Upper-semicomputable sumtests for lower semicomputable semimeasures

Bruno Bauwens

Abstract

A sumtest for a discrete semimeasure $P$ is a function $f$ mapping bit-strings to non-negative rational numbers such that

$$\sum P(x)f(x) \leq 1.$$  

Sumtests are the discrete analogue of Martin-Löf tests. The behavior of sumtests for computable $P$ seems well understood, but for some applications lower semicomputable $P$ seem more appropriate. In the case of tests for independence, it is natural to consider upper semicomputable tests (see [B. Bauwens and S. Terwijn, Theory of Computing Systems 48.2 (2011): 247-268]).

In this paper, we characterize upper semicomputable sumtests relative to any lower semicomputable semimeasures using Kolmogorov complexity. It is studied to what extent such tests are pathological: can upper semicomputable sumtests for $m(x)$ be large? It is shown that the logarithm of such tests does not exceed $\log |x| + O(\log(2)|x|)$ (where $|x|$ denotes the length of $x$ and $\log(2) = \log \log$) and that this bound is tight, i.e. there is a test whose logarithm exceeds $\log |x| - O(\log(2)|x|)$ infinitely often. Finally, it is shown that for each such test $e$ the mutual information of a string with the Halting problem is at least $\log e(x) - O(1)$; thus $e$ can only be large for "exotic" strings.

KEYWORDS: Kolmogorov complexity – universal semimeasure – sumtest

1 Introduction

A (discrete) semimeasure $P$ is a function from strings to non-negative rational numbers such that $\sum P(x) \leq 1$. A sumtest $f$ for a semimeasure $P$ is a function...
from bitstrings to non-negative reals such that
\[ \sum_{x} P(x) f(x) \leq 1. \]

Sumtests provide a rough model for statistical significance testing for a hypothesis about the generation of data \([7, 8, 9]\). This hypothesis should be sufficiently specific so that it determines a unique semimeasure \(P\) describing the generation process of observable data \(x\) in a statistical experiment. The value \(f(x)\) plays the role of significance in a statistical test. The condition implies that the \(P\)-probability of observing \(x\) such that \(f(x) \geq k\) is bounded by \(1/k\). If \(f(x)\) is high, it is concluded that either a rare event has occurred or the hypothesis is not consistent with the generation process of the data.

For many hypotheses, such as the hypotheses that two observables are independent, not enough information is available to infer a unique probability distribution. In such a case, one might consider a set of semimeasures that are consistent with the hypothesis. We say that \(m^H\) is universal in a set of semimeasures if \(m^H\) is in the set and if for each \(P\) in the set there is a constant \(c\) such that \(P \leq c \cdot m^H\). (If \(P \leq c \cdot Q\) for some \(c\), we say that \(Q\) dominates \(P\).) It might happen that the class of lower semicomputable \(P\) consistent with a hypothesis has a universal element \(m^H(x)\) and that each such element satisfies
\[ P(x) \leq O \left( 2^{-K(P)} \right) m^H(x), \]
here, \(K(P)\) is the minimal length of a program that enumerates \(P\) from below. This happens for the hypothesis of independence of two strings, or directed influences in time series (see \([1, Proposition 2.2.6]\)). In this case, if observed data results in a high value of a sumtest for \(m^H\), it can be concluded that either:

- a rare event has occurred,
- the hypothesis is not consistent with the generation of the data,
- the data was generated by a process that is only consistent with semimeasures of high Kolmogorov complexity.

Unfortunately, the use of approximations of universal tests seems not to be practical, and this interpretation is rather philosophic. However, one can raise the question whether sequences of improving tests (for example general tests for independence) reported in literature tend to approach some ideal limit. This motivates the question whether there exist large sumtests in some computability classes, and whether they have universal elements.

Let \(\mathcal{P}^\uparrow\) be the class of lower semicomputable semimeasures. In algorithmic information theory, it is well known that this class has a universal element \(m\) \([7]\), and \(m\) can be characterized in terms of prefix Kolmogorov complexity: \(\log m(x) = K(x) + O(1)\) for all \(x\) (see \([7]\) for more background on Kolmogorov complexity). For all computable \(P\) it is also known that:

- no universal element exists in the class of computable tests,
• no upper semicomputable test exists that dominates all lower semicomputable ones,

• a universal sumtest in the class of lower semicomputable tests is $m(x)/P(x)$.

In [3], these statements are studied for lower semicomputable $P$. In this case all of the results above become false: some semimeasures have a computable universal test (for example $f(x) = 1$ is a universal test for $m$ even among the lower semicomputable tests), upper semicomputable tests can exceed lower semicomputable ones, (every $P \in \mathcal{P}^1$ has an unbounded upper semicomputable test [3] Proposition 5.1), thus also $m$, and for some $P$ no universal lower semicomputable test exists [3] Proposition 4.4.

