The Classical Theory of Probability and the Principle of Indifference

Abstract: The classical theory of probability bases all probability distributions on assigning equal probabilities to equally possible alternatives. This theory has been claimed to be both inconsistent and circular.

Related to the classical theory is the principle of indifference. This principle states that if we have no reason to think one alternative is more likely than another – then they should be assigned an equal probability. It is largely because of the paradoxes raised against this principle that the classical theory is perceived to be inconsistent.

In this paper, I examine the relations between the classical theory and the principle of indifference in light of the inconsistency claims raised against both. I claim that these claims are mostly unjustified, especially with regards to the classical theory. Finally, I claim that the classical theory is not circular.

A. The classical theory of probability[1]

The historical birth of probability in the 17th century was centred primarily on gambling and games of chance. These settings were a natural background for the development of what is nowadays known as the classical theory of probability.

The classical theory begins by considering a set of equally possible alternatives and assigning each of them an equal probability. The probability of any other event is then calculated as the ratio between the alternatives falling under this event and the original set of alternatives. For example, when throwing a fair cubic die there are six basic equipossible cases, namely the die falling on each of its six faces. We then can deduce the probability of a more complicated event such as "the die will fall on a face with an even number", by noticing that half of the original alternatives fall under this event, and thus the probability is one half.

The most basic principle of the classical theory is, then, assigning an equal probability to a set of equally possible alternatives, and classical theorists commonly used this "equal possibility" terminology[2]. This principle is often referred to as "the classical

[1] I am grateful to Reshef Agam-Segal, Dafi Agam-Segal, and Hemdat Lerman for their helpful comments on this paper.

[2] It is commonly attributed to Laplace (early 19th century), but it can be traced back to much earlier thinkers, such as Leibniz (17th century), see [Hacking], pp. 122-133

definition", conveying the idea that "probability", according to the classical theory, is basically *defined* according to the principle[3].

Many view the classical theory as unfeasible. Here is, for example, a standard comment on the theory:

"*The modern mathematical treatment of probability owes its origins to Pascal's treatment of games of chance, and the classical equipossibility theory arises most naturally in this context… But paradoxes arise where there are different, equally possible candidates for the set of equally possible outcomes. And in defining probability in terms of equal possibility the theory runs into circularity"*[4]

The main objection against the theory can thus be divided in two claims[5]:

A.) <u>The theory is inconsistent</u> – There are many paradoxes that apparently show situations where different ways of phrasing what the initial set of equally possible cases are, will lead to different, incompatible probability distributions. (I shall discuss some of these paradoxes in section E).

B.) <u>The theory is circular</u> - The term "equally possible" simply means "equally probable" – so by defining equal probability in terms of equal possibility the theory runs into a vicious circle.

In this paper, I shall try to claim that both these objections are unjustified.

---

[3] It is important to note, though, that the basic principle can only be viewed as defining the term "equal probability". To use it as a complete definition of "probability" we must at least require that the axioms of probability be added.

[4] [OCP], p 720.

[5] From a modern perspective, there are two additional objections against the theory: It cannot apply to probability considerations that start with a fundamental set of cases that are not equally possible (such as tossing a biased coin), and it can not apply to infinite probability spaces.
I think it is an open question whether the classical theory can be naturally expanded to deal with these cases, but in any case I will not treat these objections in this paper. This seems to me to be a reasonable strategy, because both these objections question the scope of the theory, and not – as the objections raised above – its very feasibility.

B. Some preliminary notes on the philosophical interpretation of probability

Before I proceed, I would like to make some preliminary notes on the use of the word "probability" in this paper. What do I mean by "probability"? Obviously, the mathematical axiomatic approach to probability will not do in this context. All the axiomatic approach claims is that functions satisfying certain axioms, also satisfy certain mathematical properties. It makes no claim regarding how to interpret probability claims, or what they have to do with possibilities and likelihood, and therefore cannot serve as an interpretation of probability when discussing the feasibility of the classical theory.

I must therefore turn to the question of the philosophical interpretation of probability. However, this question is part of a vast philosophical debate, far beyond the scope of this paper. I will therefore try to adopt a relatively neutral approach, distinguishing between the different interpretations when necessary.

