Simple Probabilities are Balanced

Samuel Epstein[*]

September 27, 2023

Abstract

Simple probabilities will have balanced distributions with respect to Kolmogorov complexity. We prove a lower bound on their measure over simple strings.

It has been proven that large sets of strings are exotic if they all have similar complexities. By exotic, we mean their encoding has high mutual information with the halting sequence. Similarly if one probability over infinite strings gives large measure to sequences with low deficiency of randomness with respect to a second probability, then it is exotic. In this paper, we look at probabilities over strings of length n, and prove that they must give measure to simple strings. This result also appears in the black holes section of the Algorithmic Physics manuscript at http://www.jptheorygroup.org.

1 Tools

U is the reference universal Turing machine. $\mathbf{K}$ is the prefix free Kolmogorov complexity. $\mathbf{m}$ is the algorithmic probability. The information that the halting sequence $\mathcal{H}$ has about a string $x \in \{0,1\}^*$ is $\mathbf{I}(x;\mathcal{H}) = \mathbf{K}(x) - \mathbf{K}(x|\mathcal{H})$. The deficiency of randomness of a string x with respect to a probability p, is $\mathbf{d}(a|p) = -\log p(a) - \mathbf{K}(a)$.

Theorem 1 **([Eps23a])** *For probability p over $\{0,1\}^*$, $D \subset \mathbb{N}$, $|D| = 2^s$, $s < \max_{a\in D} \mathbf{d}(a|p) + \mathbf{I}(D;\mathcal{H}) + O(\log \mathbf{I}(D;\mathcal{H})) + \mathbf{K}(s) + O(\log \mathbf{K}(s,p))$.*

Theorem 2 **([Eps23b])** *For probability p over $\mathbb{N}$, computed by program q,*
$\Pr_{a\sim p}[\mathbf{I}(a;\mathcal{H}) > \mathbf{K}(q) + m] \overset{*}{<} 2^{-m}$.

2 Results

Theorem 3 *There is a $c \in \mathbb{N}$ where relativized to computable probability p over $\{0,1\}^n$, for $m > c$, $p\{x : \mathbf{K}(x) \leq m\} > 2^{m-n-c-3\mathbf{K}(m)}$.*

Proof. Without loss of generality, p can be assumed to have a range in powers of 2. By the assumptions of the theorem, there is an encoding of a program for p (and thus also n) on the auxilliary tape, and thus $\mathbf{K}(p) = O(1)$. Assume not, then there exist $m \in (\mathbf{K}(p) + c, n)$ such

[*]JP Theory Group. samepst@jptheorygroup.org

that $p\{x : \mathbf{K}(x) \leq m\} < 2^{m-n-c-3\mathbf{K}(m)}$, where c solely depends on the universal Turing machine. Suppose $\max\{p(x) : \mathbf{K}(x) > m\} \geq 2^{m-n-c-3\mathbf{K}(m)}$. Then

$$\mathbf{K}(p) + O(1) > \mathbf{K}\left(\arg\max_x p(x)\right) > m > \mathbf{K}(p) + c,$$

causing a contradiction, for choice of c dependent on U. Sample $2^{n+3\mathbf{K}(m)+c-m-2}$ elements D without replacement. p^* is the probability of D, where $\mathbf{K}(p^*) <^+ \mathbf{K}(m)$. Even if every element x chosen has $p(x) = 2^{m-n-c-3\mathbf{K}(m)-1}$, the total p mass sampled is not greater than

$$2^{m-n-c-3\mathbf{K}(m)-1}2^{n+c+3\mathbf{K}(m)-m-2} \leq 2^{-3}.$$

The probability q that all $x \in D$ has $\mathbf{K}(x) > m$ is

$$\begin{aligned}
q &> \left(1 - 2^{m-n-3\mathbf{K}(m)-c}/(1-2^{-3})\right)^{2^{n+c+3\mathbf{K}(m)-m-2}} \\
&> \left(1 - 2^{m-n-c-3\mathbf{K}(m)+1}\right)^{2^{n+c+3\mathbf{K}(m)-m-2}} \\
&> \left(1 - 2^{n+c+3\mathbf{K}(m)-m-2}2^{m-n-c-3\mathbf{K}(m)+1}\right) \\
&= 1/2.
\end{aligned}$$

Thus, by Theorem 2, $\Pr_{S\sim p^*}[\mathbf{I}(S;\mathcal{H}) > \mathbf{K}(p^*) + m] \overset{*}{<} 2^{-m}$. So by probabilistic arguments, there exists $D \subset \{0,1\}^n$, where for all $x \in D$, $\mathbf{K}(x) > m$ and $\mathbf{I}(D;\mathcal{H}) <^+ \mathbf{K}(p^*) <^+ \mathbf{K}(m)$. So by Theorem 1, applied to D and the uniform measure U_n over strings of length n,

$$\begin{aligned}
n - m + c + 3\mathbf{K}(m) <& \max_{a\in D}\mathbf{d}(a|U_n) + \mathbf{I}(D;\mathcal{H}) + O(\log \mathbf{I}(D;\mathcal{H})) + \mathbf{K}(n{+}c{+}3\mathbf{K}(m){-}m \\
&+ O(\log \mathbf{K}(p^*, n{+}c{+}3\mathbf{K}(m){-}m)) \\
n - m + c + 3\mathbf{K}(m) <& n - m + 2\mathbf{K}(m) + \mathbf{K}(c) + O(\log \mathbf{K}(c,m)) \\
c + \mathbf{K}(m) <& \mathbf{K}(c) + O(\log \mathbf{K}(c,m)).
\end{aligned}$$

which is a contradiction for c dependent solely on the universal Turing machine U. □

References

[Eps23a] Samuel Epstein. On the Existence of Anomalies. *CoRR*, abs/2302.05972, 2023.

[Eps23b] Samuel Epstein. The Kolmogorov Birthday Paradox. *Theoretical Computer Science*, 963, 2023.