Generalizations, Idealizations, and Induction

An Investigation of Probability Theory Through Historical Paradigms

Chelsea DeMark
Advisor: John Nolan, Department of Mathematics & Statistics
General University Honors
Spring 2012

ABSTRACT:

The purpose of this study is to investigate the historical foundations of probability theory in order to develop a comprehensive understanding of modern techniques employed in the field. To accomplish this objective, examinations were conducted concerning David Hume's philosophical problem of Induction and potential solutions via Verificationism or Falsificationism. Next, the study evolved into an exploration of early interpretations of probability theory. The interpretations; Classical, Frequency, Subjective, and Propensity, were subsequently evaluated for their validity under the criterions of meaning discussed. Upon establishing the history of probability theory, the study moved to an investigation of modern probability as developed from the frequency interpretation. It was found that the mathematical institution as a whole largely disregards the problem of induction, locally insofar as it is not applicable to the idealized theoretical nature of probability. The problem of induction is meant to be an exegesis of the scientific process; as modern notions of probability and statistical inference stand, they are tools for quantifying science rather than scientific methods themselves. Having overcome the limitations imposed by the problem of induction, the study then investigated modern methods and concepts utilized in probability, and the cohesiveness of these approaches. The frequency interpretation of probability as discussed in earlier sections of the paper serves ultimately as a unifying force. However, the study finds that an explicit definition of probability is not necessary for or helpful to the success of the discipline, and as such objectivity is able to prevail via a subjective choice in definition, as is demonstrated in the paper's conclusion.

TABLE OF CONTENTS

I. Introduction

This essay will employ the historical foundations of probability theory in order to gain a comprehensive understanding of modern techniques employed by the field. To begin, an exploration of the problem of induction as posited by David Hume is necessary, to be followed by an investigation of the theories for which this problem acted as a catalyst for development. Once we understand the tenets of these so-called criterions for meaning in science, we will have the necessary foundations for examining probability interpretations and tools, such as distributions and hypothesis testing.

II. The problem of induction

There are two objects of human reasoning, relations of ideas and matters of fact. The former concern those *a priori* notions that are either intuitively or demonstratively certain, like algebra and arithmetic. When deliberating on ideas of this nature, we start with a set of undeniable premises and move forward deductively, every step justified in a rigorous manner such that no doubt can be maintained concerning the truth of an idea. The latter type objects, matters of fact, are verified only by human experience and examination- that the sun arose this morning and that the sun will rise again tomorrow are examples. 1748 saw the publication of *An Enquiry Concerning Human Understanding,* in which David Hume questioned the rational basis by which inductive generalizations about matters of fact are justified:

> These two propositions are far from being the same, *I have found that such an object has always been attended with such an effect,* and, *I foresee, that other objects, which are, in appearance, similar, will be attended with similar effects.* I shall allow, if you please, that the one proposition may justly be inferred from the other: I know in fact, that it always is inferred. But if you insist, that the

> inference is made by a chain of reasoning, I desire you to produce that reasoning. The connexion between these propositions is not intuitive. There is required a medium, which may enable the mind to draw such an inference, if indeed it be drawn by reasoning and argument. What that medium is, I must confess, passes my comprehension; and it is incumbent on those to produce it, who assert, that it really exists, and is the origin of all our conclusions concerning matter of fact. (Hume 114)

Induction is an amplitative form of logic that takes premises accepted to be true (that the sun rose above the horizon this morning and every day previous, for as long as humankind has been around to observe such an event) and derives a conclusion that has content beyond that contained within said premises (that the sun will rise tomorrow and will rise every day in the foreseeable future); induction is what is required to jump from the first of Hume's propositions to the second. The issue therein is that arguments using induction are not necessarily truth preserving-- they may, for all we know, have false conclusions.

Hume claimed that the leap required for relating the unobserved future to the observed past made by inductive arguments is based on belief in a cause-effect relationship. The supposed power by which cause brings about effect is deemed to be a causal connection, something that is unobservable, despite the fact that all notions of cause and effect are derived *a posteriori*.[1] Hume's argument concludes that such a causal connection does not exist in actuality, but rather is projected onto sequences of events by human perception. This projection may attain varying degrees of validity based on situational uniformity. By considering past observations, a cause is determined to produce an effect when the latter is consistently observed to follow the former. It is thus implicitly assumed in statements concerning matters of fact that nature is uniform across

[1] Justification dependent on empirical evidence & existence

[2] Author's emphasis.

[3] The conclusion exactly opposite of that which an inductive inference would posit. Given the

all dimensions. This assumption, however, is unfounded; without logical contradiction, it is possible to imagine a future in which a cause brings about a different effect than experience has led us to expect (that the sun does not rise tomorrow), in which case the prior success of human projections regarding causality are obsolete. Science and probability rely upon induction for their advancement, as they take the conclusions of past experiments and predict that, under sufficiently identical circumstances, those same conclusions will be drawn in the future. Hume's argument however is that, "we have no logical basis for placing *any* confidence in *any* scientific prediction. From this moment on, for all we can know *every* scientific prediction might fail" (Salmon, 58).[2] He even goes so far as to assert that the declaration of scientific predictions as probable is unacceptable because uniformity in nature cannot be assumed; it is therefore possible that the predictive success that science and other fields of human prophecy have enjoyed in the past will not continue in the future.

III. Answers to the Problem of Induction

a. Verificationism of the Logical Empiricists

Almost two centuries later, philosophers and mathematicians alike were still enamored by the implications of the "rational skepticism" identified by Hume concerning the operations of scientific enterprise. The logical empiricists (also known as logical positivists) of the Vienna Circle made it their purpose to generate a solution of sorts in the late 1930's. The verificationist concept of science, brainchild of these endeavors, is largely based on the Verifiability

---

[2] Author's emphasis.

Principle of Meaning (also known as the Verificationist Criterion of Meaning). This principle states that a claim made in science is meaningful if and only if it is grammatically correct, *and* either empirically verifiable *or* analytically true. By "empirically verifiable," it is meant that there is some way to determine that the given statement is true or false based on observation. "Analytically true," on the other hand, requires the claim under investigation be valid in virtue of the definitions understood within its premises. Mathematics involved with algebra or calculus are wholly comprised of statements of this nature, but scientific endeavors do not often utilize them except to generate hypotheses for further inference.

David Hume was undeniably in favor of the empiricist camp, as is obvious from some of the stances he defends in *An Enquiry Concerning Human Understanding*. He speaks at one point of determining the validity of volumes we keep in our libraries, giving the test: "let us ask, does it contain any abstract reasoning concerning quantity or number? No. Does it contain any experimental reasoning concerning matter of fact and existence? No. Commit it then to the flames: for it can contain nothing but sophistry and illusion" (Hume, 211). Referencing Hume's conclusions on empiricist matters, verificationists agree that the only way to acquire meaningful knowledge that is not intuitively obvious is through observation: "arguments from experience are supposed to be derived entirely from sense and observation, through which we learn what has actually resulted from the operation of particular objects and can infer from this what their results will be in the future" (Hume, 20).

