. Carnap recognizes (1950, pp. 163) that there are several distinct notions of probability. Here, I first investigate his argument for the two notions he recognizes, examining his claims for the merit of each. I then investigate whether he has in fact exhausted the possibilities. Carnap calls the two notions "probability$_1$ and "probability$_2$". The first of these he calls "logical probability" and the second he claims is "relative frequency."

His book is primarily about the first of these, so he spends much time explicating it. Since I am most interested in the quantitative notions of probability (even if my verdict is eventually that pseudoquantities are being used), I focus on his third explicatum of probability_1. This is said to be (1950, pp. 163, italics in original):

> "a *quantitative* concept of confirmation, the *degree of confirmation* ('*h* is confirmed by *e* to the degree *r*.')"

How, then, does Carnap argue that degree of confirmation is a probabilistic notion? The first step in his argument is to show that qualitatively speaking there are notions of (degrees of) confirmation of hypotheses by evidence etc. I regard this as relatively uncontroversial. Carnap rightly points out that this is insufficient to make his probability_1 into a quantitative notion. He also recognizes that it is not prima facie absurd to consider that the strength of support given h by e is 5 (recognizing the importance of the <u>range</u> of the appropriate function). He then <u>assumes</u> several hypotheses about belief strength (somehow related to degrees of confirmation) that go a long ways to "molding it into" probabilistic terms. He postulates: (i) credence is to be measured by nonnegative numbers ≤ 1 and (ii) that if two hypotheses are L-exclusive (logically incompatible) the support given to e by the two together is the sum of the two taken separately. Since (he says) $h \vee \neg h$ is L-true and is hence certain, the strength of support of it on any evidence is therefore 1 by postulate (i) above. From this, he concludes that the credences of both h and $\neg$h are 1/2. Note that Carnap is implicitly assuming here that credence($\neg$h) = 1-credence(h). There (as yet) has been no motivation for <u>this</u> postulate (nor is it even acknowledged). The other two postulates are not justified at this stage either and Carnap even admits they are arbitrary. Carnap then generalizes these results to the case of n mutually exclusive and independent options. How one makes sure one has all the options and that they are mutually exclusive is not stated. Some motivation is found in his next section, which considers the idea that probability_1 represents a fair betting quotient. This uses the usual Dutch Book arguments. As I have remarked previously, these sorts of arguments cannot justify the use of probability completely. Carnap, to his credit, however, has recognized that some of his postulates that get him to the use of probability are arbitrary (i.e. not supported except by their "fruitfulness.")

Carnap then spends a great deal of time relating probability_1 to probability_2 and thus credences to frequency conceptions. I shall examine the latter briefly now. A general way of stating Carnap's understanding of the connection between these two notions is that one should have one's probability_1s correspond to an estimate of the appropriate relative frequency. He does successfully disentangle the difference between a hypothesis and the potential individuals and classes it refers to, and then goes on to clarify the notion of an estimate. This latter word is the key in the connection. He regards this estimation procedure as the cornerstone of his "inductive logic" and we need not examine this in great detail here. However, we can still explore his definition of estimate. On page 169, *estimate* is defined in terms of probability_1 and the notion of a mean. As Carnap reminds us, we still have no notion of *degree of confirmation*, so all of these notions are still somewhat hanging in the air. Next,Carnap, to his credit, however, has recognized that some of his postulates that get him to the use of probability are arbitrary (i.e. not supported except by their "fruitfulness.") note that not all cases of probabilty_1 form estimates of probability_2 according to Carnap. This relation only occurs when the class of the appropriate unknowns is sufficiently large. There is no indication at this stage of his work what "sufficiently large" is to be. We are then told that this estimate relation is like any of the relations between an empirical concept and its corresponding inductive concept. This is obscure. Taken strictly literally, Carnap is thus saying that probability_2 is also somewhat subjective. This is because he regards the relationship to be one between <u>concepts</u> rather than between say a concept and something in the world. This may be just a slip on his part so we need not dwell on it much. However, it is important to consider the possibility that events (in the ontological sense[2]) are not really the appropriate domain

[2] I shall in general use "event" this way, despite its possible confusion with the non-ontological sense of the word sometimes used in probability theory. The merit of these decisions will be discussed later.

of the probability_2 functions and instead a conceptual representation of them are. Since Carnap does not dwell on this point, and explains that his conception is somewhat like Reichenbach's (which we shall meet later), it seems that Carnap did indeed have an ontological interpretation in mind. (Of course, he would not have ever put it in those terms, but that is of little consequence.)

However, we can gain a better appreciation of Carnap's understanding of probability in both senses if we notice how he suggests we formulate certain inferences between the two. We turn to this now, leaving elucidation of his notion of *degree of confirmation* to another day[3]. Carnap suggests in order to perform the inference from (ii) to (iii) (pp. 173, italics in original):

> "(ii) 'The *probability* that any future throw of this die will yield an ace is 1/6.'
>
> (iii) 'If a sufficiently long series of throws of this die is made, the *relative frequency* of aces will be 1/6.'"

we need to instead replace statement (iii) with the analogous ones below. We can infer from (ii) to (iv) or to (v), Carnap says (pp. 174, italics in original):

> "(iv) 'The *estimate of the relative frequency* of aces in any future series of throws of this die is 1/6.'
>
> (v) 'The probability_1 of the prediction that the relative frequency of aces in a future series of throws of this die will be within the small interval $1/6 \pm \varepsilon$ is high (and can even be brought as near to 1 as wanted) if the series is made sufficiently long.'"

The original inference could not be done because (ii) was "logical" and (iii) factual. The new inferences, by contrast, are able to succeed. This is because they involve estimates of appropriate frequencies. Carnap is right here: one cannot <u>infer</u>

[3] As it happens, some regard this latter notion as being useless (because the degree of confirmation of any hypothesis on his conception is exactly zero [See, e.g., Bunge 1998]), and so would infect the notion of his understanding of probability. But one of the merits of Carnap's work is that it is very "modular". It seems one could in principle replace the faulty notion and continue to use the notions elucidated by it with that replacement in mind.

from a "logical" statement to one about the world directly. One's hypotheses in any science or every day life always have a margin of error. This is what (iv) and (v) have; (v) even gives this margin of error (ε). Carnap finally also deals with one might call "reflexivity," an issue many authors who adopt an understanding of probability with a domain of propositions do not consider. Carnap claims that (pp. 175) that the probability$_1$ of a statement involving probability$_1$ is logically true. This is because (so he claims) given evidence, the probability$_1$ of a hypothesis is determined uniquely. Nowhere, however, are we given general procedures for doing so, so there is no reason to believe this claim, alas. (Equally unhelpfully: if Bunge (see footnote 2, above) is correct, probability$_1$s are determined uniquely, but all are zero.)

We have thus seen that Carnap's two notions of probability are those connected to credences and frequencies and that one can interrelate them. We have also met in passing a few problems with his conception, but we shall return to these later in the works of others who criticize him. In particular, he overlooks the propensity interpretation of probability.

