The descriptive complexity of the set of Poisson generic numbers

V. Becher, S. Jackson, D. Kwietniak, and B. Mance

Department of Mathematics
University of Adam Mickiewicz

https://arxiv.org/abs/2305.10529

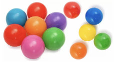
May 25, 2023

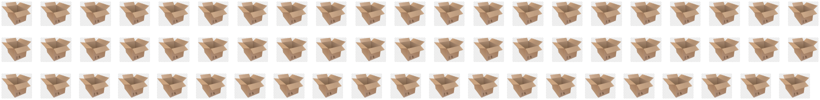
Poisson generic real numbers

Poisson generic real numbers were introduced by Zeev Rudnick in 2002. These numbers satisfy a condition stronger than normality. V. Becher's talk from a year ago goes into these in greater detail, but we will define them and describe some of the basic ideas.

The Poisson distribution

Consider the random allocation of N balls in K bins.





If N is smaller than K, a lot of bins will be empty and we expect some with exactly one ball, fewer with exactly two, still fewer with exactly three. . . .

The Poisson distribution

Consider N balls and K bins.
The probability p that a bin is allocated is $1/K$.
The expected proportion of bins with exactly i balls, for $i = 0, 1, 2, \ldots$

$$\chi(i) = \binom{N}{i} p^i (1-p)^{N-i}.$$

When N and K go to infinity but $N/K = \lambda$ is a fixed constant

$$\chi(i) \text{ converges to } e^{-\lambda}\frac{\lambda^i}{i!},$$

the Poisson probability mass function with paramenter λ.

Ω	alphabet with b symbols, $b \geq 2$.
$\Omega^{\mathbb{N}}$	infinite sequences of symbols in Ω
Ω^k	words of length k, for each $k \geq 1$

Let μ be the probability on Ω, with equal probability to each symbol.

Random allocation of N balls in $K = b^k$ bins

In the probability space (Ω^k, μ^k) the initial segment of length N of an infinite sequence can be seen as N almost independent events of placing words of length k, each one with equal probability b^{-k}.

Poisson generic sequences

For $x \in \Omega^{\mathbb{N}}$ and $\omega \in \Omega^k$, let the indicator function

$$I_j(x,\omega) = \mathbb{1}_{\{x[j,j+k)=\omega\}}.$$

Let $Z^{\lambda}_{i,k}(x)$ be the proportion of words of length k that occur exactly i times in the first $\lfloor \lambda b^k \rfloor$ symbols of $x \in \Omega^{\mathbb{N}}$,

$$Z^{\lambda}_{i,k}(x) = \mu^k \left(\omega \in \Omega^k : \sum_{1 \le j \le \lambda b^k} I_j(x,\omega) = i \right).$$

Example for $\Omega = \{0,1\}$, $b = 2$, $\lambda = 1$, $k = 3$, $b^k = \lambda b^k = 8$,

$$x = 10000100\ldots$$

For $i = 0$, $Z^{\lambda}_{i,k}(x) = 4/8$ (witnesses $011, 101, 110, 111$)
For $i = 1$, $Z^{\lambda}_{i,k}(x) = 2/8$ (witnesses $001, 010$)
For $i = 2$, $Z^{\lambda}_{i,k}(x) = 2/8$ (witnesses $100, 000$)
For $i \geq 3$, $Z^{\lambda}_{i,k}(x) = 0$

Poisson generic sequences

Definition (Zeev Rudnick)

Let λ be a positive real number. A sequence x in $\Omega^{\mathbb{N}}$ is λ-Poisson generic if for every non-negative integer i,

$$\lim_{k\to\infty} Z^{\lambda}_{i,k}(x) = e^{-\lambda}\frac{\lambda^i}{i!}.$$

A sequence is Poisson generic if it is λ-Poisson generic, for all positive λ. Let $\mathcal{P}_b$ be the set of Poisson generic real numbers.

The Borel Hierarchy

The setting:

The Borel Hierarchy

The setting:

Let X be a Polish space: a separable completely metrizable topological space.

The Borel Hierarchy

The setting:

Let X be a Polish space: a separable completely metrizable topological space. That is, X is homeomorphic to a complete metric space that has a countable dense subset.

