What is Conditional Probability?
In Defense of Lowe's Definition(s)

Hilmi Demir

Abstract: In the standard and traditional view, the concept of conditional probability is defined with what is known as the ratio formula: the probability of B given A is the ratio between the probability of A and B and the probability of A. It is well known that this definition does not match the conceptual and mathematical expectations that we have from conditional probability, especially for the probability values at the limits. Thus, as pointed out by several philosophers such as Popper and Hájek, it is fair to conclude that we have yet to have a satisfactory definition for the concept of conditional probability. E.J. Lowe, in a debate with Dorothy Edgington, proposed two different definitions of conditional probability, and unfortunately his definitions have gone unnoticed in the literature. In this paper, my main aim is to renew interest in Lowe's definitions. I achieve this aim by showing that E.J. Lowe's definitions have great potential in providing us with a satisfactory definition of conditional probability.
Keywords: conditionals, Philosophy of Probability, Kolmogorov, probabilistic independence

Özet: Standard Olasılık kuramında bir olayın bir diğer olaya koşullu olasılığı rasyo formülü olarak bilinen bir formül ile tanımlanmaktadır. Bu formüle göre B olayının A olayına koşullu olasılığı (A ve B) olayının olasığının sadece A olayının olasılığına bölünmesi ile bulunan değerdir. Bu standard tanımın özellikle limitlerdeki olasılık değerleri için kavramsal ve matematiksel beklentilerimizi karşılamadığı bilinen bir durumdur. Aralarında Popper ve Hájek gibi isimlerin de bulunduğu birçok felsefecinin de belirttiği gibi, bu durumdan elimizde tatmin edici bir koşullu olasılık tanımı olmadığını çıkarsamak yanlış olmayacaktır. E.J. Lowe, Dorothy Edgington ile girdiği bir tartışma bağlamında koşullu olasılığın iki alternatif tanımını önermiştir. Ne yazık ki, literatürde bu tanımlara gereken önem verilmemiştir. Literatürdeki bu eksikliği gidermeyi hedefleyen bu makalenin genel amacı, Lowe'ün önerilerinin tatmin edici bir koşullu olasılık tanımı sunma potansiyeline sahip olduğunu göstermektir.
Anahtar Kelimeler: şartlı önermeler, Olasılık Felsefesi, Kolmogorov, olasılıksal bağımsızlık

I. Introduction

Despite its long history and prevalent usage in many different fields, the

Demir, Hilmi. (2016). What is Conditional Probability? In Defense of Lowe's Definition(s). *Kilikya Felsefe Dergisi* 2. pp. 1-15.

concept of conditional probability has yet to have a satisfactory definition. According to the standard view, which was provided by Kolmogorov in 1933, conditional probability is defined with the following formula, which I call **SF** – short for the "standard formula."

> **(SF)** $P_A(B)=(P(A\&B))/(P(A))$ provided that $P(A)>0$
> [A is called the conditioning event and B is called the conditioned event. A and B are consistently used for the conditioning event and the conditioned event, respectively, throughout the paper.]

This formula leads to several counterintuitive consequences, especially because of not being defined where the probability of the conditioning event is 0[1]. For example, it does not define $P_A(A)$ where A is an impossible event. Furthermore, not all events with zero probability are impossible events. There are possible events with zero probability, and, for such events, the conditional probability is undefined under **SF**. This zero-denominator problem, which may be seen as a small glitch at first, is a serious conceptual shortcoming. Let me quote what Popper says about the problem:

> Systems [that define conditional probability with **SF**] are …very weak: it can occur in [such] systems that p(a,b)=r is a meaningful formula[2], while at the same time, and for the same elements p(b,a)=r is not meaningful, i.e. not properly defined, and not even definable, because p(a)=0. But a system of this kind is not only weak; it is also for many interesting purposes inadequate: it cannot, for example, be properly applied to statements whose absolute probability is zero, although this application is very important: universal laws, for example, have, we may here assume, zero probability (emphasis original; Popper 1959, p. 334).

