PREVALENCE THEORY: A MEASURE-THEORETIC APPROACH TO LARGE SETS IN INFINITE-DIMENSIONAL SPACES

SAMUEL NICOLAY

This article has been accepted for publication in the American Mathematical Monthly.

Abstract. The concept of prevalence provides a rigorous framework for identifying "large" sets in infinite-dimensional spaces, extending measure-theoretic ideas of negligibility beyond finite-dimensional contexts. This paper introduces the foundational definitions of prevalence, exploring its key properties and illustrating its relevance through examples in functional spaces. We also investigate the interplay between prevalent sets and Baire categories, emphasizing that while these notions share certain structural parallels, they are fundamentally distinct.

1. Introduction.

What is a "small set"? The simplest way to approach this notion is likely to enumerate the properties that such a set should possess, as a basis for its definition. The most natural properties are as follows:

($\mathcal{P}1$) A small set has an empty interior.
($\mathcal{P}2$) Any subset of a small set is itself small.
($\mathcal{P}3$) Any countable union of small sets remains small.
($\mathcal{P}4$) Any translate of a small set is also small.

Clearly, a minimal amount of structure is required to formulate these conditions. Let us assume we are working in a vector space. In finite dimensions, the Lebesgue measure, denoted $\mathcal{L}$, provides a natural notion of smallness: one can equate small sets with negligible sets (in the sense of Lebesgue measure). Conventionally, a property is said to hold almost everywhere if the set of points where it fails is negligible.

It is natural to ask whether an analogous notion of small set, based on a suitable measure that would serve as a counterpart to Lebesgue measure in defining negligibility, could be defined in infinite-dimensional spaces so that properties ($\mathcal{P}1$)–($\mathcal{P}4$) continue to hold. However, this hope is obstructed by ($\mathcal{P}4$). Indeed, if μ is a translation-invariant measure defined on the Borel sets of an infinite-dimensional Banach space, then there exists a ball with measure zero. This violates property ($\mathcal{P}1$), showing that translation invariance is too restrictive a condition [1].

It suffices, however, for property ($\mathcal{P}4$), referred to as quasi-invariance, to hold. A measure μ defined on the Borel sets of a vector space is said to be quasi-invariant if, for every Borel set A of measure zero, its translates also have measure zero: $\mu(A) = 0$ implies $\mu(A+x) = 0$ for all x in the space. Unfortunately, this condition is also too restrictive: if a measure defined on the Borel sets of a separable infinite-dimensional Banach space is σ-finite, regular, and quasi-invariant, then it is identically zero.

The idea proposed by Brian Hunt, Tim Sauer, and James Yorke [1] was to explore a property equivalent to having a null Lebesgue measure, one that could reasonably extend to infinite-dimensional spaces. A comprehensive survey of the resulting theory of prevalence is provided in [2].

2. Prevalence.

The concept of prevalence is underpinned by the following result:

Theorem 1. *A Borel set A in Euclidean space $\mathbb{R}^n$ is negligible in the sense of Lebesgue measure if and only if there exists a probability measure μ, defined on the Borel σ-algebra with compact support, such that $\mu(A+x)=0$ for all $x \in \mathbb{R}^n$.*

Proof. For the necessary condition, consider, for any Borel set A, the measure $\mu(A)=\mathcal{L}(A\cap B)/\mathcal{L}(B)$, where B denotes the unit ball centered at the origin. The support of μ is contained in the closure of B, which is compact.

Conversely, an application of the Fubini –Tonelli theorem to the function

$$f:\mathbb{R}^n\times\mathbb{R}^n\to\{0,1\}\quad (x,y)\mapsto \chi_A(x+y),$$

where χ_A denotes the indicator function of A, yields the identity

$$\int \mu(A-y)\,dy=\int \mathcal{L}(A-x)\,d\mu(x).$$

Now, the first integral vanishes by assumption, and the second can be rewritten as

$$\int \mathcal{L}(A-x)\,d\mu(x)=\int \mathcal{L}(A)\,d\mu(x)=\mu(\mathbb{R}^n)\mathcal{L}(A)=\mathcal{L}(A),$$

which proves that $\mathcal{L}(A)=0$. ☐

Building on Theorem 1, we arrive at the following definition:

Definition 2. A measure μ, defined on the Borel σ-algebra of a completely metrizable topological vector space, is said to be transverse to a Borel set A if:

($\mathcal{C}1$) There exists a compact set K such that $\mu(K)\in(0,\infty)$.
($\mathcal{C}2$) $\mu(A+x)=0$ for every point x in the space.

