Notre Dame Journal of Formal Logic
Volume 57, Number 3, 2016

Hyperreal-Valued Probability Measures Approximating a Real-Valued Measure

Thomas Hofweber and Ralf Schindler

Abstract We give a direct and elementary proof of the fact that every real-valued probability measure can be approximated—up to an infinitesimal—by a hyperreal-valued one which is regular and defined on the whole powerset of the sample space.

When we measure the probability of events, we assign numbers to these events in accordance to how likely they are. Standard probability theory assigns real numbers to events, but there are well-known problems with using real numbers as the measures of probability. One of them is that measure 0 events do not form a homogeneous class; that is to say, there seem to be differences in probability among events which get assigned the same measure of their probability, namely, the lowest possible measure 0. To illustrate with a standard example, let Ω be any nonempty set. Let us randomly pick an element of Ω. What is the chance that a given element $a \in \Omega$ gets chosen? If Ω is finite, then the answer should be $\frac{1}{n}$, where n is the number of elements of Ω. But what if Ω is infinite? If the measure of probability is a real number between 0 and 1, then the answer has to be 0, since it should be lower than $\frac{1}{n}$ for each n. But 0 is also the measure of the probability of the impossible event of a being picked as well as not picked. These events seem to differ in their probability, since one of them might well be the one that happens, while the other one for sure will not.

To measure probability in a way that respects this difference we thus need to employ numbers other than the real numbers as measures of probability. The reason for the failure of real numbers to be able to measure probability fine enough to respect these differences is, in the end, that real numbers have the Archimedean property: Any positive real number, no matter how small, is still larger than some $\frac{1}{n}$, $n \in \mathbb{N}$.

Received March 22, 2013; accepted January 6, 2014
First published online April 6, 2016
2010 Mathematics Subject Classification: Primary 60A; Secondary 20E
Keywords: probability, hyperreal numbers, measure theory
 10.1215/00294527-3542210

To have finer probability measures we need to employ non-Archimedean number systems instead. Hyperreal numbers are non-Archimedean extensions of the real numbers. Hyperreal numbers in particular contain *infinitesimals*: positive numbers smaller than $\frac{1}{n}$, for all $n \in \mathbb{N}$. A hyperreal-valued probability measure employs hyperreal numbers instead of real numbers as measures of probability. But can we be assured that this will always help? Can we always replace a real-valued probability measure with a *regular* hyperreal-valued one, that is, one that gives measure 0 only to the impossible event? By "replace" we mean that for every event X, the hyperreal-valued probability of X is to be *infinitely close to* (i.e., the absolute value of their difference is an infinitesimal) the real-valued probability of X. The answer to this question is affirmative: for any given real-valued probability measure there is a regular hyperreal-valued one that approximates it up to an infinitesimal.

This result is not new. It is established, for example, in work on nonstandard measure theory (see Henson [3], Cutland [2]). And it follows from work on the connection of conditional probability functions and nonstandard probability theory (see Krauss [4], McGee [5]). In this paper we propose a new and completely elementary proof of this fact. While the known proofs mentioned above rely on general results in measure theory or model theory and are sometimes indirect, we give a direct proof using only elementary methods, relying not even on the ultraproduct construction, but only on the compactness theorem. This does not prove the result from weaker assumptions, but it gives a simpler and more direct proof.

Let Ω be any infinite set (the sample space), and let F be a σ-algebra on Ω (the event space), that is, $F \subset \mathcal{P}(\Omega)$ with F closed under complements, countable unions, and $\Omega \in F$. A *real-valued probability measure* is a function μ from F into $[0, 1] \cap \mathbb{R}$ such that

(1) $\mu(\Omega) = 1$,
(2) if $X_1, \ldots, X_i, \ldots$ are countably many pairwise disjoint subsets of Ω, then

$$\mu\Big(\bigcup_{i\in\mathbb{N}} X_i\Big) = \sum_{i\in\mathbb{N}} \mu(X_i).$$

The triple (Ω, F, μ) is a *standard probability space*. A probability measure is *regular* just in case $\mu(X) > 0$ for all $X \neq \emptyset$, and *uniform* just in case, for all $a, a' \in \Omega$, $\mu(\{a\}) = \mu(\{a'\})$. Since the real numbers form an Archimedean field, there can be no uniform and regular real-valued probability measure on an infinite sample space. No positive real number is small enough to be the measure of a singleton set. To get that we need to measure probability with a non-Archimedean field.

