Rendiconti
Accademia Nazionale delle Scienze detta dei XL
Memorie di Matematica
112° (1994), Vol. XVIII, fasc. 1, pagg. 117-132

MONICA MUSIO - ROBERT LUTZ (*)

A Finitisation of the Finitely Additive Probability Theory Using non Standard Analysis (**)

SUMMARY. — In this work we prove, within the frame of Non Standard Analysis, that the finitely additive probability theory of B. de Finetti is equivalent to the elementary probability theory on finite spaces. This equivalence reduces the solution of various classical problems to purely combinatorial constructions. We use it to get a new insight in comparison of zero-probabilities, the extension of conditional probability laws, equiprobability on infinite sets etc... In particular, we get a finitely additive probability law on the power set of the Euclidean real line, which is invariant under all isometries. In order to get the paper self-contained, we give in the appendix a brief account on E. Nelson's axiomatic setting of Non Standard Analysis called *Internal Set Theory*.

Una finitizzazione della teoria delle probabilità fintamente additiva tramite l'Analisi Non Standard

SOMMARIO. — In questo lavoro ci proponiamo di dimostrare, con l'ausilio dell'Analisi Non Standard, che la teoria della probabilità finitamente additiva di B. de Finetti è equivalente alla teoria elementare delle probabilità in uno spazio finito. Tramite questa equivalenza si possono ridurre varie soluzioni di problemi classici a pure costruzioni combinatorie. Inoltre alcuni problemi come il confronto di probabilità nulle, l'estensione di leggi di probabilità condizionali, l'equiprobabilità in insiemi infiniti, vengono risolti in questa nuova ottica. In particolare abbiamo ottenuto una legge di probabilità finitamente additiva, nell'insieme delle parti della retta reale, invariante per le isometrie. Per rendere l'articolo self-contained presentiamo in appendice una breve sintesi della *Internal Set Theory* (versione assiomatizzata dell'Analisi Non Standard di E. Nelson).

The aim of this paper is to prove that the finitely additive probability theory of B. de Finetti has the same scientific content as the elementary finite probability theory. To

(*) Indirizzi degli Autori: M. Musio: Università «La Sapienza», Dipartimento di Statistica, Probabilità e Statistiche applicate, P.le Aldo Moro 5, I-00185 Roma; R. Lutz: Université de Haute Alsace, Faculté des Sciences et Techniques, Laboratoire de Mathématiques, 4, rue des frères Lumière, F-68093 Mulhouse-Cedex, France.

(**) Memoria presentata il 26 settembre 1994 da Giorgio Letta, uno dei XL.

this end, we use the full strength of Non Standard Analysis. As a by-product, we get a new insight in various classical problems, e.g. the comparison of zero-probabilities, extension properties, and equiprobability on subsets of the real line. In some sense, we give here a precise answer to the following remark of B. de Finetti:

«Let us just mention that the consideration of probability as a non-Archimedean quantity would permit us to say, if we wished, that 'zero probabilities' are in fact «infinitely small» (actual infinitesimals), and only that of the impossible event is zero. Nothing is really altered by this change in terminology, but it might sometimes be useful as a way of overcoming preconceived ideas. It has been said that to assume that $0+0+0+0+\ldots+0+\ldots=1$ *is absurd, whereas, if at all, this would be true if "actual infinitesimals" were substituted in place of "zero". There is nothing to prevent one from expressing things in this way, apart from the fact that it is a useless complication of language, and leads one to puzzle over 'les infiniment petits»* [F2].

In fact, it is now well known that such «things» may be easily formalised within the frame of Non Standard Analysis. One of the main applications of NSA is to establish a direct link between continuity and finiteness using the concept of *shadow*, which replaces the classical limit procedures without the difficulties they involve.

It is known that the limit of a sequence of σ-additive probability measures on the same space is not in general a σ-additive measure, while the limit of a sequence of finitely additive probability laws is finitely additive [Re2]. The corresponding statement within NSA is that the shadow of a finitely additive law is a standard finitely additive law. This opens the way to the finitisation of the whole finitely additive probability theory that we shall develop in the sequel. Meanwhile, E. Nelson has given a partial finitisation of the σ-additive theory; indeed he has proved in his «Radically elementary probability theory» [N2] that every stochastic process may be replaced without loss of probabilistic information by a nearly elementary process on a finite space with a finite set of times. Our finitisation is even more radical: we prove that every finitely additive conditional probability law on any algebra of events may be replaced without loss of probabilistic information by a regular simple law on some finite algebra. In this paper we will used the adjective «simple» in comparison to «conditional».

1. - Some recalls from elementary probability theory

Let Ω be a set, S a finite subalgebra of the set of events $\wp(\Omega)$ and π a probability law defined on S. We call this law *regular* in case $p(A) \neq 0$ for all $A \in S-\{\emptyset\}$.

An element $A \in S-\{\emptyset\}$ is called *prime* iff any $B \in S-\{\emptyset\}$ such that $B \subset A$ is equal to A.

Each element $B \in S-\{\emptyset\}$. contains at least one prime element. Moreover, if $E_1, \ldots, E_n$ are all the prime elements of S, then one has the canonical decomposition $B = \bigcup_{1}^{\bar{n}} B \cap E_i$, where $B \cap E_i = E_i$ or $B \cap E_i = \emptyset$. Thus a regular law p is completely defined by $\bar{n}$ positive numbers $p(E_i)$ with sum 1, through the formula $p(B) = \sum_{1}^{\bar{n}} p(B \cap E_i)$.

Recall that every finite subset G of $\wp(\Omega)$ generates a finite subalgebra which is the least subalgebra of $\wp(\Omega)$ containing G.