If the space is further assumed to be separable (thus forming a Polish topological vector space), condition ($\mathcal{C}1$) is automatically satisfied for any finite measure.

Definition 3. A Borel set A in a completely metrizable topological vector space is called shy if there exists a measure transverse to A. More generally, a set is shy if it is contained in a shy Borel set. The complement of a shy set is called prevalent.

In Definition 2, the measure is not required to be a probability measure with compact support. This flexibility allows the consideration of Lebesgue measure on finite-dimensional subspaces of the ambient space. It can be shown that every shy Borel set admits a finite, compactly supported transverse measure.

It can also be proven that if A is shy, then:

- Any subset of A is also shy.
- Any translate of A is shy.
- A has an empty interior.
- λA is shy for any $\lambda>0$.

- The complement of A is dense in the space.

Furthermore, using measure convolution, it can be shown that a countable union of shy sets remains shy. Finally, every countable set is shy, and in infinite dimensions, compact sets are shy. In particular, properties $(\mathcal{P}1)$ through $(\mathcal{P}4)$ are satisfied.

Thus, shy sets possess the necessary properties to be associated with the notion of small sets. It is worth noting that in Euclidean space, a set is shy if and only if it is contained in a negligible Borel set. In this sense, prevalence can be viewed as a generalization of the concept of "almost everywhere".

To conclude this section, we briefly mention Jens Christensen's work on Haar-null sets, which offers a slightly different perspective on small sets. Rather than working within completely metrizable topological vector spaces, Christensen focuses on Polish abelian topological groups [3, 4]. This approach predates that of Hunt, Sauer, and Yorke; however, the latter were initially unaware of Christensen's work [5]. In this framework, shy sets are referred to as Haar-null sets. Christensen does not aim to derive results on prevalence; instead, he employs Haar-null sets as a tool. In particular, he demonstrates that if a Polish abelian topological group G is not locally compact, then the compact subsets of G are Haar-null.

3. A working example: Cauchy's functional equation.

Cauchy's functional equation [6] is given by

$$f(x+y)=f(x)+f(y).$$

A function f that satisfies this equation is called an additive function. On the real numbers $\mathbb{R}$, it is straightforward to show that any additive function is linear with respect to rational numbers. Specifically, for any $x\in\mathbb{R}$ and any rational number r, we have $f(rx)=rf(x)$. Consequently, if f is continuous at even a single point, it must be linear: $f(x)=f(1)x$ for all $x\in\mathbb{R}$. Furthermore, additive functions are linear if they are Lebesgue measurable, monotonic on an interval, or bounded on an interval.

However, assuming the axiom of choice, Hamel demonstrated the existence of additive functions that are not linear [7]. Since every vector space has a basis, there exists a subset $\mathcal{H}$ of $\mathbb{R}$, called a Hamel basis, such that any real number x can be uniquely expressed as

$$x=\sum_{j\in J} r_j x_j,$$

where $(x_j)_{j\in J}$ is a finite subset of $\mathcal{H}$ and each r_j is rational. For any $x_j\in\mathcal{H}$, an additive function f must satisfy $f(rx_j)=rf(x_j)$ for any rational r, so the restriction of f to $x_j\mathbb{Q}$ is given by

$$f: x_j\mathbb{Q}\to\mathbb{R} \quad x\mapsto x\frac{f(x_j)}{x_j}.$$

Consequently, for any real number x expressed in the Hamel basis $x=\sum_{j\in J} r_j x_j$, we have $f(x)=\sum_{j\in J} r_j f(x_j)$. Thus, defining f on $\mathcal{H}$ completely determines the function. In particular, f is linear if and only if $f(x_j)/x_j=f(x_k)/x_k$ for all indices j and k. Non-linear solutions are highly irregular; in fact, their graphs are dense in $\mathbb{R}^2$.