The setting:

Let X be a Polish space: a separable completely metrizable topological space. That is, X is homeomorphic to a complete metric space that has a countable dense subset.

Examples:

$$\mathbb{R}, 2^{\mathbb{N}}, b^{\mathbb{N}}, 2^{\mathbb{N}\times\mathbb{N}}$$

The Borel Hierarchy

In any topological space X, the collection of Borel sets $\mathcal{B}(X)$ is the smallest σ-algebra containing the open sets. They are stratified into levels, the Borel hierarchy, by defining $\boldsymbol{\Sigma}^0_1 =$ the open sets, $\boldsymbol{\Pi}^0_1 = \neg\boldsymbol{\Sigma}^0_1 = \{X - A \colon A \in \boldsymbol{\Sigma}^0_1\} =$ the closed sets, and for $\alpha < \omega_1$ we let $\boldsymbol{\Sigma}^0_\alpha$ be the collection of countable unions $A = \bigcup_n A_n$ where each $A_n \in \boldsymbol{\Pi}^0_{\alpha_n}$ for some $\alpha_n < \alpha$. We also let $\boldsymbol{\Pi}^0_\alpha = \neg\boldsymbol{\Sigma}^0_\alpha$. Alternatively, $A \in \boldsymbol{\Pi}^0_\alpha$ if $A = \bigcap_n A_n$ where $A_n \in \boldsymbol{\Sigma}^0_{\alpha_n}$ where each $\alpha_n < \alpha$. We also set $\boldsymbol{\Delta}^0_\alpha = \boldsymbol{\Pi}^0_\alpha \cap \boldsymbol{\Sigma}^0_\alpha$, in particular $\boldsymbol{\Delta}^0_1$ is the collection of clopen sets. For any topological space, $\mathcal{B}(X) = \bigcup_{\alpha<\omega_1} \boldsymbol{\Sigma}^0_\alpha = \bigcup_{\alpha<\omega_1} \boldsymbol{\Pi}^0_\alpha$. All of the collections $\boldsymbol{\Delta}^0_\alpha$, $\boldsymbol{\Sigma}^0_\alpha$, $\boldsymbol{\Pi}^0_\alpha$ are pointclasses, that is, they are closed under inverse images of continuous functions.

The Borel Hierarchy

In any topological space X, the collection of Borel sets $\mathcal{B}(X)$ is the smallest σ-algebra containing the open sets. They are stratified into levels, the Borel hierarchy, by defining $\mathbf{\Sigma}^0_1 =$ the open sets, $\mathbf{\Pi}^0_1 = \neg\mathbf{\Sigma}^0_1 = \{X - A\colon A \in \mathbf{\Sigma}^0_1\} =$ the closed sets, and for $\alpha < \omega_1$ we let $\mathbf{\Sigma}^0_\alpha$ be the collection of countable unions $A = \bigcup_n A_n$ where each $A_n \in \mathbf{\Pi}^0_{\alpha_n}$ for some $\alpha_n < \alpha$. We also let $\mathbf{\Pi}^0_\alpha = \neg\mathbf{\Sigma}^0_\alpha$. Alternatively, $A \in \mathbf{\Pi}^0_\alpha$ if $A = \bigcap_n A_n$ where $A_n \in \mathbf{\Sigma}^0_{\alpha_n}$ where each $\alpha_n < \alpha$. We also set $\mathbf{\Delta}^0_\alpha = \mathbf{\Pi}^0_\alpha \cap \mathbf{\Sigma}^0_\alpha$, in particular $\mathbf{\Delta}^0_1$ is the collection of clopen sets. For any topological space, $\mathcal{B}(X) = \bigcup_{\alpha<\omega_1} \mathbf{\Sigma}^0_\alpha = \bigcup_{\alpha<\omega_1} \mathbf{\Pi}^0_\alpha$. All of the collections $\mathbf{\Delta}^0_\alpha$, $\mathbf{\Sigma}^0_\alpha$, $\mathbf{\Pi}^0_\alpha$ are pointclasses, that is, they are closed under inverse images of continuous functions.

For example, $\mathbf{\Sigma}^0_2$ consists of F_σ sets and $\mathbf{\Pi}^0_2$ consists of G_δ sets. $\mathbf{\Pi}^0_3$ contains the sets which are intersections of F_σ sets.

