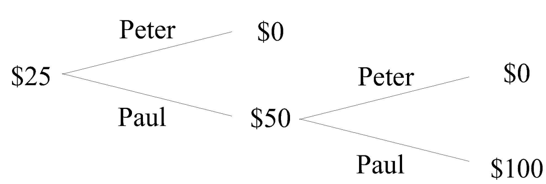
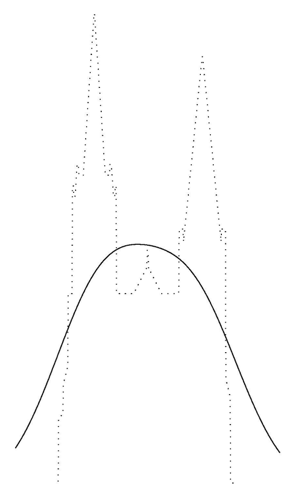
The universal measure of probabilistically nonrandom objects

Vladimir Vovk





The Game-Theoretic Probability and Finance Project

Working Paper #71

First posted September 12, 2026. Last revised September 15, 2026.

Project web site:
http://www.probabilityandfinance.com

Abstract

A finite object is called probabilistically random if it is an algorithmically random element of a finite set whose Kolmogorov complexity is relatively small. This note studies the amount of probabilistically nonrandom objects (those that are not probabilistically random) as gauged by the universal measure and without assuming any structure on the objects. The main results imply that the universal measure of probabilistically nonrandom objects is vanishingly small and establish the dependence of their universal measure on the required degree of probabilistic randomness.

A version of this note is published as an arXiv technical report: http://arxiv.org/abs/2609.13513.

Contents

1 Introduction

This note is a natural complement to [9]; both notes develop an approach to algorithmic statistics (see, e.g., [2,6–8]) inspired by Kolmogorov's original definition of stochastic objects [4, Note 5.12]. The standard definitions in algorithmic statistics assume a given structure on the space of finite objects under consideration, such as the binary strings $\{0,1\}^*$ being partitioned as the union of $\{0,1\}^n$, with the lengths of strings playing an important role in the statements of main results. Such structures are to some degree arbitrary; e.g., we may consider other increasing sequences of stopping times, different from the constant stopping times $0,1,2,\dots$. Besides, structureless settings are attractive due to their simplicity.

The previous note [9] discusses a structureless version of standard results on the universal measure of (α,β)-stochastic objects, where an object is (α,β)-*stochastic* if it has a model Ω of complexity at most α in which its randomness deficiency is at most β. (Formal definitions will be given later in the note.) While the standard results have been interpreted as saying that "the probability of obtaining a nonstochastic object in a random process is negligible" [7, Sect. 4.6], the main result of [9] says that (α,β)-nonstochastic objects are abundant: while the universal measure of (α,β)-nonstochastic objects tends to 0 as $\alpha\to\infty$ (β plays hardly any role), it does so extremely slowly.

A fundamental characteristic of a finite object ω in algorithmic statistics is its best fit function β_ω (to use the terminology of [8, (II.7)]); for each $\alpha>0$, $\beta_\omega(\alpha)$ is the smallest randomness deficiency of ω w.r.t. its model Ω of complexity at most α. Each function β_ω is monotonically decreasing to 0, to within an additive constant (it may take small negative values).

A popular alternative to the notion of the best fit function is that of the structure function for a finite object ω. Historically, structure functions appeared in print even before best fit functions [3], although there is evidence [1, Sect. 3.3, p. 858] that Kolmogorov defined best fit functions already in 1973 in his talk at an information theory conference in Tallinn; it is definitely true that he defined structure functions in that talk [4, the photo in Sect. 2]. (Cover et al. [1, p. 858] say that Kolmogorov "proposed a variant of" best fit functions, so it is also possible that those authors consider structure functions to be a variant of best fit functions.)

At the Tallinn conference Kolmogorov posed the problem of showing that the best fit functions can take approximately arbitrary shapes subject to the restriction of being monotonically decreasing to 0 [1, p. 858]. This was shown to be indeed the case by V'yugin in 1987 [10, Corollary] in a restricted setting and by Vereshchagin and Vitányi in 2004 [8, Corollary IV.9] in general; see also [6, Theorem 17.5].

