THOUGHT
A JOURNAL OF PHILOSOPHY

THE NORTHERN INSTITUTE OF PHILOSOPHY

Thought ISSN 2161-2234

ORIGINAL ARTICLE

Cardinality Arguments Against Regular Probability Measures

Thomas Hofweber

University of North Carolina at Chapel Hill

Cardinality arguments against regular probability measures aim to show that no matter which ordered field $\mathbb{H}$ we select as the measures for probability, we can find some event space F of sufficiently large cardinality such that there can be no regular probability measure from F into $\mathbb{H}$. In particular, taking $\mathbb{H}$ to be hyperreal numbers won't help to guarantee that probability measures can always be regular. I argue that such cardinality arguments fail, since they rely on the wrong conception of the role of numbers as measures of probability. With the proper conception of their role we can see that for any event space F, of any cardinality, there are regular hyperreal-valued probability measures.

Keywords probability; measurement; cardinality; hyperreal; regularity

DOI:10.1002/tht3.125

1 Introduction

A well-known concern about measuring probability with real numbers is that sometimes events that happen nonetheless have the lowest possible measure of probability, that is 0. Such measures seem to neglect differences in probability, since they assign the same measure of probability to events that do not seem to be equally probable. Since 0 is the lowest possible measure of probability it is also the measure of impossible and contradictory events, like it's raining and also not raining. But an event that happens seems to be more likely to happen than an event that is contradictory. It is thus natural to think that such measures are not fine enough when we hope to measure every difference in probability. Maybe these differences are not important for many practical purposes, and so maybe real-valued probability measures are unproblematic in practice. Still, one might hope to do better, even though we have learned to live with this situation despite such concerns. We should not disregard the prima facie plausibility that there is a difference in probability between an event that happens and an event of lowest possible probability. When we try to capture all the differences in probability then our measures might well have to be finer. And we know, at least in outline, that we can do better. We can have finer probability measures by making them not real-valued, but hyperreal-valued probability measures. Although this basic idea is well known, it has met some resistance.[1] One of the arguments that hyperreal-valued probability measures won't save us is the

Correspondence to: E-mail: hofweber@unc.edu

topic of this paper. It is in essence a cardinality argument that aims to show that employing hyperreal numbers instead of real numbers as measures of probability just pushes the issue back to higher cardinalities. Eventually we need to accept that events with the lowest possible measure of probability can happen nonetheless. This style of argument was given recently by several authors, including Hájek (2013), Easwaran (2014) and especially Pruss (2013). I hope to show in this article that these cardinality arguments are mistaken.

It is not hard to see how the problem arises with real-valued probability measures.[2] The real numbers have the Archimedean property, which means that for any positive number r, no matter how small, and any other number s, no matter how large, there is a natural number n such that multiplying r with n is larger than s. This in particular means that any positive real number, no matter how small, is still larger than $\frac{1}{n}$ for sufficiently large n. Thus it can only be added to itself finitely many times before the result gets larger than 1. Therefore, no positive real number is small enough to be the measure of the probability of infinitely many equally likely events. In such a case all of these events need to get measure 0, even when one of them will happen. What we thus need, it would seem, is to measure probability differently, employing numbers that do not have the Archimedean property. We need numbers that are not only arbitrarily finitely small, but also numbers that are infinitely small, while still larger than 0.

As is well known, there are such numbers. There are non-Archimedean extensions of the real numbers that preserve all the first-order properties of the real numbers. In fact, there are arbitrarily large such extensions of the real numbers, and even in the same cardinality they can be non-isomorphic. Such non-Archimedean extensions of the real numbers are generally called *hyperreal numbers*. Since these hyperreal numbers share their first-order properties with the real numbers they still form an ordered field. We will call a non-Archimedean ordered field that shares the first-order properties with the real numbers a *hyperreal field*. Hyperreal fields seem to have just the properties that we would have wanted in the real numbers when we measure probability, and so it is natural to suggest that we simply measure probability with hyperreal numbers instead of the real numbers.[3]