In order to give the discussion some unified framework, let me use a relatively neutral definition of probability based on that of Gnedenko[6]: Gnedenko discusses a set of background conditions G, and an event A. The three basic cases are:

1.) G is sufficient for the occurrence of A – in this case we will say that (given the conditions) A is *necessary*.
2.) G is sufficient to prevent the occurrence of A – in this case we will say that A is *impossible*
3.) Neither (1) nor (2) are true – in this case we will say that A is *random*.

According to this picture, the theory of probability attempts to give some (numerical) measure of the "level of randomness" of A. But what do we mean by the "level of randomness"? It seems to me that we can adopt the following view: The level of randomness measures the degree of connection or necessity between G and A. The different approaches to probability differ as to what kind of connection is assumed: A physical connection will yield the propensity theory; a circumstantial connection of empirical co-occurrence will yield the relative frequency theory; A logical connection

[6] See [Gnedenko], pp. 18-25.

will yield the logical theory; and a psychological connection will yield the subjective theory.

How does the classical theory fit into this picture? Some view the classical theory to be an additional interpretation to those mentioned above[7]. I think this view is misled: The classical theory is on its own quite neutral, and can be seen as compatible with most, if not all, other interpretations.

The classical theory bases the concept of probability on the concept of possibility. This still leaves open the question of what concept of possibility we are assuming: a physical possibility? Logical possibility? Psychological possibility?

In particular, we should note that the basic duality of probability interpretations between *ontological interpretations* and *epistemic interpretations* is already present in the concept of possibility[8]. Roughly speaking, epistemic interpretations of probability, view probability as concerned with human knowledge or belief. Ontological interpretations, view probability as concerned with the state of objects in the external world, independent of human beings.

Prima facie, it seems that both probability and possibility statements are sometimes understood in an epistemic and sometimes in an ontological way. For example, statements like "it is probable that the queen of England is now drinking tea", or "it is possible that it is now raining outside" are usually read as epistemic statements, while statements such as "the probability of this die to fall on a 1 is…" or "it is possible for the queen to drink some tea today", are read as ontological[9].

Other statements seem ambiguous between the two readings: "It is possible that it will rain tomorrow", may be used either to say that (in a non-deterministic world) there is

[7] See for example, [Fine], p. 9.
[8] In fact, Hacking claims that the basic duality in probability is historically inherited from the basic duality of possibility, and this inheritance has happened exactly because the classical theory based the concept of probability on the concept of possibility (see [Hacking 2] and [Hacking], pp. 122-133.
[9] This distinction is related to the distinction between *de dicto* and *de re* readings of statements in the following way: *de re* readings of possibility (probability) statements are always ontological, while epistemic interpretations must always be read as *de dicto* statements. Whether the converse relations hold, is open to debate.

a physical (or metaphysical) possibility that it will rain, or to say that regardless of metaphysical possibilities, for all we know, it may either rain or not. A parallel ambiguity occurs for the case of probabilities.

Note that this distinction should not be confused with the distinction between *subjective* and *objective* interpretations of probability. Subjective interpretations are interpretations that allow for two individuals faced with the same evidence, to assign different probabilities to the same event. Objective interpretations do not allow this, and are hence objective in the sense of being *inter-subjective*. Thus the logical interpretation, for example, is an epistemic, but objective interpretation.

C. The principle of indifference

Related to the classical theory, is the principle of indifference. Roughly, this principle states that given a set of alternatives, if we have no reason to think some are more likely than others then we should assign them all an equal probability. This principle is interesting in this context because, as we shall see, the claims for the inconsistency of the classical theory are directly related to claims that the principle is inconsistent.

A close look at the literature shows that there are at least half a dozen non-equivalent formulations of the principle[10]. I would like to focus on the difference between types of these formulations: formulations in which the principle *recommends* us to accept a certain probability distributions, and formulations in which the principle *claims* that a certain probability distribution is true.

Under the first type, we find formulations such as the following: "Assume that two alternatives are equally probable if you do not have any reason not to do so".[11] Such formulations present *action guiding principles*. They do not claim that in the cases mentioned the probabilities *are* equal – they simply direct us to *assume* so. In these cases, the principle can at best be viewed as a rationality principle.