The results about possible shapes of best fit functions are qualitative and do not indicate how common different shapes are. In this note we will be interested in how common they are under the universal measure in the structureless setting. It turns out that under the universal measure the most common shapes are those characteristic of probabilistically random objects, which were introduced in print by Kolmogorov in 1974 [3] in the language of structure functions. In terms of best fit functions, an object ω is probabilistically random (вероятностно случаен) if $\beta_\omega(\alpha)$ becomes small, which we will interpret as $O(1)$, already for a relatively small value (сравнительно небольшом значении) of α. Since best fit functions are monotonically decreasing, β_ω will stay small to the right of α as well. In our main results, Corollaries 3, 4, 5, and 6, we will interpret α being "relatively small" as $\alpha = \theta K(\omega)$ (for a small $\theta > 0$), $\alpha = K(\omega)^\theta$ (for $\theta \in (0,1)$), $\alpha = (\log K(\omega))^\theta$ (for $\theta > 1$), and $\alpha = \theta \log K(\omega)$ (also for $\theta > 1$), respectively, where $K(\omega)$ is the complexity of ω. For the objects ω of complexity $K(\omega) \geq a$, the universal measure of probabilistically nonrandom objects will shrink, as $a \to \infty$, to 0 exponentially fast, stretched-exponentially fast, superpolynomially fast, and polynomially fast, respectively.

In the standard structured setting, common shapes of best fit functions can be deduced from the preponderance, under the universal measure, of stochastic objects, which can be defined as the objects for which the graphs of their best fit functions pass close to the origin. In this case the best fit function quickly drops almost to zero and then stays close to zero. Therefore, the common objects under the universal measure are probabilistically random. The situation in structureless algorithmic statistics is radically different, since now nonstochastic objects are abundant [9, Theorem 1]. Nevertheless, the main result of this note still says that objects that are not probabilistically random are in a negligible minority under the universal measure.

There is a certain tension between the statement of [9] that nonstochastic objects are abundant and the statement of this note that there are few probabilistically nonrandom objects. As interpreted here, probabilistically random objects are a special case of (α, β)-stochastic objects ω corresponding to $\beta = O(1)$ and $\alpha = f(K(\omega))$ for a slowly growing f. The main difference between our current setting and [9] is that now we concentrate on the objects ω satisfying $K(\omega) \geq a$ for large a (in the next section we will see, in Proposition 1, that overall there are many such ω).

There are several possible settings for our topic. First, as already mentioned, Kolmogorov's structure function is a popular alternative to the best fit function of a finite object, and another alternative is the MDL function (see, e.g., [8, II.8 and II.10]). Second, in our definitions we can use plain Kolmogorov complexity C or prefix complexity K. For simplicity and concreteness, we concentrate on best fit functions and prefix complexity.

Let $\mathbb{N} := \{1, 2, \dots\}$ (more generally, $\mathbb{N}_m := \{m, m+1, \dots\}$, with the lower index omitted when $m = 1$), log be binary logarithm, K be prefix complexity, and $\mathbf{m}$ be the universal semimeasure on $\mathbb{N}$, as defined in, e.g., [5]. For $A \subseteq \mathbb{N}$, define $\mathbf{m}(A) := \sum_{\omega \in A} \mathbf{m}(\omega)$; with this extension, we also refer to $\mathbf{m}$ as the *universal measure* (since formally it becomes a measure). Let $\tau : \mathbb{N} \to (0, \infty)$ be the tail of $\mathbf{m}$: $\tau(t) := \mathbf{m}(\{n : n > t\})$. The words "increasing" and "decreasing" will be understood in the wide sense allowing intervals of constancy; we will drop "monotonically" from now on. The symbols O, o, and Θ will be used only in informal explanations and, occasionally, in proofs where their meaning is unambiguous.

