On the Existence of Anomalies, The Reals Case

Samuel Epstein*

February 23, 2023

Abstract

The Independence Postulate (IP) is a finitary Church-Turing Thesis, saying mathematical sequences are independent from physical ones. Modelling sbservations as infinite sequences of real numbers, IP implies the existence of anomalies.

1 Introduction

An outlier is an observation that is set apart from a population. There are many reasons that such anomalies occur, including measurement error and human error. However recent results have shown that outliers are ingrained into the nature of algorithms and dynamics. In [Eps21], anomalies were proven to occur in sampling algorithms. In [Eps22a], anomalies were proven to exist in the outputs of probabilistic algorithms. They were also proven to be emergent computable ergodic dynamics on the Cantor space. In [Eps22b] anomalies were shown to emergent in a more general (but still computable) class of dynamics. These results were extended into computable metric spaces in the paper [Eps23d], where a new characterization of uniform tests was proven. Furthermore, oscillations in algorithmic thermodynamic entropy were proven.

But what about measurements of systems that are too complex to be considered algorithmic? One example is the global weather system. One can attest to the fact that there are many strange formations that occur! To show that anomalies occur, one can use the Independence Postulate [Lev84, Lev13]. The Independence Postulate is a finitary Church-Turing thesis, postulating that certain finite and infinite sequences cannot be easily be found with a short "physical address". In [Eps23b], the Independence Postulate was used to show that observations, a.k.a. infinite sequences of natural numbers, that do not have outliers have high physical addresses. In other words, observations with no outliers cannot be found in nature.

In this paper, we extend these results to observations modeled by infinite sequences of reals. This enables a more natural modelling of phenomena such as fluid dynamics, etc. This paper reproduces the proof of infinite sequences in [Eps21], but without using left-total machines, which require a lengthy explaination.

2 Conventions

The function $\mathbf{K}(x|y)$ is the conditional prefix Kolmogorov complexity. The mutual information between two strings $x, y \in \{0,1\}^*$, is $\mathbf{I}(x:y) = \mathbf{K}(x) + \mathbf{K}(y) - \mathbf{K}(x,y)$. For probability p over $\mathbb{N}$, randomness deficiency is $\mathbf{d}(a|p,b) = \lfloor -\log p(a) \rfloor - \mathbf{K}(a|\langle p\rangle, b)$ and measures the extent of the refutation against the hypothesis p given the result a [G21]. $\mathbf{d}(a|p) = \mathbf{d}(a|p,\emptyset)$. The amount of information that the halting sequence $\mathcal{H} \in \{0,1\}^\infty$ has about $a \in \{0,1\}^*$, conditional to $y \in \{0,1\}^*$ is $\mathbf{I}(a;\mathcal{H}|y) = \mathbf{K}(a|y) - \mathbf{K}(a|y,\mathcal{H})$. $\mathbf{I}(a;\mathcal{H}) = \mathbf{I}(a;\mathcal{H}|\emptyset)$. We use $<^+ f$ to denote $< f + O(1)$ and $<^{\log} f$

*JP Theory Group. samepst@jptheorygroup.org

to denote $<f+O(\log(f+1))$. For a mathematical statement A, let $[A]=1$ if A is true and $[A]=0$, otherwise. The chain rule gives $\mathbf{K}(x,y) =^+ \mathbf{K}(x|y,\mathbf{K}(y))+\mathbf{K}(y)$. The following definition is from [Lev74] .

Definition 1 (Information) *For infinite sequences $\alpha,\beta\in\{0,1\}^\infty$, their mutual information is defined to be* $\mathbf{I}(\alpha:\beta)= \log\sum_{x,y\in\{0,1\}^*} 2^{\mathbf{I}(x:y)-\mathbf{K}(x|\alpha)-\mathbf{K}(y|\beta)}$.

The Independence Postulate (**IP**), [Lev84, Lev13], is an unprovable inequality on the information content shared between two sequences. **IP** is a finitary Church Turing Thesis, postulating that certain infinite and finite sequences cannot be found in nature, a.k.a. have high "physical addresses".