A *hyperreal field* $\mathbb{R}^*$ is a non-Archimedean extension of the real numbers that has the same first-order properties as the real numbers. The elements of a hyperreal field we also call *hyperreal numbers*. Since hyperreal fields do not satisfy the least upper bound principle, the notion of an infinite sum cannot be carried over straightforwardly from real numbers to hyperreal numbers. How a more general additivity principle should be formulated for hyperreal-valued probability measures is not completely settled, although there are a variety of possibilities (see Benci, Horsten, and Wenmackers [1] for one approach). Consequently, we only require a hyperreal-valued probability measure to be finitely additive. We can define a nonstandard probability space and a hyperreal-valued probability measure as follows.

Definition We call $(\Omega, \mathcal{P}(\Omega), \mu)$ a *nonstandard probability space* if and only if Ω is an infinite set and there are hyperreal numbers $\mathbb{R}^*$ such that $\mu: \mathcal{P}(\Omega) \to [0, 1] \cap \mathbb{R}^*$ satisfies the following statements.

(1) $\mu(\Omega) = 1$.
(2) If $X \subset \Omega$ and $X \neq \emptyset$, then $\mu(X) > 0$.
(3) If $k \in \mathbb{N}$ and $X_1, \ldots, X_k \subset \Omega$, where $X_i \cap X_j = \emptyset$ for all $i \neq j$, then $\mu(\bigcup_{i=1}^{k} X_i) = \sum_{i=1}^{k} \mu(X_i)$.

μ in a nonstandard probability space is a hyperreal-valued probability measure. By our definition, a hyperreal-valued probability measure is regular. Note that the event space is not merely any σ-algebra on Ω, but the whole powerset of Ω. Our main goal now is to give an elementary proof of the central result connecting standard and nonstandard probability spaces, which says that any real-valued probability measure can be approximated up to an infinitesimal by a hyperreal-valued one. This in particular implies that we can always have a regular probability measure on the whole powerset of any sample space.

Theorem *Let $(\Omega, F, \bar{\mu})$ be a standard probability space. There is then some $\mathbb{R}^*$ and $\mu: \mathcal{P}(\Omega) \to \mathbb{R}^*$ such that $(\Omega, \mathcal{P}(\Omega), \mu)$ is a nonstandard probability space and for $X \in F$, $\mu(X)$ is infinitely close to $\bar{\mu}(X)$.*

Proof Let us fix $(\Omega, F, \bar{\mu})$. We will use a simple compactness argument. We enrich the usual first-order language for an ordered field with constants "$\mu(\dot{X})$" for every $X \subset \Omega$ (for the measure of X we are looking for) as well as by constants $\dot{x}$ for all elements x of $\mathbb{R}$.

In this language, let Γ be the smallest class of formulas with the following properties. Γ contains the theory of

$$\big(\mathbb{R}; 0, 1, <, +, \cdot, (x: x \in \mathbb{R})\big),$$

and

(i) "$\mu(\dot{\Omega}) = 1$" $\in \Gamma$;
(ii) if $X \subset \Omega$ and $X \neq \emptyset$, then "$\mu(\dot{X}) > 0$" $\in \Gamma$;
(iii) if $k \in \mathbb{N}$ and $X_1, \ldots, X_k \subset \Omega$, where $X_i \cap X_j = \emptyset$ for all $i \neq j$, then, writing $X = \bigcup_{i=1}^{k} X_i$, "$\mu(\dot{X}) = \sum_{i=1}^{k} \mu(\dot{X}_i)$" $\in \Gamma$;
(iv) if $X \subset \Omega$ and $X \in F$, say $\bar{\mu}(X) = x \in \mathbb{R}$, then for every $n \in \mathbb{N}$, "$|\mu(\dot{X}) - \dot{x}| < \frac{1}{n}$" $\in \Gamma$.

It suffices to verify that Γ is consistent. In a model of Γ, μ is a finitely additive probability measure (by conditions (i) and (iii)), which is regular (by (ii)), defined on all of $\mathcal{P}(\Omega)$ (by (ii)), and approximates our given real-valued measure $\bar{\mu}$ up to an infinitesimal (by (iv)). In order to show that Γ is consistent, we verify that if $\bar{\Gamma} \subset \Gamma$ is finite, then there is a model of $\bar{\Gamma}$ whose universe is $\mathbb{R}$ and which interprets all the symbols except for the "$\mu(\dot{X})$" in the standard way. Let us thus fix a finite $\bar{\Gamma} \subset \Gamma$.