We have seen that Carnap worked heavily in the "probability of hypotheses" area, i.e. attributed a notion of probability to propositions. Another major influence on the "probabilities of propositions" school has been Keynes. He also provides some unique arguments for his understanding of probability. Keynes' position is that probability is a measure of the rational belief in a hypothesis (1948 [1921], pp. 4), taken in propositional form. However, he claims that this measure is not subjective:

> "When once the facts are given which determine our knowledge, what is probable or improbable in these circumstances has been fixed objectively and is independent of our opinion."

Does Keynes succeed in showing that the above view is correct? To his credit, he asserts that the rationality in question is human rationality. However, apart from this consideration there is very little else present to justify this viewpoint. On this consideration, it is important to realize that Keynes does claim that the probability relation is primitive, and cannot be analyzed in terms of any other notions. In this sense, his Treatise can be looked at as containing an implicit definition.

Keynes also dispenses with the frequency interpretation of probability, saying (pp. 95) that it is an "ordinary correct criticism" of the view that it departs from ordinary usage of the word. Nevertheless, he also feels that one can infer probabilities in his sense from frequencies. However, Keynes' understanding of probability is also noteworthy in another way. Unlike some more recent thinkers, he does not think that the numerical probability functions we are considering are defined on the complete domain in question. In other words, for Keynes, not every proposition has a numerical probability.

However, Keynes' work suffers from the fact that he takes it as given there are only two possible notions of probability, and having dealt with one he is left with the other. Nor does he argue for propositions having numerical credences or show how they might arise in any great detail. There is some material on the latter point when he analyzes and refines the so-called "principle of indifference." This gives some initial credences - or so he says - when the possibilities (of what is never quite made clear - in particular whether they are ontological or epistemological) associated with a situation are mutually exclusive and exhaustive (pp. 65).

However, it is not clear how to obtain probabilities in Keynes' sense in other situations. It is entirely possible that Keynes would say that his understanding of probability is not applicable in such situations. This strikes me as being a rather large impoverishment of the theory; it thus bears further investigation. A final few remarks on Keynes' understanding of probability are thus in order: despite his insistence that his notion is objective, this claim is betrayed on page 75, when he gives an example using some data and does not explain where the given probabilities came from. He owes us this, because as yet he has not shown how to obtain them. He has talked about mutually exclusive cases, but these do not easily yield probabilities such as 1/20 as he uses here. There is also a slip of a rather important (and, as we shall see, recurring) sort at the end of this page (italics in original):

> "The typical case, in which there may be a *practical* connection between weight and probable error, may be illustrated by the two cases following of balls drawn from an urn. In each case, we require the probability of drawing a white ball; in the first case we know that the urn contains black and white in equal proportions; in the second case the proportion of each colour is unknown and each ball is as likely to be black as white. It is evident that in either case the probability of drawing a white ball is 1/2, but that the weight of the argument in favour of this conclusion is greater in the first case."

The slip here is between propositions and what they refer to. If (as Keynes has said) probability functions have as domain the set of propositions, one simply cannot say, as he does, "the probability of drawing a white ball is 1/2." This is because drawing a white ball is not a proposition. It can be referred to by a proposition, and that may well (at least *prima facie*) have credence 1/2. But, then Keynes (and others who follow with similar understanding) is left with the following problem: how do we connect credences to things in the world? When the cases are mutually exclusive, as they are here (at least in the first case), we do have an earlier postulate. But even with this postulate, it is important to make this distinction. One good reason is that of the second case in the above quotation. What does Keynes mean by "as likely"? That the proposition: "the ball I will draw from the urn is white" has credence 1/2? This is where the subjectivism comes in, and this is thus, where the critics of the subjectivist interpretations of probability get their "toe in the door." With these issues now on the table, I shall leave Keynes' work and move on.

Returning to a more recent writer, I now I examine the works of Isaac Levi. As we have seen, some of Levi's work in decision theory and related fields relies on some notions of probability and, there is some confusion as to which. It thus bears investigating some of his other works in order to sort this out. I first make use of a more recent book of his (as this will allow for any considerations raised by *Hard Choices* referenced above): *The Fixation of Belief and its Undoing* (1991) (hereafter, *Fixation*). I shall also make use of the more original presentations of his understanding of probability, as these are referenced in *Fixation* and in *Hard Choices.*

In *Fixation* (pp. 166-7), Levi remarks that his understanding of probability makes use of the earlier works of Carnap and Jeffreys. However, unlike either of these two, Levi is aware that the sets involved in these probability functions in question may not yield a unique (quasi)ordering. More important for our purposes is the connection between these credences and the likelihood of error in the expansion of full belief. We must return to an earlier work of Levi in order to understand why he thinks propositions are the sorts of "things" to have probabilities in the first place. There is very little in this work that examines this notion more than in the respects I have already stated. We must thus turn to his *The Enterprise of Knowledge* in this light. However, before leaving *Fixation*, it is important to note that he does not discuss other notions of probability and thus does not deal with problems of reflexivity we have met previously in Carnap and Keynes. We do not find in this work any discussion about how to attribute credences to statements of the probabilistic sort in science or technology, and hence how to revise these in the manner the book discusses.

The Enterprise of Knowledge is one of Levi's early books on epistemology and thus (as we have seen) grounds many of what he has written since. It thus merits some investigation to cull what this influential thinker means by "probability." In this work, he distinguishes two uses of probability. One is the subjective notion, related to the ones we have seen previously (pp. 3). The other he prefers to call "chance" (pp. 230), and is a more objective notion. It is related to frequentist and propensity-ist notions we have met before/will meet later. Let us examine Levi's conception of each in turn.

Subjective probability is described as follows (pp. 3):

> "Judgements of subjective or credal probability are intimately related to evaluations of hypotheses with respect to serious possibility."

Levi then explains "serious possibility" should be read "serious possibility of truth. Thus, this usage of probability has domain whatever normally bears truth values. These are propositions (or perhaps sentences). We are then left with the usual questions: what are Levi's arguments for this viewpoint, and does he meet the possible objections to it? He bypasses the question about numerical representation on page 88, saying as much (i.e., Levi takes it for granted that credences can be given numerical values of the kinds he needs). Of course, there are a few postulates provide to obtain the required numerical representation, but their specific "quantitativeness" (i.e. that they must produce a range [0,1] etc.) is not argued for. For instance (pp. 111):

> "Let there be n hypotheses exclusive and exhaustive relative to K. Inductive logic should not mandate assigning equal Q values of 1/n to all n hypotheses; but it should not forbid a credal state which does so."

Levi also says that he is not a strict Bayesian (pp. 152), showing how with some arguments from considerations of bookies, so we cannot necessarily make even use of the typical Bayesian arguments for the issues we are discussing. An interesting feature of his discussion of bookies, however, lies in its waffling between the events described by a sentence (or proposition) and the events themselves. Does one bet on the event or the (say) truth of the proposition? This is not made clear. There is to be a connection between these two, and this provides some of the argument in favour of attributing probability to propositions. He feels the connection is necessary in order to have some sort of "epistemic consistency." What exactly this is again not spelled out. We are told, however, that this consistency is enough to ground Levi's views (pp. 261):

> "I see no reason for searching for some rationale addition to this."