IP*: Let α be a sequence defined with an n-bit mathematical statement, and a sequence β can be located in the physical world with a k-bit instruction set. Then $\mathbf{I}(\alpha:\beta)<k+n+c$ for some small absolute constant c.*

Lemma 1 ([Eps23c]) *For probability p over $\mathbb{N}$, $D\subset\mathbb{N}$, $|D|=2^s$, $s<\max_{a\in D}\mathbf{d}(a|p)+\mathbf{I}(D;\mathcal{H})+O(\log\mathbf{I}(D;\mathcal{H})+\log\mathbf{K}(p))$.*

Lemma 2 ([Eps22a]) *For partial computable f, $\mathbf{I}(f(x):\mathcal{H})<^+\mathbf{I}(x;\mathcal{H})+\mathbf{K}(f)$.*

3 Sets with Low Randomness Deficiencies

A continuous probability P over $\{0,1\}^\infty$ is identified with a function $P:\{0,1\}^*\to\mathbb{R}_{\geq 0}$, where $P(\emptyset)=1$ and $P(x)=P(x0)+P(x1)$. Randomness deficiency can be extended to continous probability measures with the following definition.

Definition 2 *The randomness deficiency of $\alpha\in\{0,1\}^\infty$ with respect to computable continuous probability measure P is $\mathbf{D}(\alpha|P)=\sup_n -\log P(\alpha[0..n])-\mathbf{K}(\alpha[0..n]|\langle P\rangle)$. The term $\langle P\rangle$ is a program to compute P.*

Let $\Omega=\sum\{2^{-\|p\|}:U(p)\text{ halts}\}$ be Chaitin's Omega, $\Omega_n\in\mathbb{Q}_{\geq 0}$ be be the rational formed from the first n bits of Ω, and $\Omega^t=\sum\{2^{-\|p\|}:U(p)\text{ halts in time }t\}$. For $n\in\mathbb{N}$, let $\mathbf{bb}(n)=\min\{t:\Omega_n<\Omega^t\}$. $\mathbf{bb}^{-1}(m)=\arg\min_n\{\mathbf{bb}(n-1)<m\leq\mathbf{bb}(n)\}$. Let $\Omega[n]\in\{0,1\}^*$ be the first n bits of Ω.

Lemma 3 *For $n=\mathbf{bb}^{-1}(m)$, $\mathbf{K}(\Omega[n]|m,n)=O(1)$.*

Proof. For a string x, let $BB(x)=\inf\{t:\Omega^t>0.x\}$. Enumerate strings of length n, starting with 0^n, and return the first string x such that $BB(x)\geq m$. This string x is equal to $\Omega[n]$, otherwise let y be the largest common prefix of x and $\Omega[n]$. Thus $BB(y)=\mathbf{bb}(\|y\|)\geq BB(x)\geq m$, which means $\mathbf{bb}^{-1}(m)\leq\|y\|<n$, causing a contradiction. $\square$

Lemma 4 *For continuous probability P over $\{0,1\}^\infty$, $Z \subset \{0,1\}^\infty$, $|Z| = 2^s$, $s <^{\log} \max_{\alpha\in Z} \mathbf{D}(\alpha|P) + \mathbf{I}(\langle Z\rangle : \mathcal{H}) + O(\log \mathbf{K}(P))$.*

Proof. We relativize the universal Turing machine to s, which can be done due to the precision of the theorem. Let $Z_n = \{\alpha[0..n] : \alpha \in Z\}$ and $m = \arg\min_m |Z_m| = |Z|$. Let $n = \mathbf{bb}^{-1}(m)$ and $k = \mathbf{bb}(n)$. Let p be a probability over $\{0,1\}^*$, where $p(x) = [\|x\| = k]P(x)$ and $\langle p\rangle = \langle k, P\rangle$. Using $D = Z_k$, Lemma 1 relativized to k produces $x \in Z_k$, where

$$
\begin{aligned}
s &<^{\log} -\log P(x) - \mathbf{K}(x|k,P) + \mathbf{I}(Z_k;\mathcal{H}|k) + O(\log \mathbf{K}(P,k|k)) \\
&<^{\log} -\log P(x) - \mathbf{K}(x|P) + \mathbf{K}(Z_k|k) + \mathbf{K}(k) - \mathbf{K}(Z_k|k,\mathcal{H}) + O(\log \mathbf{K}(P)).
\end{aligned}
$$