The Borel Hierarchy

A fundamental result of Suslin says that in any Polish space $\mathcal{B}(X) = \mathbf{\Delta}^1_1 = \mathbf{\Sigma}^1_1 \cap \mathbf{\Pi}^1_1$, where $\mathbf{\Pi}^1_1 = \neg\mathbf{\Sigma}^1_1$, and $\mathbf{\Sigma}^1_1$ is the pointclass of continuous images of Borel sets. Equivalently, $A \in \mathbf{\Sigma}^1_1$ iff A can be written as $x \in a \leftrightarrow \exists y\ (x, y) \in B$ where $B \subseteq X \times Y$ is Borel (for some Polish space Y). Similarly, $A \in \mathbf{\Pi}^1_1$ iff it is of the form $x \in A \leftrightarrow \forall y\ (x, y) \in B$ for a Borel B. The $\mathbf{\Sigma}^1_1$ sets are also called the *analytic* sets, and $\mathbf{\Pi}^1_1$ the *co-analytic sets*. We also have $\mathbf{\Sigma}^1_1 \neq \mathbf{\Pi}^1_1$ for any uncountable Polish space.

The Borel Hierarchy

$$
\begin{tikzcd}[column sep=small]
 & \mathbf{\Sigma}^0_1=\text{open} \arrow[dr, dash] & & \mathbf{\Sigma}^0_2 = F_\sigma \arrow[dr, dash] & & \mathbf{\Sigma}^0_3 \arrow[dr, dash] & & \mathbf{\Sigma}^1_1 = \text{analytic} \arrow[dr, dash] & \\
\mathbf{\Delta}^0_1=\text{clopen} \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^0_2 \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^0_3 \arrow[ur, dash] \arrow[dr, dash] & & \cdots \;\cdot\; \mathbf{\Delta}^1_1 = \mathcal{B}(X) \arrow[ur, dash] \arrow[dr, dash] & & \cdots \\
 & \mathbf{\Pi}^0_1=\text{closed} \arrow[ur, dash] & & \mathbf{\Pi}^0_2 = G_\delta \arrow[ur, dash] & & \mathbf{\Pi}^0_3 \arrow[ur, dash] & & \mathbf{\Pi}^1_1 = \text{coanalytic} \arrow[ur, dash] &
\end{tikzcd}
$$

The Borel Hierarchy

A basic fact is that for any uncountable Polish space X, there is no collapse in the levels of the Borel hierarchy, that is, all the various pointclasses $\mathbf{\Delta}^0_\alpha$, $\mathbf{\Sigma}^0_\alpha$, $\mathbf{\Pi}^0_\alpha$, for $\alpha < \omega_1$, are all distinct. Thus, these levels of the Borel hierarch can be used to calibrate the descriptive complexity of a set. We say a set $A \subseteq X$ is $\mathbf{\Sigma}^0_\alpha$ (resp. $\mathbf{\Pi}^0_\alpha$) *hard* if $A \notin \mathbf{\Pi}^0_\alpha$ (resp. $A \notin \mathbf{\Sigma}^0_\alpha$). This says A is "no simpler" than a $\mathbf{\Sigma}^0_\alpha$ set. We say A is $\mathbf{\Sigma}^0_\alpha$-*complete* if $A \in \mathbf{\Sigma}^0_\alpha - \mathbf{\Pi}^0_\alpha$, that is, $A \in \mathbf{\Sigma}^0_\alpha$ and A is $\mathbf{\Sigma}^0_\alpha$ hard. This says A is exactly at the complexity level $\mathbf{\Sigma}^0_\alpha$. Likewise, A is $\mathbf{\Pi}^0_\alpha$-complete if $A \in \mathbf{\Pi}^0_\alpha - \mathbf{\Sigma}^0_\alpha$.

Examples

In $\mathbb{R}$, $(a, b) \in \mathbf{\Sigma}^0_1$, $[a, b] \in \mathbf{\Pi}^0_1$, and $\mathbb{R} \in \mathbf{\Delta}^0_1$.

Examples

In $\mathbb{R}$, $(a, b) \in \mathbf{\Sigma}^0_1, [a, b] \in \mathbf{\Pi}^0_1$, and $\mathbb{R} \in \mathbf{\Delta}^0_1$.