A bit later (pp. 262), he runs a "dutch book"-like argument to show some features of credal consistency. But having correct credal states, according to Levi, requires knowing that certain features of a situation are stochastically irrelevant (pp. 266). Thus there is something objective in his view. Still, this does not say that one should have credence in the sentence describing an event to the same degree as the propensity of the event it describes to occur.

Confounding evidence to his own view as a whole is not discussed much by Levi, alas. Like others, his work is "inference to the best explanation. But, Levi is more frank than others concerning the applicability of his account. Consider the admission that he doesn't even know the extent of its applicability (pp. 214):

> "I am far from prepared to claim that agents do actually satisfy the requirements of my theory with any great regularity. I simply do not know."

Other authors have not considered the "factual adequacy" angle very much, so Levi's admission is at least to be praised on that account. But, the admission (to some extent) calls into question the views he has developed previously.

Our final consideration with Levi is to simply note he has anti-frequentist arguments. These are relatively standard (as we shall see below with other work in this area) and need not be rehearsed. He targets von Mises specifically. I thus leave Levi and move on to the next philosopher to consider.

Van Fraassen has also written extensively on matters pertaining to interpretations of probability. Here, I investigate several of his works which discuss the foundational question I am addressing. In *Laws and Symmetry* (1989) he discusses the connections between two views of probability (pp. 82):

> "Probability has two faces. On its subjective side, probability is the structure of opinion. But, when physics today tells us the probability of decay of a radium atom - for example - it does not in the first instance purport to say something about opinion, or to give advice, but to describe a fact of nature. This fact being a probability, we are looking upon probability's objective side - physical probability, or objective chance.
>
> There must be a connection between the two. Given that the objective chance is thus and so, my opinion must follow suit, and I must align my expectations accordingly. This summary of the connection between the two is generally called Miller's Principle: ..."

According to van Fraassen, Miller's Principle is the principle that the credence towards a state (or perhaps an event) A ought to be the same as the objective probability of A obtaining (occurring). Van Fraassen finds this approach somewhat dubious. What is most important for our purposes is his approach to understanding what objective chance <u>is</u>. Van Fraassen favors a frequency interpretation, albeit one over the long history of the universe. He claims that this generates paradoxes, but does not consider the possibility that the frequency assumption (rather than, as he believes, assumptions about laws) is the erroneous one. Obviously, this is not the place to evaluate claims about lawfulness, but it is important to note the possible oversight anyway.

Much later in his book, van Fraassen motivates the version of the subjective interpretation he favors by starting off with qualitative expressions of credence towards certain propositions. From there he moves to ordinal accounts (high, low, etc.) via means of phrases such as "my probability for rain is very high." He then introduces a scale from zero to one to render such judgments in numerical terms. No justification for this particular scale is given. He does, however, point out that he is not supposing that people have in their heads a numerically precise probability function. This immediately prompts the question: if that is so, how precise is it? Things in the world do not have a precise length (due to thermal fluctuations and so forth), but they do to a great approximation, and so the usefulness of assuming they do have such a precise length is demonstrated. By analogy, van Fraassen owes the reader a demonstration of the precision of the subjective probability functions. Our first indication is on page 157, during his discussion of calibration. This notion is illuminated by an example of someone who is said to be perfectly calibrated. Someone who is perfectly calibrated is someone who's credence exactly matches the relevant frequencies. All of this culminates in the usual "Dutch Book" (DB) argument. Again, however, there is no indication Van Fraassen is aware that the DB argument cannot <u>ground</u> the notions he wants, only (at best) supply properties of them.

He does show that this approach, with conditioning, can converge to an appropriate frequency in the case of a biased die. This is a strange special case to pick; surely we would want instead an illustration of how credences can be revised in the light of information of many kinds. He also does not address the question of initial credence. Ought one have a non-zero initial credence towards any proposition so that one can successfully conditionalize using it later? But, Van Fraassen does not see as important the need for grounding the origin of our beliefs rather than grounding their changes (pp. 171), so this possible oversight somewhat understandable (or at least, internally consistent). It is an open question, however, whether this is a good move in the scientific context. His "pragmatist" viewpoint can perhaps be applied to cases of every day life where our typical experiences do make it likely we are roughly correct much of the time. But, this is surely not the case with science. We are owed an argument showing the continuity of common sense and experience with science in this respect.

More of Van Fraassen's views on the nature of probability can be found later in this same work in his discussions of probabilistic symmetries. Here, Van Fraassen takes pains to show that his understanding of probability is not the "logical" one developed by Carnap and others. The essential difference lies in his approach to symmetry arguments and the so called "principle of indifference." Appeals to deeper symmetries allow him to develop "principle of indifference"-like principles for assigning initial probabilities to hypotheses (and to expectations of frequencies!) but Van Fraassen does not consider the possibility that one could again find ways to exploit his characterization in way analogous to the original principle. Even more curious is his discussion of the Buffon needle dropping problem. He shows how indifference and symmetry yield the conclusion that a relevant probability is $1/\pi$. He then claims that this result has been vindicated by experiment. But, how does a frequentist understanding allow for that finding? No countable number of trials would ever yield this value. It is this <u>impossible</u> (even in principle) for one to be perfectly calibrated in cases like these. More on this consideration below, in the discussion of probability in physics.

Van Fraassen's use of probabilities and elucidation thereof is not limited to *Laws and Symmetry*. He has written several articles on epistemology which are also "probabilist" in flavour. Let us take a brief look at one these to see whether they shed any light on the issues of this paper. In "Fine-Grained Opinion, Probability and the Logic of Full Belief" (1995) we find a frank statement of some of his views. The probability of a proposition is a real number in the unit interval. The self-contradictory proposition receives probability 0 and (the) tautology 1. Alas these are taken for granted and not defended in any way except implicitly by the possible fruits of the paper itself and any related works.

These considerations also raise an interesting question about probabilities associated with propositions. As have seen, Van Fraassen sets the probability of the tautology to 1. However, which propositions are tautologous is relative to a logical system. For instance, in intuitionistic logic, propositions of the form $P \vee \neg P$ are <u>not</u> tautological as they are in classical logic. So, what is $Pr(P \vee \neg P)$? This creates problems for indifference and symmetry arguments of the kind we meet later. I shall not return to this, as Van Fraassen (nor anyone else that I am aware of) has not considered this possibility. It is important to realize that merely stating that probability theory presupposes classical logic will not do here. This objection does not carry weight because there is a difference between the logic used in constructing the theory and the logic in the language that is supposedly the domain of the probability functions in question. One would need an argument to show that these are necessarily the same. (Further, if pressed, it might be possible to develop probability theory from an intuitionistic mathematics anyway.)