SF has even more puzzling implications for mutual independence

[1] Probabilities can be defined over events, sets, or propositions. Each choice may have differing implications on how to interpret probabilities. Such issues, however, do not have any effect on my argumentation in this paper. The choice I made in this paper is to stick with the choice of the theoretician that is being mentioned. For example, when analyzing Kolmogorov's view, I use events as the applicanda of probabilities, and similarly I use propositions when analyzing Hájek's, Edgington's, and Lowe's ideas. When presenting the proofs in Section IV, my choice is a bit more liberal. I use the terminology that is more apt for the reasoning presented in each proof.

[2] p(a,b) is Popper's notation for the probability of a given b.

of events. According to the independence criterion based on the formula (A and B are independent if and only if *(A&B) = (A)P(B))*, any event with probability of 1 or 0 ends up being independent from itself. This is counterintuitive, because identity is the ultimate case of probabilistic dependence (Hájek, 2010; Fitelson and Hájek, 2014). Such problems led to a search in the literature for different ways of interpreting/defining conditional probability with the purpose of resolving those problems. There have been many such attempts, which roughly follow two different routes. The first is to reverse the direction of analysis in the theory of probability. In the traditional Kolmogorovian approach, single-event probabilities[3], *P(A)*, are taken to be primitive, and conditional probabilities are defined in terms of those primitive single-event probabilities. Scholars such as Popper (1959) and Hájek (2003), among many others, suggested taking conditional probabilities, i.e. $P_A(B)$, as primitive and defining single-event probabilities derivatively by $P_T(B)$ where T represents necessity. The second route taken in the literature is to understand **SF** as a theorem that follows directly from a definition of conditional probability. Obviously, this route requires coming up with a new definition for conditional probability. My main aim in this paper is to renew interest in the second route by focusing on E.J. Lowe's work.

Dorothy Edgington and E.J. Lowe had a lively debate that started in the mid-1990s; its main topic was how to explicate conditional statements and judgments, which is not my concern here, but within the debate E.J. Lowe offered two different definitions for conditional probability. Those definitions form the main content of my paper. In his 1996 *Mind* article, Lowe first pointed out a conceptual mistake that is prevalent in the literature on conditional probability. He claims that conditional probability is not a kind of probability different than single-event probabilities; rather, it is a conditional statement about single-event probabilities. He then defined conditional probability with a conditional statement; the definition that he gave there faced serious criticisms from Edgington (1996). In a somewhat late response, Lowe retracted his 1996 definition and offered a new one in his 2008 *Analysis* article. I claim that Lowe's 2008 definition with a slight revision becomes consistent with our conceptual and mathematical expectations of conditional probability. In addition to this main claim, I also offer a different interpretation of Lowe's 1996 definition that blocks

[3] Single-event probabilities are called 'absolute probabilities' and conditional probabilities are called 'relative probabilities' in Popper's terminology.

Edgington's criticisms and becomes complementary to the revised version of Lowe's 2008 definition. My defense of Lowe's definitions consists of accomplishing the following four tasks:

i) Identify and analyze the problems about the standard formula.
ii) Analyze Lowe's 1996 definition and offer a different interpretation that blocks Edgington's criticisms.
iii) Provide a list of desiderata for a satisfactory definition of conditional probability.
iv) Prove that the revised version of Lowe's 2008 definition has great potential for satisfying the desiderata mentioned in (iii).

If I successfully accomplish these tasks, my overall conclusion will be that there is a satisfactory definition for the concept of conditional probability that is provided by Lowe and unfortunately has gone unnoticed in the literature. If this definition is included in Kolmogorov's theory of probability, **SF** becomes a theorem of Kolmogorov's formalization as desired by many such as De Finetti[4]. This may be considered as a humble contribution to Kolmogorov's theory of probability since **SF** properly interpreted is not a definition but rather a theorem, contrary to what Kolmogorov claims. This is not to claim that Kolmogorov made a mistake in his formalization. Rather, it means that in some cases conceptual clarification may not be at the top of the priority list of a mathematician, but the same luxury, as pointed out by Popper, Hájek, and Lowe, among many others, is not available to philosophers.