2 Main results

Let us refer to elements of $\mathbb{N}$ as *objects* (to emphasize that our results and discussions are applicable to other spaces of finite objects, such as $\{0,1\}^*$, in computable bijection with $\mathbb{N}$). A finite set $\Omega \subseteq \mathbb{N}$ containing an object ω is a *model* for ω; we are interested in simple models, in the sense of $K(\Omega)$ being small, for which the *randomness deficiency*

$$d(\omega \mid \Omega) := \log |\Omega| - K(\omega \mid \Omega)$$

is small. For each $\omega \in \mathbb{N}$, the *best fit function* $\beta_\omega : [0, \infty) \to \mathbb{R} \cup \{\infty\}$ is defined by

$$\beta_\omega(\alpha) := \min_{\Omega : K(\Omega) \leq \alpha} d(\omega \mid \Omega),$$

Ω ranging over the models for ω; as usual, $\min \emptyset := \infty$. The effective range of α is $[0, K(\omega) + O(1)]$, since $\beta_\omega(K(\omega) + O(1)) = O(1)$ (witnessed by $\Omega := \{\omega\}$, whose complexity exceeds $K(\omega)$ by at most an additive constant). It is clear that the function β_ω is decreasing, piecewise-constant, and right-continuous (namely, it is constant on each interval $[n, n+1)$, $n \in \mathbb{N}_0$).

We are interested in possible shapes for β_ω for complex ω (although in this note we will only see primitive flat shapes). First let us see how many such ω we have overall. In the following proposition and later on we let $c_1, c_2, \ldots$ stand for positive constants (either absolute or depending on values or functions regarded as constant).

Proposition 1. *For some absolute constant $c_1 > 1$ and for all $a \in \mathbb{N}$,*

$$\tau(a)/c_1 \leq \mathbf{m}\left(\{\omega \in \mathbb{N} : K(\omega) \geq a\}\right) \leq c_1 \tau(a).$$

This very simple estimate, proved, e.g., in the companion note [9, Appendix B], shows that there are quite a few ω with $K(\omega) \geq a$. Their universal measure tends to 0, of course, as $a \to \infty$, but it does so very slowly (see, e.g., [9, Lemma 2]). Now let us see that the vast majority of those objects are probabilistically random as measured by the universal measure. First let us state a "master theorem" (Theorem 2, proved in Sect. 4) and then specialize it to more intuitive statements (Corollaries 3–6).

Theorem 2. *Let g be a computable increasing function $g : \mathbb{N} \to [0, \infty)$ with $g(k) \leq k$ for all $k \geq c_g$. There are positive constants c_2 and c_3 such that, for all $a \in \mathbb{N}$,*

$$\mathbf{m}\left(\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(g(K(\omega))) > c_2\}\right) \leq c_3 \sum_{k=a}^{\infty} 2^{-g(k)}. \tag{1}$$

Corollary 3. *Let $\theta \in (0, 1)$ be a computable number. There are positive constants c_2 and c_4 such that, for all $a \in \mathbb{N}$,*

$$\mathbf{m}\left(\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(\theta K(\omega)) > c_2\}\right) \leq c_4 2^{-\theta a}. \tag{2}$$

Proof. Let $g(k) := \theta k$ in Theorem 2. When evaluating the right-hand side of (1), we sum a geometric series with ratio $2^{-\theta}$. $\square$

Corollary 3 asserts the exponential decay of the universal measure in a. As this describes the universal measure of the (α, β)-nonstochastic objects with $\alpha := \theta K(\omega)$, it can be regarded as a version of the standard results containing $2^{-\alpha}$ as the leading term (see, e.g., [10, Corollary and Theorem 3] and [7, Propositions 5 and 22]).

Since β_ω is decreasing, (2) implies

$$\mathbf{m}\left(\{\omega \in \mathbb{N} : K(\omega) \geq a, \exists \theta' \geq \theta : \beta_\omega(\theta' K(\omega)) > c_2\}\right) \leq c_4 2^{-\theta a}. \tag{3}$$

Applying this to a small θ, we see that the $\mathbf{m}$-share of the objects that are not probabilistically random (in a crude sense) shrinks to zero exponentially fast as $a \to \infty$, in sharp contrast with the total $\mathbf{m}$-share of objects of complexity at least a.