Although this basic idea might seem compelling, there has been a remarkable resistance to adopt it. Some of this resistance is due to the unfamiliarity of hyperreal numbers. Although they have the same first-order properties as the real numbers, they don't have the same properties in general; in particular not the same second-order properties. Especially important is that the least-upper-bound principle doesn't hold for hyperreal numbers, even though it holds for real numbers. This principle is crucial in standard definitions of infinite sums, which in turn are relied upon in the definition of countable additivity, a requirement in the standard, real-valued, approach to probability measures due to Kolmogorov. But there is no in principle obstacle to specifying a corresponding notion in a novel way, and there is even the hope, ruled out in the real-valued case, to have not only countable additivity, but arbitrary additivity. This situation simply calls for more technical work.[4] With this issue unresolved we will only require a hyperreal-valued probability measure to be finitely additive, as does Pruss in (2013). There are also slightly more philosophical objections, in particular by Timothy Williamson.[5] And finally there is one group of arguments against hyperreal-valued probability measures which are our focus here. These arguments are all in essence cardinality arguments.

2 The cardinality arguments against regular hyperreal-valued measures

Suppose we replaced the real numbers as the measures of probability with some particular system of hyperreal numbers. In the simplest case we could replace the real numbers with some non-Archimedean extension of the real numbers that has the same cardinality as the real numbers. Can we be assured that in this new setup we will never be forced to assign 0 to an event that can happen? With the intended interpretation of the event space this would mean: Can we be assured that the only member of the event space which has measure 0 is the empty set? A probability measure is called *regular* just in case only the empty set has measure 0. Probability measures are first and foremost mathematical objects, functions from set to numbers. Clearly, some probability measures are regular, while others are not. The measure that assigns $\frac{1}{2}$ to the only event *e* in the sample space is regular, the measure that assigns 0 to *e* is not. The *principle of regularity* is a constraint on what a probability measure has to be like in order to be a correct measure of probability. It states that all measures of probability have to be regular in order to be correct. The principle of regularity is thus a further requirement, besides the uncontroversial ones listed above, on what a probability measure has to be like. Depending on what one takes probability to be more precisely, one might have different reasons for why a measure of probability has to be regular. If probability is subjective probability then there might be one reason to insist on regularity, if it is objective probability there might be another, if it is something else then there might be a third. We will not be concerned with arguments that probability measures should satisfy the further constraint of regularity. Instead we will be concerned with arguments that hope to show that regularity in general can't be had, and that thus the principle of regularity makes a demand that, in general, can't be met: No matter what a probability measure is hoping to measure in the end, it is bound to fail to live up to the standards of the principle of regularity. In particular, cardinality arguments against regular hyperreal-valued probability measures hope to show that probability measures are bound to fail to live up to the standard of regularity eventually. These arguments don't aim to show that there are no regular probability measures—clearly there are such measures—but rather that probability measures in general are bound to fail to live up to the standard the principle of regularity demands, that is when we measure probability we can't demand that such measures are only correct when they are regular.

The idea of these arguments is to consider what would happen if the event space got very large, in particular larger than the size of the set of measures. Suppose that the size of the hyperreal field employed in the probability measure is κ, which could be the size of the continuum or larger.[6] Then we can consider an event space with a sample set of size larger than κ, that is to say, an event space where we have more than κ-many basic outcomes. A cardinality argument against hyperreal-valued probability measures will proceed to show that for such event spaces regularity fails. Furthermore, even if we started with different hyperreal numbers, of a larger cardinality λ, the argument would simply push the issue higher up, toward a sample space of size larger than λ. Whatever hyperreal numbers one might pick to replace the real numbers as measures of probability, eventually one would have to violate the goal of having one's probability measures be regular. In the case of real-valued probability measures this will happen already for small infinite event spaces. For hyperreal-valued probability measures the same issue will arise, but possibly only higher up, for larger event spaces. Thus the real issue isn't resolved, only pushed back. In particular, going hyperreal-valued isn't going to solve the problem that motivated this move in the first place, or so the cardinality arguments go.