The element of the Verifiability Principle of Meaning that allows for the avoidance of the problem of induction is also the principle's greatest flaw. Under the principle, a claim that is currently considered false can be alternatively verified as true because, regardless of observations made in the past, it is possible that in the future a contrary observation will be made. Thus, the Verifiability Principle of Meaning makes every scientific law suspect to meaninglessness given that future observations could be contradictory. The typical rules of engagement for inductive inference, under such circumstance, are disregarded: all prior evidence does not point to the counterinductive[3] conclusion drawn (Salmon, 63). Therefore, the principle seems to create more problems for science than it conquers, as can be observed in the additional problem of its circularity. Applying the Verifiability Principle of Meaning to itself, the claim "something is meaningful if and only if it is empirically verifiable or analytically true" is realized to be an oxymoron—the statement itself cannot be empirically verified, nor is it analytically true under any rational definition of the term. There are no premises given in the principle for the justification of such a statement.

Going back to the empirical verifiability of a statement, the logical empiricists left the question of how the truth of a proposition is to be discovered in practice. Nature is wrought with uncertainty, but the verifiability principle of meaning makes no mention of the confidence we must have in our conclusions about the observable truth. It is certainly easy enough to take note that there are, at the moment, five books on the table, or that the coffee is 160 degrees

[3] The conclusion exactly opposite of that which an inductive inference would posit. Given the knowledge that the sun has risen every day throughout all of history, induction would say that the sun will rise again tomorrow. Counter induction would say that the sun will not rise tomorrow.

Fahrenheit. But in the latter case, how can we be positive that the thermometer we used to measure the temperature is properly calibrated? We could test it, but how accurate can our methods of verification be? The lack of specificity in verificationism occludes its acceptability.

b. The Falsificationism of Karl Popper

In an attempt to dissolve the problem of induction as posited by Hume, and in vehement reaction against the verificationist criterion of meaning championed by the logical empiricists of the Vienna Circle, Karl Popper introduced the falsificationist concept of science in the 1950's. He asserted that Hume "proved decisively that induction cannot be justified, [...] and that science, if it is to be a rational enterprise, must do without it,"[4] thus moving forward to find a new way of justifying scientific statements (Salmon 64). Popper believed that there is no logical method for the development of new ideas; there likely exists an irrational, strictly creative element in the thought process of those who construct the foundations and direction for a new train of thought. Thus, we cannot search for validation of scientific enterprise insofar as the initial development of most theories concerning scientific knowledge is irrelevant to their logical analysis. We must search, rather, for the vindication of such enterprise. To clarify terminology, Herbert Feigl acknowledged that when "we speak of 'justification' we may have reference to the legitimizing of a knowledge-claim; or else we may have in mind the justification of an action" (Feigl 674). The first of these circumstances describes justification via validation, showing that the principle in question is a derivation of other accepted, more basic principles. The second justification Feigl speaks of is via vindication, showing that the theory in question "serves some purpose for which it is designed" (Salmon 61).

Popper agreed with Hume that the total or complete confirmation of a scientific hypothesis is impossible- additionally offering that all theories to be considered scientific must first undergo a severe process of *justificatio actionis*[5], whereby the conclusions drawn from the hypothesis are tested for logical consistency, scientific character, potential fruitfulness[6], and durability under the demands of scientific practice. The designer/arbiter of a theory must aim to disprove the object of their creation with unfaltering sternness, not to coddle it. A theory is thus falsifiable if and only if there exists the logical possibility of the assertion it makes being proven false, the falsifiability of a statement being that which makes it scientific. Under this definition, scientific language is more stringent than that used by laymen, with the hope of creating a scientific tendency toward undeniable perspicuity. The conclusions drawn via the rigorous testing a hypothesis undergoes are then juxtaposed with accepted statements in science to deduce further predictions, which are deemed acceptable or unacceptable by use of practical tests and applications. If tests of these secondary predictions are satisfactory, then the theory has no reason to be discarded. Such "acceptance," or rather lack of refusal, may only be temporary—the propositions and predictions made can always be discredited and overthrown by future contradictory results. A theory that withstands "detailed and severe tests and is not superseded by another theory in the course of scientific progress [...] is

[5] This is Latin for justification of actions, or vindication. Opposed with *justificatio cognitionis,* meaning justification of knowledge, or validation.
[6] "Whether the theory would constitute a scientific advance should it survive the various tests" (Cover, 430).

corroborated" (Cover, 431). Conversely, if a prediction is found unacceptable with regard to testing, then it is falsified, making all or part of the theory from which it was drawn equivalently falsified. A question to be addressed later, which was not thoroughly undertaken by Popper, concerns how a prediction is deemed as acceptable or unacceptable. In science there are often no clear-cut lines; exact measurements are hard to come by via experimentation; and so with limited accuracy we must make room for error when judging the plausibility of predictions. Thus we must ask, is there some continuous or discrete interval for results corresponding to the sufficient accuracy of a prediction? Modern probability would say yes.

While Popper did not advocate for acceptance of theories, he did announce that the use of well-corroborated theories for the purposes of practical prediction is entirely rational. Scientific progress would be impractical, or even impossible, if it did not utilize vindicated theories in the development of new ones. Thus, at some point we must take for granted the contingent validity of the basis upon which science stands and assume it is not porous, solely for the purpose of progress. Concepts developed from corroborated theories which themselves become corroborated add to the acceptability of their basis. On the other hand, if a theory that builds upon a well-corroborated basis is falsified, this indicates to the scientist that either the new theory is flawed or the predictive success of the basis from whence the theory came is limited, and thus the basis itself is flawed.

That the entirety of scientific enterprise can be undermined by a single negative result is the "naïve falsificationist" aspect of Popper's paradigm. We cannot hope to conclusively falsify, without any doubt, a given theory, as this would be dogmatic. It is naïve, therefore, to assume that rejection and falsification are inconsolably intertwined. If falsificationism is to be justifiable with regard to science, absolute rejection cannot be maintained. A positive heuristic may be used in order to avoid this dilemma: scientific progress attained should devolve[7] until an error in reasoning is identified; otherwise the theory's entire history is at risk of being considered falsified.

c. The Sophisticated Falsificationism of Imre Lakatos

Imre Lakatos' philosophy developed out of a collection of failed theories. The theory of verificationisim developed by the logical positivists was ultimately rejected as flawed by the source of their inspiration, Wittgenstein. Thomas Kuhn, who also wrote extensively on the history of scientific revolutions, created a theory that attempted to explain paradigm shifts followed by periods of normal science, but which ultimately allowed for lapses into relativism (Blaug, 48).[8] The normative theory of Karl Popper, whose ideas most closely resemble those of Lakatos, fails in that it lacks historical explanatory power. Clearly the competing theories of Lakatos' time fail in some significant capacities, prompting him to lay out three essential propositions defining his sophisticated falsificationism, and thus a shift to a progressive research program. First, a new theory must have "excess empirical content" over existing theory and explain some "novel facts" that are either "improbable" or "forbidden" by the existing theory. Second, the explanatory power of the existing theory is contained within the new theory. Third, the greater explanatory power of the new theory is corroborated by

[7] As in backwards evolution, a reversal in progress until an error is identified.