Corollary 4. *Let $\theta \in (0, 1)$ be computable. There are positive constants c_2 and c_5 such that, for all $a \in \mathbb{N}$,*

$$\mathbf{m}\left(\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(K(\omega)^\theta) > c_2\}\right) \leq c_5 2^{-a^\theta + (1-\theta)\log a}.$$

In Corollary 4, the exponential decay of Corollary 3 becomes stretched exponential.

Proof of Corollary 4. Now $g(k) := k^\theta$ in Theorem 2. The summand in the series on the right-hand side of (1) is decreasing, so the sum is at most

$$2^{-a^\theta} + \int_a^\infty 2^{-x^\theta}\,\mathrm{d}x,$$

and substituting $u = x^\theta$ turns the integral into

$$\frac{1}{\theta}\int_{a^\theta}^\infty 2^{-u} u^{1/\theta - 1}\,\mathrm{d}u \leq \frac{2}{\theta \ln 2} 2^{-a^\theta} a^{1-\theta} \tag{4}$$

for large a (small a can be absorbed by c_5). The inequality in (4) is an instance of a general fact (also used in the proof of the following corollary): if $f > 0$ is differentiable with $(\ln f)' \leq -\delta < 0$ on $[A, \infty)$, then $f(u) \leq f(A)e^{-\delta(u-A)}$ there, and so

$$\int_A^\infty f(u)\,\mathrm{d}u \leq \frac{f(A)}{\delta}.$$ $\square$

Corollary 5. *Let $\theta > 1$ be computable. There is a positive constant c_2 such that, from some $a \in \mathbb{N}$ on,*

$$\mathbf{m}\left(\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega((\log K(\omega))^\theta) > c_2\}\right) \leq a^{1-(\log a)^{\theta - 1}}.$$

Corollary 5 asserts a weaker rate of decay of the universal measure of probabilistically nonrandom objects, namely superpolynomial decay.

Proof of Corollary 5. Our computation here is as in the proof of Corollary 4, but now $g(k) := (\log k)^\theta$, and we use the substitution $u = (\log x)^\theta$. The summand in the series on the right-hand side of (1) is still decreasing, so the sum is at most

$$2^{-(\log a)^\theta} + \int_a^\infty 2^{-(\log x)^\theta} \, \mathrm{d}x, \tag{5}$$

and the substitution turns the integral (the other addend in (5) will be negligible) into $\frac{\ln 2}{\theta} \int_A^\infty f(u) \, \mathrm{d}u$, where $A := (\log a)^\theta$ and $f(u) := 2^{-u+u^{1/\theta}} u^{1/\theta - 1}$. Since $\theta > 1$,

$$(\ln f)'(u) = \ln 2 \left(-1 + \frac{1}{\theta} u^{1/\theta - 1} \right) + \frac{1/\theta - 1}{u} \leq -\frac{\ln 2}{2}$$

for large u. Therefore, for large a the integral in (5) is at most

$$\frac{2}{\theta} f(A) = \frac{2}{\theta} (\log a)^{1-\theta} 2^{-(\log a)^\theta + \log a}.$$

Now

$$2^{-(\log a)^\theta + \log a} = a^{1-(\log a)^{\theta-1}},$$

and for large a the factor $\frac{2}{\theta}(\log a)^{1-\theta}$ absorbs both the first addend in (5) and the constant c_3 of Theorem 2, which gives the bound as stated. □

Corollary 6. *Let $\theta > 1$ be computable. There are positive constants c_2 and c_6 such that, for all $a \in \mathbb{N}$,*

$$\mathbf{m}\left(\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(\theta \log K(\omega)) > c_2\}\right) \leq c_6 a^{1-\theta}.$$

The rate of decay in a in Corollary 6 is even slower, namely polynomial.