Since the cardinality of $\mathcal{H}$ is the same as the cardinality of the continuum $\mathfrak{c}$, there are $2^{\mathfrak{c}}$ functions $f:\mathcal{H}\to\mathbb{R}$, while the number of linear solutions is only $\mathfrak{c}$.

From this perspective, linear solutions to Cauchy's equation are "rare". To provide deeper insight into the functional landscape beyond mere cardinality, we employ the measure-theoretic framework introduced above. Notably, even on $\mathbb{R}$ the length of a set is unrelated to its cardinal. For example, the triadic Cantor set has the cardinality of the continuum but is negligible with respect to the Lebesgue measure (and thus shy).

To demonstrate that the set of linear solutions is shy, consider the narrowed space $C_0(\mathcal{H})$ of continuous, bounded functions $f : \mathcal{H} \to \mathbb{R}$, equipped with the standard norm. An additive function f can be associated with the function $x_j \mapsto f(x_j)/x_j$ defined on $\mathcal{H}$. To establish that the set of linear solutions is small among additive functions, we can show that the vector subspace S of constant functions, considered as a subset of the $C_0(\mathcal{H})$, is shy. This follows because S is a one-dimensional vector space.

Let $T = \{\lambda \cos : \lambda \in \mathbb{R}\}$, and define the measure $\mu(B) = \mathcal{L}(\varphi^{-1}(B))$ on the Borel sets of $C_0(\mathcal{H})$, where $\varphi : \mathbb{R} \to T$ is the map $\varphi(\lambda) = \lambda \cos$ and $\mathcal{L}$ is the Lebesgue measure. For any $g \in C_0(\mathcal{H})$, we have:

$$\begin{aligned}\mu(S+g) &= \mathcal{L}(\lambda \in \mathbb{R} : \lambda \cos - g \in S\}) \\ &= \mathcal{L}(\lambda \in \mathbb{R} : g = \lambda \cos + C \text{ for some } C \in \mathbb{R}\}) = 0,\end{aligned}$$

which is sufficient to conclude that the set S within $C_0(\mathcal{H})$ is shy.

The approach outlined above can be generalized as follows:

Definition 4. Let X be a completely metrizable vector space and S be a finite-dimensional subspace of X with dimension n. A Lebesgue mesure on X supported on S is a measure of the form

$$\mathcal{L}_S(A) = \mathcal{L}(\varphi^{-1}(A \cap S)),$$

for any Borel set A, where $\varphi : \mathbb{R}^n \to S$ is an isomorphism.

The choice of a different isomorphism does not alter the negligible sets of $\mathcal{L}_S$.

Definition 5. A probe of a subspace $A \subset X$ is a finite-dimensional vector subspace $S \subset X$ for which there exists a Lebesgue measure supported on S that is transverse to a Borel set that contains $X \setminus A$.

Proposition 6. *If a subset $A \subset X$ admits a probe, then A is prevalent.*

Let us demonstrate that any proper Borel vector subspace Y of X is shy. Let $A = X \setminus Y$. We aim to construct a probe for A. Any vector subspace S spanned by a vector $a \in A$ satisfies the conditions. Define

$$\varphi : \mathbb{R} \to S \quad \lambda \mapsto \lambda a.$$

For any $x \in X$, we have

$$\mathcal{L}_S(Y + x) = \mathcal{L}(\{\lambda \in \mathbb{R} : \lambda a \in Y + x\}) = 0,$$

since the set $\{\lambda \in \mathbb{R} : \lambda a - x \in Y\}$ contains at most one element.