Although there are several different versions of such cardinality arguments, one especially general and strong one is given in Pruss (2013). Pruss shows that no matter how large a hyperreal field we employ in the measurement of chance, there is some event space such that there is no regular probability measure from this space into that field. The argument, in essence and in the version were we allow ourselves the axiom of choice, is simply this.[7] Suppose that $\mathbb{H}$ is a hyperreal field of size κ, and Ω is a sample space of larger size, say κ^+. Using the axiom of choice we can obtain a well-ordering $<_\Omega$ of Ω. It will have order type α for some ordinal $\alpha \geq \kappa^+$. Each element $e \in \Omega$ thus has a place in this ordering as some e_β with $\beta \leq \alpha$. Define for each $\beta \leq \alpha$ the set $\Omega_\beta = \{e \in \Omega | e <_\Omega e_\beta\}$, that is the set of all members of Ω that come before e_β in our well-ordering. Now, when $\lambda < \beta$ then $\Omega_\lambda \subsetneq \Omega_\beta$ and thus Ω_β can be decomposed into two non-empty subsets Ω_λ and $\Omega_\beta - \Omega_\lambda$. Since these subsets are non-empty both have to have positive measure (by regularity) in any probability measure defined on them, and thus (by finite additivity and the construction) for any two ordinals β and λ, with $\beta > \lambda$, the measures of Ω_β and Ω_λ have to be different, with $\mu(\Omega_\beta) > \mu(\Omega_\lambda)$. This means that we need κ^+-many different hyperreal numbers as the measures of the probability of all the Ω_β, but by assumption we only have κ-many available, since $\mathbb{H}$ is of size κ. κ^+-many such Ω_β thus need to get assigned to 0 in such a probability measure, and so regularity fails, and fails badly.

Pruss's cardinality argument shows that for any given hyperreal field $\mathbb{H}$, no matter how large, there is some event space *F* such that there is no regular probability measure from *F* into $\mathbb{H}$. Therefore, no matter which hyperreal numbers we pick to be the measures of probability, regularity is bound to fail somewhere. There is thus little hope to save regularity by using hyperreal numbers instead of real numbers, or so goes the argument. As Alan Hájek puts it: "Pruss's result clinches the case. Regularity cannot be sustained with anything resembling Kolmogorov's axiomatization […]" (Hájek 2013, p. 22). I respectfully disagree.

3 Flexibility and measurement

Although the technical result on which the cardinality arguments are based on are clearly correct, the arguments don't show what they hope to show. The reason is simply a mistaken conception of the role of hyperreal numbers in the measurement of probability. The cardinality arguments rely on that one and the same hyperreal field needs to be employed in the measurement of probability of any event space. It assumes that we pick

$\mathbb{H}$ first, and then we find a large enough F such that there is no regular probability measure from F into $\mathbb{H}$. But this is not how we have to do it, and in general it is not what we do in other cases of measurement. We need to remind ourselves that numbers are measures of probability, mathematical objects suitable to represent or mirror certain probabilistic features of events. We should not require that all event spaces have their probabilities measured with the same numbers. Thus we should not require that there is a fixed hyperreal field that functions as the measures for any event space, even ones that are constructed using that field or some features of it, like its cardinality. To overcome the obstacles we encounter with real-valued probability measures we need to not simply replace the real numbers with some fixed hyperreal field, once and for all, but rather to replace the real numbers with a hyperreal field suitable for the particular task at hand, possibly different ones for different tasks. We can say that probability measurement is *rigid* just in case the same measures need to be employed in all cases of measurement, and *flexible* just in case we can employ different measures in different cases, including different hyperreal fields of different cardinality. Pruss's argument and other similar cardinality arguments implicitly rely on a rigid conception of measurement. However, measurement in general, and measurement of probability in particular, should be flexible. We should first look at the space of events F whose probability we hope to measure, and then employ an *F-suitable* (hyperreal) field as the measures. A (hyperreal) field is suitable, as we will see, just in case it is large enough. It will be sufficient that its cardinality is greater than that of the event space. We are focusing here on hyperreal fields as measures, but with flexible measurement we can, of course, also accept the real numbers as measures when they are suitable, as well as other measures altogether which are not fields. Whether or not probability measurement is taken to be flexible or rigid in general is the crux here, and measurement in general should always be flexible. I would like to illustrate this briefly first by considering what measurement does, and secondly by giving other examples of flexible measurement analogous to our main concern.