---

[10] For possible formulations see [Fine],p. 167, [van Fraassen], p. 299, [Gnedenko], p. 38, [Howson &Urbach] p.38.

[11] see [Fine] p. 167

Under the second type we can find formulations such as this: "A set of alternatives will have equal probability if there is symmetry between the relevant evidence in favour of each of them"[12]. Such formulations purport to establish what the probability distribution actually is[13]. It seems that these formulations are more presumptuous – for according to them, the principle can provide us with new knowledge.

The main objection raised against the principle of indifference is that it leads to paradoxes, and is therefore inconsistent. But what exactly does this claim amount to? I think the answer differs for the two types of formulations of the principle discussed above.

<u>Objective probability formulations</u>: In these cases, the problem of inconsistency arises, if it will turn out that indeed there is some situation in which the principle directs to accept contradictory probability distributions.

The discussion of the consistency with regards to this type of formulations requires a very cautious consideration of how exactly the principle is presented. It also requires a cautious consideration of what is the interpretation of probability assumed. For example, understanding the principle as providing objective logical probabilities seems much more feasible than as providing objective physical probabilities: This can be seen if we consider a second objection raised against the principle of indifference: that it produces knowledge out of ignorance. Such an objection seems much graver, if we view the principle as yielding *physical* probabilities based on *epistemic* ignorance.

The above considerations led to an extensive philosophical debate regarding the consistency of different complicated and delicate formulations of the principle – a debate which I will not further discuss.

<u>Action guiding formulations</u>: How can such formulations be inconsistent? One reply may be that there is a situation, in which the principle does not determine a unique

[12] see [Keynes], p. 60 – this is a praphrase rather than a direct quote.

[13] Such a claim can be made even with regards to the subjective interpretation to probability, though in that case, in order to make sense of the claim "the probability of e is x" (rather than "…is x for subject s"), a further argument will be needed to show why *all* rational subjects will have an equal probability assignment for the specified event. It is therefore unlikely that the proponents of the subjective theory will adopt the principle of indifference (under this reading) universally.

course of action – i.e. it may direct us to two incompatible ways of acting. Understood this way, these formulations are at least as venerable to inconsistency as principles of the previous type.

However, I do think such cases to be correctly characterized as inconsistent. It seems to me perfectly reasonable that an action guiding principle will give me some recommendation as to how to act. It may not suggest a unique course of action, nor does it have to be the only action guiding principle I am using.

More specifically: We may view the principle of indifference not as directing me to a unique probability function, but as limiting my considerations to a restricted set of probability functions – namely, those that can be obtained by some application of the principle. In so far that I have applied the principle, I have acted rationally – regardless of whether there was another alternative way of applying the principle. Nor must this be the only rationality principle I should follow. If someone pointed a gun at me, and directed me to assume that the probability of event A is ½, the rational course of action is probably following her directions rather than following the principle of indifference!

There is of course, much room for discussing why the principle is rational. This discussion may be philosophical as well as involving theories of economics, psychology and game theory, but is not part of our present concern.

D. The relation between the classical theory and the principle of indifference

The classical theory of probability is often conceived as deeply related, or even as yet another formulation of the principle of indifference[14].

This relation seems quite natural, when we remember that the classical theory is based upon the principle that equally possible cases are equally probable. Like the principle of indifference, the classical theory offers a sufficient condition for a set of alternatives being equally probable[15].

---

[14] See for example [van Fraassen], p. 293, [Hacking], p. 122, 126 and [Fine], p. 19.

[15] This relation, is only natural when considering the "objective probability formulations" of the principle, and it is to these formulations that I will refer in this section.

Some have viewed the principle of indifference as an attempt to provide the classical definition with content – i.e. to provide some criteria for applying the definition[16]. The main motivation for such a move seems to be the following: The classical theory is viewed as *circular* (for equal possibility *is* equal probability). The principle of indifference is supposed to give a better definition of what we mean by "equally possible", thus rescuing the classical theory from its circularity.