EXTENSION LEMMA: *Let S be a subalgebra of a finite algebra of events S^* and π a regular simple law on S. Then there is a regular law π^* on S^* which extends π.*

PROOF: Let $E_1, \ldots, E_n$ be the prime elements of S, and call v_i the number of prime elements of S^* contained in E_i. For each prime element K of S^* contained in E_i, define $\pi^*(K) = \pi E_i) / v_i$. Clearly this procedure yields a regular law on S^* which extends π. Of course, other extensions exist.

2. - FINITELY ADDITIVE CONDITIONAL PROBABILITY LAWS: THE MAIN PROBLEM

Recall that after de Finetti [F1] (see also [Re2]), the condition of coherence for conditional probabilities leads to the following definition, where may be infinite:

DEFINITION: *Let $H \subset A \subset \wp(\Omega)$ be algebras of events. A conditional (finitely additive) probability law p^* on the set $A \mid H = A \times (H - \{\emptyset\})$ of conditional events is a real valued function such that, for all $F \in H - \{\emptyset\}$, $K \in H - \{\emptyset\}$, $A \in A$, $B \in A$,*

(i) $p^*(A|H) \geq 0$.

(ii) *$H \subset A$ implies* $p^*(A|H) = 1$.

(iii) *If $A \cap B \cap H = \emptyset$, then* $p^*(A \cup B|H) = p^*(A|H) + p^*(B|H)$.

(iv) *If $K \cap H \neq \emptyset$, then* $p^*(A \cap H|K) = p^*(A|H \cap K) p^*(H|K)$.

To each conditional law p^*, we may associate the *simple probability law* p on (Ω, A) defined by the formula $p(A) = p^*(A|\Omega)$. The law p is finitely additive, that is $p(A \cup \cup B) = p(A) + p(B)$ whenever $A \cap B = \emptyset$.

From condition (iv), we get $p(A \cap H) = p^*(A|H) p(H)$. If $p(H) = 0$, this formula defines p^* from p. This is the case for all $H \neq \emptyset$ whenever p is *regular*.

On infinite spaces, most laws are not regular.

The conditional extension problem reads as follows:

Given a simple law p on A, and a subalgebra of «hypothesis» $H \subset A$, construct a conditional law p^* on the set $A \mid H$ of conditional events, such that $p = p^*(\cdot|\Omega)$.

Using tools of NSA, we shall give an alternative solution to extension problems of this kind, which have been extensively studied by B. de Finetti and his followers (see [Re2]). We use a «finitisation procedure» that we shall now develop.

3. - Finite reductions and the finitisation procedure

See the appendix for some essential topics of Non Standard Analysis in the setting of Nelson's Internal Set Theory.

Definition 3.1: *Let A be a standard algebra; a* reduction *S of A is a subalgebra of A which has the same standard elements as A. If S is finite, we call it a* finite reduction *of A.*

Remark: If S is standard, then S and A are both standard sets with the same standard elements. Hence by the transfer axiom, they are equal. Thus proper reductions have to be non standard.

Main tool 1: *Each standard algebra A admits a finite reduction.*

Proof: As a consequence of the idealisation axiom, there is a finite subset G of A which contains all standard elements of A. Then the subalgebra S of A generated by G is a finite reduction of A.

Remark: The simplicity of this proof is misleading. Indeed, it uses the full strength of the idealisation axiom scheme within NSA. Deepness may agree with simplicity! Notice that, as the choice of G is not unique, there may be different finite reductions of A.

Main tool: 2: *Let S be a reduction of an algebra of events $A \subset \wp(\Omega)$ and let π be a simple regular law on S. Then there is one and only one standard conditional law p^* on $A|A$ such that for all standard A and H, $A \in A$, $H \in A$-$\{\emptyset\}$, one has $p^*(A|H) = {}^{o}(\pi(A \cap H)/\pi(H))$.*

Proof: Define π^* on $S|S$ by the formula $\pi^*(A|H) = \pi((A \cap H)/\pi(H))$. This law satisfies all the requirements of definition 2.1. Since $S|S$ has the same standard elements as the standard set $A|A$, the shadow p^* of π^* is the unique standard law on $A|A$ such that for A and H standard, $p^*(A|H) = {}^{o}(\pi(A \cap H)/\pi(H))$ (see consequence (iii) of standardisation in the appendix). Using this formula, the verification (after transfer) that p^* is a conditional law is immediate.

We call p^* *the extended shadow of π*. Notice that the link between π and p^* given by the last formula works only for standard A and H. For non standard events, the formula is no longer true, but by transfer all internal properties of p^* may be verified on the standard events, so that we never have to compute p^* on non standard conditional events.

There is no analogous statement for σ-additive probability laws. In general, the shadow of such a law is not σ-additive. An equivalent classical formulation is the well known fact that the limit of a convergent sequence of σ-additive laws is, in general, not σ-additive.

REMARK: We may compute $p^*(A|H)$ even in case $P(H) = 0$, where p is the standard simple law associated to p^*. Indeed, this means that $\pi(H)$ and $\pi(A \cap H)$ are infinitesimals, but $\pi(H)$ is not 0. Thus, by direct comparison of infinitesimals, we read on π the rich stratification of the standard zero-probability events. We may also understand why $\sum \pi(A_i)$ can be non-infinitesimal even if all the $\pi(A_i)$ are infinitesimals: this can happen if the number of terms is infinitely large.

As a first application of these tools, let us prove.

THEOREM 3.1: *For each set Ω, there is a conditional law on $\wp(\Omega)|\wp(\Omega)$.*

PROOF: By transfer, we may suppose that Ω, and hence also the algebra $\wp(\Omega)$, is standard. By main tool 1, there is a finite reduction S of $\wp(\Omega)$. Call its prime elements E_i and, choosing positive numbers $\pi(E_i)$ with sum 1, define a simple regular law π on S as in §1. Then main tool 2 yields a conditional law on $\wp(\Omega)|\wp(\Omega)$.

