A Chain Rule for the Randomness Deficiency Function

Samuel Epstein*

October 23, 2023

Abstract

This paper applies the addition equality theorem for algorithmic entropy in [Ǵ21] to the Cantor space and generalizes it to arbitrary joint distributions. This application implies that randomness deficiency of infinite sequences obeys the chain rule, analogous to the Kolmogorov complexity case. This is a generalization of van Lambalgen's Theorem. In addition, a dual integration trick shortens the proof of the original theorem.

1 Introduction

Prefix free Kolmogorov complexity, $\mathbf{K}$, obeys the chain rule, with for $x, y \in \{0,1\}^*$,

$$\mathbf{K}(x,y) =^+ \mathbf{K}(x) + \mathbf{K}(y|x, \mathbf{K}(x)).$$

In this paper, we apply the addition equality theorem for algorithmic entropy in [Ǵ21] to the specific case of infinite sequences and generalize to arbitrary joint distributions. The chain rule could potentially be proven for arbitrary joint distributions over computable metric spaces [HR09] using probability kernels. We also shorten the original proof using an integration trick. The consequence to this is a chain rule for the randomness deficiency function $\mathbf{D}$ for computable probabilities μ over infinite sequences, where $\mathbf{D}(\alpha|\mu, x) = \sup_n -\log \mu(\alpha[0..n] - \mathbf{K}(\alpha[0..n]|x)$. A measure $\mu\nu$ over the space $\{0,1\}^\infty \times \{0,1\}^\infty$ is computable if $\mu\nu(x,y)$ is uniformly computable over $x, y \in \{0,1\}^n$. Its marginal is μ and the conditional is $\nu_x = \nu(\cdot|x)$. The randomness deficiency over the space $\{0,1\}^\infty \times \{0,1\}^\infty$, is $\mathbf{D}(\alpha, \beta|\mu\nu) = \sup_n -\log \mu\nu(\alpha[0..n], \beta[0..n]) - \mathbf{K}(\alpha[0..n], \beta[0..n])$. The discrete case for $\mathbf{d}(x|p) = -\log p(x) - \mathbf{K}(x)$ is trivial. The result detailed in this paper is as follows.

Theorem. *Relativized to probability $\mu\nu$ over $\{0,1\}^\infty \times \{0,1\}^\infty$,*

$$\mathbf{D}(\alpha, \beta|\mu\nu) =^+ \mathbf{D}(\alpha|\mu) + \mathbf{D}(\beta|\nu_\alpha, (\alpha, \lceil \mathbf{D}(\alpha|\mu) \rceil)).$$

This is a generalization of van Lambalgen's Theorem, which states (α, β) is ML random iff α is ML random and β is ML random with respect to α. An open question is whether $\mathbf{D}$ follows the linear inequalities that parallel Shannon entropy $\mathcal{H}$, as Kolmogorov complexity was shown to do [HRSV00]:

Conjecture. *Given $\{\alpha_1 \dots \alpha_n\} \in \{0,1\}^{\infty n}$, and random variables $\{\beta_1, \dots, \beta_n\}$ is it the case that over all $W \subseteq \{1, .., n\}$, and $\lambda_W \in \mathbb{R}$,*

$$\sum_W \lambda_W \mathbf{D}(\alpha^W) \leq 0 \Leftrightarrow \sum_W \lambda_W \mathcal{H}(\beta^W) \geq 0?$$

*JP Theory Group. samepst@jptheorygroup.org

2 Results

For the nonnegative real function f, we use $<^+ f$, $>^+ f$, and $=^+ f$ to denote $< f+O(1)$, $> f-O(1)$, and $= f \pm O(1)$. The Kolmogorov complexity of a lower computable function f is $\mathbf{K}(f)$, the size of the shortest program that enumerates it. As shown in [Ǵ21], $2^{\mathbf{D}(\alpha|\mu)} \stackrel{*}{=} \mathbf{t}_\mu(\alpha)$ where $\mathbf{t}_\mu$ is a universal lower computable μ-test. Furthermore, a modification of the proof Theorem 2.3.4 in [Ǵ21] to the $\{0,1\}^\infty \times \{0,1\}^\infty$ space can be used to show that $2^{\mathbf{D}(\alpha,\beta|\mu\nu)} \stackrel{*}{=} \mathbf{t}_{\mu\nu}(\alpha,\beta)$, where $\mathbf{t}_{\mu\nu}$ is a universal lower computable test over $\{0,1\}^\infty \times \{0,1\}^\infty$. For measure μ and lower continuous function f over $\{0,1\}^\infty$, we use the notation $\mu^x f(x) = \int_{x\in\{0,1\}^\infty} f(x)d\mu(x)$. Throughout this section, the universal Turing machine is assumed to be relativized to the joint probability $\mu\nu$. Thus there is an $O(1)$ sized program that can compute the marginal $\mu(x\{0,1\}^\infty)$ uniformly in $x \in \{0,1\}^*$, and the conditional $\nu(y\{0,1\}^\infty|x)$, uniformly in $x, y \in \{0,1\}^n$. The operator $\mu^x \nu^y_x$ is equivalent to $\mu\nu^{xy}$.

Proposition 1 $\log \nu^y_x 2^{\mathbf{D}(x,y|\mu\nu)} <^+ \mathbf{D}(x|\mu)$.