However, the principle of indifference turns out to be a poor solution, for it is (viewed by many as) inconsistent, and therefore by using it to explicate the classical theory – the theory becomes inconsistent as well. According to this view then, the classical theory is *either* inconsistent *or* circular.

But must we accept this picture? I think not. My claim is that we should view the classical theory as independent, and deal with the circularity claim without appealing to the principle of indifference. The principle can, at best, supply us with some mode of establishing the equipossibility of cases, but it isn't necessarily the only mode, nor must it be viewed as playing a part in defining the classical theory.

Nor should the principle and the theory be equated. Under all its formulations the principle of indifference talks about some kind of epistemic ignorance or indifference regarding our possibilities. The same is true of the classical theory, only if we understand the modality stated in it as epistemic as well. As we have seen – this is not the only interpretation of the theory: it may be interpreted as relating either to epistemic or to ontologically equally possible cases.

Moreover, even if we do understand the modality in the classical theory as epistemic, we will accordingly get (as the theory is after all a definition) an epistemic notion of probability. The principle of indifference, on the other hand, can be interpreted as referring to either epistemic or ontological probabilities.

---

[16] see for example [Keynes], p. 13.

E. Some case studies

The discussion regarding the feasibility of the principle and the theory is centred on specific cases of their application. In the remainder of the paper I wish to present three of these cases, and discuss to what extent they actually do make a case for or against the principle and the theory.

Case 1: tossing a symmetric die

The following example is often taken as a positive case, which supposedly shows a useful and correct application of the principle of indifference. It may help us understand why so many have found the principle compelling in the first place.

Suppose that we toss a symmetric cubic die. What is the probability that the die will fall on the number 5? The standard reply is, of course, 1/6. But how have we reached this conclusion?

One reply will be that we have established this by empirical testing[17]. However, such a reply seems to me obviously wrong. Most of us are willing to accept the probabilities regarding the fair die without conducting numerous experiments. More over, the number of tosses in any such experiment is finite – while the probability (even if taken as a relative frequency claim) makes a claim about an infinite limit. This means that such empirical results can at best be viewed as scientific evidence, but cannot explain the amazing degree of certainty with which we seem to accept the claim. Finally, consider a dodecahedron shaped die[18] (a symmetrical die with twelve faces). Even if you have never seen such a die, you are probably quite certain that the probability of the die falling on each of its faces is 1/12.

It seems then, that we can conclude the probabilities for the die using *a priori* reasoning. But how can we accomplish that, if not by using the principle of indifference? Since all the faces are symmetric, we have no reason believe that the die will fall on one face rather than on another. The principle of indifference therefore tells us to give the die's falling on each of the faces an equal probability.

---

[17] Indeed, similar results for coin tossing are corroborated by empirical testing -see [Kerrich] (Kerrich has conducted an experiment of tossing a coin 10,000 times, while he in captive during world word II).
[18] This example is also given in [Strevens]

The principle of indifference seems to appear here with all its glamour. Using (epistemic) indifference we were able to conclude (physical) probabilities – and this result seems both useful and corroborated by empirical testing! Nonetheless, I wish to claim that it is *not* the principle of indifference that justifies our probabilities with regard to the die.

Strevens[19] presents two arguments in favour of this claim. First, he claims, the principle if indifference is not the right *kind* of principle to apply in this case, for it is an action-guiding rationality principle, but the probability claim regarding the die is a claim about true physical probabilities. As we have seen, this claim is true only when assuming the action guiding formulations of the principle. Furthermore, it seems to me that the force of the die case lies in the very fact that it argues in favour of the stronger, "objective probability" formulations of the principle.

Strevens' second argument is more relevant: Suppose that we toss a symmetric die one million times, and it falls on the face marked with "1" all of these times. Supposedly, we have obtained some important information with regards to the die, and we cannot continue claiming that all its faces are epistemically symmetric as far as we are concerned. Nonetheless, even after learning these results, we will not change our initial probability claim. Either we will assume that the die is not, as we originally thought, *physically* symmetric, or (if we are absolutely sure that it is), we will simply conclude that a very improbable even has actually happened. Our initial belief that a symmetric die falls with an equal probability on each of their faces will not change.