Some relevant results related to prevalence are as follows:

Proposition 7 ([8])**.** *Given $p \in (0, \infty)$, define ℓ^p as the space of sequences $(x_j)_{j\in\mathbb{N}}$ such that $\sum_{j\in\mathbb{N}} |x_j|^p < \infty$ (in particular, ℓ^1 is the space of absolutely summable sequences). For every $p \in (1, \infty)$, the set of sequences in ℓ^p that do not belong to ℓ^1 is prevalent in ℓ^p.*

Proposition 8 ([9])**.** *The set of continuous functions on* $[0,1]$ *that are nowhere differentiable is prevalent in the space of real-valued continuous functions on* $[0,1]$*.*

Proposition 9 ([10])**.** *Let* $C^\infty([0,1])$ *denote the space of functions that are infinitely differentiable on* $[0,1]$*. The set of functions on* $[0,1]$ *that are nowhere Gevrey-differentiable forms a prevalent subset of* $C^\infty([0,1])$*. In particular, the set of nowhere analytic functions is also prevalent in* $C^\infty([0,1])$*.*

4. Baire Categories.

The concept of first-category sets was introduced by René Baire in his work on the non-uniform convergence of series [11].

Definition 10. A subset of a topological space is said to be meager if it is contained in a countable union of closed sets with empty interior. The complement of a meager set is called residual. A set that is not meager is referred to as a second-category set.

Definition 11. A topological space is called a Baire space if every countable intersection of dense open sets is dense.

A space is a Baire space if and only if the only meager open set is the empty set. Locally compact Hausdorff spaces and completely metrizable spaces are Baire spaces.

The elementary theory of meager sets immediately leads to the following result:

Proposition 12. *In a Baire vector space, meager sets satisfy properties (*$\mathcal{P}1$*) through (*$\mathcal{P}4$*).*

Thus, meager sets provide another notion of “small” sets.

One may naturally ask about the relationship between shy sets and meager sets. Unfortunately, these notions are distinct. Let us provide two examples; further examples can be found in [1, 8, 12], among others.

The set of Liouville numbers

$$\{x \in \mathbb{R} \setminus \mathbb{Q} : \forall n \in \mathbb{N}\ \exists p, q \in \mathbb{Z} \text{ such that } q > 1 \text{ and } |x - \frac{p}{q}| < \frac{1}{q^n}\}$$

is residual and shy.

In the Banach space c_0 of sequences that converge to 0, the positive cone

$$\{(x_j)_{j\in\mathbb{N}} \in c_0 : x_j \geq 0\ \forall j\}$$

is meager but not shy.

References

[1] Hunt BR, Sauer T, Yorke JA. Prevalence: A translation-invariant “almost every” on infinite-dimensional spaces. Bull Amer Math Soc. 1992;27:217-38.

[2] Hunt BR, Kaloshin VY. Prevalence. In: Broer HW, Hasselblatt B, Takens F, editors. Handbook of Dynamical Systems. Elsevier; 2010. p. 43-87.

[3] Christensen JPR. On sets of Haar measure zero in Abelian Polish groups. Israel J Math. 1972;13:255-60.

[4] Christensen JPR. Measure theoretic zero sets in infinite dimensional spaces and applications to differentiability of Lipschitz mappings. Publ Dép Math Lyon. 1973;10:29-39.

[5] Hunt BR, Sauer T, Yorke JA. Prevalence: An addendum. Bull Amer Math Soc. 1993;28:306-7.

[6] Kuczma M. An Introduction to the Theory of Functional Equations and Inequalities. Gilányi A, editor. Springer; 2009.

[7] Hamel G. Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung: $f(x+y)=f(x)+f(y)$. Math Ann. 1905;60:459-62.
[8] Ott W, Yorke JA. Prevalence. Bull Amer Math Soc. 2005;42:263-90.
[9] Hunt BR. The Prevalence of Continuous Nowhere Differentiable Functions. Proc Amer Math Soc. 1994;122:711-7.
[10] Bastin F, Esser C, Nicolay S. Prevalence of "nowhere analyticity". Stud Math. 2012;210:239-46.
[11] Baire R. Sur les fonctions de variables réelles. Ann Mat Pura Appl. 1899;3:1-123.
[12] Oxtoby JC. Measure and Category. Graduate Texts in Mathematics. Springer; 1980.