논리연구 13-1(2010) pp. 53～81

A Third Way to Interpretations of Probability*

Kyoung-Eun Yang

【Abstract】 This essay attempts to suggest a third way to interpret probabilities based on diagnoses of weaknesses and strengths of the objective and the subjective interpretations. While the objective interpretations capture important intuitions in employing probability, it can provide the definitions of probability only in a circular way. Although the subjective interpretations seem to avoid this problem, it also fails to understand how one can extract certain information from the ignorance of a given system. It will be suggested that these problems can be bypassed if we consider probabilities as theoretical structure, which provide significant generalizations about the world.

【Key words】 the objective interpretation of probability, the subjective interpretation of probability, the pluralist interpretation of probability, probability as theoretical structure

---

* 접수완료 : 2009. 12. 28 심사 및 수정완료 : 2010. 2. 19

1. Introduction

This essay attempts to suggest a third way to interpret probabilities based on diagnoses of the problems of the objective and the subjective interpretations. The objective interpretations, although necessary for understanding scientific theories, fail to provide non-circular definitions of probability. The subjective interpretations, while solving the above problem, instead fall short of comprehending how one can extract certain information from the ignorance of a given system. We will see that these impasses can be resolved if we see probabilities as theoretical structure, which provide significant generalizations about the world. It will be argued that a new approach enables us to circumvent the shortcomings of traditional interpretations of probability, and appreciate why these interpretations nevertheless appear to be convincing.

2. The Interpretation of Probability and Its Evaluation Criterion

An interpretation of probability provides an analysis of the primitive terms of the formal system of probability axioms. It is an attempt to comprehend what it means to apply mathematical axioms to probability claims in regards to the events. Philosophers are interested mainly in Kolmogorov axioms as the canonical axiomatization of probability, which can be formulated as follows.

Given a non-empty set Ω, an algebra on Ω can be defined as a set F of subsets of Ω that is closed under complementation and union with respect to Ω. P is a probability function from F to the real numbers, if and only if it satisfies following the three properties: (1) non-negativity: $0 \leq P \leq 1$, (2) normalization: $P(\Omega) = 1$, and (3) finite additivity: $P(A \vee B) = P(A) + P(B)$ for all $A, B \in F$ such that $A \& B = ø$. (Kolmogorov 1933) These axioms, rather than specifying how to assign probabilities in diverse experimental setups, merely provide constraints on how probabilities can be assigned. In actual cases, probabilities are assigned on the basis of estimates obtained from past experiences, analyses of conditions which generate outcomes, and assumptions underlying the experiments. An interpretation of probability, which provides the meaning of the mathematical axioms of probability, is related to the choice of a single or a combination of factors that influence the way probabilities are assigned.

Salmon (1966) suggests three criteria to determine the validity of given interpretations: admissibility, ascertainability, and applicability. As for admissibility, Salmon claims that "an interpretation of a formal system is admissible if the meanings assigned to the primitive terms in the interpretation transform the formal axioms, and consequently all the theorems, into true statements." (Salmon 1966, 64) This criterion requires a certain interpretation to satisfy the mathematical relationship specified by the probability axioms. Given that Kolmogorov's axioms are the canonical formalization of probability, it seems that admissibility requires probability to satisfy the axioms. In this way, admissibility makes interpretations relevant to a specific mathematical structure of probability. In contrast, the criterion of ascertainability is related with the condition for probabilities to be related with a real aspect of world. Ascertainability demands that "there be some method by which, in principle at least, we can ascertain values of probabilities." (ibid., 64) According to this criterion, interpretations of probability will be worthless if they do not provide specific methods to determine what values the probabilities are. Lastly, the criterion of applicability is concerned with the relations between formal mathematics and the practice of science. Applicability states that the interpretation is representative of the real practice of using probability. In order to show the force of this criterion, Salmon employs Bishop Butler's well-known aphorism, "probability is the very guide of life." (ibid. 64) All these three criteria provide checklists for a specific interpretation to be a valid one. These criteria will be employed in order to evaluate the objective and the subjective interpretations from the following section on.