We do, however, find a discussion of continuous magnitudes and probability on pages 350-351 of this article. This consideration is of what he calls a "transfinite lottery paradox"; it and a discussion of full belief (i.e. should we have credences in contingent statements as high as the credence towards the tautology) lead him to a discussion of two place probability. In the course of this discussion, Van Fraassen concludes that absolute probability (the one place probability we have been considering so far and for the most part elsewhere in this paper) ought to be regarded as a derivative notion, with two place probability as basic. This is in opposition to the usual practice as one might guess by looking at the samples of other work elsewhere in the current paper. It is thus, a rather large price to pay to avoid the paradoxes Van Fraassen has been discussing. This development does allow him to elucidate some traditional epistemological ideas, such as one of the personally *a priori* (pp. 355 ff.). These are those propositions that are not epistemically different from the tautology. If one regards this as useful notion to elucidate, this result thus lends support to the merits of the kind of probabilism Van Fraassen espouses. Of course, the converse is true too. No further developments in this paper are much on our theme.

I thus turn to Richard Jeffrey's understanding of probability. I examine his paper "From Logical Empiricism to Radical Probabilism" (1993). He explains that probabilities apply to propositions, and they are expressions of judgments along the lines of the Cartesian affirmation and denial. He points out that whether we take the interval of probabilities to be $[0,\infty]$ or $[0,1]$ makes no difference in the end. The first amounts to an "odds" presentation. I.e., the probability ∞ amounts to infinite odds. As he points out in a footnote (#2, pp. 129), one can map probability (p) as ordinarily understood onto odds (o) by $o = p/(1-p)$. Curiously, he recognizes that the gambling approach to understanding probability only works for some propositions (pp. 121):

> "Probability 1 corresponds to infinite odds, 1:0. That's a reason for thinking in terms of odds: to remember how monumentous it may be to assign probability 1 to a hypothesis. It means you'd stake your all on its truth, if it is the sort of hypothesis you can stake things on."

Since some hypotheses one cannot (or will not) stake things on, do they not thereby have a probability? Or can they just not have probability one? Here, it is not clear what Jeffrey means.

Jeffrey calls his position "radical probabilism" and contrasts his view to the "rationalistic Bayesianism" he claims Keynes, Carnap and others have adopted. Jeffrey's position is said to be a synthesis (of a sort) between rationalistic approaches and empiricist ones. This is important for our present concern, as it tells us the possible ground for his attribution of probabilities - a blend (as he might have said), rather than a sum, of experience and reason. A merit of his approach, or so he says, is that it allows the agent to have other options for communicating influences of experience, such as by so-called "Bayes factors." This possible extension to the "standard" Bayesian (subjectivist, domain is propositions) approach is supposed to show by a sort of "judge it by its fruits argument" the merits of Jeffrey's views.

Throughout the above discussion, we have seen that there is some understanding that probability is primarily to be understood in one of two families of ways. However, there is a third conception in the literature that has gained some currency. This is what K. Popper, M. Bunge, and D. H. Mellor call "propensity." Mellor's (1995) work is largely based on Popper's in the respects the current work is concerned with and in the interest of manageability of the current work, I shall leave it aside.

Popper was apparently the first to popularize[4] the propensity interpretation in his influential *The Logic of Scientific Discovery* (1999 [1959]). In this work, Popper considers three notions of probability. One is the propensity interpretation. He suggests this is an elucidation of the pretheoretic idea of "degree of randomness." This terminology is not exactly satisfactory. To say that a toss of an ace on a six sided die has a degree of randomness 1/6, say, does not seem too felicitous - it is unclear what "degree of randomness" is supposed to mean. He argues for the propensity interpretation in several ways. One is by showing that it explains the frequentist's intuition about frequency of random events. For instance, if I consider the class of events of tossing a fair six sided die, this class has a given distribution that is increasingly likely to be near a uniform distribution as the class increases in size. Popper also argues against the notion of probability as degree of confirmation by showing that it leads to a paradox (pp. 391).

Bunge took up the propensity interpretation and has spent some

[4] Bunge (1996) attributes the invention of this view to Poisson. I have not taken the time to verify this.

time defending it. Bunge's argument for the propensity interpretation is largely by exclusion. In other words, he attempts to show the credence approach and the frequency approach are mistaken, and thus show that the third approach is thereby the correct one by default. I shall draw upon his remarks in two relatively recent works (1998, 1996); his earlier work (1983) makes similar points. Let us see how this goes. The first of his central claim is that any factual use of probability centers around the concept of randomness: as he puts it (1996, pp. 36):

> "No randomness, no probability."

This postulate needs some support; however, it does jibe well with our intuitive notion of at least some of the species of probability.

Another of Bunge's arguments references the well-known results of Tversky and Kahneman. As Bunge notes, Tversky and Kahneman's results tend to show that people's actual credences towards propositions as a matter of fact do not behave like probabilities.

Bunge thus takes Kahneman and Tversky's results as refuting the notion that degrees of belief are probabilistic. Note also that he considers the possible move to an "as if" (i.e. that agents act "as if" they "attempted to" maintain probabilities of propositions) position untenable, as it renders the hypothesis unfalsifiable. There are several possible rejoinders to Bunge's objections. An obvious one is an appeal to normativity. In other words, someone might claim that although it is descriptively false that credences obey the Kolmogorov axioms, these axioms would remain, though only as a normative standard. Bunge partially rejects this suggestion. He has no objection to people scientifically investigating the extent to which people succeed or fail to live up to this norm. However, he does not believe that one should take action based on considerations of subjective probabilities precisely because they are arbitrary. We have seen that although many thinkers have proposed rigorous standards for the change of probabilities, there are very few who propose adequate standards for the initial values of probability.

Is this latter point of Bunge's justified? We have seen above that there is very little attempt to do so. Some recent textbooks attempt to develop such standards. We have met one of these previously; Skyrms' *Choice and Chance* is one such. Another is the recent *Argument, Critical Thinking, Logic and the Fallacies* (Woods, Irvine and Walton 2000). These proceed in a similar way; Skyrms' book has the merit that it considers the domains of the appropriate probability functions. Alas, the justification for attributing numeric credences to propositions is by the Dutch Book argument (pp. 186-7). As I noted above in the discussion of Gärdenfors, the Dutch Book argument can only (at best[5]) ground the probabilistic interpretation if one thinks that propositions are the appropriate domain of a probability function. This is because the existence of the credences in question have to be postulated first. The Dutch Book argument does not say whether or not this should be the case. Skyrms does not show awareness of this point. Nor can the Dutch Book argument undermine the arbitrariness of selecting credences; it only tells one (again, at best) how one's credences should interrelate.

Further, Bunge has several arguments against the frequency interpretation. The first of these is a response to those who point out that we do not have access to propensities in an experiment. For instance, if we have a series of coin tosses that come out {H, H, T, T, H, T, H, T, T, T} we only have access to the two frequencies (H=4/10; T=6/10). Bunge points out that this approach relies on operationalism - the philosophy of science that ways of measuring just <u>are</u> the properties of something (e.g. that length is what is measured by a ruler, etc.). This criticism would obviously have no weight with someone who is a committed operationalist (Bunge does argue against operationalism generally elsewhere, but the current work is not the place to review these arguments). However, he has another argument (borrowed from Ville [1939] and Feigl) that the frequency interpretation of probability is incorrect because it makes illegitimate use of the notion of a

[5] In the interests of making this paper manageable, I am ignoring claims to the effect that the Dutch Book argument need not be accepted by rational agents (e.g., Kyburg 1991).

limit. This argument goes as follows: take any sequence of coin tosses: there is a definite proportion of heads and tails (say, 4/10 heads). This value does not approach a mathematical limit in the technical sense of limit. In fact, as time goes on it becomes increasing unlikely that the frequency of exactly 1/2 is observed. We may approach a "normal distribution" with greater and greater accuracy, but that is not the same claim.

