Weak Law of Large Numbers (contd.)

We show that every K-incompressible string obeys WLLN.

The major tool we use is "deficiency of randomness"

Overall Strategy:

If a long string does not satisfy the Weak Law of Large numbers (i.e. it contains either too many ones or too many zeroes)

then that string has "a short description".

↓

If a string does not satisfy WLLN, then there is a Martin-Löf test which captures the string

Definition. (deficiency of randomness)

Consider positive real functions on strings. i.e. functions of the form $d : \Sigma^* \to [0, \infty)$. The function is called a "deficiency of randomness function" or a "Martin-Löf test"

(1) $\forall n \ \forall k$ $$\sum_{\substack{x \in \Sigma^n \\ d(x) > k}} \Pr(x) \le 1/2^k$$ (inverse exponential falloff)

(2) d is "lower semicomputable", which we define on the next page.

Definition

A function $f: \{0,1\}^* \to \mathbb{R}$ is lower semicomputable if

there is a total computable function $\hat{f}: \{0,1\}^* \times \underbrace{\mathbb{N}}_{\text{patience parameter}} \to \mathbb{Q}$

such that $\forall x \in \{0,1\}^*$

(1) $\forall n$ $\quad \hat{f}(x,n) \leq \hat{f}(x, n+1) \leq f(x)$ [Monotonically increasing]

(2) $\lim_{n \to \infty} \hat{f}(x,n) = f(x)$. [Convergence in the limit]

e.g. The algorithmic probability $m(x)$ is a lower semicomputable function

Where we're headed: If a string x violates WLLN, then there is a deficiency function d s.t. $d(x)$ is large. If $d(x)$ is large, then x will be Kolmogorov compressible.

Contrapositive: If $K(x)$ is high, then x satisfies WLLN.

We will form 2 deficiency functions.

(1) The first is for WLLN

(2) The second is for K-incompressibility.

Let's do the (2)nd:

idea.

"length − $K(x)$"

$$d_0(x) = |x| - K(x \mid |x|).$$

If x is incompressible, $K(x \mid |x|) \approx n$, then $d_0(x) \approx n - n \approx 0$.

If $K(x \mid |x|) = \log(|x|)$, then $d_0(x) = n - \log n$.

We verify that d_0 is indeed a "deficiency of randomness" function (equivalently an ML-test):

(1) $$\sum_{\substack{x \in \Sigma^n \\ d_0(x) > k}} \Pr(x) = \sum_{\substack{x \in \Sigma^n \\ K(x) \leq n-k}} \Pr(x) \leq \frac{1}{2^k}.$$ (by counting short programs)

(2) K is upper semicomputable, hence d_0 is lower semicomputable.

Thus d_0 is a Martin Löf test. $\square$

Deficiency for WLLN