Strevens goes on to give an alternative explanation to why we do assign an equal probability to each of the faces. His claim is that our reasons rely on very specific knowledge of the physical qualities of the die, together with our knowledge of certain features of our physics.

[19] see [Strevens]

I will not go into the details of his account but I think the general direction is correct. Our initial set of conditions *G* includes the fact that we are tossing a physically symmetric die. The necessity we are considering is a physical necessity. The reason we know the die will fall with equal probability on each of its faces is that we know the die is symmetric with regards to certain physical qualities (such as geometric shape, and weight distribution), and our physics is such that it will respect such symmetries.

These symmetry considerations must not be confused with the principle of indifference. The principle uses only our *epistemic* symmetry to conclude the probabilities. The different use of such epistemic considerations can be seen in the following example: Imagine that the outcome of a "toss" is determined by the will of a demon. Suppose we know nothing about the priorities and preferences of the demon. Will we be equally inclined to assume that the probability of the die "landing" on each of its faces is equal? Even if the answer is positive, we will not do so with the same certainty as in the previous case. Moreover, (to apply Strevens' test), if our die "falls" on "1" in a million consecutive "tosses", I think we will see this as evidence that our demon has a special attachment to the number "1", and we will assign to it a higher probability than to the other faces.

I have concluded that the die example does not show an application of the principle of indifference. As opposed to this, I think it can definitely be seen as an application of the classical theory. In this case, the classical theory tells us that physically equally possible cases should be assigned an equal (physical) probability. Of course, the classical theory does not *prove* that our initial cases are equally possible, but nor should it do so. That is, after all, the task of physics.

Case 2: The second Ace problem

Although this problem is not a classic "principle of indifference paradox", I have chosen to present it, mainly because I think its solution illuminating to this discussion.

The puzzle goes as following: Alice has a pack containing three cards: A black ace, a red ace and a two. Alice arbitrarily chooses two out of three of these cards. There are three possibilities to what Alice chose, so Bob applies the principle of indifference (or, for that matter, the classical theory) and concludes that the probability of Alice having each of these configurations is 1/3. In particular, he concludes that the probability that Alice has both aces is 1/3.

Suppose that now Alice (sincerely) says to Bob: “I have the red ace” (denote this statement by *R*). Bob now reduces the possibilities to two, and concludes (using the principle of indifference) that the probability that Alice has both aces is ½. Alternatively, supposed that Alice actually says “I have the black ace” (denote this statement by *B*) - using a similar consideration, Bob concludes that the probability that Alice has both aces is ½.

The puzzle is therefore this: Even before Alice made her statement (*B* or *R*) Bob knew that she must have had at least one of the aces. The situation of this ace being either black or red is completely symmetric with regards to the question of whether Alice has both aces or not, thus the colour of the ace should not provide any relevant information. Why then does Bob change his probability of the two aces case after Alice makes her statement?

To emphasize the problem: Suppose that Alice actually says nothing to Bob. Bob can make the following thought experiment: “Since Alice has at least one ace, she can sincerely make at least one of the statements *B* and *R*. In either case, I will be willing to change my probability of the two ace case to ½.” He therefore concludes that even though Alice has said nothing, the probability of Alice’s having two aces is actually ½ and not 1/3 as he originally thought. Different applications of the principle, then, lead Bob to accept different probability assignments to the same problem.

A deeper analysis of the problem[20] reveals a different solution. The key to solving this paradox is to bring out the question of *which protocol* (or action strategy) is Alice following when she makes her statement. Let us examine two possible protocols:

<u>Protocol A</u>: If Alice has the black ace, she makes statement *B*. Only if she doesn’t, she makes statement *R*. In the latter case (assuming that Bob knows the protocol), Bob

[20] Arguments by Joe Halpern are taken from lecture notes of his course “reasoning about uncertainty” – currently yet unpublished.

knows that the probability of both aces is zero. In the former case, Bob can conclude that she has both aces with probability ½. If this is the protocol used, the symmetry is broken, and Alice's statement *is* informative.

Protocol B: If Alice has only one ace, she makes the only possible true statement of the two. Otherwise (if she has both aces), she tosses a fair coin, and according to the result of the toss makes either statement *R* or *B*. A simple analysis of the cases shows that whichever statement Alice makes, the probability of her having both aces remains 1/3. While the symmetry between the two cases remains, Bob does not change his initial probability.