Let $\{X_1, \ldots, X_n\}$ be the set of all $X \subset \Omega$ such that "$\mu(\dot{X})$" occurs in a formula from $\bar{\Gamma}$. We may assume without loss of generality that $X_1 = \Omega$. For every $I \subset \{1, \ldots, n\}$, let us write

$$Y_I = \bigcap_{i \in I} X_i \Big\backslash \bigcup_{j \notin I} X_j .$$

Then $\{Y_I : I \subset \{1, \ldots, n\}\}$ is a partition of Ω, and for every i, $1 \leq i \leq n$, $\{Y_I : i \in I \subset \{1, \ldots, n\}\}$ is a partition of X_i. The Y_I thus give us a finite base from which every X_i can be generated as a union of elements in the base. We need to assign positive real numbers to each "$\mu(\dot{X}_i)$" (for $X_i \neq \emptyset$) that satisfy the finitely many equations of the form of (iii) and (iv) that are in $\bar{\Gamma}$. It is tempting to define such a number based on how many elements of the base are required to build X_i, what the smallest $\frac{1}{n}$ is that occurs in $\bar{\Gamma}$ in an equation of kind (iv), and how many nonempty X_i were assigned measure 0 by $\bar{\mu}$. But $\bar{\mu}$ might not be defined on X_i, since it is only defined on $X \subset \Omega$ with $X \in F$, whereas μ needs to be defined on all of $\mathcal{P}(\Omega)$. We will write "$\bar{\mu}(X) \downarrow$" for $X \in F$, that is, the fact that $\bar{\mu}(X)$ is defined, or equivalently, X is $\bar{\mu}$-measurable. In order to find values for our "$\mu(\dot{X}_i)$" we need to replace our Y_I with $\bar{\mu}$-measurable Y_I^*, which we will define as the smallest $\bar{\mu}$-measurable expansion of Y_I by other elements of our base as follows.

For every $I \subset \{1, \ldots, n\}$, let us denote by Y_I^* the *smallest* Y of the form

$$Y = Y_I \cup Y_{I_1} \cup \cdots \cup Y_{I_m},$$

where $m \in \mathbb{N}$, $I_i \subset \{1, \ldots, n\}$ for every i, $1 \leq i \leq m$, and $\bar{\mu}(Y)$ is defined. (We allow $m = 0$, i.e., $Y = Y_I$.) Note that Y_I^* is well defined, as $\Omega = X_1$, $\bar{\mu}(\Omega) \downarrow$, and the intersection of finitely many $\bar{\mu}$-measurable sets is $\bar{\mu}$-measurable, so that we may equivalently write Y_I^* as

$$\bigcap \{Y = Y_I \cup Y_{I_1} \cup \cdots \cup Y_{I_m} : m \in \mathbb{N} \wedge \forall i \left(I_i \subset \{1, \ldots, n\}\right) \wedge \bar{\mu}(Y) \downarrow\}.$$

Let us write $\mathcal{F}$ for the set of all Y_I^*, where $I \subset \{1, \ldots, n\}$. It is easy to see that $Y_I^* = \emptyset$ if and only if $Y_I = \emptyset$.

Let $Y_I^*, Y_{I'}^* \in \mathcal{F}$, where $I, I' \subset \{1, \ldots, n\}$. Suppose that $Y_I^* \cap Y_{I'}^* \neq \emptyset$. There is then some $J \subset \{1, \ldots, n\}$ such that $Y_J \subset Y_I^* \cap Y_{I'}^*$. As $\bar{\mu}(Y_I^*) \downarrow$ and $\bar{\mu}(Y_{I'}^*) \downarrow$, we must have $Y_J^* \subset Y_I^* \cap Y_{I'}^*$. If $Y_I \cap Y_J^* = \emptyset$, then $Y_I^* \setminus Y_J^*$ is a $\bar{\mu}$-measurable set of the right form which is properly contained in Y_I^*, which contradicts the choice of Y_I^*. Hence $Y_I^* \subset Y_J^*$. Symmetrically, we get $Y_{I'}^* \subset Y_J^*$, and thus $Y_I^* \cup Y_{I'}^* \subset Y_J^* \subset Y_I^* \cap Y_{I'}^*$; that is, $Y_I^* = Y_{I'}^*$.

We have verified that for all I and I', I with $I' \subset \{1, \ldots, n\}$, if $Y_I^*, Y_{I'}^* \in \mathcal{F}$ and $Y_I^* \cap Y_{I'}^* \neq \emptyset$, then $Y_I^* = Y_{I'}^*$. In other words, $\mathcal{F}$ is a partition of Ω into (finitely many) $\bar{\mu}$-measurable sets.