Proof of Corollary 6. This time $g(k) := \theta \log k$, and so we can bound the sum on the right-hand side of (1) as

$$\sum_{k \geq a} k^{-\theta} \leq a^{-\theta} + \int_a^\infty x^{-\theta} \, \mathrm{d}x = a^{-\theta} + \frac{a^{1-\theta}}{\theta - 1} \leq \frac{\theta}{\theta - 1} a^{1-\theta}. \qquad \square$$

All four corollaries can be stated in the form (3). They assert that there are few probabilistically nonrandom objects as gauged by the universal measure; e.g., Corollary 5 says that the probabilistically nonrandom objects are superpolynomially few and Corollary 6 says that the probabilistically nonrandom objects are polynomially few. The corollaries trade off the strictness of the requirement of probabilistic randomness against the preponderance of probabilistically random objects.

3 Optimality of the rates

As in the previous section, we start with a theorem that looks complicated, and Corollaries 8–11, which are simpler and more intuitive, can be regarded as the main results of this section.

Theorem 7. *Let $g : [1, \infty) \to [0, \infty)$ be an increasing function. There are constants c_7, c_8, and c_9 such that, for*

$$m_g(a) := \min \{m \in \mathbb{N} : g(m + a + c_7 \log a) + c_8 \leq m\} \tag{6}$$

and for all $a \in \mathbb{N}_2$,

$$\mathbf{m}\left(\left\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(g(K(\omega))) = \infty\right\}\right) \geq 2^{-m_g(a) - c_9 \log a}. \tag{7}$$

As usual, $\min \emptyset$ in (6) is understood to be ∞, in which case the right-hand side of (7) is understood to be 0 (and so the inequality (7) holds trivially in this case). For a proof of the theorem, see Sect. 5. But now let us state more explicit results corresponding to Corollaries 3–6.

Corollary 8. *Let $\theta \in (0, 1)$. There is a constant c_{10} such that, for all $a \in \mathbb{N}_2$,*

$$\mathbf{m}\left(\left\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(\theta K(\omega)) = \infty\right\}\right) \geq 2^{-\frac{\theta}{1-\theta} a - c_{10} \log a}.$$

According to Corollary 8, the rate of decay of the universal measure in Corollary 3 is optimal up to replacing θ by $\frac{\theta}{1-\theta}$; and we are mainly interested in small θ.

Proof of Corollary 8. Set $g(k) := \theta k$. The defining inequality in (6) reads $\theta(m + a + O(\log a)) + O(1) \leq m$. $\square$

Corollary 9. *Let $\theta \in (0, 1)$. There is a constant c_{11} such that, for all $a \in \mathbb{N}_2$,*

$$\mathbf{m}\left(\left\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(K(\omega)^\theta) = \infty\right\}\right) \geq 2^{-a^\theta - a^{2\theta - 1} - c_{11} \log a}. \tag{8}$$

The exponent $-a^\theta - a^{2\theta-1} - c_{11} \log a$ on the right-hand side of (8) is $-a^\theta - O(\log a)$ for $\theta \leq 1/2$ and $-a^\theta(1 + o(1))$ in general; so for $\theta \leq 1/2$ the bounds in Corollaries 4 and 9 are in close agreement, the universal measure being $2^{-a^\theta \pm O(\log a)}$.

Proof of Corollary 9. Now $g(k) := k^\theta$. Trying $m = a^\theta + E$ in the defining inequality in (6) and using

$$(a + R)^\theta \leq a^\theta + \theta R a^{\theta - 1}$$

with $R = a^\theta + E + c_7 \log a$, we can see that the defining inequality holds as soon as

$$\theta a^{2\theta - 1} + \theta E a^{\theta - 1} + \theta c_7 a^{\theta - 1} \log a + c_8 \leq E,$$

and so $E = a^{2\theta - 1} + O(1)$ suffices. $\square$

Corollary 10. *Let $\theta > 1$. There is a constant c_{12} such that, for all $a \in \mathbb{N}_2$,*

$$\mathbf{m}\left(\left\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega((\log K(\omega))^\theta) = \infty\right\}\right) \geq a^{-(\log a)^{\theta-1} - c_{12}}.$$

The two bounds in Corollaries 5 and 10 agree up to a factor of $a^{O(1)}$, the universal measure being $a^{-(\log a)^{\theta-1} \pm O(1)}$.