3. The Frequency Interpretation and Its Problem

Given that the intuition of probability comes from attempts to understand how often a specific event occurs among a certain population of events, the frequency interpretation seems to reflect the real practice of science that uses probability. In the case of coin tosses, the probability of the occurrence of a head P(H) can be interpreted as the relative frequency of the occurrence of heads. Under this interpretation, the axioms of probability are satisfied in that P(H) is between 0 and 1, and P(H or ¬H) = P(H) + P(¬H). Hence, we can say that both applicability and admissibility hold in this interpretation. But when taking ascertainability into account, we can clearly see problems. In the case of a finite number of coin tosses, although many more than half of the tosses may turn out to be head, we still maintain that P(H) is 0.5 under an assumption that the coin looks fair. When there is a difference between the actual occurrence of relative frequency and the values of probability that are decided by assumptions such as fairness, we do not consider relative frequencies as the criteria deciding the values of probability. Of all possible values of the actual frequency of the events, we have no basis on which to either select or reject any values as the probability of the events occurring. Since there is not necessarily a connection between the actual frequency and the value of probability, actual frequency fails the criteria of ascertainability. (Nagel 1939)

At this point, we can attempt to establish the connection by modifying the conception of frequency. Although the real relative frequency interpretation, which claims that P(H) can be interpreted as the actual frequency of heads, cannot secure the correspondence between probabilities and frequencies, we cannot say that the intuition based on probability claims that depend on the frequency of occurrence is a complete failure. This implies that the actual frequency interpretation needs to be adjusted to incorporate ascertainability, while preserving admissibility and applicability. By maintaining the insight of relative frequencies while discarding the constraint of reality, one can modify the frequency interpretation. Instead of realized frequencies, we can sustain hypothetical relative frequencies. In the case of a finite number of coin tosses, the actual occurrence of heads needs not be exactly half of all trials. This interpretation claims that it will be enough to say that P(H) is 0.5 if the limit of the occurrence number of heads converges to 0.5. (Venn 1876, Reichenbach 1949, von Mises 1957)

Using this interpretation, it seems that the problem of ascertainability is circumvented, since we can establish the connection between probabilities and frequencies by means of the law of large numbers. The law is applied under the condition that each trial is independent and has the same probability of success. The law then states that as the number of trials increases, the observed relative frequency of a specific outcome approaches to the probability given by the Binomial probability model. Mayo, a supporter of the frequency interpretation, suggests the law of large numbers is an answer to the problem of ascertainability of probabilistic models: “a certain pattern of regularity emerges when they are applied in a long series of trials. ⋯ The pattern of regularity concerns the relative frequency with which specified results occur. The regularity being referred to is the long-run stability of relative frequencies. ⋯ Our warrant for such a conceptual representation is captured by the law of large numbers (LLN).” (Mayo 1996, 165-6) She claims that the regularity comes from 'empirical fact,' that is, the result of a large number of carefully carried out trials. In this way, frequentists suggest that the law of large numbers can link frequencies with probabilities.

It is questionable, however, whether Mayo's attempt to solve the problem of ascertainability is successful. Since the key idea of linking frequency and probability is to employ LLN, we need to characterize exactly the role of LLN that she has in mind. She claims that since the empirical law of large numbers holds in many phenomena, mathematical theory of probability can be used to model such phenomena: "The pattern of regularity concerns the relative frequency with which specified results occur. The regularity being referred to is the long-run stability of relative frequencies. ⋯ if in repeatedly carrying out a series of random experiments of a given kind we find that they always conform to the empirical law of large numbers, then we can use the calculus of probability to make successful predictions of relative frequencies." (ibid., 165-7) Her claim is essentially that the LLN is an empirical law that establishes the connection between empirical facts and mathematical models.