[8] Quote: "As I read him, Lakatos is as much appalled by Kuhn's lapses into relativism as he is by Popper's ahistorical if not antihistorical standpoint"

empirical evidence (Lakatos '70, 116). In addition, Lakatos offers the idea of a hard core and protective belt to explain how science transitions through time. The hard core of a scientific research program (SRP) consists of those features that are essential. If the assumptions within the hard core are rejected, there is good reason to reject the SRP. The protective belt, on the other hand, is the extension of the hard core into other areas of interest. The concept of negative and positive heuristics helps to define the relationship between the hard core and protective belt, as best explained by Lakatos:

> While the negative heuristic specifies the 'hard core' of the programme, the positive heuristic consists of a partially articulated set of suggestion or hints on how to develop the 'refutable variants' of the research program, how to modify, sophisticate, the protective belt. (Lakatos '68, 171)

Lakatos' theory allows for consistent scientific progress since, "there is no falsification before the emergence of a better theory," which is to say that the successor of a theory can still have flaws, but must be better than the previous one in order to be worthy of implementation in the first place (Lakatos '70, 120).

IV. Probability Interpretations in relation to Criterions of Meaning

As plausible solutions to the problem of induction developed, it became necessary to evaluate relevant scientific practices with respect to these methods. As we will subsequently discuss, interpretations concerning the nature of probability and chance began to develop centuries before Hume originally discussed the problem of induction. As such, these interpretations were debated philosophically but their scientific acceptability remained undetermined. With the arrival of verificationism and falsificationism came the development of methods for judging these interpretations in relation to soundness, scientific meaning, and overall legitimacy. The neoteric discipline of probability theory did not fall explicitly under the header of mathematics since the objects under consideration were beyond relations of ideas. Thus statements made in the ever-developing interpretations required *a posteriori* reasoning and empirical verification. The synthetic nature of these statements made them subject to Hume's skepticism, making probability theory a prime candidate for investigation under criterions of meaning. At the same time, it seems the confirmation of statements made within science requires a form of quantification, an estimation of the likelihood that the statements are true which can only be spoken of coherently by probability. Answers to the problem of induction can thus help qualitatively guide the development of probability theory, and likewise a developed notion of probability can help quantify the level of meaning achieved by scientific statements under paradigms such as verificationism or falsificationism. For instance, when a statement or theory is not rejected following the tests required by falsificationism, the statement or theory is corroborated. Is it more likely then, to be true? And if so, how much more likely? These are questions that can be answered with the establishment of a normative probability, or guidelines thereof. With consideration of the previously discussed criteria of meaning, we may ask the following:

> What criteria are appropriate for assessing the cogency of a proposed interpretation of probability? Of course, an interpretation should be precise, unambiguous, non-circular, and use well-understood primitives. But those are really prescriptions for good philosophizing generally; what do we want from our interpretations *of probability*, specifically? (Hajek)

In response to the essential lattermost question posed in the *Stanford Encyclopedia of Philosophy*, we should require that an interpretation be *admissible* under some formal axioms, and also under the theorems deductively obtained from them. We will discuss a structure of this sort in section **V.b.: The Probability Calculus** below. Second, we should require the interpretation be *ascertainable* in that it should be calculable; there should be a method by which we can figure out what the values of probabilities under the interpretation are. Last, we should implore that the interpretation be *applicable,* since the demand for an understanding of chance and probability historically emanates from the desire to quantify empirical situations, to integrate an intellectual construct into the physical universe (W. Salmon, 64).

V. Early History of Probability

We now know that any useful concept of probability must obey certain basic rules and axioms; when giving interpretations of probability by specifying the relationship between mathematical theory and corresponding elements of reality, we are concerned with the meaning of statements made under such paradigms with respect to fundamental necessities. Additionally, we must be interested in the conditions necessary for determining the truth-value of these assertions. In our analysis of historical and modern interpretations, we will primarily identify admissibility by virtue of their satisfaction of these elementary conditions. Only after an interpretation has been vindicated of violations may we consider its philosophical tenability. To be clear, “there is no single formal system that is ‘probability’, but rather a host of such systems” (Hajek). The interpretations we are about to discuss each have their own footholds for legitimacy, but it must be understood that they are not universally, and some not even locally, applicable.

a. Gambling as the Catalyst of Probability Theory

An early desire to understand probability arose from gambling; people, notably Blaise Pascal and Pierre de Fermat, wanted to understand the nature of chance and be able to properly divide stakes. Their correspondence in 1654 utilized combinatorial methods in order to enumerate favorable and equally likely cases, thus developing the foundation of probability theory in (Hald,44). In the exchanges between these two famous mathematicians, "they discuss the value of a throw in a dice game in which the player undertakes to throw a 6 (i.e. at least one 6) in eight throws" (Hald, 55).

Let player I be the person hoping to obtain a 6, and let player 2 be their opponent. The value of a throw as described by Fermat is the amount of the stake that would be awarded player I should the game cease mid-play, before the total number of turns are actualized (i.e. 8 throws are made). This value is expectation of a throw, E(*a,b*), where *a* is the number of throws made thus far and *b* is the difference between *a* and the total number of throws *n* in the game, or in other words the number of throws separating player II from a win. If we let the stake in question equal 1, E(*a,b*) is also player I's probability of winning. Pascal computed $E(a,b)=\sum_{i=a}^{a+b-1}\binom{a+b-1}{i}\left(\frac{1}{2}\right)^{a+b-1}$. In modern terms, his problem would be solved using the cumulative distribution function (CDF) for a Binomial distribution with *n* trials and *p*=(1/6). Fermat alternatively described $E(a,b)=\sum_{i=0}^{b-1}\binom{a-1+i}{a-1}\left(\frac{1}{2}\right)^{a+i}$, which can be interpreted in modern terms as a

Negative Binomial distribution where *n* is the number of possible trials, and the desired number of successes is greater than or equal to 1, with $p=(1/6)$. These early methods for divvying stakes provided an expandable basis upon which later generations of probabilists were able to build a theory.

b. The Probability Calculus

The axiomatization of probability was not developed until the 20th century, when Andrey Kolmogorov published *The Foundations of the Theory of Probability* in 1933. However, many of the interpretations developed before the advent of these standards are, at the least, minimally compatible with the following axioms. If we understand E as the set of elementary events *x, y, z*… and the family *F* of subsets of E, whose members are chance events, then we have:

I. *F* is closed with respect to unions, intersections, and complements
II. *F* contains the set E
III. To each set A of *F*, a nonnegative real number P(A) is attached. This number P(A) is called the probability of the event A.
IV. P(E)=1
V. If A and B are disjoint [sets in *F*], $P(A \cup B) = P(A) + P(B)$

(Von Plato, 217)

How the probabilities P(A) in the above axioms are determined is dependent on the interpretation, or in modern notions distribution, under consideration. We will henceforth keep the Kolmogorov axioms in mind when discussing early notions of probability.

c. Classical Probability

Existing in a paradigm seemingly incompoissible[9] with that of the subjective theory to be subsequently discussed, classical probability instantiates an objective competitor in the trial for dominant logical perspicuity. Evolving into a substantial theory around 1810, the classical tradition in probability was championed for over a century and into the 1930's. After the development of calculus, the approaches posited by Pascal and Fermat were furthered- Pierre Simon de Laplace had developed the tools necessary to conceptualize the classical interpretation of probability associated with the questions posed by previous generations. He did so with the help of his principle of indifference, stating that given a set of elementary events, and no reason to prefer any event in the set to any other, each of these events may be labeled as equally possible. The classical interpretation thus states that the probability of an outcome is the ratio of favorable cases to the number of equally possible cases. Under this interpretation, the only necessary knowledge of probability can be gathered *a priori*.[10] There is no need for experiments to be performed; only the object of interest, be it a dice or a coin, requires evaluation in order that a correct probability statement be given.

While Laplace's developments in probability theory were exceptional for the time, the interpretation at hand has some irreconcilable flaws. In a situation with a complex physical nature beyond the understanding provided by the classical interpretation of probability, such as infinite sets of events, our knowledge of the

[9] In the Kuhnian sense of the term, meaning signifiers take on different signification within the two paradigms/languages, making it so they cannot be discussed coherently in a concurrent fashion.
[10] Justification independent of experience, inherent to statement.

situation is inhibited; we are thus unable to calculate a probability in the classical sense. Unless the human mind is a machine compatible with determinism, as this Laplacian conceptualization seems to infer, any a priori knowledge of certain complex circumstances would trigger infinite regress, taking probability beyond the scope of human capabilities. The notion of *deus ex machina*[11] is preconceived in Laplace's development of the theory. Additionally, the classical interpretation is not defined to be compatible with a set of events that does not have equally likely outcomes, a fundamental problem of which "Laplace was fully aware" (Salmon 74). An example of this issue lies in the probabilities associated with a loaded die or a weighted coin. The applications of classical probability put "undue emphasis on the assumption of equally possible events," and as we will see below, often result in inconsistencies (Von Mises, 79)

There are ambiguities involved concerning how to apply the principle of indifference that overshadow the issues that arise out of the overly simplistic nature of the classical interpretation. We are told that in a situation where we have no reason to consider one outcome more likely than another, we must declare all outcomes as equally possible. However, this instruction may have multiple interpretations in any given situation, and thus will consistently generate inconsistencies in probability values. An example of this situation can be observed in the "sloppy bartender," also known as Bertrand's Paradox: whenever a 3:1 martini is ordered, meaning 3 parts gin and 1 part vermouth, the bartender will mix a drink proportioned somewhere between 2:1 and 4:1. The principle of indifference

[11] Latin for "god as machine," implies determinism

assumes .5 probability that the drink will be between 2:1 and 3:1, and .5 probability between 3:1 and 4:1. However, interpreting the proportions as factions, the results are slightly different. A 2:1 martini has 20/60 vermouth and 4:1 has 12/60. The principle of indifference assumes .5 probability that the drink is between 20/60 and 16/60, and .5 probability that the drink is between 16/60 and 12/60. Comparing the two methods of inferring probability, we see that 16/60 is not equal to the 3:1 ratio we originally considered the half way mark between 2:1 and 4:1, which has 15/60 vermouth. Therefore, the results using different applications of the principle of indifference deliver inconsistencies.

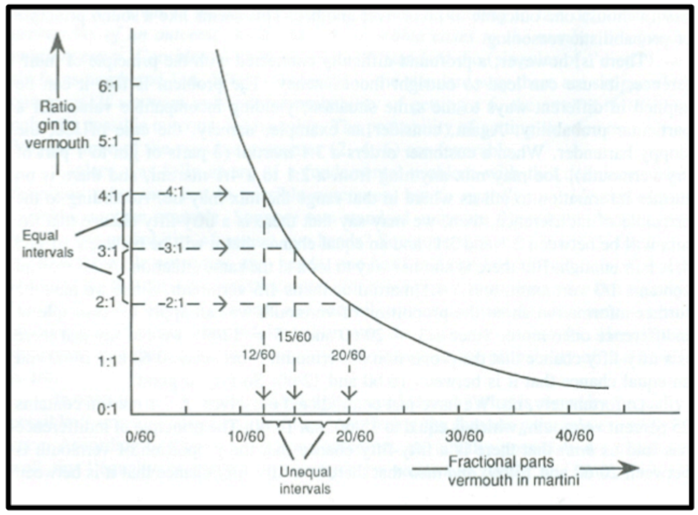

Figure 1: A graphical interpretation of the Sloppy Bartender example (M.H.Salmon, 76).

Despite the fact that the classical interpretation fails to satisfy the basic requirements an interpretation of probability should, we can glean some useful ideas from its tenets. The principle of indifference seems to be the source of much turbulence for the interpretation, but we can take away the concept of a ratio of favorable events to equally possible outcomes as a tool for future interpretations. As it stands, the capricious objectivity of the classical interpretation is not a useful mechanism for understanding chance. If there is going to be inconsistency in the values posited by a theory, it should not be as a result of *a priori* values.

Should the values be determined from a subjective standpoint, then it is all we can do to avoid inconsistency on the basis of the irrationality of individuals. Many mathematicians and philosophers found there to be a legitimate claim to a personal notion of probability, resulting from the fact that the proponents of objective theories struggled to validate their claims.

d. Frequentist Probability

The first great mind to conceptualize a version of the frequentist interpretation of probability was Aristotle, defining "the probable as that which generally happens: 'A probability is a generally approved proposition: what men know to happen or not to happen, to be or not to be, for the most part thus and thus'" (Madden, 167). Frequentist probability was incredibly influential in the development of modern probability, and "it was more or less tacitly assumed as *the* interpretation of probability by the main proponents of mathematical probability in the 1920's. No attention was paid to the idea of subjective probabilities" (Von Plato, 238). We will consider the infinite relative frequency interpretation of probability, which bridges mathematics and *a posteriori* experimentation. Theoretically, say the outcomes of a countably infinite ensemble of events are known by trials conducted. If $X_1, X_2, \dots, X_n$ represent *n* possible outcomes, the probability of an outcome $P(X_i)$ is the limiting frequency $f_i = \frac{y_i}{m}$ with which X_i appears in an ensemble of size m, known as the outcome's relative frequency. Thus, the limiting frequency of an outcome has value p if for any arbitrarily small number δ>0 there exists some finite number N=$N(\delta)$ such that for all m >N, the values of f_i do not deviate from p by more than δ. In symbols:

$$P(X_i) = \lim_{m\to\infty} f_i = \lim_{m\to\infty} \frac{y_i}{m}.$$

The relative frequency of an outcome within any randomly selected finite subensemble of size *m* is equal to the number of specific outcomes in the subensemble favorable to the desired outcome, divided by the total number *m* of outcomes in the subensemble[12]. The limiting frequency, as previously mentioned, is the ratio achieved as *m* approaches infinity.