Proof of Corollary 10. In this proof $g(k) := (\log k)^\theta$. Suppose $R = O((\log a)^\theta)$ (we are interested in R of the form $R = m + c_7 \log a$). It is always true that $\log(a+R) \leq \log a + R/(a \ln 2)$ and hence, by the mean value theorem under our assumption $R = O((\log a)^\theta)$,

$$(\log(a+R))^\theta \leq \left(\log a + \frac{R}{a \ln 2}\right)^\theta \leq (\log a)^\theta + O\left(\frac{(\log a)^{2\theta-1}}{a}\right).$$

We can see that the defining inequality in (6) holds with $m = (\log a)^\theta + O(1)$. □

Corollary 11. *Let $\theta > 1$. There is a constant c_{13} such that, for all $a \in \mathbb{N}_2$,*

$$\mathbf{m}\left(\left\{\omega \in \mathbb{N} : K(\omega) \geq a, \beta_\omega(\theta \log K(\omega)) = \infty\right\}\right) \geq a^{-c_{13}}. \tag{9}$$

The bounds in Corollaries 6 and 11 match each other only in a very weak sense: both are polynomial. The right-hand side of (9) is particularly non-specific; it is just the generic polynomial rate.

Proof of Corollary 11. Set $g(k) := \theta \log k$. Now the defining inequality in (6) reads $\theta \log(m + a + c_7 \log a) + c_8 \leq m$, which $m = \theta \log a + O(1)$ satisfies since $\log(a + O(\log a)) = \log a + o(1)$. □

4 Proof of Theorem 2

This proof uses standard techniques and merely applies them in a new way. A key role is played by the sets $D_k := \{\omega \in \mathbb{N} : K(\omega) \leq k\}$. It is well known that $|D_k| \leq 2^{k-K(k)+O(1)}$; see, e.g., [5, Theorem 64]. We will use the technique of "bounded complexity lists" (see, e.g., [2] and [7, Sect. 4]), and we start by fixing a jointly computable way of enumerating all elements of D_k given k. Without loss of generality we assume that g only takes integer values, $g : \mathbb{N} \to \mathbb{N}_0$. Without loss of generality we also assume $a \geq c_g$; indeed, $a < c_g$ can be absorbed by the constant c_3 in (1).

Set $k := K(\omega)$ for a given object $\omega \in \mathbb{N}$. We will split D_k in the order in which it is enumerated into blocks of equal size 2^b and a remainder of size less than 2^b taking $b := k - g(k)$ (here we are using the condition $g(k) \leq k$). As Ω we take the block containing ω (assuming it exists); let us check that $K(\Omega) \leq g(k) + c_{14}$. Indeed, to describe Ω, we first describe k, which requires $K(k)$ bits but also determines $g(k)$ and b; then the ordinal number of Ω as a block is determined by a number at most

$$2^{k-K(k)+c_{15}}/2^b = 2^{k-K(k)+c_{15}}/2^{k-g(k)} = 2^{g(k)-K(k)+c_{15}}, \tag{10}$$

which requires at most $g(k) - K(k) + c_{16}$ bits given k. Overall, describing Ω requires at most $g(k) + c_{14}$ bits. (If the last exponent in (10) is negative, there are no blocks and ω is guaranteed to be in the remainder, which case is also covered by the argument below.)

Let us check that ω is a random element of Ω:

$$\begin{aligned} d(\omega \mid \Omega) = \log|\Omega| - K(\omega \mid \Omega) \le \log|\Omega| - K(\omega) + K(\Omega) + c_{17} \\ \le k - g(k) - k + g(k) + c_2 = c_2. \end{aligned}$$

Now we can deduce essentially the statement of the theorem: the ω satisfying $K(\omega) = k \ge a$ and $\beta_\omega(g(k) + c_{14}) > c_2$ are among the remainder and so their overall universal measure is less than $2^b 2^{-k+c_{18}} = 2^{-g(k)+c_{18}}$; it remains to sum over all such k.