In the measurement of probability, we assign numbers to events in accordance with how likely, or probable, they are. We employ numbers from a particular number system to capture facts about probability, and to mirror certain relations of probability among the events as relations among the numbers which are assigned to them as the measures of their probability. So conceived the numbers used as measures are tools. They are employed for a certain purpose. When measurement is conceived this way it should not be required that the same numbers are used in each case. The numbers are tools for representing facts connected to probability of the events in question. But why should we require, in advance, that the same tool be employed in each case? Why should we bind ourselves to use a particular tool in all cases unless we had reason to think that the work that needs to be done can be done with that tool in all cases? Such a demand would not only be unjustified in the case of the measurement of probability, it isn't met in other cases of measurement either.

Consider another case of measurement: cardinal numbers as the measures of the size of a collection. On the rigid conception of measurement there should be one system of such numbers that measures all possible cases of sizes of collections. But then we can find analogue cardinality arguments to refute any candidate for being these measures of size, using the proposed collection of such measures to construct a counterexample. When it comes to finite collections we can simply use the natural numbers. But they won't be good enough for all cases, in particular not the collection of all natural numbers. Maybe we thus need to use the cardinal numbers in the set theoretic sense, as all the initial ordinals? But what about the collection of all of them? No single number in that collection is large enough to be the measure of the size of the collection of all such numbers. Any attempt to find such measures can be refuted, analogously to the cardinality arguments against regular probability measures, as not being general enough. We can always use the collection of the measures proposed to find a case of measurement where they are not good enough. What is wrong with all this is not the use of natural numbers or set theoretic cardinal numbers as the measures of size, but the rigid conception of measurement. On the flexible conception of measurement the natural numbers are perfectly good as the measures of size of finite collections, the cases where they work perfectly. This can be so even if natural numbers are not identical to the finite ordinals, but merely isomorphic to them. The natural numbers are perfect in the cases where they work, even though they won't work in all cases. For other cases we need to employ the set theoretic cardinal numbers or something like them. They won't work for all cases either, but this doesn't show that they weren't the right ones for the cases where they do work.

One reaction could be to stick with rigid measurement and deny that we should distinguish different infinite sizes. The collection of possible measures of size, once and for all, is then just the collection $\{0, 1, 2, ..., \infty\}$. This collection can have its size measured as well with a member of this collection, it is simply of size ∞. But this would deny distinctions which should be acknowledged. To deny distinctions of size to save rigidity in measurement would get things the wrong way round. In our measurements, we hope to capture the distinctions which are there. Measurement thus must live up to what is to be measured. The situation is essentially the same with regular hyperreal-valued probability measures. If we stick with the real numbers, once and for all, then we need to deny that there are any distinctions among "infinitesimal probabilities." Their measures are infinitely close to 0, and thus they all must be 0, the only number among the real numbers infinitely close to 0. But if we are flexible and employ hyperreal-valued measures then we can accept the distinctions where they are, but we can't employ one set of measures for all possible cases of measurement. We need to embrace the differences where they are and to be flexible in measurement.

4 Kolmogorov redone

Although a proponent of the cardinality arguments, Alan Hájek observed, in (2013, p. 22), that a defender of regularity might pick different hyperreal fields as measures, depending on what event spaces we hope to measure. However, Hájek objects to such an approach not because of any considerations about measurement or the role of measures as such, but because this won't give us anything like Kolmogorov's axiomatization of a probability measure. As Hájek puts it: "If there is to be that sort of freedom in the range of probability functions, the theory will have to look very different from Kolmogorov's" (Hájek 2013, p. 23). Here there are two separate issues. First there is the question of what additivity principle we should require for hyperreal-valued probability measures. All of our above discussion only assumed finite additivity, including Pruss's result. Since the real numbers, but not the hyperreal numbers, satisfy the least-upper-bound principle, it is easy to define countable additivity on the real numbers, but no similar definition carries over to the hyperreal numbers. What a stronger additivity principle for hyperreal numbers should look like is actively discussed,[8] but not directly tied to our question about flexible or rigid measurement. Second, limiting ourselves to finite additivity we can easily restate Kolmogorov's axioms in a way congenial to flexible measurement, contrary to Hájek's objection. Instead of having a probability measure be a triple $<\Omega, F, \mu>$ where μ is a function from F into $\mathbb{R}$, we take it to be a triple $<\Omega, F, \mu>$ where μ is a function from F into an *F-suitable* hyperreal field. Here a F-suitable hyperreal field can simply be one that is of at least the size of the powerset of Ω: $2^{|\Omega|}$. All other requirements can remain the same, assuming only finite additivity, as we did throughout. We can now see that this gives us all we want.