Here we have no *a priori* condition on the resulting law. In more restrictive cases, we need the following

SIMPLE FINITISATION LEMMA: *Let S be a finite reduction of the standard algebra A. Then each standard simple law p on A is the shadow of some regular simple law π on S.*

PROOF: Let S be a finite reduction of A. Call E_i, $i = 1, \ldots, n$ the prime elements of S. At least one of the $p(E_i)$, say $p(E_n)$, is non-zero since $1 = p(\Omega) = \Sigma p(E_i)$. Choose $n - 1$ positive infinitesimal real numbers ε_i such that $\Sigma\varepsilon_i \simeq 0$ and $\Sigma\varepsilon_i < p(E_n)$.

Define $\pi(E_i) = p(E_i) + \varepsilon_i$ for $1 \leq i \leq n - 1$ and $\pi(E_n) = p(E_n) - \Sigma\varepsilon_i$.

Extend π by additivity to the algebra S. Then, for each standard event A in A,

$$\pi(A) = \sum_1^n \pi(E_i \cap A) = \Sigma p(E_i \cap A) + \eta \quad \text{with } |\eta| \leq \Sigma\varepsilon_i .$$

Thus $\pi(A) = p(A) + \eta$ and hence ${}^\circ(\pi(A)) = p(A)$ as requested.

THEOREM 3.2: *Each simple law p on an algebra $A \subset \wp(\Omega)$ extends to a conditional law p^* on $A \,|\, A$.*

PROOF: By transfer, we may suppose p and A standard and look for a standard extension p^*. Main tool 1 yields a finite reduction S and the simple finitisation lemma a law π on S whose extended shadow p^* is a solution of the problem.

REMARK: In the proof of the lemma, the choice of the finite reduction S is not unique. Moreover the choice of the ε_i has a deep influence on the value of $\pi(A \cap$

$\cap H) / \pi(H)$. For instance, suppose that E_1 and E_2 are standard, with $p(E_1) = p(E_2) = 0$. Then for $A = E_1$ et $H = E_1 \cup E_2$ we get $p(A|H) = {}^\circ(\varepsilon_1 / (\varepsilon_1 + \varepsilon_2))$ which depends on the arbitrary positive number $\varepsilon_2 / \varepsilon_1$. Thus the conditional extension of p is generally not unique.

Theorem 3.2 concerns the construction of a conditional law from a simple law in case some hypothesis have 0-probability. But if the conditional law is soon given on some $A \mid H$ where H is a subalgebra of A, the simple finitisation lemma is not sufficient to get a law π which respects this supplementary constraint. Thus we need a

Conditional finitisation lemma: *Each standard conditional law p^* on a standard $A \mid H$ is the extended shadow of some regular simple law π on an arbitrary finite reduction S of A.*

Proof: Define the $(S \cap H)$-valued sequence (G_n) by the inductive condition $G_0 = \Omega$, G_{n+1} = the union of all $K \in S \cap H$ such that $K \subset G_n$ and $p^*(K|G_n) = 0$.

As $p^*(G_n|G_n) = 1$, the sequence is strictly decreasing as long as $G_n \neq \emptyset$ and $F_n = G_n - G_{n+1}$ is a non empty element of the finite algebra $S \cap H$. Hence $G_n = \emptyset$ for n large enough. Call k the last index for which $G_n \neq \emptyset$. Then $F_k = G_k$ and, if $n \leqslant k$, for every non empty event $K \in S \in H$ such that $K \subset F_n$, we have $p^*(K|G_n) > 0$. Call μ_n the least of these positive numbers, and $\mu > 0$ the least of all the μ_n. Notice that μ is smaller than 1 and may be infinitesimal.

Thus, for all $H \in S \cap H$, and all $n \leqslant k$ such that $H \cap F_n \neq \emptyset$ we have

$$p^*(H|G_n) = p^*(H \cap F_n|G_n) + p^*(H \cap (\Omega - G_n|G_n) + \\ + p^*(H \cap G_{n+1}|G_n) = p^*(H|G_n) \geqslant \mu .$$

Each F_n can be partitioned in two elements F_n^0 and F_n^+ of S, where F_n^0 is the union of all elements K of S contained in F_n such that $p^*(K|G_n) = 0$. Hence $p^*(F_n^+|G_n) = 1$.

Call $\nu(A)$ the number of prime elements of S contained in A. The function ν is clearly additive on disjoint unions. Let ξ be a positive infinitesimal and $\varepsilon = \mu\xi$.

For each $n \leqslant k$ let η_n be a positive infinitesimal such that $\nu(F_n^0)\eta_n < \varepsilon$. Now we can define a law π on S by the formula

$$\pi(A) = \sum_{n=0}^{k} [p^*(A|G_n)(1 - \nu(F_n^0)\eta_N) + \nu(A \cap F_n^0)\eta_n] \varepsilon^n - p(A)(\varepsilon + \varepsilon^2 + \dots + \varepsilon^k).$$

It is clear that π is additive and that $\pi(\Omega) = 1$. The conditions above imply that If $\pi(A) = 0$, then $p^*(A|G_n) = 0$ and $n(A \cap F_n^0) = 0$ for all n; hence $A \cap F_n^+ = A \cap \cap F_n^0 = \emptyset$ for all n, which implies $A = \emptyset$. Thus π is a regular law on S. Now consider a standard pair $(A, H) \in A \mid H$.

Then $(A, H) \in S \mid S \cap H$ and we have to prove that $p(A \cap G) / \pi(H) = = p^*(A|H)$.