Bunge is not (apparently) aware of Skyrms' attempt to resuscitate the frequency interpretation from this objection. Let us look at Skyrms' attempt (pp. 201-205, 212-213). He manages to show that <u>some</u> sequences of events do exhibit a long run limiting frequency, but correctly points out that some do not. He claims that the general solution is difficult and so does not deal with it, appealing to von Mises and Reichenbach.

It is also important to note that Skyrms is also aware of the propensity interpretation and agrees with Bunge, Popper, Mellor (etc.) that it is an interpretation useful in science. He suggests it is useful to formulate what he calls "laws of nature."

Do von Mises and Reichenbach deal with the "limit" objection I mentioned above? They do not seem to be aware of Ville's argument, although by the time of the english translations of their works, Ville's work was available. Reichenbach (1949) even references it!

I shall thus turn to von Mises' account of probability. His understanding is relatively straightforward. He believes that the only understanding of probability is one of limiting frequency of attributes (properties, or perhaps the events that produce said properties) in an infinite collective. "Collective" here is a technical term and is also used in Ville's account, as we shall see. Von Mises <u>does</u> bite the bullet and claim that he is using "limit" in the sense it is used in analysis. These considerations are expounded as follows (1957 [1928], pp. 15, italics in original):

> "We shall say that a collective is a mass phenomenon or a repetitive event, or simply, a long sequence of observations for which there are sufficient reasons to believe that the relative frequency of the observed attribute would tend to a fixed

> limit if the observations were indefinitely continued. This limit will be called *the probability of the attribute considered within the given collective*"

Thus, it is necessary for von Mises to show that there are such collectives. It is quite reasonable, as he does, to show that attributing probability outside of the context of a correct collective can lead to improper results. This occupies the next few pages of the book. On page 23, he introduces the concept of "randomness" which plays a large role in elucidating the appropriate notion of collective and hence of probability. In turn, the notion of place selection (pp. 24) is the key to understanding randomness. He writes (pp. 24-25, italics in original):

> "In this way we arrive at the following definition: A collective appropriate for the application of the theory of probability must fulfill two conditions. First, the relative frequencies of the attributes must possess limiting values. Second, these limiting values must remain the same in all partial sequences which may be selected from the original one in an arbitrary way. Of course, only such partial sequences can be taken into consideration as can be extended indefinitely, in the same way as the original sequence itself. Examples of this kind are, for instance, the partial sequences formed by all odd members of the original sequence, or by all members for which the place number in the sequence is the square of an integer, or a prime number, or number selected according to some other rule, whatever it may be. The only essential condition is that the question whether or not a certain member of the original sequence belongs to the selected partial sequence should be settled *independently of the result* of the corresponding observation, i.e., before anything is known about this result."

Von Mises then says that the limiting values of the relative frequencies in a collection must be independent of all possible place selections. With this on the table, he can finally tackle the question of the existence of collectives with the properties he needs. We are assured that there are, he says, by experimental results. He points out that numerous people have gone to Monte Carlo with a "system" and nevertheless gone home poorer. The existence of collectives of his kind he says is demonstrated in the same way that the law of conservation of energy is demonstrated - by direct induction from many cases. He claims that just as there is no deductive argument to conservation of energy, there is no deductive argument per se for the existence of collectives. In some sense this is correct - one cannot conclusively show that some thing or some property exists from reason alone. But, one can show that under certain plausible assumptions it will. It <u>is</u> unfortunate that von Mises selected a conservation law, however, as these can be shown to be equivalent to other principles that one might have postulated that are certainly transempirical[6]. Perhaps, then, it might have been possible to use a similar approach in the context of von Mises' concerns.

Reichenbach's approach is rather similar (in fact, he claims that von Mises defended the frequentist approach against several objections). His approach, however, does not make use of a randomness postulate. His definition of probability is then (1949, pp. 69, originally in italics):

> "If for a sequence pair x¡y¡ the relative frequency $F^n(A,B)$ goes towards a limit p for $n \rightarrow \infty$ the limit p is called the probability from A to B within the sequence pair."

Thus, to understand Reichenbach's definition we must first understand sequence pair. A sequence pair is simply a pair of outcomes one from one class and one from another. More critical for the purposes of the present paper is that Reichenbach does mean "limit" in its sense in analysis. Specifically, we must analyze his arguments for the existence of this limit. On page 70 of his book, Reichenbach correctly points out that we don't know whether or not this limit exists in part because we always only have access to a finite initial sequence of the total infinite sequence needed. He regards this extrapolation procedure to be a (perhaps, the only) species of inductive inference. He also suggests that we can know that this probability exists if we have access to a defining equation of the sequence. This is odd on two grounds. First, it seems to defy the common intuitions about probability (about uncertainty, for instance). Second, it clashes immediately with views of other frequentists such as von Mises, because von Mises' views concerning place selection rule this out. (Unless one could <u>only</u> obtain a defining equation of a sequence of trials after the fact - place selection does not work *post facto*.)

[6] Stenger (2000) draws attention to the results of Noether which show that (e.g.) the law of conservation of energy is equivalent to time translation symmetry.

This disagreement in itself is not a problem, but it is difficult to square with Skyrms' appeal to both of them.

Later, Reichenbach does discuss the inductive procedure in question. He assumes (pp. 445) that the sequence in question does have a limit of the appropriate sort. Here, we are told that one, in absence of the information to the contrary, assume that the rest of the sequence contains the same frequencies as the initial part one has observed. This is not quite as odd as it sounds; Reichenbach allows "cross inductions" and similar procedures to further refine the posited probability. One also is told to use the narrowest reference class available, though it is not clear how one is to this. This has intuitive plausibility, but no guidelines on how to use it are discussed in any great detail.

These methods are refined in discussions of "advanced knowledge" where we meet Fisher's methods, etc. These require knowledge of a distribution of sorts. This is interesting, because it runs into the issue I have raised concerning transcendental numbers and the frequency interpretation.

A final note concerning Reichenbach is in order. Reichenbach asserts that we attribute probabilities to sentences, though does not give much in the way of supporting argument. He does argue against Keynes' understanding of this viewpoint. Reichenbach claims that Keynes' rejection of "events" was a mistake. He believes one needs them in order to attribute probability to sentences. Reichenbach's analysis of the probability of sentences is relatively simple: if one has a frequency of events statement, one can write out a directly corresponding sentence with the same probability. For example, suppose one has induced (as Reichenbach would say) that the probability that a die will throw a 6 is 1/6. Then "This die will throw a 6." has probability 1/6 according to Reichenbach.