The argument so far gives us an extra additive constant in the statement of Theorem 2 ($g(K(\omega)) + c_{14}$ in place of $g(K(\omega))$). It can be eliminated by increasing b by an additive constant: using blocks of size 2^b with $b := k - g(k) + c_{19}$, c_{19} being large as compared with c_{14}, divides their number by $2^{c_{19}}$, so that describing Ω now requires at most $g(k)$ bits, while $d(\omega \mid \Omega)$ and the universal measure of the remainder grow by at most c_{19} and the factor $2^{c_{19}}$, respectively, which is absorbed by c_2 and c_3. (And if there are no blocks, we have no Ω to describe.)

5 Proof of Theorem 7

The proof uses the busy-beaver-type function

$$B(k) := \max\{n \in \mathbb{N} : K(n) \le k\} \tag{11}$$

(with $\max \emptyset := 0$, as usual); this is an increasing function that grows extremely quickly. The main component of the proof of Theorem 7 is the following lemma.

Lemma 12. *There exists a constant c_8 such that, for every object $\omega \in \mathbb{N}$ and every $\alpha \ge 0$, we have $\beta_\omega(\alpha) = \infty$ whenever $\omega > B(\alpha + c_8)$.*

Proof. It is clear that we can choose c_8 such that $K(\max \Omega) \le K(\Omega) + c_8$ for all Ω. The statement of the lemma follows from there being no models Ω for $\omega > B(\alpha + c_8)$ of complexity $K(\Omega) \le \alpha$: indeed, for any model Ω for ω,

$$\begin{aligned} \omega > B(\alpha + c_8) \Longrightarrow \max(\Omega) > B(\alpha + c_8) \\ \Longrightarrow K(\max \Omega) > \alpha + c_8 \Longrightarrow K(\Omega) > \alpha. \quad \square \end{aligned}$$

The intuition behind Lemma 12 is that a model for ω must in particular name a number at least as large as ω, so no object has models cheaper than the least complexity of a number equal to or exceeding it. The idea of the proof of Theorem 7 is to extend a busy-beaver number by random bits obtaining many numbers at a controlled complexity.

Now we can prove the theorem. Let $m := m_g(a)$, and for $y \in \{0, 1, \dots, 2^{a+1} - 1\}$ put $\omega_y := B(m) \cdot 2^{a+1} + y$ (intuitively, extend the binary representation of $B(m)$ by $a+1$ random bits). These 2^{a+1} objects are distinct and exceed $B(m)$, so $K(\omega_y) > m$ by the definition (11).

Now let us bound the complexity of ω_y. Concatenating a shortest description for $B(m)$ (of length $K(B(m)) \leq m$, again by the definition of B), a shortest description for a, and y written in exactly $a+1$ bits gives a prefix-free description for ω_y, so (assuming $a \geq 2$)

$$K(\omega_y) \leq m + a + c_7 \log a. \tag{12}$$

On the other hand, fewer than 2^a of the ω_y have $K(\omega_y) \leq a - 1$, so at least $2^{a+1} - 2^a = 2^a$ values of y satisfy $K(\omega_y) \geq a$; let us call these y *suitable*; they will witness (7).

For a suitable y, we have, by (12), the monotonicity of g, and the definition (6) of $m_g(a)$,

$$g(K(\omega_y)) + c_8 \leq g\bigl(m + a + c_7 \log a\bigr) + c_8 \leq m;$$

therefore,

$$\omega_y > B(m) \geq B(g(K(\omega_y)) + c_8)$$

and hence, by Lemma 12, $\beta_{\omega_y}(g(K(\omega_y))) = \infty$.

On the other hand, by (12),

$$\mathbf{m}(\omega_y) \geq 2^{-m-a-c_7 \log a - c_{20}};$$

summing over the at least 2^a suitable y gives at least $2^{-m-c_9 \log a}$.