Call r the integer such that $H \subset G_r$ and $H \not\subset G_{r+1}$. Then $H \cap F_r \neq \emptyset$ and hence $p^*(H|G_r) \geq \mu$. From this and the conditions on ε and the η_n we get $\pi(A \cap H) = = p^*(A \cap H|G_r) + \alpha p^*(H|G_r)$ and $\pi(H) = p^*(H|G_r)(1+\beta)$ where α and β are infinitesimals. Hence the expected result follows from the formula

$$p^*(A \cap H|G_r)/p^*(H|G_r) = p^*(A|H \cap G_r) = p^*(A|H).$$

Thus p^* is the extended shadow of π.

Theorem 3.3: *Each conditional law p^* on $A\,|\,H$ extends to a conditional law on $A\,|\,A$.*

Proof: Transfer, choose a finite reduction of A and use the conditional lemma. Then the extended shadow of π to $A\,|\,A$ has the expected property.

A natural question is to extend a law from a subalgebra A to the whole algebra $\wp(\Omega)$, or from $A\,|\,H$ to $\wp(\Omega)|\,\wp(\Omega)$. This problem has been solved by de Finetti for simple laws in 1939, and for conditional laws by Regazzini [Re2]. Here we use a direct procedure, based on the Extension lemma of §1.

Theorem 3.4: *Every conditional law p^* on $A\,|\,H$ extends to a conditional law on $\wp(\Omega)|\,\wp(\Omega)$.*

Proof: By transfer suppose Ω, A, H, and p^* standard. Consider a finite reduction S^* of $\wp(\Omega)$; then $S = A \cap S^*$ is a finite reduction of A since all standard elements of A are in S^*. The conditional finitisation lemma yields a regular law π on S whose extended shadow to $A\,|\,H$ is p^*. This law extends to a regular law π^* on S^* (see §1). The extended shadow p^{**} of π^* to $\wp(\Omega)|\,\wp(\Omega)$ has the expected property.

Remark: As announced in the introduction, these results show that the finitely additive probability theory and the elementary finite probability theory have exactly the same scientific content. The link between both is «general nonsense» in which there are no probabilistic concepts, but only consequences of the principles of NSA.

Here below we will give another application of finite reductions. It concerns the misleading formulation of the strong law of large numbers.

4. - About the strong law of large numbers

4.1. Let Ω be a finite set, p a probability law on $\wp(\Omega)$ and $\Phi = \Omega^{\mathbb{N}}$ the set of all Ω-valued infinite sequences. The set Σ of cylinders, that is all products $A \times \Phi$ with $\Omega^n \supset A$ for some n, is a subalgebra of $\wp(\Phi)$. We extend p to the product law on Σ which is the only finitely additive law such that $p(\{\omega_1, \ldots, \omega_n\} \times \Phi) = p(\{\omega_1\}) \ldots p(\{\omega_n\})$.

Write $f_k(\omega : \phi)$ for the frequency of an $\omega \in \Omega$ within the k first terms of a sequence $\phi \in \Phi$. For each positive real number μ and for each pair of integers such that

$m < n$ consider the cylinder $A_{mn}(\mu, \omega, \phi) = \{\phi \in \Phi : p(\omega) - \mu \leq f_k(\omega : \phi) \leq p(\omega) + \mu, \forall k, m \leq k \leq n\}$.

The strong law of large numbers claims that *for all* $\mu > 0$ *and* $\varepsilon > 0$, *there is a rank* r *such that* $r < m < n$ *implies* $p(A_{mn}(\mu) > 1 - \varepsilon$.

In the theory of Kolmogorov, there is a unique σ-additive extension p^* of p to the σ-algebra Σ^* generated by Σ. For this special extension, the strong law has the following elegant consequence: $p^*(B(\omega, \phi)) = 1$, where $B(\omega, \phi) = \{\phi \in \Phi : \lim f_n(\omega : \phi) = p(\omega)\}$.

This formulation gives the impression that the behaviour of the «good» statistical sequences (that is the sequences which would be generated by a random device with independent issues) is such that the frequencies tend to the corresponding probabilities when the length of the sample tends to infinity. And the strong law tells us that nearly all sequences are «good», a behaviour that we may expect from actual randomness and which gives confidence in the frequential interpretation of probability.

4.2. However his interpretation is specific to the special extension p^* of p. But other extensions, which are only finitely additive, have the same legitimity concerning statistical experiences (which are always restricted to cylinders). However for such laws (which have the advantage to be extendible to the whole algebra $\wp(\Phi)$) the probability of $B(\omega, \phi)$ behaves as bad as possible (see [Re3]).

THEOREM 4.2: *For each real number* $\alpha \in [0, 1]$, *there is an extension* p' *of* p *to* $\wp(\Phi)$ *such that* $p'(B(\omega, \phi)) = \alpha$.

Thus if we choose such a law with $\alpha = 0$, we conclude that nearly all sequences are «bad» ... and that finitely additive probability has nothing to do with randomness!

Let us understand this through the finitisation technic of § 3 which shows clearly the degrees of freedom shared by finitely additive extensions even if they are defined on the whole algebra $\wp(\Phi)$.

LEMMA 4.2: *Let* Φ *be a set,* A *a subalgebra of* $\wp(\Phi)$, p *a simple law on* A, $\alpha \in [0, 1]$ *and* $A \in \wp(\Phi)$. *Suppose that* A *and its complement* A^c *have non empty intersections with each element of* A. *Then there is an extension* p' *of* p *to* $\wp(\Phi)$ *such that* $p'(A) = \alpha$.

PROOF: By transfer, we may restrict the proof to the case where the constants of the statement Φ, A, p, α, A are standard, and we look for a standard law p'.

