Empiricism, Probability, and Knowledge of Arithmetic

Sean Walsh
Department of Logic and Philosophy of Science
University of California, Irvine
swalsh108@gmail.com
http://www.swalsh108.org

Progic 2013
Munich September 18, 2013

The thesis that I want to examine today is this:

> arithmetical knowledge may be legitimately extended by confirmation just as it may be by proof.

This thesis is one component of a *broader empirical account of arithmetical knowledge*, according to which:

> subsequent to some appropriately empirical acquisition of knowledge of quantifier-free truths like $7 + 5 = 12$,
>
> one may legitimately extend this knowledge to knowledge of more complex arithmetical truths by way of confirmation.

My goal today is to defend this thesis against 2 objections of the form: this kind of probability is *too close* to arithmetical truth.

Among notions of confirmation, I focus on probabilistic variants:

> Hypothesis h is *confirmed* by evidence e relative to background knowledge K if $P(h|e \,\&\, K) > P(h|K)$

Here P is a probability assignment, a function $P : \mathrm{Sent}(L) \rightarrow \mathbb{R}$ which satisfies, for all φ, ψ in the signature L of arithmetic:

> (P1) $P(\varphi) \geq 0$,
> (P2) $P(\varphi) = 1$ if $\models \varphi$,
> (P3) $P(\varphi \vee \psi) = P(\varphi) + P(\psi)$ if $\models \neg(\varphi \,\&\, \psi)$

where $\models$ is the usual consequence relation from first-order logic.

So we're conceiving of probabilities being assigned to fine-grained entities like sentences as opposed to coarse-grained entities like sets of possible worlds.

Among arithmetical axioms, I focus on the *Peano axioms*, given by:

Robinson's Q, a finite set of axioms that imply all the addition and multiplication tables, e.g. $7 + 5 = 12$. It has axioms like $\forall\, x\; s(x) \neq 0$.

Mathematical Induction, a principle that articulates a canonical means by which to establish a universal hypothesis about natural numbers. It says: if zero has a property F and $n + 1$ has F whenever n has F, then all natural numbers have F.

The thesis that I want to examine today is this:

> arithmetical knowledge may be legitimately extended by confirmation just as it may be by proof.

Among notions of confirmation, I focus on probabilistic variants:

> Hypothesis h is *confirmed* by evidence e relative to background knowledge K if $P(h|e \,\&\, K) > P(h|K)$

The thesis that I want to examine today is this:

> arithmetical knowledge may be legitimately extended by confirmation just as it may be by proof.

Among notions of confirmation, I focus on probabilistic variants:

> Hypothesis h is *confirmed* by evidence e relative to background knowledge K if $P(h|e \,\&\, K) > P(h|K)$

Two special cases of the thesis are then:

I. Let Robinson's Q be written $h \equiv \forall\, \overline{x}\; F(\overline{x})$ and consider evidence of form $e \equiv \bigwedge_{i=1}^{N} F(\overline{a}_i)$. Then one may justifiably infer from e to h on the basis of justification in e and h's being confirmed by e.

II. Let $h \equiv \forall\, x\; G(x)$ and $e \equiv [G(0) \,\&\, \forall\, n\, (G(n) \rightarrow G(n+1))]$. Then one may justifiably infer from e to h against $K \equiv$ Robinson's Q on the basis of justification in e, K and h's being confirmed by e relative to the background of K.

Two motivating thoughts on the significance of the thesis:

First, it accounts for our knowledge of arithmetical axioms, using a source of justification which is routinely and efficaciously employed in other parts of our ordinary reasoning, thus serving to partially dissipate a skeptical concern about the possibility of mathematical knowledge.

Second, the parallel between mathematical knowledge and scientific knowledge (broadly construed) is widely accepted in many parts of philosophy of mathematics. However, the one place in philosophy of mathematics that has traditionally resisted any form of empirical treatment is *arithmetic*– here, by contrast, logicist and fictionalist approaches have been dominant.

The approach that I adopt obviously requires the specification of some class of probability assignments. Today, I want to focus on two objections, which have the form:

> this kind of probability assignment, and hence this kind of confirmation, is *too close* to arithmetical truth.

If there was such conceptual proximity, then:

> any claim to this general approach being broadly empirical in character would be severely undermined, *and*
>
> any skeptical doubts that we could come to be justified in believing arithmetical truths would be equally strong doubts that we could come to be justified in believing that such-and-such sentences had such-and-such probability.

Outline

I. Introduction

II. Avoiding Counting Assignments

III. Avoiding Alignment of True and Probable

IV. Summary plus Further Questions

So one kind of probability assignment is a counting assignment:

$$P(\varphi) = a_1 \cdot T_1(\varphi) + \cdots + a_n \cdot T_n(\varphi)$$

wherein $a_1 + \cdots + a_n = 1$ and $T_1, \ldots, T_n$ are in $[K]$, the space of complete consistent extensions of our background knowledge K.

So one kind of probability assignment is a counting assignment:

$$P(\varphi) = a_1 \cdot T_1(\varphi) + \cdots + a_n \cdot T_n(\varphi)$$

wherein $a_1 + \cdots + a_n = 1$ and $T_1, \ldots, T_n$ are in $[K]$, the space of complete consistent extensions of our background knowledge K.