6 Conclusion

Corollaries 3–6, and Theorem 2 in general, require α (the argument of β_ω) to be at least about $\log K(\omega)$: the series $\sum_k 2^{-g(k)}$ in Theorem 2 converges for $g(k) = \theta \log k$ if and only if $\theta > 1$, which is exactly the range covered by Corollary 6. The case $\alpha = \theta \log K(\omega)$ with $\theta \leq 1$ is, however, also interesting. For example, if ω is a binary data sequence of a fixed length n with negligible $K(n)$ (e.g., $n = B(K(n))$) generated from a regular statistical model with p parameters, we expect it to satisfy $K(\omega) = \Theta(n)$ and to be (α, β)-stochastic with $\alpha \approx \frac{p}{2} \log n$ and $\beta = O(1)$; Corollary 6 says nothing when $p \leq 2$. In particular, it would be interesting to know what shapes of β_ω are possible with non-negligible universal measure in the range $\alpha \leq \log K(\omega)$.

In the case of $\alpha = \Theta(\log K(\omega))$ the difference between sets and probability measures as models becomes essential (see, e.g., [7, Proposition 2]). This makes extending this note's results to probability measures as models an interesting direction of research. Considering alternatives to best fit functions, first of all structure functions but also MDL functions, as defined in [8, (II.8) and (II.10)], is also likely to lead to illuminating results.

Acknowledgments

Claude Opus 5 and Fable 5.1 have been used in exploring proof ideas, which I reviewed carefully, and in checking the note. I take full responsibility for all claims and statements made here, including mathematical statements and their proofs.

References

[1] Thomas M. Cover, Péter Gács, and Robert M. Gray. Kolmogorov's contributions to information theory and algorithmic complexity. *Annals of Probability*, 17:840–865, 1989.

[2] Péter Gács, John T. Tromp, and Paul Vitányi. Algorithmic statistics. *IEEE Transactions on Information Theory*, 47:2443–2463, 2001.

[3] Andrei N. Kolmogorov. Сложность алгоритмов и объективное определение случайности. *Успехи математических наук*, 29(4):155, 1974. Abstract of a talk before the Moscow Mathematical Society. Meeting of 16 April 1974. English translation: [7, Sect. 3.5].

[4] Alexey Semenov, Alexander Shen, and Nikolai Vereshchagin. Kolmogorov's last discovery? (Kolmogorov and algorithmic statistics). *Theory of Probability and its Applications*, 68:582–606, 2024. Also published as arXiv:2303.13185 [math.LO], October 2023.

[5] Alexander Shen, Vladimir A. Uspensky, and Nikolai Vereshchagin. *Kolmogorov Complexity and Algorithmic Randomness*. American Mathematical Society, Providence, RI, 2017.

[6] Nikolai Vereshchagin and Alexander Shen. Algorithmic statistics revisited. In Vladimir Vovk, Harris Papadopoulos, and Alexander Gammerman, editors, *Measures of Complexity: Festschrift for Alexey Chervonenkis*, chapter 17, pages 235–252. Springer, Cham, 2015.

[7] Nikolai Vereshchagin and Alexander Shen. Algorithmic statistics: forty years later. In Adam Day, Michael Fellows, Noam Greenberg, Bakhadyr Khoussainov, Alexander Melnikov, and Frances Rosamond, editors, *Computability and Complexity: Essays Dedicated to Rodney G. Downey on the Occasion of His 60th Birthday*, volume 10010 of *Lecture Notes in Computer Science*, pages 669–737. Springer, Cham, 2017.

[8] Nikolai Vereshchagin and Paul Vitányi. Kolmogorov's structure functions and model selection. *IEEE Transactions on Information Theory*, 50:3265–3290, 2004.

[9] Vladimir Vovk. The universal measure of nonstochastic objects. Technical Report arXiv:2608.27098 [math.LO], arXiv.org e-Print archive, August 2026. For the latest version, see gtfp.net, Working Paper 70.

[10] Vladimir V. V'yugin. On the defect of randomness of a finite object with respect to measures with given complexity bounds. *Theory of Probability and its Applications*, 32:508–512, 1987.