Let S be a finite reduction of $\wp(\Phi)$; and F the finite subalgebra $S \cap A$. Call $F_1 \dots F_k$ the prime elements of F and $S_1 \dots S_n$ these of S. Then every F_i is the disjoint union of some of the S_j. To define a law π on S whose shadow is a standard law p' on equal to p on A, it is sufficient to share for each i the number $p(F_i)$ between all the S_j which are contained in F_i. As A is standard, $A \in S$; hence A is a union of some of the S_j. From the hypothesis we get that every F_i contains one of them, and the same is true for A^c.

Choose k non negative real numbers α_i of sum α such that for every i, $\alpha_i \leq p(F_i)$. Then distribute the α_i on the S_i which constitute $A \cap F_i$ and $p(F_i) - \alpha_i$ on these which constitute $A^c \cap F_i$. This defines a law π on S whose shadow satisfies the additional condition $p'(A) = \alpha$.

Now it is easy to see that the sets $B(\omega, \phi)$ satisfy the hypothesis of the lemma, since they meet each cylinder (one may always continue a given finite sequence such that the frequency of ω tends to a given value), but are not contained in a cylinder (since the n first terms of a sequence with given frequency are arbitrary). This proves Theorem 4.2.

Further applications of finite reductions (e.g. to the study of random variables and stochastic processes) will be dealt with in another paper. For the moment, we confine us to investigate the power of another finitisation tool which may be useful to construct a law that agrees with some additional structure of the set Ω.

5. Pointwise reduction and equiprobability on number fields.

5.1. In this section we construct conditional probability laws on some infinite $\wp(\Omega)|\wp(\Omega)$, which remain invariant under the action of some transformation group on the set Ω. This concerns mainly the number sets N, Z, Q, R with their groups of translations. We need another finitisation technic, which concerns the points of Ω and not a subalgebra of $\wp(\Omega)$. This is

Main tool 3: *Let Ω be a standard set, F a finite subset of Ω which contains all standard elements of Ω. Let $\phi: F \to]0, 1]$ be a strictly positive function such that*

$$\sum_{x \in F} \phi(x) = 1. \text{ Define } \pi: \wp(\Omega) \to [0, 1] \quad \text{by } \pi(A) = \sum_{x \in A \cap F} \phi(x) \text{ for } A \subset \Omega.$$

Then π is a simple law on $\wp(\Omega)$ but it is not regular (because $\pi(A) = 0$ in case $A \cap F$ is empty).

If $H \in \wp(\Omega)$-$\{\emptyset\}$ is standard, then $H \cap F \neq \emptyset$, since H contains some standard element; t h u s $\pi(H) \neq 0$. Then there is one and only one standard conditional law p^* on $\wp(\Omega)|\wp(\Omega)$ such that, for any standard A and H, $p^*(A|H) = {}^\circ(\pi(A \cap H)/\pi(H))$.

Proof The existence of p^* follows from the standardisation axiom, while the uniqueness and the verification of the properties follow from the transfer axiom.

We call this *point-wise reduction procedure* an (F, ϕ)*-construction*.

We may choose F and ϕ in order to satisfy some additional conditions on p^* or on the associated simple law p.

We consider here three important examples of such choices that solve equiprobability problems. The two first ones have been solved previously by E. Regazzini [Re1] using a classical limit procedure and de Finetti's extension lemma.

5.2. Example 1: *Construct a law on $\wp(Z)$ which is invariant under all translations*

By transfer, a standard law is invariant under all translations iff it is invariant under the standard translations applied to standard subsets of Z. Take two infinitely large positive integers m and n, and consider $F = \{-m, \ldots, 0, \ldots, n-1\}$ with the constant function $\phi(x) = 1/(m+n)$. The resulting standard law p is invariant under all standard translations. Indeed, for A standard, if a (resp. b) is the cardinal number of $F \cap A$ (resp. $F \cap t(A)$), then the absolute value of the difference $b - a$ after a translation t by a standard integer τ is less than τ. Hence $\pi(t(A)) - \pi(A)$ is infinitesimal and thus $p(t(A)) = p(A)$.

Notice that if $m/n \simeq 1$, then p is also invariant under the symmetry around any point. The induced law $p^*(\cdot | N)$ is a simple law on N which is invariant under all positive translations. Taking $m = n$, we call this standard law p_n and compute the probability of some subsets of N to see how it depends on the choice of n.

Standard singletons (hence, by transfer, all singletons) have zero probability. Thus, by additivity, finite subsets have zero probability.

If A is the set of even integers and B the set of the odd ones, we get $p_n(A) + p_n(B) = 1$ and $p_n(A) = p_n(B)$, since B is the image of A under translation by 1. Hence $p_n(A) = 1/2$. For the same reason the set of multiples of an integer k has probability $1/k$.

Here is an infinite set with zero probability: $D_2 = \{2^k \cdot k \in N\}$. In fact, if $2^s \leq n < 2^{s+1}$, then $\mathrm{card}(D \cap F) = s+1$ and $p_n(D) = {}^\circ((s+1)/n) = ((s+1)/2^s \cdot 2^s/n) = 0$ since s is infinitely large. The same is true for D_a, a standard. By transfer, every D_a has zero probability for any law p_n.

Let us consider an example of a standard «exotic» subset E of N such that $p_n(E)$ depends on n.

Consider the union E of all $A_k = \{4^k, 4^{k+1}+1 \ldots, 4^k + 4^k - 1\}, k \in N$. If $n = 4^b$ for some integer b, we get $p_n(E) = 1/3$. If $n = 2.4^b$ the probability is $p_n(E) = 2/3$. In the general case the value lies between $1/3$ and $2/3$.

Observe here the structure of the conditional probabilities that you get from the laws π_n in case the hypothesis have zero probability for the shadow laws p_n.