The philosophical concern relates to true arithmetic $Th(\mathbb{N})$, the set of all sentences true on the standard model of arithmetic.

So one kind of probability assignment is a counting assignment:

$$P(\varphi) = a_1 \cdot T_1(\varphi) + \cdots + a_n \cdot T_n(\varphi)$$

wherein $a_1 + \cdots + a_n = 1$ and $T_1, \ldots, T_n$ are in $[K]$, the space of complete consistent extensions of our background knowledge K.

The philosophical concern relates to true arithmetic $Th(\mathbb{N})$, the set of all sentences true on the standard model of arithmetic.

If $T_1, \ldots, T_n$ included true arithmetic $Th(\mathbb{N})$, then one would want to know what it is about our relationship to true arithmetic that grants it this preferred status.

While if not, then one would want to know what about $T_1, \ldots, T_n$ make them reliable indicators of truth. Why these and not others?

So one kind of probability assignment is a counting assignment:

$$P(\varphi) = a_1 \cdot T_1(\varphi) + \cdots + a_n \cdot T_n(\varphi)$$

wherein $a_1 + \cdots + a_n = 1$ and $T_1, \ldots, T_n$ are in $[K]$, the space of complete consistent extensions of our background knowledge K.

The philosophical concern relates to true arithmetic $Th(\mathbb{N})$, the set of all sentences true on the standard model of arithmetic.

If $T_1, \ldots, T_n$ included true arithmetic $Th(\mathbb{N})$, then one would want to know what it is about our relationship to true arithmetic that grants it this preferred status.

While if not, then one would want to know what about $T_1, \ldots, T_n$ make them reliable indicators of truth. Why these and not others?

. . . This dilemma suggests that we should avoid appeal to counting assignments . . . But how can we do that?

A counting assignment P induces a probability measure $\widehat{P} : \mathrm{Borel}([K]) \to [0,1]$ on the Borel subsets of the space $[K]$ by:

$$\widehat{P}([\varphi]) = a_1 \cdot T_1(\varphi) + \cdots + a_n \cdot T_n(\varphi)$$

wherein $[\varphi]$ is the clopen set $\{T \in [K]s : T \models \varphi\}$. If P is a counting assignment, then note the elementary consequence:

$$\widehat{P}(\{T_i\}) = \lim_{\ell \to \infty} \widehat{P}([\chi_{T_i \upharpoonright \ell}]) = a_i > 0$$

where $\varphi_1, \ldots, \varphi_n, \ldots$ is fixed enumeration of $\mathrm{Sent}(L)$ and

$$\sigma \mapsto \chi_\sigma \equiv \bigwedge_{\sigma(i)=1} \varphi_i \wedge \bigwedge_{\sigma(i)=0} \neg\varphi_i$$

An *atom* of a probability measure $\widehat{P} : [K] \to [0,1]$ is a theory T in the space $[K]$ such that $\widehat{P}(\{T\}) > 0$.

An *atom* of a probability measure $\widehat{P} : [K] \to [0,1]$ is a theory T in the space $[K]$ such that $\widehat{P}(\{T\}) > 0$.

A *continuous* probability measure $\widehat{P} : [K] \to [0,1]$ is one which has no atoms.

So one way to avoid counting assignments P is to restrict attention to those P such that $\widehat{P}$ is continuous.

But how can we obtain access to such a P? Epistemic access in mathematics comes in many forms: one way is axioms, another way is computation.

An *atom* of a probability measure $\widehat{P}:[K]\rightarrow[0,1]$ is a theory T in the space $[K]$ such that $\widehat{P}(\{T\})>0$.

A *continuous* probability measure $\widehat{P}:[K]\rightarrow[0,1]$ is one which has no atoms.

Ostensibly, being continuous is very hard to test since its official definition requires checking something about every theory.

But by using König's Lemma, it turns out that a probability measure $\widehat{P}:[K]\rightarrow[0,1]$ is *continuous* if and only if

$$\forall\,\epsilon>0\;\exists\,\ell\;\forall\;|\sigma|=\ell\;\widehat{P}([\chi_\sigma])<\epsilon$$

Since $P(\psi)=\widehat{P}([\psi])$, we thus have axioms for P which guarantee that it is not a counting assignment.

Consider the partial order, where the order is by extension:

$$\mathbb{P} = \{P : S \to \mathbb{Q} : S \text{ finite algebra} \subseteq \mathrm{Sent}(L) \ \& \ P \models P1\text{-}P3\}$$

This Π^0_1-partial order has the following Δ^0_2-dense sets:

$$D_n = \{P \text{ in } \mathbb{P} : \mathrm{dom}(P) \text{ contains } \varphi_n\}$$
$$D_\epsilon = \{P \text{ in } \mathbb{P} : \exists\, \ell\ \mathrm{dom}(P) \supseteq \{\varphi_1, \ldots, \varphi_\ell\} \ \& \ \forall\ |\sigma| = \ell\ P(\chi_\sigma) < \epsilon\}$$

So there's Δ^0_2-generic $G \subseteq \mathbb{P}$ and its union is a Δ^0_2-definable $P : \mathrm{Sent}(L) \to [0,1]$ such that $\widehat{P} : [K] \to [0,1]$ is continuous.