II. Analysis of the Standard Formula

Kolmogorov, in his *Foundations of Probability*, provided a formalization of the theory of probability. After stating his five axioms, he gave the standard formula as the definition of conditional probability. It is quite to the point to ask whether the formula really is a definition or not. First of all, as Hájek (2003) points out, the standard formula is not a stipulative definition of a made-up operation, because the concept of conditional probability is not a technical notion such as a zero-sum game or categorical imperative. The concept of conditional probability had been in use in ordinary language long before any philosopher or mathematician attempted to give a definition for it. Ordinary language locutions such as the likelihood of *B* in the light of *A*, informed by *A*, relative to *A*, or upon discovering *A*

[4] Reported in Lowe 1996.

are used to state conditional probability statements. Such locutions are used to form conditional probability statements in more theoretical work as well. Bayes' Proposition 5 is a clear example of this: "If there be two subsequent events, the probability of the second b/N and the probability of both together P/N, and *it being first discovered that the second event has also happened*, the probability I am right is P/b" (emphasis added; Bayes, 1764, prop 5).

These remarks are sufficient enough to show that any given definition for conditional probability cannot be considered as a stipulative one. It must be understood as a definition of a concept that has been in use in ordinary discourse for a long time and, thus, the given definition must be consistent with the intuitions that we have about the concept. If this is true, then one can assess the success of a suggested definition by seeing if the implications of the definition are consistent with intuitively expected results.

Let us put the standard formula to this assessment test. When the probability of the condition A is greater than zero, the standard formula works very well. First of all, as proven by Williams (1980), the standard formula gives us the most minimal revision in the probability of B under the assumption that event A has happened. In other words, the change in probability values implied by the standard formula is no more or less than what is required upon learning new evidence. Despite this success, however, the standard formula fails severely when the probability of the condition is 0. Our intuitions tell us that $P_A(A)$ is 1 regardless of the value of the probability of A. Similarly, $P_A(\text{not-}A)$ must be zero no matter what $P(A)$ is. However, the standard formula leaves these values undefined when $P(A)$ is 0. This problem is well known and might be dismissed as a technicality because of the seeming oddness of conditioning on impossible events. But the problem is more general than just impossible events. In Kolmogorov's theory, all *practically* impossible events are assigned zero probability. Thus, some zero-probability events are theoretically possible events. Here is what Kolmogorov says: "For an impossible event $A, P(A) = 0$. But the converse is not true. $P(A) = 0$ does not imply the impossibility of A . In this case, event A is practically impossible." (Kolmogorov, 1950/1933, p. 5).

Such practically impossible but theoretically possible events are assigned the probability of 0. Following Hájek's (2010) terminology, I shall call such events improbable events. Let me give a simple example of such an event: suppose that a fair coin is tossed infinitely many times. The probability that it lands heads every time is 0. Because of that, **SF** leaves the following conditional probability undefined: the probability of the coin landing heads every time given the event that it landed heads every time. But, conceptually, it is clear that this conditional probability must be 1. It seems to me that this example is enough to show that conditioning on events with 0 probability is a serious problem for the standard formula.

Another problem that afflicts the standard formula is the mutual independence of random events in the theory of probability. Before explaining the independence problem, it will be useful to mention how crucial mutual independence is for the theory of probability. According to Kolmogorov, his formalization of the theory of probability is equivalent to the "theory of additive set functions". So, in a sense, there is no need for a separate theory of probability. The main thing that led to the formation of the theory of probability is the concept of mutual independence. He says: "Historically, the independence of experiments and random variables represents the very mathematical concept that has given the theory of probability its peculiar stamp" (Kolmogorov, 1950/1933, p. 8).

In Kolmogorov's formalization, the mutual independence of events is defined as follows. Events A is are mutually independent if and only if

$$P(A_1\&A_2\& \dots \&A_n)=P(A_1)P(A_2) \dots P(A_n)$$

For the ease of presentation, I will use the instantiation of this definition for the mutual independence of two events, i.e. $(A\&B) = (A) P(B)$. This definition together with the standard formula implies the other well-known formulation of mutual independence: A and B are independent from each other if and only if $P_A(B)=P(B)$. These two criteria work well for the events that have a probability value strictly between 0 and 1. This is why it has been used in empirical fields without much questioning, since probability values of 0 and 1 are not cases one runs into in such fields. However, events with probability values of 0 and 1 are conceptually important, and unfortunately Kolmogorov's definition fails to capture our intuitions about mutual independence in such cases. For example, an event with the probability value of 0 or 1 ends up being independent of itself since $P(A\&A) = P(A)P(A)$ in both cases. Not only that, but furthermore, according to this criterion, any event with probability of 0 is independent of its negation, anything that entails it, and anything that it entails (Hájek, 2010; Fitelson and Hájek, 2014).