For instance, if A and H are standard and finite, H non-empty, then $p_n(A|H) = \mathrm{card}(A \cap H)/\mathrm{card}(H)$ since $A \cap H \cap F = A \cap H$ et $H \cap F = H$.

A good example with infinite subsets is $p_n(D_4 | D_2) = 1/2$.

5.3. Example 2: *Construct a law on $\wp(Q)$ which is invariant under all rational translations.*

Choose an infinitely large integer n and consider the set $F = \{z/n! \text{ with } z \in Z$ and $|z| \leq (n+1)!\}$. Choose again ϕ constant, i.e. $\phi(x) = 1/\mathrm{card}(F)$. The same argument as above shows that the resulting law is invariant under all rational translations and point-symmetries.

Call $p_1 = p^*(\cdot \mid Q_1)$ the induced law on $Q_1 = Q \cap [0, 1]$. For every standard subinterval of Q_1, one has $p_1([\alpha, \beta]) = \beta - \alpha$. This follows from $p_1([\alpha, \beta]) = {}^{\circ}(\pi([\alpha, \beta]) / \pi([0, 1])) = {}^{\circ}(b - a + 1 / 1 + n!) = \beta - \alpha$, where $\alpha = a / n!$ and $\beta = b / n!$. By transfer, we get that *the probability of each subinterval of Q_1 is equal to its length.* This property is much weaker than the invariance by translation and does not depend on the value of n.

Notice that *all bounded subsets of Q have zero probability.*

Indeed, by transfer, restrict the proof to standard subsets. Then the number $b - a + 1$ of rationals $z / n!$ between two standard rationales $\alpha = a / n!$ and $\beta = b / n!$ is infinitesimal relative to $2(n + 1)!$, which gives zero probability for the shadow-law. Again, one has here a rich conditional structure among zero-probability subsets. There are also subsets whose probability depends on n.

Clearly one gets uniform induced laws on the set of decimal or dyadic numbers.

A more difficult case is

5.4. Example 3: *Construct a law on $\wp(R)$ which is invariant under all real translations.*

Let $G = \{g_1, \ldots, g_k\}$ be a finite subset of $[0, 1]$ which contains all standard elements of $[0, 1]$. Call $Z.(G)$ the Z-sub-modules of R generated by G. Each element x of $Z(G)$ has at least one (non-unique) representation as $x = \sum n_i g_i$ with $n_i \in Z$. Choose an infinitely large integer ω and consider the subset F of those element in $Z(G)$ which have at least one representation with all coefficients $n_i \in [-\omega, \omega]$. Then F is finite since $\mathrm{card}(F) \leq (2\omega + 1)^k$. Moreover F contains G and also $G \pm n$ for all standard integers n. Hence F contains all standard real numbers. As previously, take ϕ to be constant on F.

We claim that *the (standard) simple law p associated to this (F, ϕ)-construction is invariant under all translations and point-symmetries.*

In fact, let t_s be the translation by a non-zero element g_s of G. It leaves the infinite modulus $Z(G)$ invariant but not the set F. Thus we try to compare $\mathrm{card}(t_s(F) - F \cap \cap\, t_s(F))$ with $\mathrm{card}(F)$.

For each $x = \sum n_i g_i$ where $n_i \in [-\omega, +\omega] \cap Z$ and $t_s(x) \notin F$, one has $n_s = \omega$. If not, then $-\omega \leq n_s < \omega$; hence $t_s(x) = x + g_s$ would have all its coefficients in $[-\omega, +\omega] \cap Z$, and consequently would be an element of F.

Thus if $t_s(x) \notin F$, the $2\omega + 1$ distincts elements $x, x - g_s, \ldots, x - 2\omega g_s$ of $Z(G)$ have a representation with all coefficients in $[-\omega, +\omega] \cap Z$, hence belong to F. If an other element $x' \in F$ satisfies $t_s(x') \notin F$, then we claim that the two sets $\{x, x - g_s, \ldots, x - 2\omega g_s\}$ and $\{x', x' - g_s, \ldots, x' - 2\omega g_s\}$ don't intersect.

In the opposite case, there would be two distinct integers a et a' between 0 and $2n$

such that $x - ag_s = x' - a'g_s$; if $a < a'$, this would imply $t_s(x) \in F$ since $t_s(x) = x' - (a' - a - 1)g_s$ with $0 \leqslant a' - a - 1 \leqslant 2\omega - 1$. In the same way $a' < a$ would imply $t_s(x') \in F$; both cases contradict the hypothesis.

Hence to each $x \in t_s(F) - F \cap t_s(F)$ corresponds a subset of F with $2\omega + 1$ elements and the subsets associated to distinct x don't intersect. As a result we get

$$\mathrm{card}\,(t_s(F) - F \cap t_s(F)) \leqslant \mathrm{card}\,(F) / (2\omega + 1).$$

For the same reason, we get $\mathrm{card}\,(t_s^{-1}(F) - F \cap t_s^{-1}(F)) \leqslant \mathrm{card}\,(F) / (2\omega + 1)$. As t_s is one-to-one from $\boldsymbol{R}$ to $\boldsymbol{R}$, we deduce from this that $\mathrm{card}\,(t_s(F) \cap F) \leqslant \mathrm{card}\,(F) / (2\omega + 1)$.

Let A be a standard subset of $\boldsymbol{R}$. Then $\pi(A) = \mathrm{card}\,(A \cap F) / \mathrm{card}\,(F) = \mathrm{card}\,(t_s(A \cap F)) / \mathrm{card}\,(F) = \mathrm{card}\,(t_s(A) \cap t_s(D)) / \mathrm{card}\,(F)$ and $\pi(t_s(A)) = \mathrm{card}\,(t_s(A) \cap F) / \mathrm{card}\,(F)$.