However, it seems that her argument misleads us by confusing the status of empirical laws. What is the empirical law of large numbers? Mayo considers the empirical law as coming from 'the long-run stability of relative frequencies.' It is definitely not a mathematical theorem such as the Binomial model that Mayo has in mind. Nor is it an empirical fact since no one can make trials of coin tosses forever to determine what the relative frequency of heads will be. It seems that what she attempt to say is that since the empirical law of large numbers holds in many phenomena, the mathematical law of large numbers can be applied to model such phenomena. (Uchii, 2001) In this spirit, she take an example of coin tossing case: “the accordance between observed relative frequencies and probability is repeatedly hold, … [t]he Binominal model, therefore, is an excellent model of this coin tossing mechanism and can be used to estimate the expected relative frequencies.” (Mayo 1996, 167) So, the law is rather an assumption that connects a mathematical theorem and an empirical fact. (Uchii, 2001) Given that the law of large number does not surpass the status of a theoretical assumption, we cannot say that Mayo definitely solves the problem of ascertainability since her argument supporting the frequency interpretation begs question.

What about the other way, that is, can we count on the mathematical theorem of the law of large numbers for the solution to the problem of ascertainability? By means of this mathematical theorem, we might attempt to rationalize empirical facts, and relate the concept of frequency with the concept of probability. In this case, however, we have to confront the big question of how mathematical structure can represent the world. Even if this big problem could be solved, we should only face another problem. The real problem with the mathematical theorem of the law of large numbers, as Skyrms points out, is its failure to provide a non-circular account of how probability claims should be interpreted. (Skyrms 1980) The theorem in case of coin tosses states the chance that the percentage of events H diverges from P(H) by a fixed positive e, goes to zero as the number of trials n approches to infinity, for every e. In other words, P(H) = 0.5 if and only if P(H = 0.5 + e | n time coin tosses) converges to 1 as n goes to a large number, for every e. Yet this is a circular characterization of probability. The concept of probability appears on both sides of the 'if and only if statement.' (Sober 1993) So the probability of the occurrence of a specific event is defined by appealing to other probability claims. If we want a non-circular definition of probability, this mathematical theorem by no means offer an account of how probability statements should be understood under the relative frequency interpretation.

We can see, then, that neither empirical nor mathematical law is successful in providing the solution to the problem of ascertainability. Given that the law of large numbers is the key strategy to secure ascertainability, the frequency interpretation has the problem of justification in regard to ascertainability.

The problem of ascertainability within the frequency interpretation becomes most serious, when we attempt to understand the probability of a single case by means of the frequency interpretation. The problem of ascertainability gets worse as the number of trials, for example, of coin tosses, is reduced to significantly small figure since the case, however weak it is, loses even the support of LLN. Accordingly, when considering the interpretation of the probability of a single case, the problem of ascertainability develops into the most severe one. From the perspective of the frequency interpretation, we cannot say that a single case is probable or improbable since probability is characterized in a collective sense. But it seems that regardless of how many times it will happen, the probability claims of occurring a unique event is still meaningful. As argued above, it is not clear that we can definitely understand the probability claims even within the coin tosses by means of its frequency.

Yet let's suppose that we somehow find a way to comprehend the frequency interpretation within coin tosses. The fact that this specific case supports the frequency interpretation, however, still does not mean that we can capture all probability claims from the perspective of the frequency interpretation. Since this special case is artificially constructed to make repeated experiments possible, the coin tosses seem to be designed under the assumption of the frequency interpretation. But with an argument within this rather idiosyncratic case, we cannot make a legitimate generalization regarding less idealized experiments or real probabilistic events, which are no less prevalent in real practices of science. And probability claims of a single case appears very often within the context of scientific theories, we can see that the problem of ascertainability is closely related with the problem of applicability.

4. The Propensity Interpretation and Its Shortcoming

Given that the relative frequency interpretation fails to provide a link between the occurrence of events and probability, we might suspect that the propensity interpretation offers a solution to this problem. The propensity interpretation offers the connection between the occurrence of events and probability by means of dispositional properties. (Popper 1959) When in the coin tossing case we make a probability statement that P(H) is 0.5, what makes this link true is a certain property called propensity of a coin to land heads when tossed. It seems that this concept of propensity can resolve the problem of characterizing probability in a non-circular way.