A problem with this method for finding relative frequencies is choosing a subensemble out of the infinite ensemble, as there are infinitely many options. The choice must be randomly executed; many methods have been derived in order to simplify selection of random ensembles, like Exclusion of Successful Gambling Systems (ESGS) as proposed by Richard Von Mises. This tactic allows random ensembles to be formed using a rule dictating the selection of elements from the infinite ensemble. For example, from an ensemble consisting of the outcomes of rolling a 6-sided die infinitely many times, a subensemble may be constructed by selecting the values that come after 2 and 3 are rolled consecutively: given the leading string of terms of an infinite ensemble {1, 4, 6, 2, 2, 3, **6**, 3, 4, 5, 6, 2, 4, 5, 3, 2, 4, 6, 2, 3, **5**, 1, 1, 3, …} the first two values of the subensemble would be 6 and 5. If 100 distinct rules are utilized to construct 100 distinct[13] subensembles from one main infinite ensemble, the idea is that the relative frequencies of a certain element

[12] The finite relative frequency interpretation is a simplified version of this, and can be used when dealing with finite situations. The infinite variation is necessary in order to solve more complex probabilities.
[13] Although a subensemble may be constructed by the consecutive implementation of more than one rule.

will be approximately the same in each,[14] all representative of the infinite ensemble's relative frequency. This approach to selecting a random sample does not take into consideration conditional probabilities- how are we to ensure that our method for selection does not affect the selection itself (using the selection method above, how do we know that there is the same likelihood of throwing a 4 or a 6 after obtaining 2 and 3 consecutively?). To relate a finite version of frequentist probability and the ESGS to modern notions, we may think of the main ensemble as a population and the subensemble as a random sample. Using various rules to procure a number of random samples, the relative frequencies of a particular event in each sample should not differ to dramatically, from each other or from the population frequency.

We identify an additional problem regarding the frequentist interpretation of probability: it is subject to the problem of induction. Because the theory is based entirely on empirical evidence, it requires induction for projecting the past onto the future. Consider that we have no knowledge of the infinite ensemble's diversity past any empirically observable point; it is thus not logically inconsistent to presume that if we collected a subensemble from the tail end of our ensemble, it would produce a much different limiting frequency than that which is obtained via a subensemble derived from earlier portions.

e. Subjective Probability

i. A Multiplicity of Interpretations

[14] Assuming the subensembles are finite, yet large enough to provide an accurate sample.

Assuming nature is consistent when interpreting probabilities is useful, and some argue necessary when making claims in sciences that require an objective stance. This is often the case in the natural sciences and mathematics, in that we speak of idealizations and abstractions. Disciplines under the social sciences such as psychology, economics, and political science often demand an alternate approach that takes into account the specific details of unique circumstances. The scope of this uniqueness can be so narrow as to include only the beliefs of an individual, or so wide as to include the beliefs of a culture. The subjective side of probability "has something to do with the degree of conviction with which an individual believes in one proposition or another," the 'something' mentioned being indeterminate; constraints placed on the beliefs maintainable by individual are loose. Thus the undogmatic nature of subjective probability is conducive towards dialogical, democratic conversations (M.H.Salmon, 81). Disagreements between two individuals, groups, or parties regarding the probability of a situation "can be treated merely as a datum, rather than as a cause of conflict. […] We need not get hung up on deep and abstract (and possibly artificial) issues in order to proceed to use this theory of probability in testing hypothesis, in making statistical inferences, […] and so on" (Kyburg, 72).

Unlike the purely "physical" notion of probability utilized by the frequentist interpretation, subjective probability can be regarded under a multiplicity of constructions; it can be considered as a descriptive theory of decision-making, a descriptive theory of degrees of belief, a normative theory of decision-making, or a normative theory of degrees of belief. Normative theories relate to an imperative or ideal, instructing subscribers as to how they **ought** to behave under certain circumstances. Descriptive theories, on the other hand, describe things within the world as it currently exists. Theories concerning decision-making predict behavior under uncertainty, while those concerning degrees of belief speak to the dispositions of subjects towards taking certain bets.

Subjective approaches have been championed by prominent voices in the social sciences, but this philosophical popularity cannot speak to the validity and usefulness of the theory in practice. The probability calculus is the sole overarching mandate in the regulation of probability assignments, making it easy to assume agreement with other, more rigorous concepts of probability. However, the subjective dialectic allows for a wide range of interpretations of this calculus, and so there is ambiguity concerning what numbers under the theory represent depending on which of the four preceding constructions a theory falls into, if any at all. The branch of subjective probability associated with idealized, rational beliefs is known as epistemic probability.

ii. De Finetti and Epistemic Probability

Before Bruno de Finetti developed his version of epistemic probability in 1931, the notion of personal convictions having an effect on likelihood had been rendered practically obsolete by the dominating physical conceptions of probability. When de Finetti arrived on the scene, he introduced "subjectivism as a way out of an impasse into which he thought the classical epistemic notion of probability had been led," a way to escape the determinism of Laplace (Von Plato, 241). In order to do this, he developed a theory that restricted "the scope of uncertainty judgments to genuine observables," and used these measurements of exact empirical events to quantify a subjective uncertainty about unknown quantities, like the number of future successes (Dawid, 45).

An important rule in subjective probability is that a rational agent should avoid a combination of beliefs that would lead to inconsistency by aligning their beliefs with the probability calculus provided in the Kolmogorov axioms. The motivating force behind this requirement is known as a Dutch book, a set of bets that would guarantee a loss on the behalf of an agent should they decide to take them all. The Dutch book argument as formulated by F.P. Ramsey in 1926 purports that "rational agents will have partial beliefs that obey the probability calculus. [...] Having credences which violate the probability calculus ensures the existence of a Dutch book" (Eagle, 31).