Now $t_s(A) \cap F = [t_s(A) \cap (t_s(F) \cap F)] \cup [t_s(A) \cap (F - t_s(F) \cap F)]$ (disjoint union) and $t_s(A) \cap t_s(F) = [t_s(A) \cap (t_s(F) \cap F)] \cup [t_s(A) \cap (t_s(F) \cap F)]$ (disjoint union). Thus $\pi(A) = \mathrm{card}\ [t_s(A) \cap (t_s(F) \cap F)] / \mathrm{card}\,(F) + \xi$ and $\pi(t_s(A)) = \mathrm{card}\ [t_s(A) \cap (t_s(F) \cap F)] / \mathrm{card}\,(F) + \zeta$ with $0 < \xi < 1/(2\omega + 1)$ and $0 < \zeta < 1/(2\omega + 1)$. Hence $\pi(t_s(A)) = \pi(A)$.

Since each translation t by a standard element of $[0, 1]$ is one of the translations t_s, this proves that the shadow p of π is invariant under all these translations (transfer to get $p(t_s(A)) = p\{A\}$ for all A). As p is standard, by transfer it is also invariant under all translations by an element of $[0, 1]$. By iteration, we get invariance under all translations. Moreover, as F is invariant under symmetry around 0, we get that p is invariant under all isometries of $\boldsymbol{R}$.

In this construction, the law p depends on the choice of G and ω. Nevertheless there are subsets A of $\boldsymbol{R}$ for which $p(A)$ can be computed directly from the invariance property.

For instance, suppose that there is an infinite sequence $t_1, \ldots, t_n \ldots$ of translations such that the subsets $t_i(A)$ are mutually disjoint. Then $p(\bigcup_1^n t_i(A)) = np(A) < 1$ for every integer n, hence $p(A) = 0$. From this we get $p(\boldsymbol{Q}) = 0$ (iterate the translation by an irrationnal number). Also each bounded subset of $\boldsymbol{R}$ has zero probability.

The above construction can be easily extended to $\boldsymbol{R}^d$ where the dimension d is standard. By transfer, this proves the existence of a finitely additive law on each euclidean space which is defined on all subsets and invariant under all translations.

Appendix

A brief account on NSA.

We give here the elements of NSA which are essential to read our paper. This theory was created in the early sixties by A. Robinson [Ro] who developped it as a consequence of Model theory. In 1977, E. Nelson [N1] gave an axiomatic formulation of NSA which we use in this paper. In his formal theory called Internal Set Theory (IST), he extends the language and the axiomatic of the Zermelo-Fraenkel (ZF) set theory. One of the possibilities of this extended theory is to formalise the infinitesimal concepts that Leibniz introduced in 1670.

Technically, the language of IST contains the language of ZF, whose formulas are called *internal* , and a new monadic predicate called «st» (read *standard*). The formulas which contain somewhere the predicate «st» are called *external*. The axiomatic is constituted by all the axiom schemes of ZF restricted to internal formulas and by three additional schemes called *transfer*, *idealisation* and *standardisation*, which regulate the semantics of *standard*. We give these schemes with some important consequences. To this end we use the following abbreviations:

$\forall^{st}x$ for $\forall x\, st(x)$ (read *for every standard x*)

and $\exists^{st}x$ for $\exists x\, st(x)$ (read *there is a standard x*).

Transfer scheme: *For each internal formula $A(x, t_1, \ldots, t_k)$ with $k+1$ free variables, one has the axiom:*

$$\forall^{st}t_1 \ldots \forall^{st}t_k[\forall^{st}xA(x, t_1, \ldots, t_k) \Rightarrow \forall xA(x, t_1, \ldots, t_k)].$$

This axiom ensures that, for all standard values of the parameters $t_1, \ldots, t_k$, the property $A(x, t_1, \ldots, t_k)$ is true iff it is true for every standard value of x.

By contraposition, if $B(x, t_1, \ldots, t_k)$ is the negation of $A(x, t_1, \ldots, t_k)$, then, for all standard values $t_1, \ldots, t_k$, there exists an x satisfying the property B iff there exists a standard x with this property. In particular, if $B(x, t_1, \ldots, t_k)$ is only satified for one element x_0, this element is necessarily standard; for instance, every uniquely defined object within classical analysis is standard. So are the number sets $N, Z, Q, R, C \ldots$ and the explicit numbers $0, 1, 2, \ldots, 1654327, \pi, e$, the functions exp, log, sin, cos,....

Notice that to prove within IST an internal theorem $\forall x A(x, t_1, \ldots, t_k)$, where the parameters have fixed values, the transfer scheme allows to restrict the proof to the standard values of x, provided all the parameters $t_1, \ldots, t_k$ are restricted to standard values.

Idealisation scheme: *For each internal formula $B(x, y, t_1, \ldots, t_k)$ with free variables $x, y, t_1, \ldots, t_k$, one has the axiom:*

$$\forall^{st}t_1 \ldots \forall^{st}t_k[\forall^{st}Z, Z \text{ finite}, \exists x\forall y \in ZB(x, y, t_1, \ldots, t_k)] \Leftrightarrow [\exists\xi\forall^{st}yB(\xi, y, t_1, \ldots, t_k)].$$

When the parameters are fixed to standard values, this axiom sheme yields an ideal element ξ which is related to all standard y, provided the binary relation B satisfies the first part of the statement.

The most important consequences of the idealisation scheme are the following:

(i) *A set E is standard and finite iff all its elements are standard.*

Hint: take as $B(x, y, E)$ the formula «$x \in E$ and $y \in E$ and $x \neq y$».

(ii) *The set of integers* N *contains infinitely large elements, that is integers which are larger than all standard integers.*

Hint: take as $B(x, y, N)$ the formula «$x \in N$ and $y \in N$ and $x \geq y$».