However, although this move provides a possible way of solving part of the problem, it still fails to grasp the link between the occurrence of events and the probability of all probabilistic cases. In this scheme, there are two ways to examine what propensity a certain probability system, such as coin tossing, has: we can make a finite number of trials to get relevant evidence (Gilles 2000), or we can figure out the physical structure of the coin which determines the occurrences of heads (Popper 1959). The former method capturing the link between the occurrence and probability does not avoid the problem associated with the frequency interpretation. For in this case the concept of propensity does not add anything new to the concept of frequencies, since the propensity of P(H) still depends on the frequencies of P(H). What an interpretation of probability aims to do is to give a meaning of probability in a non-circular way; that is, it does not use the concept of probability itself. This kind of definition, however, is not provided by an interpretation based on relevant evidence of the propensity of a probability system. (Sober 1993)

In this sense, the alternative approach, which characterizes the probability of a specific event by examining its physical constitution, seems to be successful in defining probability in a non-circular way. For the propensity stems from the property that causes certain events to happen. According to Popper, "[c]ausation is just a special case of propensity: the case of a propensity equal to 1." (Popper 1990) Causality, then, seems to provide us with a wherewithal which links the occurrence with probability without depending on the concept of frequency. According to the propensity interpretation, the concept of probability is then completely captured by causal relations between cause and effect.

Aside from the problem of defining causality, however, this approach can be criticized for its lack of its generality to embrace significant parts of probability statements. Concerning applicability, there are cases that cannot be explained through this interpretation. Many probability claims do not describe any such causal relation. Gillies (2000) summarizes a shortcoming of the idea that propensities are generalisation of causes: "Causes have a definite direction in time. But situation is very different with probabilities. For event A, B, we usually have that if P(A|B) is defined, then so is P(B|A). Probabilities have symmetry where causes are asymmetrical. It thus seems that propensity cannot after all be a generalisation of cause." (Gillies 2000, 143) The probabilistic case of how often the cause of certain events occurs given its effects cannot be understood within the framework of propensity based on causal relations.

Sober (1993) exhibits an example from the philosophy of biology, "we can talk of the probability that an offspring will be heterozygote if its parents are heterozygote. Here, the parental genotype causes the genotype of the offspring. But we can also talk about the opposite relationship: the probability that an individual's parents were heterozygote, given that the individual itself is heterozygote. Offspring genotypes do not cause the genotypes of parents." (Sober 1993, 63) The propensity interpretation, although providing the connection between the occurrence and probabilities of certain cases, is short of becoming a legitimate generalization of all kinds of probability. Accordingly, both objective interpretations of probability, frequency and propensity interpretation, fall short of comprehending all cases of probabilities.

5. The Classical '*A Priori*' Interpretation

So far we have discussed the problems associated with the objective interpretations of probability. By asking what we have learnt from the problems of the objective interpretations, we may then find the clues of alternative interpretive schemes. In spite of its plausible intuition, the frequency interpretation fails to show the link between the occurrence of certain events and its probability. The way of overcoming this problem, i.e. providing propensity interpretation based on causal relations, turns out to be only partially successful.

According to Sklar (1995), the major difficulty in providing a non-circular definition of probability arises from its reductionist methods, which attempts to define probabilities by means of other concepts which does not contain the concept of probability. The failure of the frequency interpretation can be accredited to the difficulty of reducing probabilities to empirical data, while the propensity interpretation cannot completely understand probabilities only by means of causal relations. Sober (1993) points out that the problems of characterizing probabilities empirically stem from the attempts to define probabilities by means of their evidence. Although the frequency of certain events can be admitted as evidence of probability, this does not necessarily mean that it also provides the definition of the concept. Given that the problem of the frequency interpretation comes from the empiricist approach, which attempts to elevate the status of evidence to the definition of probability, we can see that the empiricist way of reducing the concept of probability to something else, which does not presuppose the concept of probability, needs modification.

The propensity interpretation based on causality is proposed to overcome this problem. It is supposed to connect the occurrence of certain events with its probability by means of dispositional properties. But as shown in the previous section, only some probability statements can be reduced to causal claims.