De Finetti's position was that the only stipulation that should be enforced in determining the rationality of beliefs is that they satisfy the Dutch book argument. To further understand this position, consider the psychometric[15] study developed by de Finetti in conjunction with L.J. Savage: assume the agent in question is a miser desiring only money, a desire which does not change regardless of how rich or poor they may become. The agent is given a list of propositions $(X_1, \dots, X_n)$ and is offered $1 if they will designate a corresponding sequence of real numbers $(p_1, \dots, p_n)$, called the player's previsions. The agent, should they accept the deal, must repay a portion of their dollar once the truth-values of the X_i's have been revealed. The size of this forfeiture is determined by a predetermined scoring function S(p,w), where $w(X_i)$ is the truth value of X_i,

[15] Psychometrics is the field of study concerned with theory and technique involved in the psychological measurements of knowledge, abilities, attitudes, etc.

equal to either 0 or 1, at world w, where w is indexed by a set of possible worlds W. This function assigns repayment values up to $1 to each pair using quadratic-loss rules of the form:

$$S(p, w) = \sum_i \lambda_i [\mathrm{w}(\mathrm{X}_i) - \mathrm{p}_i]^2$$

Where $\lambda_1, \dots, \lambda_n$ are weights that sum to 1. If we consider the case where all the weights are equal, i.e. $\lambda_i = 1/n$ for all i, then this quadratic loss formula forces "any minimally rational [agent] to report previsions that obey the laws of probability, [and] they reveal the beliefs of expected utility maximizers because an [agent] who aims to maximize her *expected* payoff will invariably report a prevision for each proposition that coincides with her degree of belief for it" (Joyce 93). A lemma posited by de Finetti concerning the game, which formalizes the above statement, is as follows:

> In a prevision game scored by a quadratic-loss rule S, every prevision sequence p that violates the axioms of probability can be canonically associated with a sequence p^* that obeys the probability axioms and which *dominates* p in the sense that $S(p,w)>S(p^*,w)$ for all worlds w. (Joyce 93)

Thus, an agent whose sole concern is their love of money will always obey the probability calculus, because failing to do so would result in unduly throwing away money. The sequence of real numbers $(p_1, \dots, p_n)$ assigned by the agent are the degrees of belief, and if they are coherent then they act as probabilities.

With regard to the above example, it becomes clear that de Finetti's subjective probability is legitimate in the case where clear rules are identified. Critics of de Finetti claim that the Dutch book argument's pragmatism renders it irrelevant to epistemological probability theses. The argument presupposes that agents, rational or otherwise, have a practical reason for holding certain degrees of belief and will, out of prudence, stick to them regardless of the true state of the world. The notion of epistemic probability, however, implicitly concerns "the accuracy of the [agent's] opinions as representations of the world's state" (Joyce 95). Ralph Kennedy and Charles Chihara, notable critics of the Dutch book argument, claim that the real numbers assigned by the agent "cannot be degrees of belief since considerations relevant to judging the rationality of setting particular [betting quotients] in a [competitive betting situation] are different from those relevant to judging the rationality of having particular degrees of belief" (Davidson, 415). If we interpret the study above with regard to this argument, we see that the situation is idealized. The goal of the competitive situation and the associated betting quotients correspond exactly to the pragmatic demands of the agent in question- this is not always the case.

It seems as if the objection purported by Kennedy and Chihara is sustainable, but we may suggest as many probabilists have:

> Imprudence, while not *constitutive* of epistemic failings, often *reliably indicates* them. [...] It is reasonable to think that *systematic* deficiencies in practical reasoning *that do not depend on the truth or falsity of the reasoner's beliefs*, like the tendency of probabilistically inconsistent misers to throw away money, are symptoms of deeper flaws. (Joyce 96)

If this counter objection is valid, then the Dutch book argument is simply a device for demonstrating epistemic irrationality via pragmatism.

iii. Bayesian Confirmation Theory

Taking de Finetti's Dutch book argument to be an acceptable form of subjective probability, since it conforms to the Kolmogorov axioms, we look to the Bayesian prior probabilities that influenced the theory, and which are supported by it. Reverend Thomas Bayes was an English mathematician in the

1700's, famous for the theorem that bears his name; the branch of epistemology also bearing his name did not evolve until several centuries after his time, following the axiomatization of probability. Bayes' defines conditional probability as follows:

Given events A and B, the probability of event A occurring given that event B occurs, denoted $P(A|B)$, is given by:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

If we reverse this conditionalization, we find:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Solving for $P(A \cap B)$ we obtain:

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A)$$

A manipulation of this equivalence leads to Bayes' theorem, given as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

The model of Bayesian confirmation theory lets A in the above formulation represent a hypothesis, and B represent evidence. Thus, the evidence B confirms the hypothesis A if $P(A|B) > P(A)$, that is, if the probability of the hypothesis given the evidence is greater than the probability of the hypothesis alone. We should note that surprising evidence, given by a small P(B), confirms more than evidence that is easily forecasted, since a smaller denominator makes for a larger overall product.

To interpret confirmation subjectively, take P(A) to be the degree of belief in A, a hypothesis, before evidence is given- also known as A's prior probability. Then $P(A|B)$ is the posterior degree of belief accounting for the evidence, and $\frac{P(B|A)}{P(B)}$ is the support provided for the hypothesis by the evidence. To speak of

Bayesian confirmation in an objective manner, we must "emphasize the extent to which prior probabilities are rationally constrained;" if there is no information with regard to the likelihood of evidence, or rather its prior probability, then we must assume that all potential evidence is equiprobable (Hajek).

To exemplify the difference between subjective and objective Bayesian confirmation, say we are fishing in a pond filled with white and orange koi, but have no information regarding the distribution of white koi versus orange koi. Before actually catching any fish, the subjective version of the theory says that the designation of the prior probability of catching a white koi is unrestricted so long as it satisfies probabilistic coherence. The objective approach, however, evokes a version of the principle of indifference, claiming that since we have no knowledge of the diversity of the pond's koi population, we must assign a prior probability of ½ to both P(W) and P(O), the prior probabilities of catching a white or an orange koi, respectively. This assignment ensures that prior probabilities are invariant with regard to a change of label, be it W or O. The only restriction objectivists pose is this notion of invariance, since few objectivist Bayesians are extreme enough to claim that prior probabilities can be known a priori, as the principle of indifference would have us believe.

f. Propensity Probability

Like the frequentist interpretation of probability, the propensity view contrives an interpretation of probability that designates it as a physical notion- a disposition, tendency, or propensity of a certain type of physical situation to result in a certain type of outcome, "or to yield a long run relative frequency of such an outcome" (Hajek). While Charles Saunders Pierce was the first to conceptualize an elementary version of the propensity interpretation, it was fully articulated later by Karl Popper, the champion of falsificationism. Popper's motivation in defining the propensity view was to elucidate single-case probability attributions, the likes of which being observable in quantum mechanics. Under Popper's definition, "a probability *p* of an outcome of a certain type is a propensity of a repeatable experiment to produce outcomes of that type with limiting relative frequency *p*" (Hajek). This means that given a repeatable experiment, say rolling a die, we say that the die has a probability 1/6 of landing with the 3-side up if repetition of the experiment has a propensity to produce a sequence of results in which the limiting frequency of the 3-side is 1/6.