This result gives a solid foundation to the infinitesimal calculus of Leibniz. In fact there are infinitely large real numbers (these with infinitely large integral part) and thus by inversion infinitesimal real numbers which satisfy the classical rules concerning the two operations. We write $x \simeq y$ for «x is infinitely near y». Notice that the sum of an infinitely large number of infinitesimals may take any value, while the sum of a non-infinitely large number or infinitesimals is infinitesimal.

(iii) *There is a finite set F such that every standard x is an element of F.*

Hint: Take as B the relation «x is finite and $y \in x$».

Such a finite set F cannot be standard: in fact, if F would be standard, then its intersection E with the set N of integers would be a finite standard set. Hence all its elements would be standard. Call k its last element. Then $k + 1$ is standard, hence belongs to E, a contradiction.

One of the most misleading features of NSA is the risk to use illegal subset formation, that is to apply the extensionality axiom of ZF to non-internal properties. For instance, there is no «subset of standard integers», nor of «infinitely large integers» or «of infinitesimal real numbers». A positive counterpart is the

Standardisation scheme: *For each formula (internal or not) $C(x, t_1, \ldots, t_k)$ with free variables $x, t_1, \ldots, t_k$, one has the axiom:*

$$\forall t_1 \ldots \forall t_k [\forall^s E \exists^s Y \forall^s x [x \in Y \Leftrightarrow x \in E \text{ and } C(x, t_1, \ldots, t_k)]].$$

Notice that the set Y is unique, since two different solutions would have the same standard elements, which contradicts the transfer axiom scheme.

In the main text we use three important consequences of standardisation.

(i) *Each real number a such that $|a|$ is non infinitely large has a «shadow» ${}^\circ a$, that is an unique standard real number such that ${}^\circ a \simeq a$.*

Hint: Take $t_1 = a$, $E = \mathbf{R}$, and for C the formula «$x < a$». You get a standard subset Y of $\mathbf{R}$, bounded from above by the same standard number as a. The least upper bound of Y is the shadow of a.

The shadow satisfies the following properties, whenever the shadows exist: ${}^{\circ}(a+b) = {}^{\circ}a + {}^{\circ}b$, ${}^{\circ}(ab) = ({}^{\circ}a)({}^{\circ}b)$ and if $a < b$, then ${}^{\circ}a \leq {}^{\circ}b$. Moreover, if a is standard, ${}^{\circ}a = a$.

(ii) *For every subset B of a standard set E, there is a unique standard subset A of E which has the same standard elements as B.*

HINT: Apply standardisation to the formula «$x \in B$».

(iii) *With the same notation as in* (ii), *let f be a function on B such that for each element a of B, the image $f(a)$ has a shadow. Then there exists a unique standard function g on A, called the shadow of f, satisfying the relation $g(a) = {}^{\circ}(f(a))$ for every standard element a of A.*

By transfer, the internal properties of g may be proved by restriction to the standard elements of A, and thus may be deduced from properties of the initial function f. Notice that for non standard elements $a \in A \cap B$, $g(a)$ may be quite different from ${}^{\circ}(f(a))$.

A justification of NSA is the meta-theoretical result (proved by finitary arguments) that the theory IST is a conservative extension of ZF. This means that if an internal statement T is a theorem in IST, then it has also a proof within ZF, hence it is a theorem in ZF.

Thus we may legitimately use the powerful tools of NSA to prove any theorem which can be formulated in the language of classical mathematics. Often, such a proof is easier to discover than a classical one. This is due to more direct and natural external formulation of the wanted result after transfer. The «economy» depends mainly on how much idealisation you may use. In this paper—as elsewhere in the probability theory (see e.g. [N1] and [N2])—the main part is plaied by the deep consequence (iii) of the idealisation sheme.

REFERENCES

[D-R] F. DIENER - G. REEB, *Analyse Non Standard*, Hermann, Paris (1989).
[F1] B. DE FINETTI, *La prévision: Ses lois logiques, ses sources subjectives*, Ann. Inst. H.Poincaré, 7 (1937), 1-68.
[F2] B. DE FINETTI, *Teoria della Probabilità*, Einaudi, Torino (1970).
[L-G] R. LUTZ - M. GOZE, *Non Standard Analysis. A Pratical Guide with Applications*, Lecture Notes in Mathematics, 942, Springer (1981).
[M] M. MUSIO, *L'Analisi Non Standard: una rapida panoramica e qualche applicazione*, Tesi di Laurea, Cagliari (1991).
[N1] E. NELSON, *Internal set theory: a new approach to Non Standard Analysis*, Bull. Amer. Math. Soc., 83 (1977).
[N2] E. NELSON, *Radically Elementary Probability Theory*, Ann. Math. Studies, 117, Princeton U. Press (1987).

[P] L. Piccinato, *Bayesian Inference and Decision Techniques*, ed. P. Goel and A. Zellner, North-Holland (1986).

[Re1] E. Regazzini, *Coherence, Exchangeability and Statistical Models*, Atti del convegno della S.I.S., 6 Aprile 1991, Cagliari, vol 2, 101-137.

[Re2] E. Regazzini, *Finitely additive conditional probabilities*, Rend. Sem. Mat. Fis. di Milano, 55 (1985), 69-89.

[Re3] E. Regazzini - G. Zanoni, *Stochastic convergence in a finitely additive setting*, Dip.to Mat. «F. Enriques». Univ. di Milano (1988), 29-188.

[Ro] A. Robinson, *Non Standard Analysis*, North-Holland, Amsterdam (1966).

Direttore responsabile: Prof. A. Ballio - Autorizz. Trib. di Roma n. 7269 dell'8-12-1959
«Monograf» - Via Collamarini, 5 - Bologna