Instead of the objective approaches which fail to provide legitimate interpretations, we can adopt the classical ‘a priori’ interpretation, which relates probability claims to the absence of any evidence, or the presence of symmetrically balanced evidence. In this scheme, the concept of probability comes from our ignorance of what will happen. In case of coin tossing, its range of mutually exclusive and exhaustive outcome is decided a priori. Our inability to predict outcomes makes us assign equal probabilities to all the possible outcomes. This approach can be dated to the very earliest attempt to define probabilities and is referred to as the 'classical interpretation of probability.' Laplace claimed that "the theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favourable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability." (Laplace 1814, 6-7) This approach begins with the partitioning of all the equally possible outcomes. The probability of a certain event can then be assigned as the fraction of the number of the events to the total number of possibilities. This interpretation seems to provide a sensible analysis for many cases that employ probability, which are not only idealized cases, such as a game where a fair die is used, but also more complicated cases that employ ensemble theories in statistical mechanics, which provides models for other probabilistic cases.

When it comes to ascertainablity and applicability, however, the classical interpretation clearly shows its limitation. The major intuition of the classical interpretation is that all the cases of probability can be analysed with the partitions of equally probable alternatives. We can then ask the question of who has the final word in deciding what is "equally possible." Although the answer in case of a fair die case seems straightforward, this case cannot be considered as being representative of general cases. For it is artificially created by employing the classical interpretation to arbitrate the outcome of the game. In this case, the game maker pre-determines the equi-probability of the game by setting the number of alternatives to the number of faces on the die. The equi-probability can be unanimously decided based on our own convention of probable alternatives. However, within the case of less artificial and more complicated situations which often occur in social and natural phenomena, different people with different points of view may have different views about what should be considered as “equally undecided.”

At this point, the supporters of the classical interpretation might escape this problem by adopting a representative rational agent who selects a set of equally possible alternatives. However, given that we have yet no conclusive idea about what a rational agent or a representative agent exactly means (Kahneman and Tversky 1982), the classical interpretation still has the burden of providing a neutral position concerning equi-probability. Without this neutral position to decide what is equally probable, the core concept of the classical interpretation, “equally possible,” remains unclear.

In order to overcome this vagueness, the classical interpretation can be articulated by employing what Keynes (1921) calls the ‘Principle of Indifference,’ which maintains that the equal probabilities can be assigned to a set of possible outcomes in which there is no evidence supporting the possibility of one outcome over another. However, it is doubtful that this modification really succeeds in overcoming the vagueness found in the classical interpretation of probability. For the Principle of Indifference glosses over the problems of ascertainability and applicability, rather than resolving it.

There are two relevant clarifications of the principle which are worth considering: (1) in any given event there is no evidence supporting one possible outcome over another, and (2) every outcome must have symmetrically balanced evidence. These clarifications, however, still have shortcomings. (Hájek, 2002) With the first clarification it is very hard to imagine a situation where there is no evidence to support one possibility over another. Even in a die toss (which comes closest to an idealized situation) we learn to assess the bias of the die through eye-witness accounts, historical data, and theoretical knowledge of its structure. Furthermore, it becomes more difficult to find a case in the real world where a range of equally probable outcomes can be determined without recourse to past experiences and theoretical knowledge. Accordingly, we can see that we cannot apply the first clarification to real-world cases: it can only applied to artificially generated idealized situations such as a fair die, which rarely exist within the real world, the problem of applicability is manifest.

The second clarification attempts to resolve the problems confronted in the first clarification: it seeks to restore the reality of probability assessment by employing symmetrically balanced evidence to determine equally probable outcomes of an event. As we have seen from the former case, it seems that we need prior knowledge, such as historical information of a certain die's bias, to characterize the probability of real-world events. In this scheme, probability claims are defined with reference to the presence of symmetrically balanced evidence. Symmetrically balanced evidence can be formally described in terms of equality of conditional probabilities: the evidence is symmetrically balanced if and only if $P(O_1|E) = P(O_2|E) = \cdots = P(O_n|E)$, where E is the given evidence and $O_1, O_2, \cdots, O_n$ are the possible outcomes. (Hájek, 2002) At this point, however, we can see that this characterization of probability is a circular one. Probabilities are characterized based the Principle of Indifference, which in turn decided by other probabilities. Within this formalization, the classical interpretation based on the Principle of Indifference depends on the concept of probability in defining probabilities. In other words, the values of probabilities can only be determined with reference to other probabilities. We can expect this problem of circularity earlier, since the main intuition of the classical interpretation is that probabilities are based on the partition of equally probable outcomes.