Since $\mathbf{K}(k) <^+ n + \mathbf{K}(n)$, by the chain rule,

$$
\begin{aligned}
&\mathbf{K}(Z_k|k) + \mathbf{K}(k) \\
<^+ &\mathbf{K}(Z_k|k,\mathbf{K}(k)) + \mathbf{K}(\mathbf{K}(k)|k) + \mathbf{K}(k) \\
< &\mathbf{K}(Z_k,k) + O(\log n) \\
< &\mathbf{K}(Z_k) + O(\log n).
\end{aligned}
$$

So

$$
s <^{\log} -\log P(x) - \mathbf{K}(x|P) + \mathbf{K}(Z_k) - \mathbf{K}(Z_k|k,\mathcal{H}) + O(\log n + \log \mathbf{K}(P)).
$$

Since $\mathbf{K}(k|n,\mathcal{H}) = O(1)$, $\mathbf{K}(Z_k|\mathcal{H}) <^+ \mathbf{K}(Z_k|k,\mathcal{H}) + \mathbf{K}(n)$. So

$$
s <^{\log} -\log P(x) - \mathbf{K}(x|P) + \mathbf{I}(Z_k;\mathcal{H}) + O(\log n + \log \mathbf{K}(P)).
$$

By Lemma 3, $\mathbf{K}(\Omega[n]|Z_k) <^+ \mathbf{K}(n)$ so by Lemma by 2,

$$
n <^{\log} \mathbf{I}(\Omega[n];\mathcal{H}) <^{\log} \mathbf{I}(Z_k;\mathcal{H}) + \mathbf{K}(n) <^{\log} \mathbf{I}(Z_k;\mathcal{H}).
$$

The above equation used the common fact that the first n bits of Ω has $n - O(\log n)$ bits of mutual information with $\mathcal{H}$. So

$$
s <^{\log} -\log P(x) - \mathbf{K}(x|P) + \mathbf{I}(Z_k;\mathcal{H}) + O(\log \mathbf{K}(P)).
$$

By the definition of mutual information $\mathbf{I}$ between infinite sequences

$$
\mathbf{I}(Z_k;\mathcal{H}) <^+ \mathbf{I}(Z:\mathcal{H}) + \mathbf{K}(Z_k|Z) <^{\log} \mathbf{I}(Z:\mathcal{H}) + \mathbf{K}(k|Z).
$$

Now m is simple relative to Z and by Lemma 3, $\Omega[n]$ is simple relative to m and n. Furthermore k is simple relative to $\Omega[n]$. Therefore $\mathbf{K}(Z_k|Z) <^+ \mathbf{K}(n)$. So

$$
\begin{aligned}
s &<^{\log} -\log P(x) - \mathbf{K}(x|P) + \mathbf{I}(Z:\mathcal{H}) + \mathbf{K}(n) + O(\log \mathbf{K}(P)) \\
s &<^{\log} \max_{\alpha\in Z} \mathbf{D}(\alpha|P) + \mathbf{I}(Z:\mathcal{H})) + O(\log \mathbf{K}(P)).
\end{aligned}
$$

□

Through careful observation, the above lemma can even be tightened to the following corollary

Corollary 1 *For continuous probability P over $\{0,1\}^\infty$, $Z \subset \{0,1\}^\infty$, $|Z| = 2^s$, $s < \max_{\alpha\in Z} \mathbf{D}(\alpha|P) + \mathbf{I}(\langle Z\rangle : \mathcal{H}) + O(\log \mathbf{I}(\langle Z\rangle : \mathcal{H}) + \log \mathbf{K}(P))$.*

4 Observations as Reals

We model observations as infinite sequences of reals in the interval $[0,1]$, or equivalently infinite sequences γ of infinite sequences $\gamma_i \in \{0,1\}^\infty$, where each γ_i is unique. Of course, in the real world, infinite sequences of observations do not exist. But infinite sequences model processes that are potentially never ending. Let $\langle\gamma\rangle \in \{0,1\}^\infty$ be a standard encoding of γ. Let $\gamma(n) \subset \{0,1\}^\infty$be the first 2^n infinite sequences of γ. The following theorem uses the simple fact that $\mathbf{I}(f(\alpha):\mathcal{H}) <^+ \mathbf{I}(\alpha:\mathcal{H}) + \mathbf{K}(f)$, for $\alpha \in \{0,1\}^\infty$.