$\mathbb{Q} \in \mathbf{\Sigma}^0_2, \mathbb{R}\backslash\mathbb{Q} \in \mathbf{\Pi}^0_2$, and $\emptyset = \mathbb{Q} \cap \mathbb{R}\backslash\mathbb{Q} \in \mathbf{\Delta}^0_2$.

Examples

In $\mathbb{R}$, $(a, b) \in \mathbf{\Sigma}^0_1, [a, b] \in \mathbf{\Pi}^0_1$, and $\mathbb{R} \in \mathbf{\Delta}^0_1$.

$\mathbb{Q} \in \mathbf{\Sigma}^0_2, \mathbb{R}\backslash\mathbb{Q} \in \mathbf{\Pi}^0_2$, and $\emptyset = \mathbb{Q} \cap \mathbb{R}\backslash\mathbb{Q} \in \mathbf{\Delta}^0_2$.

Note that $\mathbb{Q}$ is $\mathbf{\Sigma}^0_2$-complete, $\mathbb{R}\backslash\mathbb{Q}$ is $\mathbf{\Pi}^0_2$-complete, but $\emptyset \in \mathbf{\Delta}^0_1$.

Examples

In $\mathbb{R}$, $(a, b) \in \mathbf{\Sigma}^0_1, [a, b] \in \mathbf{\Pi}^0_1$, and $\mathbb{R} \in \mathbf{\Delta}^0_1$.

$\mathbb{Q} \in \mathbf{\Sigma}^0_2, \mathbb{R}\backslash\mathbb{Q} \in \mathbf{\Pi}^0_2$, and $\emptyset = \mathbb{Q} \cap \mathbb{R}\backslash\mathbb{Q} \in \mathbf{\Delta}^0_2$.

Note that $\mathbb{Q}$ is $\mathbf{\Sigma}^0_2$-complete, $\mathbb{R}\backslash\mathbb{Q}$ is $\mathbf{\Pi}^0_2$-complete, but $\emptyset \in \mathbf{\Delta}^0_1$.

Let $X = C([0, 1])$ with the sup norm. If
$S = \{f \in X : f$ is nowhere differentiable$\}$, then $S \in \mathbf{\Pi}^1_1\backslash\mathbf{\Sigma}^1_1$ (R. D. Mauldin 1979).

Examples with normal numbers

Let $\mathcal{N}_b$ be the set of numbers normal in base b. The following resolve questions of Kechris.

Examples with normal numbers

Let $\mathcal{N}_b$ be the set of numbers normal in base b. The following resolve questions of Kechris.

The set $\mathcal{N}_b$ is $\mathbf{\Pi}^0_3$-complete (Ki and Linton 1994).

Examples with normal numbers

Let $\mathcal{N}_b$ be the set of numbers normal in base b. The following resolve questions of Kechris.

The set $\mathcal{N}_b$ is $\boldsymbol{\Pi}^0_3$-complete (Ki and Linton 1994).

The set $\bigcap_b \mathcal{N}_b$ is also $\boldsymbol{\Pi}^0_3$-complete (Becher, Heiber, and Slaman 2014).

Examples with normal numbers

Let $\mathcal{N}_b$ be the set of numbers normal in base b. The following resolve questions of Kechris.

The set $\mathcal{N}_b$ is $\mathbf{\Pi}^0_3$-complete (Ki and Linton 1994).

The set $\bigcap_b \mathcal{N}_b$ is also $\mathbf{\Pi}^0_3$-complete (Becher, Heiber, and Slaman 2014).

Moreover, $\bigcup_b \mathcal{N}_b$ is $\mathbf{\Sigma}^0_4$-complete (Becher, Slaman 2014).

Examples with normal numbers

Let $\mathcal{N}_b$ be the set of numbers normal in base b. The following resolve questions of Kechris.

The set $\mathcal{N}_b$ is $\mathbf{\Pi}^0_3$-complete (Ki and Linton 1994).

The set $\bigcap_b \mathcal{N}_b$ is also $\mathbf{\Pi}^0_3$-complete (Becher, Heiber, and Slaman 2014).

Moreover, $\bigcup_b \mathcal{N}_b$ is $\mathbf{\Sigma}^0_4$-complete (Becher, Slaman 2014).