So next we must analyze Ville's treatment of the issue of frequency to see whether von Mises and Reichenbach's attempt is successful. I turn to this now. Summarizing Ville's argument is a bit of a challenge, since his book can be regarded as a sustained argument against it. I collect a few points to consider: Ville proves the following (pp. 54, italics removed):

> "Étant donné un système S, dénombrable, de sélections, si l'on procède à une suite de tirage indépendants puvant donner lieu à l'apparition d'un événement E avec probabilité (au sens classique) constante et égale à 1 pour le résultat de cette suite dirages soit un collectif relativement à S, collectif dans lequel la probabilité (au sens de la théorie des collectifs) de l'événement E est égale à p."

I next quote from his summaries (pp. 132, italics removed):

> "Il y a contradiction, dans la théorie des collectifs, à supposer que tous les événements, probabilisables dans la théorie classique, que l'on peut associer à une même variable aléatoire, on en même temps (dans un même collectif) une probabilité bien définie."

He also asks the question von Mises appealed to the world to answer, concerning the notion of collective. He (pp. 133) does not find this approach very satisfactory. He points out that (in my paraphrase) it amounts to question begging, because we need the notion of collective to evaluate probability claims in von Mises' account anyway. Ville's argument is that von Mises has assumed that probabilities are thus and so, but we need that assumption in order to test the account.

Ville's argument concludes as follows (pp. 140):

> "... le passage à la limite au sens de l'analyse ne peut pas être utlisé pour définir la probabilité, à moins de se garder contre les contradictions par un luxe de précautions nullement imposées par la nature de la question (définition du collectif relative à un system dénombrable de sélections, exclusions, dans le cas d'une variable aléatoire, des événements correspondant a des ensembles non mesurables au sens de Jordan) dont la seule justification est précisément d'empêcher la contradiction."

It is also important to note that Ville also argued against subjectivist interpretations (apparently of all kinds, though this is not immediately clear). On page 16 he writes:

> "Si, en effet un historien dit: j'estime à 0,9 la probabilité que Jules César ait effectivement débarqué en Grande-Bretagne, et que la découverte d'un document nouveau, irréfutable, montre que cet événement se soit réalisé, nous en déduisons que le jugement de l'historien était bon. Mais, en appréciant de cette manière, nous jugeons plutôt l'historien que son opinion."

He also targets Reichenbach explicitly on page 17 with the following:

> "Nous croyons alors qu'il n'y a plus aucun moyen de juger scientifiquement une telle opinion; il n'y a aucune raison de croire une personne qui n'a pas donné de preuves de sa clairvoyance."

Ville thus has provided detailed criticisms of frequentism (and a few of subjectivist interpretations) and is not often cited by other authors, which is unfortunate. Bunge, however, is mistaken <u>if</u> he believes that Ville supported the propensity interpretation. Ville's book is largely formal in character - which is interesting, because it shows that formal considerations can refute factual hypotheses.

<u>Section 3 - Scientific and Technological Uses of Probability</u>

In this section, I shall survey several scientific and technological uses of probability. I shall examine its use in genetics, in physics (both in statistical mechanics and quantum theory), and finally in computer science.

I shall start with genetics. In this discussion, I shall make use of simple, Mendelian cases to illustrate the use of the probability concept. It is assumed that more complicated cases in genetics retain the use of the same concept. (If they do not, I regard this as unfortunate metascientific flaw, but one I need not deal with in this paper.) A typical introductory textbook (Griffiths *et al* 1996, pp. 30-31, bold in original) explains the use of probability in this field as follows (after reviewing methods that do not use the theory of probability per se):

> "Application of simple statistical rules is the third method for calculating the probabilities (expected frequencies) of specific phenotypes or genotypes coming from a cross. The two statistical rules needed are the **product rule** and the **sum rule**, [...]**"**

How should we interpret this statement for the present purpose? There are two features of interest in the above characterization. One concerns the notion of "expected frequency." We must analyze this to see whether this has anything to do with any of the frequentist notions of probability we have met so far. The second feature of interest is the grounding of the product and the sum rules. One of our overarching themes in the current paper is how to apply probability. The justification for the sum and product rules of probability will provide an example of how this is done in an area of little controversy. I take each of these issues in turn.

"Expected frequency" due to its mere use of the word "frequency" conjures up impressions of the various frequentist understandings of probability. Appearances are deceiving, however. Although Reichenbach has claimed that frequentist interpretations of probability apply in single cases, it is not clear how to do in this case. What is the appropriate reference class? We need some genetic theory explaining how to do so. This is all the more critical when we remember that not all traits assort in the Mendelian manner. But, by the time we get our theory up and running, we can state it in non-frequentist terms. We shall see how to do this now in the discussion of the two rules.

The product rule is grounded on Mendel's second law (pp. 29):

> "During gamete formation the segregation of the alleles of one gene is independent of the segregation of the alleles of another gene."

Note that this grounding makes no use of frequencies. Instead, it talks about particular events occurring (with given propensities). Segregations are individual occurrences, not properties of collectives. (For sure, one can make a collective of all occurrences of a given kind, but that is not what is at issue here.) A similar fact can be noted about the grounding of the sum rule. Furthermore, these event-oriented descriptions also occur in the neutral statement of the probability rules in this text (pp. 31):

> "The product rule states that the probability of independent events occurring together is the product of the probabilities of the individual events."

Note again the lack of reference to frequencies and further the word occurring. This is also specified in individualistic (rather than in collective) terms. Now I move on to another example of the use of probability in the sciences.

The quantum theory is often stated to be the probabilistic theory *par excellence*. But, as we have seen already, there are uses of probability in other branches of science. Some might argue that since the quantum theory is ontologically very basic, its use of probability should be the most fundamental. We have already met reasons to suppose this is suspect. Fundamental or otherwise, however, quantum theory does exhibit a use of probability. Here, I must pause to emphasize a use of probability, as some of the foundational work in quantum mechanics centers around this notion. Since this is not the place to do foundations of quantum mechanics, I shall make due with a few remarks. First, one cannot make due with just looking at the *dicta* of scientists or philosophers on the matter. I shall provide an example where a scientist says one thing didactically and develops elucidations (in particular, equations or formal statements) which do not model his view stated in ordinary language. Second, a related concern is about the philosophical pronouncements of scientists on the content of their field. Let us take each of these in turn; once that it is done I shall have cleared the way for a greater understanding of the use of probability in quantum theory.