Proof. Let $f(x) = \log \nu^y_x 2^{\mathbf{D}(x,y|\mu\nu)}$. The function f is lower computable and has $\mu^x 2^{f(x)} \leq 1$. The proposition follows from the universal properties of $\mathbf{t}_\mu$, where $2^f \stackrel{*}{<} \mathbf{t}_\mu$. □

Proposition 2 *If $i < j$, then*

$$i + \mathbf{D}(x|\mu, j) <^+ j + \mathbf{D}(x|\mu, i).$$

Proof. By the properties of $\mathbf{D}$, we have

$$\mathbf{D}(x|\mu, j) <^+ \mathbf{D}(x|\mu, i) + \mathbf{K}(j-i) <^+ \mathbf{D}(x|\mu, i) + j - i.$$

□

Definition 1 *Let $F : \{0,1\}^\infty \to \mathbb{Z} \cup \{-\infty, \infty\}$ be a lower semicomputable function. An (ν, F)-test is a function $t : \{0,1\}^\infty \times \{0,1\}^\infty \to \mathbb{Z} \cup \{-\infty, \infty\}$ that is lower semicomputable and $\nu^y_x t(x,y) \leq 2^{F(x)}$. There exists a maximal (ν, F) test, $\mathbf{t}_{(\nu,F)}$, such that $t \stackrel{*}{<} \mathbf{t}_{(\nu,F)}$.*

Proposition 3 *Let $F : \{0,1\}^\infty \to \mathbb{Z} \cup \{-\infty, \infty\}$ be a lower semicomputable function of Kolmogorov complexity $O(1)$. For all y and with $F(x) > -\infty$,*

$$\mathbf{t}_{(\nu,F)}(x,y) \stackrel{*}{=} 2^{F(x)} \mathbf{t}_{\nu_x}(y|x, -F(x)).$$

Proof. To prove the inequality $\stackrel{*}{>}$, let $g(x,y,m) = \max_{i \geq m} 2^{-i} \mathbf{t}_{\nu_x}(y|x,i)$. This function is lower computable, and decreasing in m. The function $g(x,y) = g(x,y,-F(x))$ is lower semicomputable since $-F$ is upper semi-computable. The multiplicative form of Proposition 2 implies

$$g(x,y,m) \stackrel{*}{=} 2^{-m} \mathbf{t}_{\nu_x}(y|x,m)$$
$$g(x,y) \stackrel{*}{=} 2^{F(x)} \mathbf{t}_{\nu_x}(y|x,-F(x)).$$

Since $\mathbf{t}_{\nu_x}$ is a test:

$$\nu^y_x 2^{-m} \mathbf{t}_{\nu_x}(y|x,m) \leq 2^{-m}$$
$$\nu^y_x g(x,y) \stackrel{*}{<} 2^{F(x)},$$

which implies

$$g(x,y) \overset{*}{<} \mathbf{t}_{(\nu,F)}(x,y)/\mathbf{m}(g|x) \overset{*}{<} \mathbf{t}_{(\nu,F)}(x,y)2^{\mathbf{K}(F)} \overset{*}{<} \mathbf{t}_{(\nu,F)}(x,y)$$

by the optimality of $\mathbf{t}_{(\nu,F)}$. We now consider the upper bound. Since, given fixed x, $2^{-F(x)}\mathbf{t}_{(\nu,F)}(x,y)$ is an y-test conditional on x and $-F(x)$, we have

$$2^{-F(x)}\mathbf{t}_{(\nu,F)}(x,y) \overset{*}{<} \mathbf{t}_{\nu_x}(y|x,-F(x))/\mathbf{m}(\mathbf{t}_{(\nu,F)}|x,-F(x)) \overset{*}{<} \mathbf{t}_{\nu_x}(y|x,-F(x))2^{\mathbf{K}(F)} \overset{*}{<} \mathbf{t}_{\nu_x}(y|x,-F(x)).$$

□

Theorem 1 *Relativized to probabilities μ and ν over $\{0,1\}^\infty$,*

$$\mathbf{D}(x,y|\mu\nu) =^+ \mathbf{D}(x|\mu) + \mathbf{D}(y|\nu_x,(x,\lceil\mathbf{D}(x|\mu)\rceil)).$$

Proof. Let $f(x,y) = 2^{\mathbf{D}(x,y|\mu\nu)}$. Proposition 1 implies $\nu_x^y f(x,y) \overset{*}{<} 2^{\mathbf{D}(x|\mu)+c}$. Let $F(x) = \lceil\mathbf{D}(x|\mu)\rceil + c$ for some suitable constant $c \in \mathbb{N}$. Note that if h is a lower computable function such that $\nu_x^y h(x,y) \overset{*}{<} 2^{\mathbf{D}(x|\mu)}$, then $\mu^x\nu_x^y h(x,y) \overset{*}{<} \mu^x\mathbf{t}_\mu(x) \overset{*}{<} 1$, so $h \overset{*}{<} f$, so f is a universal F-test. Proposition 3 (noting that $\mathbf{K}(\lceil\mathbf{D}(\cdot|\mu)\rceil) = O(1)$) gives

$$\mathbf{D}(x,y|\mu\nu) = \log f(x,y) =^+ F(x) + \mathbf{D}(y|\nu_x,(x,-F(x)))$$
$$\mathbf{D}(x,y|\mu\nu) =^+ \mathbf{D}(x|\mu) + \mathbf{D}(y|\nu_x,(x,\lceil\mathbf{D}(x|\mu)\rceil)).$$

□

References

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

[HR09] M. Hoyrup and C. Rojas. Computability of probability measures and martin-löf randomness over metric spaces. *Information and Computation*, 207(7):830–847, 2009.

[HRSV00] D. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin. Inequalities for shannon entropy and kolmogorov complexity. *Journal of Computer and System Sciences*, 60(2):442–464, 2000.