Such problems about the standard formula stated above, in my opinion, lead to two conclusions. First, since the standard formula does not capture our intuitions about conditional probability statements especially for the events with probability values at the limits, it cannot be considered as a satisfactory definition of the concept of conditional probability. Thus, we need to search for a satisfactory definition for conditional probability. Second, since the standard formula works very well for the probability values strictly between 0 and 1, a satisfactory definition of conditional probability must imply the standard formula, but not the other way around. In other words, given a satisfactory definition, the standard formula must be a theorem of the theory of probability together with its axioms. These two conclusions, in turn, imply the desiderata for a satisfactory definition of conditional probability that will be stated in Section IV below.

For now, let us turn to the debate between Edgington and Lowe in order to see if we can find a satisfactory definition for conditional probability in that part of the literature.

III. A Bit of History and Lowe's Definitions

Dorothy Edgington in her lengthy survey "On Conditionals" (1995), following Ramsey's footsteps, attempted to give an explication of conditional statements in terms of the conditional probability of the consequent of the statement given the antecedent. While doing that, she interpreted probabilities as degrees of belief, and a derived notion of degrees of belief-under-a supposition. Lowe in his 1996 article raised a serious charge against Edgington's attempt. His charge was that the very concept of conditional probability is defined by a conditional statement, and using it to explicate conditional statements would be circular. While presenting his charge, since the standard formula does not include a conditional statement, Lowe had to show that the standard formula is not a definition of conditional probability, or at least not a satisfactory one. After stating several shortcomings of the standard formula (the details of which are not crucial for my aims in this paper) he says the following: "I conclude that we do need to define conditional probability." In that article, Lowe suggested the following conditional statement as the definition of conditional probability[5]: "The

[5] He examines three alternative definitions and ends his analysis with this definition as the one he favors.

conditional probability of *B* given *A* is the probability which *B* has/would have if the probability of *A* is/were 1" (p. 609). Let me call this definition **L96** and use Lowe's symbolism of the definition instead of its verbal expression.

(L96) $P_A(B) =_{df} (\text{the } x)\ [\text{if } P(A)=1, \text{then } P(B)=x]$

Since this definition does not provide a way of calculating the value of x, Lowe combines this definition with the standard Bayesian assumption so that the standard formula is implied by **L96.**

> This definition results in the same value for the conditional probability of *B* given *A* as does the standard ratio-based definition [i.e. **SF**] of conditional probability…when that is taken in conjunction with the usual "Bayesian" assumption that we should update our subjective probabilities according to the principle of "conditionalization". According to the latter principle, the value that (*B*) should have when *A* is discovered to be true is the value that (*B* | *A*) [the probability of *B* given *A*] had beforehand (ibid, p. 609).

Edgington (1996), in her response to Lowe, points out two main shortcomings of the principle of conditionalization. The first one is: "Counterexamples to the principle of conditionalization would be ten a penny, were it the foolhardy recommendation that always, on learning *A*, you make your new *P(B)* equal to your old *P(B given A)*. In fact, it recommends this only if *A* is all that you presently learn which is relevant to *B*" (p. 624). Edgington is right here; one may learn more than just *A* or one might have some other background information that becomes relevant to *B* upon learning *A*. This is exactly why Lewis (1976) suggested his idea of Imaging as an alternative to Bayesian Conditionalization. In Lewis' Imaging, upon learning *A*, the probability of *not-A* is transferred to only one alternative, to the possible world that is closest to the *not-A-world*. In the principle of conditionalization, the probability of *not-A* is distributed across all remaining possible worlds in proportion to their previous probabilities[6]. In Lewis' Imaging the background knowledge about the distance of *A-worlds* to *not-A worlds* becomes relevant upon

[6] It should be obvious that Imaging and Bayesian Conditionalization give different results. Interested readers may work out the following example, which is borrowed from Lewis (1976). Suppose that there are three possible worlds with equal probabilities. The first two worlds are A-worlds, whereas the third is a not-A-world. The second world is the closest to the third one.

learning *A*. The second shortcoming that Edgington mentions is that the principle of conditionalization is a normative principle and people do not always abide by it. On the basis of these two shortcomings, she concludes that "Lowe's definition is inadequate and beyond repair." Although I agree with her analysis of the principle of conditionalization, I disagree with her conclusion. As I discuss below, there is a different interpretation of Lowe's definition that is adequate and in need of no repair.