Let

$$\mathcal{N}_b^{\perp} = \{y \colon \forall x \in \mathcal{N}_b \ (x + y) \in \mathcal{N}_b\}.$$

be the set of numbers that preserve normality in base b under addition. The set $\mathcal{N}_b^{\perp}$ is $\mathbf{\Pi}^0_3$-complete (Airey, Jackson, M. 2022). See my talk from Numeration 2018.

Examples with normal numbers

As a special case of a much more general theorem, Airey, Kwietniak, Jackson, M. proved that sets of normal numbers for b-ary expansions, continued fraction expansions, Lüroth series expansions, and others are all $\mathbf{\Pi}^0_3$-complete. See Kwietniak's talk from Numeration 2019.

Difference hierarchy

A set D is in $D_2(\mathbf{\Pi}^0_3)$ if $D = A \backslash B$ for $A, B \in \mathbf{\Pi}^0_3$. A set D is $D_2(\mathbf{\Pi}^0_3)$-hard if its compliment isn't in $D_2(\mathbf{\Pi}^0_3)$. Moverover, D is $D_2(\mathbf{\Pi}^0_3)$-complete if it is $D_2(\mathbf{\Pi}^0_3)$ and $D_2(\mathbf{\Pi}^0_3)$-hard.

Difference hierarchy

A set D is in $D_2(\mathbf{\Pi}^0_3)$ if $D = A \backslash B$ for $A, B \in \mathbf{\Pi}^0_3$. A set D is $D_2(\mathbf{\Pi}^0_3)$-hard if its compliment isn't in $D_2(\mathbf{\Pi}^0_3)$. Moverover, D is $D_2(\mathbf{\Pi}^0_3)$-complete if it is $D_2(\mathbf{\Pi}^0_3)$ and $D_2(\mathbf{\Pi}^0_3)$-hard.

Airey, Jackson, and M. showed that sets of normal, distribution normal, and ratio normal numbers for the Cantor series expansions are $\mathbf{\Pi}^0_3$ complete and all non-empty differences are $D_2(\mathbf{\Pi}^0_3)$-complete (2022). See my talk in Numeration 2019.

Difference hierarchy

A set D is in $D_2(\mathbf{\Pi}^0_3)$ if $D = A \backslash B$ for $A, B \in \mathbf{\Pi}^0_3$. A set D is $D_2(\mathbf{\Pi}^0_3)$-hard if its compliment isn't in $D_2(\mathbf{\Pi}^0_3)$. Moverover, D is $D_2(\mathbf{\Pi}^0_3)$-complete if it is $D_2(\mathbf{\Pi}^0_3)$ and $D_2(\mathbf{\Pi}^0_3)$-hard.

Airey, Jackson, and M. showed that sets of normal, distribution normal, and ratio normal numbers for the Cantor series expansions are $\mathbf{\Pi}^0_3$ complete and all non-empty differences are $D_2(\mathbf{\Pi}^0_3)$-complete (2022). See my talk in Numeration 2019.

Vandehey proved that the set of numbers that are continued fraction normal, but not normal in any base is uncountable by assuming the Generalized Riemann Hypothesis (2016).

Difference hierarchy

Jackson, M., and Vandehey (2021) improved this result by showing that this set is $D_2(\mathbf{\Pi}^0_3)$-hard NOT assuming the Generalized Riemann Hypothesis. Moreover, the set of numbers that are continued fraction normal, but not normal in a fixed base b is $D_2(\mathbf{\Pi}^0_3)$-complete. And the set of numbers normal in a base p, but not base q is $D_2(\mathbf{\Pi}^0_3)$-complete when p and q are relatively prime. See Jackson's talk in 2021.

Main result

We can finally state what we proved! Posted last week on arXiv, Becher, Kwietniak, Jackson, M. proved that $\mathcal{P}_b$ is $\mathbf{\Pi}^0_3$-complete and $\mathcal{N}_b \setminus \mathcal{P}_b$ is $D_2(\mathbf{\Pi}^0_3)$-complete.

Fun corollaries

An immediate consequence is that the set $\mathcal{N}_b \setminus \mathcal{P}_b$ is uncountable.

Fun corollaries

An immediate consequence is that the set $\mathcal{N}_b \setminus \mathcal{P}_b$ is uncountable.