Moral: there are at least some probability assignments which aren't counting assignments and which are still computationally tractable.

Outline

I. Introduction

II. Avoiding Counting Assignments

III. Avoiding Alignment of True and Probable

IV. Summary plus Further Questions

This problem arises with the simple observation that probability assignments only have rules for the propositional connectives, but are applied here to sentences with quantifiers.

Initial impetus: shouldn't there be rules for the quantifiers as well?

One natural rule is the following version of countable additivity:

$$\omega\text{-additivity: } (\mathrm{P}\omega)\ P(\forall\, x\, \varphi(x)) = \lim_N P(\bigwedge_{n=1}^{N} \varphi(s^n(0)))$$

One natural rule is the following version of countable additivity:

$$\omega\text{-additivity: } (\mathrm{P}\omega)\ P(\forall\, x\ \varphi(x)) = \lim_N P(\bigwedge_{n=1}^{N} \varphi(s^n(0)))$$

One natural rule is the following version of countable additivity:

$$\omega\text{-additivity: } (\mathrm{P}\omega)\ P(\forall\, x\, \varphi(x)) = \lim_N P(\textstyle\bigwedge_{n=1}^{N} \varphi(s^n(0)))$$

But ω-additivity forces the alignment of the true and the probable:

> Suppose that $0 < \epsilon < \frac{1}{2}$ is a low error threshold, and suppose further that P is an ω-additive probability assignment such that $P(K) > 1 - \epsilon$, where $K \equiv$ Robinson's Q. Then for all φ
>
> $$(\omega, 0, s, +, \times) \models \varphi \iff P(\varphi) > 1 - \epsilon \qquad (*)$$

One natural rule is the following version of countable additivity:

$$\omega\text{-additivity: } (\mathrm{P}\omega)\ P(\forall\, x\ \varphi(x)) = \lim_N P(\textstyle\bigwedge_{n=1}^{N} \varphi(s^n(0)))$$

But ω-additivity forces the alignment of the true and the probable:

> Suppose that $0 < \epsilon < \frac{1}{2}$ is a low error threshold, and suppose further that P is an ω-additive probability assignment such that $P(K) > 1 - \epsilon$, where $K \equiv$ Robinson's Q. Then for all φ
>
> $$(\omega, 0, s, +, \times) \models \varphi \Longleftrightarrow P(\varphi) > 1 - \epsilon \qquad (*)$$

Why is this a problem? Difficult to know things about this assignment. High priors entails low degree of confirmation.

What would a *satisfactory* solution to this problem look like?

Need reasons for thinking that we're are not committed to ω-additivity in the same way that we're are committed to the other probabilistic rules.

My solution is that the only probabilistic rules which we're committed to are those which follow from invulnerability to a Dutch Book.

This solution is satisfactory because: it turns out that ω-additivity does not follow from invulnerability.

One disadvantage of the proposed solution is that one has to admit a robust distinction between probabilistic rules and probabilistic facts.

Another disadvantage of the proposed solution is that in other settings countable additivity is advantageous.

But at least some of the work of ω-additivity can be done by the following version of additivity, which follows from invulnerability and doesn't force alignment:

$$\omega_1\text{-additivity: } (\mathrm{P}\omega_1)\ P(\textstyle\bigwedge_n \varphi_n) = \lim_N P(\bigwedge_{n=1}^N \varphi_n)$$

The relevant difference between ω-additivity and ω_1-additivity is that the n on the RHS of ω_1-additivity merely indexes the sentence and is not a variable that may be quantified over.

Outline

I. Introduction

II. Avoiding Counting Assignments

III. Avoiding Alignment of True and Probable

IV. Summary plus Further Questions

The thesis that I wanted to examine today is this:

> arithmetical knowledge may be legitimately extended by (probabilistic) confirmation just as it may be by proof.

We looked at 2 objections of the form:

> this kind of probability assignment, and hence this kind of confirmation, is *too close* to arithmetical truth.

The first objection was that counting assignments seem inexplicably attracted towards truth or potentially unreliable. The response was to articulate axiomatic and computational methods of avoiding counting assignments.

The second objection was that the probabilistic ω-rule will force alignment of truth and high probability. The response was to suggest that we're committed to only those probabilistic rules which follow from invulnerability.

Suppose P is an arithmetically definable probability assignment, and that background knowledge K includes Robinson's Q and is true on the standard model. Suppose further that P is K-regular, i.e. satisfies: $P(\varphi) = 1$ implies $K \models \varphi$.

By the diagonal lemma, there is L-sentence λ such that

$$0 < P(\lambda) < P(\lambda | P(\lambda) = 0) = 1 \tag{1}$$

So λ would be confirmed by an observation of $P(\lambda) = 0$.

However, this seems intuitively wrong: for, in general, *knowing that something has low probability is a bad reason for believing it*. Perhaps the best response is: this general truth admits exceptions– but one would like a better understanding of how broad this phenomena is.