The debate between Edgington and Lowe went silent for more than a decade. In 2008, however, Lowe made another contribution to the debate. In his article "What is 'Conditional Probability'?", after acknowledging the difficulties afflicting his previous definition **L96**, Lowe retracts that definition and proposes the following conditional statement as a replacement. I shall call this new definition **L08**.

> **(L08)** For any proposition B, $P_A(B)$ is the probability that B has if a probability of 1 is assigned to A and for any propositions C and D that entail A, the ratio (P(C))/(P(D)) is left unaltered in value.

Now, as I show below, this definition with a small revision actually meets all the required desiderata for a successful definition. Before proceeding, however, there is a further point that I would like to make. Lowe did not have to give up on his earlier definition and replace it with the new one. **L96** can be interpreted in such a way that it becomes complementary to **L08**. **L96** can be interpreted as specifying the class of all probability revisions upon new evidence. As briefly stated above, there are several different ways of probability updates; some examples from the literature are the principle of conditionalization, Lewis' Imaging, and Gärdenfors' Generalized Imaging, and one could come up many more different ways. Thus, **L96** can be interpreted as defining that class and not a particular member of that class. Under this interpretation, one does not have to add the Bayesian assumption, which was the very assumption that caused Edgington's concerns. Moreover, because of being the definition of a class, **L96** does not have to imply the standard formula, which extensionally corresponds to only one member of that class: the principle of conditionalization. Now, if **L96** defines the class of all revisions, what is **L08? L08** defines a particular member of the class. This particular member, as proven by Williams (1980), is the one that makes the most minimal revision upon new evidence[7]. To repeat, **L96** defines a class, and **L08** defines the global minimum of that class. Under this interpretation, **L96** and **L08** are complementary to each other and, thus, Edgington's assessment of **L96** as being inadequate and beyond repair no longer applies.

IV. Revision, Desiderata, and Proofs

At first glance, it looks like **L08** runs into similar problems when the probability of the conditioning event A is 0. For example, one may raise the following objection:

> **Objection:** Assume that *P(A)* is 0 and instantiate *C* and *D* in **L08** with *(A&B)* and *A*, respectively. Since both *(A&B)* and *A* entail *A*, given **L08**, the original value of the ratio of the probability of these two propositions must be left unaltered in value, but the original value is undefined since *P(A) = 0* by our assumption, and *P(A&B)* is also 0 because of that. **L08** also leaves all the ratios, which are undefined in the standard formula, as undefined. Thus, **L08** fares no better than **SF** when it comes to conditioning on zero-probability events/propositions.
>
> **Reply:** The objector's reasoning misses a crucial point that is implicitly stated in the very last part of **L08**: "…the ratio *(P(C))/(P(D))* is left unaltered in value." In order to be "left unaltered in value", the ratio must have a value to begin with. Since any ratio that is undefined does not have a value assigned to it, the ratios that are undefined in the original distribution do not fall under the ratios that are supposed to be kept with the same value. Thus, the objection does not work against **L08**.

Now, this reply may not convince the skeptic objector. In such a case, it is quite easy to revise **L08** slightly and block the objection in a more straightforward manner. We could revise **L08** by qualifying "the ratio *(P(C))/(P(D))*" with "if it is defined". For undefined ratios, whatever the antecedent of **L08** implies will apply. Our slightly revised version, **L08′**, will read as follows:

[7] To be more specific, Williams proved that Bayesian Conditionalization, which is defined by the standard formula, gives us the most minimal revision. In the following section, I show that L08 implies the standard formula, and thus L08 represents the global minimum of the class of all revisions.

(L08′) For any proposition B, $P_A(B)$ is the probability that B has if a probability of 1 is assigned to A and for any propositions C and D that entail A, the ratio $(P(C))/(P(D))$, if it is defined, is left unaltered in value.

L08′ has the potential of being a satisfactory definition of conditional probability. In order to show that, let me first state a list of desiderata for a satisfactory definition, which follows from the conclusions drawn in Section II.