In fact, Halpern shows that in any possible protocol used by Alice, the paradox will not arise. This solution emphasizes the need for careful applications of the principle of indifference, taking the protocols used into account.

This observation leaves many open questions: What should we do if we do not know the protocol? What counts as a detailed enough protocol or specification of the problem? I do not purport to have a general answer to these questions.

Rather, I wish to claim that the classical theory of probability can only be applied correctly when we actually *know* that we have a set of equally possible cases. Understood this way, the theory (or for that matter the principle) is completely consistent – though it is not always applicable. If applied recklessly, contradiction can indeed arise.

Case 3: The paradox of the balls in the urn

This paradox, suggested by Keynes[21], is one of the many classical paradoxes raised against the principle of indifference[22].

[21] See [Keynes], p. 53. I am here giving a simplified version of Keynes' paradox.

[22] Such paradoxes include Bertrands paradox, Buffon's needle problem (presented as a paradox), the book paradox and the cube factory paradox. My claim is that all these paradoxes can be analysed in a similar fashion as this one, and - time permitting, I will be happy to show how.

Suppose we have an urn, with three balls in it. All we know is that each of the balls is either white or black. What then is the probability that there are three black balls in the urn?

One application of the principle will be as follows: for each ball, there are two possible cases – either it is white or it is black. Altogether, there are eight equally possible configurations of the balls in the urn. According to the principle then, the possibility of three black balls is 1/8.

However, a different application of the principle states that there are four possible ratios between the balls: either all three are black, one black and two whites, one white and two blacks and three whites. For all we know, all these are equally possible, so the probability of three black balls is rather ¼.

These two contradicting results[23], present a serious challenge to the principle of indifference. Can this contradiction be resolved? When attempting to answer this question we must specify which notion of probability we are assuming, and which principle have we actually applied in the example:

1.) We are assuming an epistemic notion of probability, and we have assigned equal probabilities based on equal epistemic knowledge.

2.) We are assuming an ontological notion of probability, and have assigned equal probabilities based on equal epistemic knowledge.

In both of the above cases, it seems that the only way around the paradox is we try to phrase the principle in a more accurate way, showing that one of the alleged applications was not actually a proper use of the principle[24].

3.) We are assuming a notion of ontological or physical probabilities, and have assumed that (ontologically) equally possible cases are equally probable. How

[23] Interestingly, while the first solution will now be considered as the proper textbook solution, Keynes mentions the latter as the text book solution! (see [Keynes], p. 54).
[24] This direction is perused by Keynes, with, in my view, only partial success.

> can our two previous considerations fit into this framework? Here I think, we must go back to the notion of protocol discussed in the previous case study.
>
> What protocol was used in order to produce our “random urn”? I can imagine at least two: According to one, we have a ball-producing machine. Every time we press a button, the machine flips a fair coin, and according to the result, throws into our urn either a black or a white ball. Our “random urn” is a result of three consecutive presses of the button.
>
> According to the second protocol we have four folded notes, stating each of the four possible ratios. We randomly pick one, give it to the cashier in the urn shop, and in return receive an urn with the appropriate ratio of balls.
>
> The important point here is that in the former case only the first probability distribution will be correct, and in latter only the second will be correct. Without knowledge of the protocol, there is indeed no reason to prefer one solution over the other.

Note that cases (1) and (2) can be seen as applications of the principle of indifference and cases (1) and (3) as applications of the classical theory. I have claimed that in case (3) there is no paradox. The questions whether the direction for solution suggested for (1) and (2) can be adequate or not is still, I think, open.

My general tendency is to think that an idea such as (2) – of inferring physical probabilities from epistemic indifference is indeed likely to lead us into paradoxes, while the ides of using such considerations for deducing epistemic probabilities is probably more likely to be reconciled.

F. Dealing with the circularity claim

Finally, I want to turn to the claim that the classical theory is circular. This claim is based on the following argument: According to the classical theory, “equal probability” is defined in terms of ”equal possibility”. But the *only* way to explicate what we mean by “equally possible” is using “equally probable”, thus the definition is viciously circular.