Also, since $\mathcal{N}_b \setminus \mathcal{P}_b$ is D_2-$\mathbf{\Pi}^0_3$-complete, there cannot be a $\mathbf{\Sigma}^0_3$ set A such that $A \cap \mathcal{N}_b = \mathcal{P}_b$ (as otherwise, we would have $\mathcal{N}_b \setminus \mathcal{P}_b = \mathcal{N}_b \setminus A \in \mathbf{\Pi}^3_0$, a contradiction). Thus, no $\mathbf{\Sigma}^0_3$ condition can be added to normality to give Poisson genericity.

Fun corollaries

As an application, consider the following definition of weakly Poisson generic:

Say $x \in [0,1)$ with base b expansion (x_j) is weakly Poisson generic in base b if for every $\epsilon > 0$, every rational λ, and non-negative integer j, we have that for infinitely many k that $|Z^{\lambda}_{j,k}(x) - e^{-\lambda}\frac{\lambda^j}{j!}| < \epsilon$.

Note that being Poisson generic in base b implies being weakly-Poisson generic. However, being weakly-Poisson generic is a $\mathbf{\Pi}^0_2$ condition.

For every base b there is a base-b normal number which is weakly Poisson generic but not Poisson generic.

Fun corollaries

As another application, consider the following version of discrepency. Suppose f is a function assigning to each word $w \in b^{<\omega}$ and each positive integer n a positive real number $f(w,n)$. Given $x \in [0,1)$ with base b expansion (b_j), say the (w,n)-discrepancy is

$$D(x,w,n) = \left|\frac{n}{b^{|w|}} - W(x \restriction n, w)\right|,$$

where $W(u,w)$ is the number of occurrences of w in u.

We say a real number x has base b f-large discrepancy if for all w and all n we have that $D(x,w,n) > f(w,n)$. The set of x with f-large discrepancy, for any fixed f, is easily a $\mathbf{\Pi}^0_1$ set.

Fun corollaries

The set of numbers that are Borel normal to base b are exactly those for which the discrepancy of their initial segments of their expansion in base b goes to zero. We conjecture that the Poisson generic numbers in base b can not have very low discrepancy of their initial segments (for instance, the infinite de Bruijn sequences exist in bases $b \geq 3$, they satisfy that $Z^1_{1,k} = 1$ for every k, hence they do not correspond to Poisson generic numbers, and they have low discrepancy.) However, we have the following, which states that the Poisson generic reals cannot be characterized as the set of normal numbers satisfying a large discrepancy condition.

For every function f, the set of base-b Poisson generic reals is not equal to the set of normal numbers with f-large discrepancy.

Fun corollaries

There are also many other naturally occurring sets of real numbers are defined by conditions which make them $\mathbf{\Sigma}^0_3$. Examples include countable sets, co-countable sets, the class BA of *badly approximable* numbers (which is a $\mathbf{\Sigma}^0_2$ set), the Liouville numbers (which is a $\mathbf{\Pi}^0_2$ set), and the set of $x \in [0,1]$ where a particular continuous function $f\colon [0,1] \to \mathbb{R}$ is not differentiable. In all these cases, the theorem implies that either the set omits some Poisson generic number, or else contains a number which is normal but not Poisson generic. Of course, many of these statements are easy to see directly, but the point is that they all follow immediately from the general complexity result.

Wadge reduction

Let X and Y be Polish spaces and let $A \subseteq X$ and $B \subseteq Y$ along with a continuous function $f : Y \to X$ where $f^{-1}(A) = B$. Then if B is $\mathbf{\Sigma}^0_\alpha$-complete (resp. $\mathbf{\Pi}^0_\alpha$-complete), then A is $\mathbf{\Sigma}^0_\alpha$-hard ($\mathbf{\Pi}^0_\alpha$-hard).

Wadge reduction

Let X and Y be Polish spaces and let $A \subseteq X$ and $B \subseteq Y$ along with a continuous function $f : Y \to X$ where $f^{-1}(A) = B$. Then if B is $\mathbf{\Sigma}^0_\alpha$-complete (resp. $\mathbf{\Pi}^0_\alpha$-complete), then A is $\mathbf{\Sigma}^0_\alpha$-hard ($\mathbf{\Pi}^0_\alpha$-hard).

The function f reduces the question of membership in A to membership in B.