Feynman's elementary exposition of how probability is used in quantum mechanics can be found in his famous lectures on physics (Feynman *et al* 1963). Here, he shows how probability is a property of events or perhaps of things to undergo a particular event. As we have seen, there is not much difference between these two. However, Feynman's exposition clearly does not compute probabilities of sentences or propositions. It is a bit harder to see whether the probabilities in question are in the things or within us (i.e whether physics makes use of objective or subjective probabilities). There are no primitives in the equations that could possibly be referring to human beings[7], so it appears to be safe to say the quantum theory makes use of objective, propensity-interpretation randomness. Let us see how. Take a typical example of an equation in quantum theory (Feynman *et al* 1963: I, pp. 41-7):

$$P(E)=\alpha e^{\frac{-E}{kT}}$$

This equation states the probability of a harmonic oscillator possessing an energy E. The other terms are α, a proportionality constant, k is Boltzmann's constant, e the base of the natural logarithm, and T the absolute temperature. Temperatures and energies are not characteristics of an agent or his search procedure (i.e. Feynman betrays his own equations when he says that this is the chance of <u>finding that</u> a harmonic oscillator has a given energy.) As for whether it supports the frequency interpretations over the propensity interpretations, one can see that it does not appeal to the notion of error, or "within a certain limit," or any of the catch phrases of the frequency interpretation's school. It does <u>not</u> appeal to the limiting value of a collective. (It <u>refers</u> to a class of events, that is certain, but the probabilities in question are of individual events.) Further, there is no residual "within a certain fluctuation ε" or the like: the probability in question is as precise a value as the other properties involved - i.e, not:

$$P(E)=\alpha e^{\frac{-E}{kT}}\pm\varepsilon$$

I repeat (again, to forestall any objections): this consideration may not appeal to the operationalist. There is also the question of whether the notion of frequency is well defined in a potentially infinite class of cases. Ville (1939) has defended this in certain special cases, but as we have seen he has also shown that this does not thereby <u>identify</u> the probability with the

[7] This argument against the subjectivistic (mis)interpretation of quantum mechanics can be found in greater detail in two articles in Bunge, ed. 1967. Bunge's own article (Bunge 1967) and Popper's (Popper 1967) are on this subject. I do not have time to further develop such an argument here: it is simply important to note for the present paper that there are good reasons to suppose that quantum theory does not make use of subjective notions of probability in any form.

frequency.

Despite the above sorts of considerations, the conclusion that quantum physics uses objective propensities has been recently disputed by Auyang (1995). Her argument, in essence, is that she cannot conceive of the notion of probability as a property of a thing or a system, and thus, it seems safer to her (pp. 85) to interpret the probabilities in quantum theory as being frequencies. This does not appear to work for all the usual reasons we have seen; it has the further complication that one can calculate probabilities of singular events in quantum mechanics; Auyang (and Feynman) give some information on how. Adopting a frequency understanding is thus very strange. Let us turn to another branch of physics and see what conception of probability is used there.

It turns out that in statistical mechanics we can apply a similar argument as the one we have seen in the "quantum case" to the probabilistic equations in this field. For instance (Feynman et al 1963, I:43-3):

$$P(t) = e^{\frac{-t}{\tau}}$$

The above is the equation for the probability that a given molecule survives a time t without collision, where τ is the mean time between collisions. Again, these terms have nothing to do with humans or their faculties. The equation thus expresses an objective notion of probability.

Admittedly, we cannot survey all of physics for possible use of other notions of probability, but we can draw several general lessons from these two cases. One is that objective probability of events in the sense of propensity is at work here. If there are other uses to be found in physics and they are to cohere correctly with these uses, semantic bridges similar in spirit to those van Fraassen (see above, section 2) has tried to bridge between kinds of probability.

The above examples from physics also illustrate another possible argument against frequentism. This concerns the usual postulate (see above for von Mises' explicit adoption of it) that the event space is countable. If that is so, how does one understand certain probabilities calculable in physics. For instance, after a duration equal to one time constant has elapsed in the statistical mechanics case the probability that there has been no collision is:

$$P(t) = e^{\frac{-t}{\tau}} = e^{-\frac{\tau}{\tau}} = e^{-1}.$$

This value is a transcendental number; no countable number of events of the appropriate kind could yield it as a frequency. Yet, the quantitative notion of probability is indispensable to physics. von Mises claims that his understanding of "limit" is exactly as it is in understood in mechanics. But there, defining (say) velocity as a limiting value ($v =_{df} \frac{dx}{dt}$) works, because (we assume) continuity and the like - hence the limit in question is well defined. Further, it is not clear how one could have an appropriate credence to the related propositions - how does one (as a finite creature) deal with the infinite precision needed? Particularly, how does one properly condition (or the like) using it? Conditioning in this context also raises an interesting question: suppose that one obtains data to the effect that a probability of the above kind was, say, 0.367, rather than 0.367879441171442(...) as calculated. How does that result affect the <u>calculation</u>, and in particular, the other variables used in calculating it? The time constant, τ, "represents" the contribution of many other factors of the system. The "probabilities apply to propositions" schools owe us an answer to how to update our beliefs about them, in order that our probabilities match up, and our credences towards appropriate propositions in turn must be updated. Can this be done? I leave this as an open question.

In computer science, probability (by use of randomizers) is used in algorithm design. Clearly this use of probability is not subjectivistic in the sense that the algorithms' are not involved with credences towards propositions. They are also not subjectivistic in that the randomizer in question is not merely a matter of ignorance in <u>one</u> sense. As we have seen, there are two ways of understanding the notion of subjective probability. One concerns credence towards propositions; however, there is another kind which concerns what might be called "subjective chance." This comes to a head here, because of the usual charge that computers do seldom make use of "genuine" randomizers. One might ask, then, whether or not this is an objective notion of probability or a subjective notion. I claim that it is an objective notion. The usual argument that because we could predict the outcome of the algorithm if only we knew more (not much more in the case of some random number generators!) is a red herring. After all, predictability is an epistemological notion (one that relies on certain ontological features, to be sure). The issue is instead that the random number selected is independent of the data fed to the quicksort or the number to be tested for primality (for example). This invokes randomness in another sense. This understanding (i.e. as involving independence) of randomness is how Bunge (1996) shows the usefulness of developing probabilistic models in social science.

With that objection taken care of, we can now look at the remaining details of the appropriate domain of the probability function in question. A typical computerized random number generator works by first setting a new "seed" by division of an initial seed (often selected by using the number of milliseconds since system startup or another analogous value) by a large prime number. This value is then used to generate a series of numbers by division and taking appropriate moduli (to place the resulting random number into range.) This value is then used in a program in various ways. I illustrate this with the following pseudocode for the Miller-Rabin algorithm (adapted from Cormen *et al* 1996 [1990]):

```
Miller-Rabin (n,s)
for j = 1 to s
        a = Random (1, n-1)
        if Witness (a,n) then
                return COMPOSITE
return PRIME
```

We need not investigate the Witness procedure to appreciate the essential point here. The gist of the Miller-Rabin primality testing is to randomly look for "witnesses" to the compositeness of the specified number, n. If one is found, we know the number is composite. If none are found after s tries, the number is increasingly likely to be prime the larger s is. Again, there are no frequencies appealed to directly, and certainly no "logical" use of probability present in this use of the randomizer. And yet we call this a probabilistic prime number checker.

Note that this "level of analysis" is needed in order to satisfy at least some of the postulates for randomness that some of the views we have met needed. As remarked earlier, we need independence. Clearly the mere picking of a random seed is independent of that part of the program, but further we also have the desired (as noted above) independence from the actual use of the program. In the first part, the seed initialization, we have a <u>selection process</u> that suggests an event interpretation. This is typical of many scientific and technological uses of probability: even if randomness is not in the things themselves, the notion of objective probability can be still used if the randomness is found in an appropriate selection process. Finally, it is important (for those who would still deny the use of probability) that the probability calculus is used in the analysis of this algorithm. I close my brief considerations of use of probability in computer science with a discussion of this.