Theorem 1
For probability P over $\{0,1\}^\infty$, $\gamma \in \{0,1\}^{\infty\mathbb{N}}$, let $t_{\gamma,P} = \sup_n(n - \mathbf{K}(n) - \max_{\alpha\in\gamma(n)} \mathbf{D}(\alpha|P))$. Then $t_{\gamma,P} <^{\log} \mathbf{I}(\langle\gamma\rangle:\mathcal{H}) + O(\log \mathbf{K}(P))$.

Proof. By Corollary 1 applied to $\gamma(n)$,

$$n < \max_{\alpha\in\gamma(n)} \mathbf{D}(\alpha|P) + \mathbf{I}(\gamma(n):\mathcal{H}) + O(\log \mathbf{I}(\gamma(n):\mathcal{H}) + \log \mathbf{K}(P))$$
$$n - \max_{\alpha\in\gamma(n)} \mathbf{D}(\alpha|P) <^{\log} +\mathbf{I}(\gamma(n):\mathcal{H}) + O(\log \mathbf{K}(P))$$
$$n - \max_{\alpha\in\gamma(n)} \mathbf{D}(\alpha|P) - \mathbf{K}(n) <^{\log} +\mathbf{I}(\langle\gamma\rangle:\mathcal{H}) + O(\log \mathbf{K}(P))$$
$$t_{\gamma,P} <^{\log} +\mathbf{I}(\langle\gamma\rangle:\mathcal{H}) + O(\log \mathbf{K}(P)).$$

□

Let k be a physical address of γ. $\mathcal{H}$ can be described by a small mathematical statement. By Theorem 1 and **IP**, there is a small constant c where

$$t_{\tau,\gamma} <^{\log} \mathbf{I}(\langle\gamma\rangle:\mathcal{H}) + O(\log \mathbf{K}(P)) <^{\log} k + c + O(\log \mathbf{K}(P)).$$

It's hard to find observations with small anomalies and impossible to find observations with no anomalies.

5 Discussion

One avenue for future research is the relationship of outliers with different areas of physics. In thermodynamics, oscillations of entropy has been shown to occur [Eps23d]. One area of study is into presence of outliers in quantum information theory. Recently, a Quantum EL theorem has been proven [Eps23a]. Can this theorem be extended (as the EL Theorem was extended to the Outlier Theorem) to a statement saying streams of quantum qubits will contain outlying states?

References

[Eps21] Samuel Epstein. All sampling methods produce outliers. *IEEE Transactions on Information Theory*, 67(11):7568–7578, 2021.

[Eps22a] S. Epstein. The outlier theorem revisited. *CoRR*, abs/2203.08733, 2022.

[Eps22b] S. Epstein. Outliers, dynamics, and the independence postulate. *CoRR*, abs/2207.03955, 2022.

[Eps23a] S. Epstein. A Quantum EL Theorem. *CoRR*, abs/2301.08348, 2023.

[Eps23b] S. Epstein. On the existence of anomalies. *CoRR*, abs/2302.05972, 2023.

[Eps23c] S. Epstein. On the existence of anomalies. *CoRR*, abs/2302.05972, 2023.

[Eps23d] S. Epstein. Uniform Tests and Algorithmic Thermodynamic Entropy, 2023. http://www.jptheorygroup.org/doc/Oscillation.pdf.

[G2́1] Peter Gács. Lecture notes on descriptional complexity and randomness. *CoRR*, abs/2105.04704, 2021.

[Lev74] L. A. Levin. Laws of Information Conservation (Non-growth) and Aspects of the Foundations of Probability Theory. *Problemy Peredachi Informatsii*, 10(3):206–210, 1974.

[Lev84] L. A. Levin. Randomness conservation inequalities; information and independence in mathematical theories. *Information and Control*, 61(1):15–37, 1984.

[Lev13] L. A. Levin. Forbidden information. *J. ACM*, 60(2), 2013.