The essential problem of the classical interpretation originates from its *a priori* nature. This interpretation claims that probabilities can be determined *a priori* by constructing the space of probabilities based on the Principle of Indifference. It is *a priori* in that any possible outcomes are presumed to be equi-probable without being affected by what happens in nature. In case of coin tosses, we assign probabilities by employing Bernoulli theorem. The theorem states that if the probability of head is 1/2, then the probability approaches 1 that in n tosses there are approximately n/2 heads as n increases. Within the framework of the epistemic interpretation, the theorem simply specifies relative number to equi-probable alternatives. Yet it is no more than a theorem which allow us to conclude that these alternatives will occur equally often. In this process, the Principle of Indifference seeks to specific information about what will happen by *a priori* consideration. But how it is possible that we can transform our ignorance about a given system to specific information? How we assign probability in spite of ignorance? As Fine states, "If we are truly ignorant about a set of alternatives, then we are also ignorant about combinations of alternatives and about subdivisions of alternatives." (Fine 1973, 170) So, the classical interpretation has a problem of explaining how we can extract certain information from our ignorance.

6. Alternative Approaches to the Traditional Interpretation

Although the objective and the epistemic interpretations have their problems, it does not follow that we should throw away both interpretations due to these problems. By entertaining different interpretations for distinct cases, we can give up the single overarching interpretation and accept a specific interpretation case by case. We can also imagine that even within a specific case, probability claims contain both objective and epistemic aspects. It may be possible then to coalesce two distinct interpretations, the objective and epistemic interpretations, in order to make up for the shortcomings of each other.

Gillies proposes a pluralist interpretation claiming that probability statements consist of partly objective and partly epistemic components. (Gillies 2000) His view suggests that "there is not a single notion of probability, but rather several different, though interconnected, notion of probability which apply in different contexts." (ibid. 169) The probability calculus then accepts "a number of different interpretations each of which is valid in a particular area or context." (ibid. 180)

For example, the single case probability that Mr Smith aged 40 will live to be 41 can be interpreted by employing the objective interpretation along with epistemic one. By choosing a reference class which characterizes Mr Smith as certain type of people, such as Englishman aged 40 who smokes two pack of cigarettes, we can make repeated measurements of that kind of people. We can, then, introduce objective probabilities for events that are outcomes of some sets of repeatable conditions. However, the probability of this event, though objectively based, is by no means completely objective because constructing the reference class also depends on our own classification. We *construct* a set of reference class to the event depending partly on what we think relevant, for which a relative frequency can be calculated. Since these reference classes are constructed through our choice of information which we consider significant, it can be still shaded by our lack of relevant information.

Along these lines, we can split the context of the case into two different constituents, and then apply two different interpretations to each one. Accordingly, we can compensate the shortcoming of the epistemic interpretation with the objective one. With the pluralist interpretation, we can say that the objective interpretation complements the shortcoming of the epistemic interpretation, which claims that probabilities are employed to determine the outcomes of given event where there is lack of relevant information. Probability can then be interpreted as partly objective aspects plus a partly tentative placeholder before we acquire complete knowledge. This pluralistic interpretation, then, presents a means of resolving the problems associated with the subjective interpretation.

However, it is still questionable as to whether we can build a legitimate pluralistic interpretation simply by combining two problematic interpretations. It seems that this plain combination of two interpretive schemes combines their individual problems without solving them. The objective elements of the pluralist interpretation still possess the problem of ascertainability, which is inherent within the objective interpretation. Although statistical data relative to a given reference class are available, ascertainability is still a problem. For the linkage between probabilities and frequencies is still weak with respect to that reference class. Given that the pluralistic interpretation depends on the objective interpretation in its original form, unique problems of the objective interpretation cannot be remedied just by complementing the interpretation of probability with epistemic aspects.

Furthermore, without modifying the intuition of the epistemic interpretation, we still have a problem of explaining how to extract specific information from our ignorance. In this way, the plain combination of the two interpretations with their original forms preserves their weakness even after they are merged. With these two problems inherent within each separate element, the pluralistic interpretation glosses over the two problems without solving them.