(D1) The definition must be in line with our conceptual intuitions about conditional probability. A representative set of such intuitions, which is borrowed from Hájek (2003), is the following: for all values of (A) between 0 and 1, inclusive,

$P_A(A)=1$
$P_A(not\text{-}A)=0$
$P_A(F)=0$ *where* F is a necessarily false proposition and A is not,
$P_A(T)=1$ *where* T is a necessarily true proposition, i.e. *not-F* .

(D2) The definition must imply the standard formula when $P(A) > 0$
(D3) $P_A(B)$ must be defined for all values of (A).
(D4) The definition must not be implied by the standard formula.
(D5) The definition should work well with an independence criterion so that no proposition, regardless of its probability value, should turn out to be independent from itself.

Proof of (D1):

The proof here is trivial; just replace B in **L08′** with A, *not-A*, F and T, respectively. Under the assumption that $P(A)$ is 1 as **L08′** requires, $P(A)$ is 1, $P(not\text{-}A)$ is 0, $P(F)$ is 0, and $P(T)$ is 1.

Proof of (D2):

1. $P(A)>0$	Assumption
2. $(A\&B)$ entails A and A entails A	
3. $(P(A\&B))/(P(A))$ is defined	From *1* and *2*
4. $(P(A\&B))/(P(A))=(P_A(A\&B))/(P_A(A))$	From **L08′** and *3*
5. $(P(A\&B))/(P(A))=(P_A(A\&B))/1$	From *4* and **(D1)**

6. $(P(A\&B))/(P(A))=(P_A(B))/1$ From 5
7. $(P(A\&B))/(P(A))=P_A(B)$ From 6

The equation in (7) is nothing but the standard formula[8]. Thus, **L08′** implies the standard formula.

Proof of (D3):

From **(D2)**, we know that **L08′** defines conditional probabilities when *(A)* > 0. Thus, here we only need to show that **L08′** defines conditional probability when the probability of the condition is 0. As we stated, events/propositions with zero probability come in two types: impossible events/necessarily false propositions and improbable events/propositions that correspond to such events.

Type 1: The conditioning proposition, *A*, is a necessarily false proposition. This has three cases. The conditioned proposition, *B*, can be i) a necessarily false proposition, ii) a necessarily true proposition, or iii) a contingent one.

i) $P_F(F)=1$ Trivial, see **(D1)**
ii) $P_F(T)=P_F(not\text{-}F)=0$ Trivial, see **(D1)**
iii) $P_F(B)=?$

This case is a bit tricky. It is generally accepted that the value here has to be 1[9]. Here we can reason in the following way by using event terminology. If an impossible event were to happen, then any other possible event would also have happened. This reasoning, which I admit is not very strong, allows us to assign 1 to the probability.

Type 2: A is a proposition that represents an improbable event. We again have three cases.

i) $P_A(F)=0$ Trivial, see **(D1)**
ii) $P_A(T)=1$ Trivial, see **(D1)**
iii) $P_A(B)=?$

[8] Allow me to state two clarifications about the proof. First, line 3 follows from lines 1 and 2, because line 3 is a ratio and the only way for a ratio to be undefined is when its denominator is zero. We know from line 1 that the denominator of the ratio in question is greater than zero. Second, line 6 follows from line 5 simply because of the logical equivalence between (A & B & A) and (B & A) without needing any further assumption. To see this, the interested reader may apply the standard formula to the two formulas in question.

[9] See Popper, 1959.

> This case is also a bit tricky. Since A is an improbable event, it corresponds to an infinite sequence. With the same reasoning presented above, we could assign 1 to the probability. If B itself is an improbable event, i.e. corresponds to another infinite sequence, it will entail *not-A*, in which case the value will be 0 as proven in **(D1)**.

Proof of (D4):

The standard formula does not imply **L08′** since **L08′** defines conditional probabilities where **SF** leaves them undefined as proven in **(D3)**.

Proof of (D5):

> For mutual independence, I use the following criterion: A and B are mutually independent if and only if $P_A(B)= P_{(not\text{-}A)}(B)$. As we know from the literature, this criterion is equivalent to the two criteria discussed in Section II whenever (A) is strictly between 0 and 1. The problem arises for the two criteria previously discussed when (A) is 1 or 0. In such cases, events/propositions end up being independent from themselves. Given **L08′**, the criterion I use does solve this problem.