Let us now pick $\epsilon \in \mathbb{R}$, $\epsilon > 0$, such that $\epsilon < \frac{1}{n}$ for all occurrences of "$\frac{1}{n}$" in a formula of type (iv) from $\bar{\Gamma}$ and also $\epsilon < \bar{\mu}(Y_I^*)$ for all $I \subset \{1, \ldots, n\}$ such that $\bar{\mu}(Y_I^*) > 0$. Let k be the number of $Y \in \mathcal{F}$ such that $\bar{\mu}(Y) = 0$ and "$\mu(\dot{Y})$" occurs in $\bar{\Gamma}$, and let l be the number of $Y \in \mathcal{F}$ such that $\bar{\mu}(Y) > 0$ and "$\mu(\dot{Y})$" occurs in $\bar{\Gamma}$. For $Y \in \mathcal{F}$, let $\#(Y)$ be the number of nonempty subsets Y_I, $I \subset \{1, \ldots, n\}$, of Y. Let us now define, for $I \subset \{1, \ldots, n\}$,

$$\mu(Y_I) = \begin{cases} 0 & \text{if } Y_I^* = \emptyset, \\ \frac{1}{\#(Y_I^*)} \cdot \frac{\epsilon}{k} & \text{if } Y_I^* \neq \emptyset \text{ and } \bar{\mu}(Y_I^*) = 0, \\ \frac{1}{\#(Y_I^*)} \cdot (\bar{\mu}(Y_I^*) - \frac{\epsilon}{l}) & \text{if } Y_I^* \neq \emptyset \text{ and } \bar{\mu}(Y_I^*) > 0. \end{cases}$$

We then also define, for $1 \leq i \leq n$,

$$\mu(\dot{X}_i) = \sum_{i \in I \subset \{1, \ldots, n\}} \mu(Y_I).$$

It is straightforward to see that this assignment verifies that $\bar{\Gamma}$ is consistent. Since this holds for arbitrary finite $\bar{\Gamma}$, it follows, by the compactness theorem, that Γ is consistent as well, as we hoped to show. □

It is worth noting that although we motivated the need for hyperreal-valued probability measures on an infinite sample space with examples of uniform probability measures, measures where for all $a, a' \in \Omega$ $\mu(\{a\}) = \mu(\{a'\})$, no assumption is made in the theorem or the proof that μ is uniform. The result holds in general, whether or not singleton sets have the same or different probability.

Corollary *Let Ω be any infinite sample space. There is a hyperreal field $\mathbb{R}^*$ of at most cardinality $2^{|\Omega|}$ and a regular probability measure from $\mathcal{P}(\Omega)$ into $\mathbb{R}^*$.*

Proof Take some real-valued probability measure $\bar{\mu}$ defined on some σ-algebra on Ω. By the Theorem there is a hyperreal field $\mathbb{R}^*$ and a regular probability measure from $\mathcal{P}(\Omega)$ into $\mathbb{R}^*$. We can see from the proof that the size of the theory Γ is bounded by the cardinality of $\mathcal{P}(\Omega)$, and thus, by the downward Löwenheim–Skolem theorem, there is such an $\mathbb{R}^*$ of at most size $2^{|\Omega|}$. □

References

[1] Benci, V., L. Horsten, and S. Wenmackers, "Non-Archimedean probability," *Milan Journal of Mathematics*, vol. 81 (2013), pp. 121–51. MR 3046984. DOI 10.1007/s00032-012-0191-x. 370

[2] Cutland, N. J., "Nonstandard measure theory and its applications," *Bulletin of the London Mathematical Society*, vol. 15 (1983), pp. 529–89. Zbl 0529.28009. MR 0720746. DOI 10.1112/blms/15.6.529. 370

[3] Henson, C. W., "On the nonstandard representation of measures," *Transactions of the American Mathematical Society*, vol. 172 (1972), pp. 437–46. MR 0315082. 370

[4] Krauss, P., "Representation of conditional probability measures on Boolean algebras," *Acta Mathematica Academiae Scientiarum Hungaricae*, vol. 19 (1968), pp. 229–41. Zbl 0174.49001. MR 0236080. 370

[5] McGee, V., "Learning the impossible," pp. 179–99 in *Probability and Conditionals: Belief Revision and Rational Decision*, edited by E. Eells and B. Skyrms, Cambridge University Press, Cambridge, 1994. MR 1373432. 370

Acknowledgments

The authors would like to thank the Studienstiftung des deutschen Volkes for inviting us to conduct a seminar on the infinitely small and the infinitely large at the Sommerakademie Neubeuern, where the core of this paper was written. We would also like to thank the participants of this seminar for the very inspiring and enjoyable atmosphere.

Hofweber
Department of Philosophy
University of North Carolina at Chapel Hill
Chapel Hill, NC 27599
USA
hofweber@unc.edu
http://www.thomashofweber.com

Schindler
Institut für Mathematische Logik und Grundlagenforschung
Universität Münster
48149 Münster
Germany
rds@uni-muenster.de
http://http://wwwmath.uni-muenster.de/u/rds