Adherents of the Propensity interpretation of probability fall into two camps, those who allow single-case propensities that do not require the use of limiting frequencies, and those who favor long-run propensities. We will discuss only the views of the latter group, noting that these probabilists do not allow for the experiment's theoretical repetitions to extend indefinitely, as this would subject the propensity interpretation to the same conceptual issues faced by the frequentists; those propensity probabilists who favor the single-case scenario do so in order to avoid a dependence on frequencies that could easily result in this exact problem.

Defined accordingly with respect to Popper's falsificationism, long run propensity probability claims, "a fair die has a propensity — an *extremely strong* tendency — to land '3' with long-run relative frequency 1/6. The small value of 1/6 does *not* measure this tendency" (Hajek). Long run propensity may thus seem inconclusive enough to skirt the problem of induction, yet the interpretation seems to make assumptions about uniformity in nature, which Hume adamantly advised against. Furthermore, we cannot consider it an admissible interpretation of probability because it does not fully satisfy the probability calculus, unless we choose to stipulate in our definition of propensities that they obey the probability calculus. Even so, when using this interpretation in conjunction with accepted probabilistic methods like Bayes' Rule implied by the axioms of the probability calculus, we observe inconsistencies. In *Philosophy of Science,* the authors speak of a factory that produces Frisbees from two machines, Old and New. The propensity of the Old machine to produce a Defective Frisbee is $P(D|O)=.02$, and the propensity of the New machine to produce a Defective Frisbee is $P(D|N)=.01$ (M.H.Salmon 80). It is sensible to think of these values under the propensity interpretation, because we can think of the 'repeatable experiment' as the creation of Frisbees by the New or Old machine. However, if we were to reverse this conditionalization to find either $P(O|D)$ or $P(N|D)$, conceptualization of these probabilities as propensities is inconsistent with reality. The defective Frisbee either came from the new machine or the old machine; there is a definitive and a posteriori element to this "probability" that is beyond the scope of the propensity view.

VI. Probability in the Modern Era

a. Development of Modern Probability

The advent of calculus in the late 17th century made room for a concrete foundation on which to develop modern notions of probability, which was taken advantage of in the creation of discrete probability distributions. Jakob (otherwise referenced as James or Jacques) Bernoulli and his brother John (otherwise referenced as Johann, or Jean), after the publication of Leibniz's papers on Calculus, began correspondence amongst both each other and with Leibniz himself whereby "they developed the rules of differentiation and integration in the form used today" (Hald 221). This fundamental understanding of calculus allowed Jakob Bernoulli to explore the problems associated with probability theory, which he addressed in his *Arts Conjectandi*, translated as *The Art of Conjecturing,* in 1713. It was here that he considered the probability of getting *m* successes and *n-m* failures in a series of *n* trials, each with the same probability of success *p*. If a specific ordering is under consideration, he found this probability to be $p^m q^{n-m}$, and if the ordering of successes and failures is arbitrary then he found that this probability becomes $\binom{n}{m} p^m q^{n-m}$ (Hald 227). In the case where *n* is equal to 1, the calculation is the outcome of a single Bernoulli trial. Throughout his text, Bernoulli continued to develop new tools for solving the problems associated with games of chance, finding the "enumeration of equally likely and favorable cases by combinatorial methods" as well as "calculation of expectations, making systematic use of conditional expectations, and recursion" (Hald 235).

Bernoulli experiments are fundamental to our understanding of probability today. We take to defining a Bernoulli distribution, and corresponding probability mass function (PMF) as follows:

> Let X be a random variable associated with a Bernoulli trial by defining it as follows: X(success)=1 and X(failure)=0.
> That is, the two outcomes, success and failure, are denoted by 1 and 0, respectively. The PMF of X can be written as
> $$f(x)=p^x(1-p)^{1-x}, \quad x = 0,1,$$
> and we say that X has a Bernoulli Distribution. (Hogg 78).

The definition of Bernoulli trials and distributions lead to the definition of a binomial distribution, which consists of a sequence of Bernoulli trials. With the binomial distribution as a tool, probability theory was built to encompass many other distributions, each of which can be considered to constitute its own “interpretation” of probability.

It is admissible to regard each distribution, and corresponding distribution function, as a separate, valid interpretation of probability insomuch as satisfaction of the Kolmogorov axioms is implicit in each of their definitions, and each with different ranges of applicability. Modern probability theory is thus, as a collection of these distributions, able to encompass a multiplicity of interpretations.

b. Cohesion Between Modern Approaches

The parameters that determine probability interpretations, like p in the binomial probability mass function, are not fully determined by the interpretation at hand. There are stipulations on what the parameter represents, but the actual associated value has a somewhat subjective element. In theoretical applications of probability theory, we assume parameters are known in order for computations to be executed. In other circumstances, parameters are estimated using various methods, like Maximum Likelihood Estimation, Method of Moments, or Least Squares Regression Estimation. Since probability functions under the modern paradigm are only fully determined in idealizations, we may use empirical observations to falsify or corroborate these idealizations.

When we claim that the probability a coin will land heads on 8 out of 10 flips is determined by a Binomial PMF with $p=1/2$, we obtain an intellectual approximation for P(x=8: n=10). We can test this approximation by conducting a large number of trials, each consisting of 10 flips, and observing how many trials are consistent with 8 flips being observed. The idea in this empirical testing is that for more and more trials of 10 flips, the proportion of trials resulting in 8 heads will tend to the value designated under the intellectual construct of the Binomial distribution. We can furthermore use our empirical observations in these trials to determine whether the claim $p=1/2$ is valid, where p is the probability of success, or in this case the probability of landing heads. Hypothesis testing of this nature is an important development of probability theory, and the language used in the execution of methods for hypothesis testing resembles that of the criterions of scientific meaning posited in the sections above.

Answers to the problem of induction are found as useful models for structuring modern methods, yet the problem of induction itself has come to be regarded largely as extraneous or irrelevant. Humean skepticism considered, probability interpretations have always lain on shaky foundations. As interpretations were conceptualized further and applied to a world beyond idealizations, skeptics would argue that it was as if the proponents of probability theory were building a house without stopping to check that the foundation was level. All probability interpretations make an intellectual claim about the physical world, and by Hume's word we must declare: the tendency of probability theory to abuse induction for the purpose of cohesion requires unfounded assumptions about consistency in nature. On the other hand, disregarding the problem of induction is beneficial for the advancement of science, probability, and other endeavors: we cannot operate under chaos. It is only rational to assume that the world will continue to behave as it does currently. While it is not unreasonable to speculate that there are an infinite number of alternate worlds where, instead of the sun coming up, infinitely many other situations occur, this is not what we are interested in- we want to limit our speculations to states of the world consistent with the world we are in now, based on the limited information we have regarding such states: past, future, and present.