Although Gillies' attempt to combining the objective and epistemic interpretations cannot solve problems associated with either interpretation, it seems that the intuition under the pluralist interpretation, which contains aspects of both objective and epistemic attitudes, is still viable. The point is that although the combination itself is plausible, a simple combination is not sufficient to resolve problems associated with the interpretations of probability. Since, without modifying of their original intuition the weaknesses of each interpretations of probability are preserved, it is necessary to modify the both interpretations.

At this point, we need to contemplate a more sophisticated approach that considers both strengths and weaknesses of the objective and the epistemic interpretations, rather than simply adopting a hybrid one. By evaluating exactly what strengths and weaknesses of each interpretation has, we can see how a modified interpretation can be constructed. The strength of objective interpretation lies in the fact that probability claims can extract specific information from the world, although the link between probability and frequency is still weak. In contrast, when it comes to the epistemic interpretation, ascertainability is less problematic, while the question of how we can extract certain information from the ignorance of the system still needs explanation. In that case, for the interpretation of probability to be a valid one, it seems that we need to adopt both objective features that relate probability claims with real aspects of the world, that is, applicability, and epistemic ones capable of securing a method of deciding the value of probability, that is, ascertainability.

Given that the problem of epistemic interpretation i.e., how specific information can be extracted from our ignorance, we need to think about the epistemic interpretation in terms of imposing certain structure for this ignorance to be tamed. We can suggest that our epistemic process to posit probabilities is to impose certain *constraint* over the events. Our epistemic frameworks codify the way in which our ignorance can be controlled. By casting reference class through certain epistemic processes, such as symmetry considerations, we are able to extract information that represents certain aspect of world. We can say then that assigning equi-probabilities is a result of an active epistemic process rather than a passive one. In other words, this active process plays a constructive role. In this way, this epistemic process generates specific *frameworks which increase our understanding of certain aspects of the event*. Assinging probabilities, then, plays the role of 'explanation.' (Sklar 1995)

But probabilities can be on certain empirical basis by successfully explaining a given event. Along these lines, probabilities have the objective aspects by which we can access empirical information. Sober (1993) points out that it is because probabilities are employed not only for predictions but also for explanations. We can provide a shortcut explanation by substituting detailed information about causal factors affecting a specific event with probabilities. In this spirit, Sober claims that "[o]ur reason for using probability here is not that we are ignorant: we are not. We possess further information about the idiosyncratic details concerning each mating pair. These would be relevant to the task of prediction, but not necessarily to the task of explanatory description." (Sober 1993, 65) A given event can be explained by means of the generalization which probabilities catch by substituting complex causal structure. An explanatory structure can be imposed by constructing a probability space for a given system. To do so, we select significant information to be fit in the mathematical axiom of probability, while neglecting information which is irrelevant to explanation.

Probabilities employed in evolutionary theories provide a case for the contention that probability is a certain kind of generalization. Sober presents a case of Mendelian mechanism of mating pairs in which both parents are heterozygote, i.e. Aa. Although different frequencies of heterozygote offspring satisfies the Mendelian mechanism, the mating pairs are different from each other in many ways that accounts for their different frequencies. Sober provides two distinct ways of explaining the different frequencies of mating pairs; "[w]e could describe these different mating pairs one at a time and list the unique constellation of causal influences at work in each. However, another strategy is to try to isolate what these parental pairs have in common. We do this when we describe each of them as participating in a Mendelian process in which P(offspring is Aa | parents are Aa and Aa) = 0.5." (ibid. 65) In this way, we can substitute the complicated workings of reproduction of individual mating pairs with a probability claims that intend a population level generalization. If the epistemic interpretation treats probability merely as a placeholder of our ignorance, we fail to acknowledge the facts that certain information, in order to have a meaningful generalization, is intentionally chosen to be marginalized as insignificant or to be ignored as inconsequential. When we already know limited workings of the causal structure of the system, we employ probabilities in order to extract information from the event considered as coincidence.