My objection to this argument lies in rejecting the premise that the concept of "equal possibility" must be prior to the concept of "equal probability". The relations between possibility and probability can rather be viewed as somewhat parallel to the relation between "heaviness" as measured by an un-scaled balance, and "weight" as measured by a scale. Even without a notion of numerical weight, we can say that *a* is as heavy as *b*. We do not have to understand this claim as saying "the numerical weight of *a* is equal to the numerical weight of *b*". Similarly, we may think that it is equally possible for a symmetric die to fall on 1 or on 2, with out going via the claim that the probability of both events is 1/6[25].

Thus, it may be possible that the concept of "equal probability" is more primitive and independent of the concept of "equal probability". To strengthen this, note that the historical appearance of the concept "possibility" preceded the appearance of the concept "probability" by thousands of years. It is highly likely that so did the concept of "equally possible".

It may be objected that even if the above remarks show that the classical theory is not circular, it is at any rate void. If any sentence of the form "*a* is equally probable to *b*" is simply a paraphrase for "*a* is equally possible to *b*", then the classical theory seems completely idle.

Yet the above objection misses an important point. The classical theory consists not only of the classical definition, but also of using the mathematical axioms of probability to expand probability claims to other (non primitive) events. According to my understanding, the intention of this expansion is to use probability to capture an intuitive idea of "the degree of possibility"[26].

To return to metaphor of weights: Even given balances, the notion of numerical weights is not idle, They give a precise notion of the degree of "heaviness". Similarly, the classical theory assumes that we have a notion of being "equally or more possible than…". The classical theory uses this notion to define an exact numerical notion of

[25] Cf. Frege's discussion of the definition of direction: [Frege], pp. 74-80.
[26] This idea is formed by Leibnitz, who claimed that: ""*probablitas est gradus possibilitas"* see [Hacking, 122-133].

"a degree of possibility". Moreover, it does so using a powerful and vast mathematical theory, thus providing us with an effective method of calculating the probabilities of numerous events.

Conclusion

In this paper I have tried to look into the relations between the principle of indifference and the classical definition of probability, especially with regards to the inconsistency claims raised against both.

My conclusion is that taking equally possible cases to be equally probable is not necessarily inconsistent, providing that we remember the following: The classical theory does not provide us with a magic formula for producing probabilities! In order to apply these principles we must first know that a certain set of alternatives is equally possible. This knowledge is by no means trivial – whether it refers to ontological or to epistemic possibilities.

Nor is the theory circular or void. The classical theory assumes a notion of possibility, and uses it to assign probability for an initial set of equally possible cases. It then uses a powerful mathematical theory to define and calculate the probability of numerous other events. We may use this powerful mathematical tool, providing its initial application has been correct.

References

[Fine] Fine, T.L., *Theories of Probability*, Academic Press, New York, 1973.

[Frege] Frege G., *The foundations of arithmetic*, trans. by Austin J.L., Harper Brothers, NY (second revised edition).

[Gnedenko] Gnedenko B.V., *The theory of probability*, Chelsa Publishing Company, New York, 1962.

[Hacking] Hacking I. , *The Emergence of Probability*, Cambridge University Press, Cambridge (U.K.), 1975.

[Hacking 2] Hacking I., 'Equipossibility theories of probability', *British Journal for the Philosophy of Science* 22 (1971), pp. 339-355.

[Howson&Urbach] Howson, C. and Urbach, P., *Sceintific Reasoning: the Bayesian approach,* 2ed edition, La Salle: Open court press, 1993.

[Kerrich] Kerrich J.E., *An experimental introduction to the theory of probability*, Belgisk Import Compogni, Copenhagen, 1950.

[Keynes] Keynes J. M., *A treatise on probability*, Macmillan, London, 1963.

[OCP] Hendrich T. (ed.), *The Oxford Companion to Philosophy*, Oxford University Press, Oxford, 1995.

[Strevens] Strevens M., 'Inferring probabilities from symmetries', *Nous* 32(2) (1998), pp. 231-246.

[van Fraassen] van Fraassen B.C., *Laws and Symmetry* , Oxford University Press, Oxford, 1989.