5 The existence of flexible regular hyperreal-valued measures

Embracing flexibility undermines the cardinality arguments against hyperreal-valued probability measures, but it doesn't complete the story. To complete it we need to know whether we can always find some hyperreal field suitable to be the measure of probability. That is to say, we need to know whether for any event space we can find some hyperreal field such that there is a regular probability measure from that events space into that hyperreal field. This question has an affirmative answer (with references to the proofs in the following footnote):

> **Fact 1**: For any sample space Ω of any cardinality, and any algebra F on Ω there is a hyperreal field $\mathbb{H}$ of at most size $2^{|\Omega|}$ and a regular probability measure from F into $\mathbb{H}$.

This in particular means that for any event space F, there is an F-suitable hyperreal field $\mathbb{H}$ and a regular probability measure from F into $\mathbb{H}$, vindicating our restatement of Kolmogorov's setup just above. In fact, we can do much better. We can approximate any given real-valued probability measure on F up to an infinitesimal difference as below:

> **Fact 2**: For any sample space Ω of any cardinality, and any algebra F on Ω and any real-valued probability measure μ on F there is a hyperreal field $\mathbb{H}$ and a regular probability measure ν from F into $\mathbb{H}$ such that the difference between μ and ν is infinitesimal.[9]

This is in a sense as good as it gets, since such measures can not in general agree everywhere. The real-valued one is in general not regular, and so the hyperreal-valued one can't agree with it completely. Thus cardinality is not a problem for hyperreal-valued measures, only rigidity in measurement is. Once we embrace flexibility in measurement, as we should generally, then regularity in the measurement of probability is perfectly fine. For any event space we can be assured that we can find a set of measures such that there is a regular probability measure from the event space into these measures. This is what we wanted, and this is what we can have.

6 Conclusion

Cardinality arguments against regular hyperreal-valued probability measures seem to show that we can't save regularity by moving from real numbers to hyperreal numbers as measures of probability. I have argued that this is mistaken. Hyperreal-valued probability measures can always be regular as long as we realize that our approach to measurement must be flexible, not rigid. Once we accept probability measurement as flexible we can see that there always is a regular probability measure on any event space of any cardinality.

Acknowledgments

The author thanks two anonymous referees for a number of good suggestions, as well as Alan Hájek and Alexander Pruss for discussions.

Notes

1 For more discussion of various arguments against real-valued probability measures, for both objective and subjective probability, see Lewis (1983, 175f), Skyrms (1980), Hájek (2013), Hofweber (2014), and Easwaran (2014). We will not try to evaluate these motives here, but rather whether cardinality arguments show that we can't have what many think we need.

2 As is standard, a (finitely additive) real-valued probability measure is defined as follows: let Ω be a set (the sample space), and let F (the event space) be an algebra on Ω, that is F contains Ω and is closed under complements and finite unions. A probability measure is a function μ from F into $[0, 1]_{\mathbb{R}}$ satisfying finite additivity and $\mu(\Omega) = 1$. It is often also required that F be a σ-algebra (i.e., being closed under countable unions), and μ being countably additive. Our concerns already arise with finite additivity and so we will be content with it.

3 For a great survey on hyperreal numbers, see Keisler (1994).

4 See Benci et al. (2013) for one attempt to do this.

5 See his Williamson (2007). A detailed discussion of his argument is in Hofweber (2014).

6 To clarify the terminology, I take a hyperreal field always to extend the real numbers. There are countable non-Archimedean real closed fields, but they don't contain all the real numbers, and thus don't count as a hyperreal field for us here.