1. A and B are mutually independent iff $P_A(B)= P_{(not\text{-}A)}(B)$	Independence
2. $P_A(A)= 1$ for all values of $P(A)$	From **(D1)** and **(D3)**
3. $P_{(not\text{-}A)}(A)= 0$ for all values of $P(A)$	From **(D1)** and **(D3)**
4. $P_A(A)\neq P_{(not\text{-}A)}(A)$ for all values of $P(A)$	From 2 and 3
5. A is not independent from itself for all values of $P(A)$	From 1 and 4

V. Conclusion

There is something odd about treating the standard formula as the definition of the concept of conditional probability. It just gives us an equation with which values of conditional probabilities can be calculated. The technique of how to calculate the value of an instantiation of a concept is not a definition of that concept. For example, $P=F/S$ provides us only with a formula for calculating the values of pressure given a specific set of conditions in chemistry. But this is not a definition of pressure. The definition, as we all know, is the force applied to per unit area. The definition and the formula are closely related, because if the definition is a true definition of the concept, then the formula should follow from the definition directly. This is what happens in the case of pressure. Once you know the definition, there is no need to memorize the formula; it falls out of the definition with no need for any further premise. The standard formula for conditional probability is analogous to $P=F/S$; it should directly follow from the definition of the concept of conditional probability. In this paper, I have shown that a slightly revised version of Lowe's 2008 definition has this property. It not only implies the standard formula but also allows us to make sense of the concept even for probability values at the limits. If my reasoning in the two tricky cases in the proofs is accepted, then, I dare to say, we should call the revised version of Lowe's 2008 definition the definition of conditional probability.

This has at least three crucial implications.

I1. One does not need to reverse the direction of analysis in the theory of probability, i.e. take conditional probabilities as primitive and define single event probabilities via conditional probabilities, in order to solve the problems afflicting the standard formula.

I2. Since the definition itself is a conditional statement, using conditional probabilities for explicating conditional statements/judgments is in danger of circularity. It seems that Lowe is right, after all, in his assessment of Edgington's work on conditionals.

I3. Including **L08**′ in the textbooks on the theory of probability as the definition of conditional probability and proving the standard formula as a theorem will provide a more firm conceptual foundation for the theory of probability. I am not sure, however, how much attention mathematicians will pay to what philosophers say.

Hilmi Demir, Bilkent Üniversitesi, Türkiye

References

Bayes, Thomas. (1764). An Essay Towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions of the Royal Society of London* 53, pp. 37-418.

Fitelson, Brandon and Hájek, A. (2014). Declarations of Independence. *Synthese*. 10.1007/s11229-014-0559-2.

Hájek, Alan. (2003). What Conditional Probability Could Not Be. *Synthese* 137, pp. 273-323.

Hájek, Alan. (2010). *A Plea for the Improbable*. AAP presidential address.

Hájek, Alan. (2011). Conditional Probability. *Handbook of the Philosophy of Science, Volume 7: Philosophy of Statistics*, Eds. Prasanta S. Bandyopadhyay and Malcolm R. Forster. pp. 99-137.

Kolmogorov, A. N. (1933/1950). *Grundbegriffe der Wahrscheinlichkeitrechnung, Ergebnisse Der Mathematik*. Translated as Foundations of Probability. New York: Chelsea Publishing Company.

Lewis, D.K. (1976). Probabilities of Conditionals and Conditional Probabilities. *Philosophical Review* 85, pp. 297-315.

Lowe, E. J. (1996). Conditional Probability and Conditional Beliefs. *Mind* 105, pp. 603–15.

Edgington, Dorothy. (1995). On Conditionals. *Mind* 104. pp. 235–329.

Edgington, Dorothy. (1996). Lowe on Conditional Probability. *Mind* 105, pp. 617–30.

Gärdenfors, Peter. (1982). Imaging and Conditionalization. *Journal of Philosophy* 79, pp. 747-60.

Popper, Karl. (1959). *The Logic of Scientific Discovery*, Basic Books.

Williams, P.M. (1980). Bayesian Conditionalisation and the Principle of Minimum Information. *British Journal for the Philosophy of Science* 31 (2), pp. 131-44.