Cormen *et al* prove (Theorems 33.38, 33.39, pp. 842-3) that the probability of the algorithm producing incorrect output (i.e. reporting PRIME when the parameter n is actually composite) is at most 2^{-s} where (again) s is the number of times through the loop in the algorithm. They point out that this analysis is strictly speaking only correct if the number to tested is random and the use of probability to analyze the algorithm also depends on that. This selection process (as usual) is an event. If the number selected is "fixed", then probability doesn't apply. In the propensity interpretation, the event's occurring has a probability, so once an event occurs, no probability can reasonably be attributed to it.

In addition, many of the anti-proposition-probability arguments that came up in the examination of the cases in physics and elsewhere also apply here. We need not rehearse them.

Section 4 - Lessons and Conclusion

In this section, I draw seven general lessons.

The first of these is that many scientific and technological fields make use of an objective event propensity interpretation of probability. This suggests that this interpretation, which is heretofore unpopular, should be cultivated more by philosophers.

A second lesson to be drawn concerns the state of the argumentation on various sides. Simply, appealing to the early authors such as Keynes and Carnap do not settle the questions of interpretation.

Related to the previous consideration, an author wanting to make use of probability theory for some factual purpose should clearly specify what interpretation she has in mind and stick to it. Ontological heterogenity raises many puzzles and is poor scientific and philosophical practice.

Fourth, it seems that one possibility for confusion concerning the various interpretations lies in the range of the probability function. Many interesting functions have range [0,1]. It does not follow that they are elucidatable with the help of probability. For instance, in chemistry one sometimes deals with the quantity known as *mole fraction* (See Zumdahl 1993, pp. 500) This quantity has nothing to do with probability, despite its range being [0,1][8]. So our lesson could be: even if credences (for example) are analyzable into a range [0,1] it would not then follow that they could be analyzed with the help of the probability calculus. Perhaps some have thought the Dutch Book argument shows the necessity of this, but that only shows certain conditions to apply IF they are to behave like probabilities.

[8] The mole fraction of one species A in a solution is the ratio of the number of moles of A to the total number of moles of substance in the solution.

Also, as noted in the introduction section of this paper, a philosophical elucidation of probability ought to make use of the scientific (and technological) understanding of it. The third section of this paper suggests firmly that the propensity interpretation is underdeveloped. A future extension to the current work would examine claims that subjective probability is the way to go in social science. I have not had the time to investigate this beyond the limited amount I have discussed and so leave it for another time.

A related sixth point concerns the use probability in "inductive logic." None of the science and technology texts surveyed had any use for these notions of probability whatsoever. This tends to support Bunge's viewpoint about the merits (or demerits) of said interpretation, but more work is clearly indicated. (After all, it is possible that it is just an omission on the part of scientists *et al*.)

The seventh and final moral to draw from the considerations of this paper is that there is very much work to be done in the area of foundations of probability, given the lack of consensus we can observe in the works above. However, with works such as Skryrms, which do survey lots of interpretations, there is hope. Building something new out of the pieces I have put on the table here is next, but that is another project for another time.

Works Cited

Auyang, S. Y. 1995. *How is Quantum Field Theory Possible?* New York: Oxford University Press.

Bunge, M. (ed.) 1967. *Quantum Theory and Reality*. New York: Springer-Verlag.

Bunge, M. 1967. "A Ghost Free Axiomatization of Quantum Mechanics." In Bunge, M. (ed) 1967.

Bunge, M. 1983. *Treatise on Basic Philosophy*, Vol. 6: *Understanding the World.* Dordrecht: Reidel.

Bunge, M. 1996. *Finding Philosophy in Social Science.* New Haven: Yale University Press.

Bunge, M. 1998. *Philosophy of Science*. (2 vols). New Brunswick:

Transaction Publishers.
Carnap, R. 1950. *Logical Foundations of Probability*. Chicago: University of Chicago Press.
Cormen, T. H., Leiserson, C. E., Rivest, R. L. 1996 (1990). *Introduction to Algorithms*. Cambridge: MIT Press.
Feynman, R. P., Leighton, R. B., Sands, M. 1963. *The Feynman Lectures on Physics*. (3 vols.) Reading: Addison-Wesley.
Gärdenfors, P. 1988. *Knowledge in Flux*. Cambridge: MIT Press.
Griffiths, A. *et al.* 1996. *An Introduction to Genetic Analysis* (6e). New York: W. H Freeman and Company.
Jeffrey, R. 1993. "From Logical Empiricism to Radical Probabilism." In F. Stadler (ed.) *Scientific Philosophy: Origins and Developments*. pp. 121-130. Dortrecht: Kluwer.
Keynes, J. M. 1948 (1921). *A Treatise on Probability*. London: MacMillan and Company Limited.
Kyburg, H. 1991. "Normative and Descriptive Ideals." pp. 129-139 In Cummins, R. and Pollock, J. 1991. *Philosophy and AI: Essays at the Interface*. Cambridge: MIT Press.
Levi, I. 1980. *The Enterprise of Knowledge*. Cambridge: MIT Press.
Levi, I. 1986. *Hard Choices.* Cambridge: Cambridge University Press.
Levi, I. 1991. *The Fixation of Belief and its Undoing*. Cambridge: Cambridge University Press.
Mellor, D. H. 1995. *The Facts of Causation*. London: Routledge.
Popper, K. R. 1967. "Quantum Mechanics Without 'The Observer'" in Bunge, M. ed. 1967.
Popper, K. R. 1999 (1959). *The Logic of Scientific Discovery*. London: Routledge.
Reichenbach, H. 1949. *The Theory of Probability* (2e.) Los Angeles: University of California Press.
Skyrms, B. 1975. *Choice and Chance: An Introduction to Inductive Logic* (2e). Encino: Dickenson Publishing Company.
Stenger, V. 2000. *Timeless Reality: Simplicity, Symmetry and Multiple Universes*. Amherst: Prometheus Books.
Van Fraassen, B. 1989. *Laws and Symmetry*. Oxford: Clarendon Press.
Van Fraassen, B. 1995. "Fine-Grained Opinion, Probability and the Logic of Full Belief." Journal of Philosophical Logic: **24**: 349-377.
Von Mises, R. 1957 (1928). *Probability, Statistics and Truth*.

New York: MacMillan.
Ville, J. 1939. *Étude critique de la notion de collectif*. Paris: Gauthier-Villars.
Weisstein, E. W. 2001. *Eric Weisstein's World of Mathematics*. Internet - http://mathworld.wolfram.com/.
Woods, J., Irvine, A. & Walton, D. 2000. *Argument: Critical Thinking, Logic and the Fallacies*. Toronto: Prentice-Hall.
Zumdahl, S. 1993. *Chemistry* (3e). Lexington: D. C. Heath and Co.