7 Pruss is careful to show that the result in spirit does not depend on the axiom of choice, but the argument is more complex without it and requires a further assumption about the ordering of the sample space. Since the axiom of choice isn't really at issue here, at least not in my mind, I am happy to go with the simpler and stronger version. Pruss is also careful to generalize the argument so that it does not just apply to hyperreal fields, but also to measures with less structure on them. Since we will see towards the end that we can always have a

hyperreal field that saves regularity we can concentrate on hyperreal fields as measures, and sideline Pruss's more general, but also more complex, setup.

8 See Hofweber (2014) and Benci et al. (2013) for more details.

9 It should be noted here that, as in Pruss (2013), it is only assumed that the hyperreal-valued measure is finitely additive. It is, of course, also only required that the hyperreal-valued measure agrees with the given real-valued one (up to an infinitesimal) where the real-valued one is defined. I won't outline the proofs here, partly because they would make the paper unnecessarily technical, partly because they would push it over the word limit, and partly because they are not in question. There are a number of different ways to establish them. On the one hand there is work on non-standard measure theory which implies them, see, in particular, Henson (1972) and Cutland (1983). On the other hand, there are results that connect conditional probability functions, so called Popper functions or Popper-Rényi functions, to hyperreal-valued probability measures. See, in particular, Krauss (1968) and McGee (1994). Furthermore, these facts are also a consequence of the results in Benci et al. (2013), with further additivity conditions satisfied as well. Fact 1 was, as far as I can tell, first established by Otto Nikodým in the 1950s in Nikodým (1956). See also section 4 of Luxemburg (1962) for the proof. Thanks to an anonymous referee for this reference. To my knowledge and taste, the simplest and most direct proof of both facts is in Hofweber and Schindler (2014). Here we derive such hyperreal-valued probability measures directly without a detour via conditional probabilities, without relying on general results from measure theory, and without even using the ultraproduct construction.

References

Benci, V., L. Horsten, and S. Wenmackers. "Non-Archimedean Probability." *Milan Journal of Mathematics* **81**.1 (2013): 121–51.

Cutland, N. "Nonstandard Measure Theory and Its Applications." *Bulletin of the London Mathematical Society* **15** (1983): 529–89.

Easwaran, K. "Regularity and Hyperreal Credences." *The Philosophical Review* **123**.1 (2014): 1–41.

Hájek, A. *Staying Regular?* Unpublished manuscript (2013).

Henson, C. W. "On the Non-Standard Representation of Measures." *Transactions of the American Mathematical Society* **172** (1972): 437–46.

Hofweber, T. "Infinitesimal Chances." *Philosophers' Imprint* **14**.2 (2014): 1–34.

Hofweber, T. and Schindler, R. "Hyperreal-Valued Probability Measures Approximating a Real-Valued Measure." *Notre Dame Journal of Formal Logic* (2014).

Keisler, H. J. "The Hyperreal Line," in *Real Numbers, Generalizations of the Reals, and Theories of Continua*, edited by P. Ehrlich. Dordrecht: Kluwer, 1994, 207–38.

Krauss, P. "Representation of Conditional Probability Measures on Boolean Algebras." *Acta Mathematica Academie Scientiarum Hungaricae* **19**.3–4 (1968): 229–41.

Lewis, D. *Philosophical Papers*, Vol. **1**. Oxford: Oxford University Press, 1983.

Luxemburg, W. A. J. "Two Applications of the Method of Construction by Ultrapowers to Analysis." *Bulletin of the American Mathematical Society* **68**.4 (1962): 416–19.

McGee, V. "Learning the Impossible," in *Probability and Conditionals: Belief Revision and Rational Decision*, edited by E. Eells and B. Skyrms. Cambridge: Cambridge University Press, 1994, 179–200.

Nikodým, O. M. "On Extension of a Given Finitely Additive Field-Valued, Non-Negative Measure, on a Finitely Additive Boolean Tribe, to Another One More Ample." *Rendiconti del Seminario Matematico della Università di Padova* **26** (1956): 232 – 327.

Pruss, A. R. "Probability, Regularity, and Cardinality." *Philosophy of Science* **80**.2 (2013): 231 – 40.

Skyrms, B. *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*. New Haven, CT: Yale University Press, 1980.

Williamson, T. "How Probable Is an Infinite Sequence of Heads?" *Analysis* **67**.3 (2007): 173 – 80.