7. Conclusion

From aforementioned lessons learnt from strengths and weaknesses of the objective and the subjective interpretations, a third way to interpretations of probability is suggested. As for the objective interpretation, we have seen that direct links between frequencies and probabilities are weak. Yet, it is possible that this direct empirical link can be replaced by indirect link based on probability as a generalization. What is worth noting regarding the shortcomings of the subjective approach is that the interpretation of probability is not only about our ignorance of a given system, but is also related with casting specific structures within the system by projecting the mathematical structure of probability in order to provide significant generalizations. Given this way of modifying the interpretations of probability, we can see that probability claims are not solely concerned with a summary of empirical data, but with constituting certain structures that provide significant generalizations. With this perspective of seeing probability within its constructive role, we have a way of providing a subjective aspect within the objective interpretation. Considering this Kantian perspective, we can have an alternative interpretation of probability.

References

Bartha, P. and Johns, R., 2001, “Probability and Symmetry”, Philosophy of Science 68 (Proceedings), S109-S122.

Bricmont, J., 1996, “Bayes, Boltzmann and Bohm: Probabilities in Physics”, Annals of the New York Academy of Science 79, 1-21.

de Finite, B., 1937, ‘Foresight. its logical laws, its subjective sources,’ in Kyburg & Smokler 1964, pp. 53-118.

de Finetti, B., 1970, Theory of probability, Wiley.

Earman, J., 1992, Bayes or Bust, Cambridge: MIT Press.

Gillies, D., 2000, ‘Varieties of propensity,’ British Journal for the Philosophy of Science 51, pp. 807-835.

Gillies, D., 2000, Philosophical Theories of Probability, Routledge.

Galavotti, M.C., 1991, ‘The notion of subjective probability in the work of Ramsey and de Finetti,’ Theoria LVII(3), pp. 239-259.

Hacking, I., 1975, The Emergence of Probability, Cambridge University Press.

Howson, C. and Urbach, P., 1993, Scientific Reasoning: The Bayesian Approach, Open Court, 2nd edition.

Humphreys, P., 1985, “Why Propensities Cannot Be Probabilities”, Philosophical Review, 94: 557-70.

Kahneman, D., Slovic P. and Tversky, A., (eds.), 1982, Judgment Under Uncertainty. Heuristics and Biases, Cambridge: Cambridge University Press.

Keynes, J. M., 1921, A Treatise on Probability, Macmillan and Co.

Laplace, P. S., 1814, English edition 1951, A Philosophical Essay on Probabilities, New York: Dover Publications Inc.

Lewis, D., 1970, “How to Define Theoretical Terms”, Journal of Philosophy 67: 427-446.

Lewis, D., 1980, “A Subjectivist's Guide to Objective Chance”, in Studies in Inductive Logic and Probability, Vol II., University of California Press, reprinted in Lewis 1986b.

Popper, K. R. 1983, Realism And The Aim of Science, Hutchinson.

Popper, K. R. 1959, “The Propensity Interpretation of Probability”, British Journal of the Philosophy of Science 10: 25-42.

Reichenbach, H., 1949, The Theory of Probability, Berkeley: University of California Press.

Salmon, W., 1966, The Foundations of Scientific Inference, University of Pittsburgh Press.

Sklar, R., 1992, Philosophy of Physics, Westview Press.

Sklar, R., 1995, Physics and Chance, Philosophical Issues in the Foundations of Statistical Mechanics, Cambridge University Press.

Skyrms, B., 1980, Causal Necessity, Yale University Press.

Sober, E., 1993, Philosophy of Biology, Oxford University Press.

Uchii, S., 2001, The Law of Large Numbers, http://www.bun.kyoto-u.ac.jp/~suchii/ LLN.html

van Fraassen, B. C., 1989, Laws and Symmetry. Oxford:

Clarendon Press.

Venn, J., 1876, The Logic of Chance, 2nd ed., Macmillan and co; reprinted, New York, 1962.

von Mises R., 1957, Probability, Statistics and Truth, revised English edition, New York: Macmillan.

von Plato, J., 1994, Creating modern probability: its mathematics, physics and philosophy in historical perspective, Cambridge University Press.

Whewell, W., 1860, On the Philosophy of Discovery: Chapters Historical and Critical, London.

Woodward, J., 2001, A Theory of Explanation: Causation, Invariance and Intervention, Oxford: Oxford University Press

Yonsei University
Email: newtleib@googlemail.com