Regardless of our take on skepticism, we can freely admit that the modern probability "theory," or archipelago of probability interpretations, is consistent with Popper's falsificationism, at the very least. Using large, random samples as a basis for inferring properties of a population is rational if we can test the likelihood of these claims as a method of explaining the outcome of the random sample.

VII. Hypothesis Testing

Scientific explanation, how we justify the results obtained through trial and observation, is a topic deserving of its own conversation, where we would undertake an exploration of the models for such explanation and the philosophical problems they themselves encounter. For the sake of focus, we will currently disregard the many approaches a scientist, probabilist, or practitioner may take in conditionalizing the relationship between *explanans* and *explanandum*[16]; we will comment on hypothesis testing, a vindication of probabilistic claims with reference to empirical evidence.

[16] "Customarily, the thing that gets explained is called the *explanandum*. (In Latin *explanandum* means simply "the thing to be explained.") The thing that does the explaining is called the *explanans*. (again, a Latin word meaning "that which explains")" (Cover, 675).

If we were to investigate the theory behind Hypothesis testing for coherence under criterions of scientific meaning, we would find that it is compatible with the falsificationism of both Popper and Lakatos. With regard to the inferences made from a random sample to a population, in hypothesis testing we only make claims about what we **think** the structure of a population is based on the behavior of idealized samples. We make no absolute claims, and in this sense the language of hypothesis testing mirrors Popper's language in his formulation of falsificationism; we cannot *accept* a theory, we can only *corroborate* it. When using probability theory in testing a null hypothesis (H_0) against an alternative hypothesis (H_1), our options are to reject or not reject H_0. Hypothesis testing does not allow us to *accept* a hypothesis, only render it *more likely* to be true.

VIII. Conclusions

We began this paper by elucidating the problem of induction as posited by David Hume- that there is no rational basis for moving from invariable observations of past or current states to a future state of a consistent nature. Our conceptualizations of science and nature were thus found to be limited- what can we say about the future? Nothing for certain! An assortment of ambitious minds stepped forward in respective attempts to solve Hume's problem, each of them achieving a failure of the same caliber as the last. Verificationism was found to be viscously circular and inconsistent. Falsificationism was either naïve or dogmatic.

While the philosophical end of the spectrum was busy struggling with criterions of meaning and other analytic issues, we saw that probability theory evolved from a discussion about dividing stakes to a fully fledged, although controversial, topic. The question came to be, how do we interpret probabilities? Is it acceptable to have more than one interpretation? Is it acceptable to have more than one set of guidelines? To answer this question in the context of our exploration, consider the following:

> From a mathematical point of view, it is not necessary or even desirable to define probability explicitly According to the usual axiomatic procedure, probability is an undefined notion, a real number between 0 and 1 satisfying certain rules of operation from which the calculus of probability is developed by deduction. From this point of view any interpretation of probability is admissible if only the axioms are satisfied. [...] Since objective probability is an idealized relative frequency, the axioms have been chosen, implicitly in the beginning and explicitly later on, such that probabilities satisfy the basic rules of operation as relative frequencies. (Hald 246)

This opinion, expressed by Anders Hald in his own historical exploration of probability theory, mirrors those we have reached in this paper. Each interpretation we studied provided a unique method for implementing an intellectual construct in a physical world. What is essential is that a probability theory be parsimonious in restricting the applicability of interpretations that fall under it. While it is of the upmost importance that probabilities, under any interpretation, are coherent, we would also like it if the diversity of the physical universe was considered before ascribing likelihoods to physical events. By simply associating the structure of probability theory with the Kolmogorov axioms, a multiplicity of interpretations were able to flourish, each with an exclusive space in which calculations made are pertinent. We can thus objectively or subjectively approach the parameterization of the world surrounding us, after we decide the best technique for doing so.

IX. Works Cited

Blaug, M. “Kuhn versus Lakatos, or paradigms versus research programmes in the history of economics” *HOPE* Vol. 7. No 4, 1975. Print.

Cover, J.A. and Martin Curd. *Philosophy of Science: The Central Issues*. New York: W.W. Norton & Company, 1998. Popper, Karl. *The Problem of Induction*. Print.

Davidson, Barbara and Robert Pargetter. "In Defence of the Dutch Book Argument." *Canadian Journal of Philosophy* 15.3 (1985): 405-23 pp. <http://www.jstor.org/stable/pdfplus/40231424.pdf?acceptTC=true>. Online. Accessed 4/18/2012.

Dawid, A.P. "Probability, Causality, and the Empirical World: A Bayes-De Fenett-Popper-Borel Synthesis." *Statistical Science* 19.1 (2004). Print.

Eagle, Antony. Introduction. *Philosophy of Probability: Contemporary Readings*. Ed. Eagle, Antony. New York: Routledge, 2011. Print.

Feigl, Herbert. "Validation and Vindication: An Analysis of the Nature and the Limits of Ethical Arguments." *Readings in Ethical Theory*. Ed. Hospers, John and Wilfrid Sellars. New York: Appleton-Century-Crofts, Inc., 1952. Print.

Hajek, Alan. "Interpretations of Probability". 2011. *Stanford Encyclopedia of Philosophy*. 2012. <http://plato.stanford.edu/entries/probability-interpret/#ProInt>. Online. Accessed 4/18/2012.

Hald, Anders. *A History of Probability and Statistics and Their Applications before 175-*. New York: John Wiley & Sons, 1990. Print.

Hogg, Robert V. and Elliot A. Tanis. *Probability and Statistical Inference*. 8 ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2010. Print.

Hume, David, ed. *An Enquiry Concerning Human Understanding*. Great Britain Oxford University Press, 1999. Print.

Joyce, James M. "Non-Pragmatic Vindication of Probabilism." *Philosophy of Probability: Contemporary Readings*. Ed. Eagle, Antony. New York: Routledge, 2011. Print.

Kyburg, Henry E. "Subjective Probability: Criticisms, Reflections and Problems." *Philosophy of Probability*. Ed. Eagle, Antony. New York: Routledge, 2011. Print.

Lakatos, I. "Criticism and the Methodology of Scientific Research Programs" *Proceeding of the Aristotelian Society*. Vol 69, 1968. Print.

Lakatos, Imre. *Criticism and the growth of knowledge*. Cambridge University Press, 1970. Print.

Madden, Edward H. "Aristotle's Treatment of Probability." *Philosophy of Science* 24.2, 1957 167-72. Print.

Salmon, Merrilee H. *Introduction to the Philosophy of Science*. Upper Saddle River, NJ: Prentice Hall, 1992. Print.

Salmon, Wesley. *The Foundations of Scientific Inference*, University of Pittsburgh Press: Pittsburgh, 1966. Print.

Von Mises, Richard. *Probability, Statistics, and Truth*. New York: The Macmillan Company, 1957. Print.

Von Plato, Jan. *Creating Modern Probability: Its Mathematics, Physics, and Philosophy in Historical Perspective*. New York: Cambridge University Press, 1994. Print.