Consider the hypothesis of independence for bivariate semimeasures. The corresponding class of semimeasures for this hypothesis is $P(x)Q(y)$ where $P$ and $Q$ are univariate semimeasures. The subset of lower semicomputable semimeasures has a universal element given by $m(x)m(y)$. A sumtest relative to this semimeasure can be called an independence test. It is not hard to show that any lower semicomputable independence tests is bounded by a constant $\frac{1}{2}$.

In [3] it is shown that upper semicomputable independence tests exist whose logarithm equal $n + O(\log n)$ for all $n$ and pairs $(x, y)$ of strings of length $n$. Moreover, there exist a generic upper semicomputable test $u_h$ defined for all computable $h$, and this test is increasing in $h$: if $h \leq g$ then $u_h \leq u_g$. The tests obtained in this way somehow “cover” all tests: every upper semicomputable test is dominated by $u_h$ for some computable $h$. Moreover, $u_h$ has a characterization in terms of (time bounded) Kolmogorov complexity, and for fixed $(x, y)$ and increasing $h$, the value $\log u_h(x, y)$ approaches algorithmic mutual information $I(x; y) = K(x) + K(y) - K(x, y)$. Unfortunately, no universal upper semicomputable sumtest exist.

The first result of the paper is to define such a generic test $u_h$ for all lower semicomputable semimeasures in terms of Kolmogorov complexity. This generalizes the result mentioned above. The remaining goal of the paper is to investigate whether upper semicomputable tests are pathological, in particular can such tests for the universal semimeasure $m$ be large? This would imply that these tests identify “structure” which can not be identified by compression algorithms.

We show that there are sumtests for $m$ for which the logarithm exceeds $\log |x| - O(\log \log |x|)$, but for no test its logarithm exceeds $\log |x| + O(\log |x|)$. This is small compared to the logarithm of the independence tests discussed above, which equal $n + O(\log n)$ for some $(x, y)$ of length $n$. We show that the logarithm of each test is bounded by the mutual information with the Halting problem $0'$ given by $K(x) - K^0(x)$, up to an additive constant (that depends on the test). Hence, strings $x$ for which such a test is large, are unlikely to be produced in a statistical experiment. We also show that there is no universal upper semicomputable test for $m$.

1 Supposed that a sumtest $f$ for $m(x)m(y)$ is unbounded. For each $k$ one can search a pair $(x, y)$ such that $f(x, y) \geq 2^k$. For the first such pair that appears we have $\log m(x) = K(x) \leq K(x, y) \leq K(k) \leq O(\log k)$ up to $O(1)$ terms, and similar for $m(y)$. This implies $f(x, y)m(x)m(y) \geq 2^k - O(\log k)$, which contradicts the condition in the definition of sumtest.
2 A generic upper semicomputable sumtest

Let $P$ be a lower semicomputable semimeasure. For every computable two argument function $h$, we define an upper semicomputable sumtest $u_h$ for $P$. For increasing $h$ these sumtests are increasing and we show that for each upper semicomputable sumtest $f$ there exists a computable $h$ such that $u_h$ exceeds $f$ up to a multiplicative constant.

In our construction we use $m(\cdot|\cdot)$, which is a bivariate function that is universal for all lower semicomputable conditional semimeasures. Let $m(\cdot|\cdot)$ and $P(\cdot)$ represent approximations of $m(\cdot|\cdot)$ and $P(\cdot)$ from below.

For any computable $h$, let

$$u_h(x) = \inf_s \left\{ \frac{m_{h(x, s)}(x|s)}{P_s(x)} \right\}.$$ 

**Theorem 1.** $u_h$ is an upper semicomputable sumtest for $P$. For any upper semicomputable sumtest $e$ for $P$, there exist a $c$ and a computable $h$ such that $u_h \leq c \cdot e$.

**Remark.** Let $t$ be a number. We can define a generic test $u_h$ using time bounded Kolmogorov complexity, given by

$$K^t(x) = \min \{|p|: U(p) \text{ outputs } x \text{ and halts in at most } t \text{ steps}\}.$$ 

Indeed, we simply fix the approximation $m_t(x|s)$ to be $2^{-K^t(x|s)}$, and by the conditional coding theorem, this defines a universal conditional semimeasure $m(\cdot|\cdot)$.

**Proof.** Clearly, $u_h$ is upper semicomputable, so we need to show it is a sumtest. For each fixed $s$, the function $m_{h(x, s)}(x|s)/P_s(x)$ is a sumtest for $P_s$, and $u_h(x)$ is not larger; thus for all $s$:

$$\sum_x u_h(x)P_s(x) \leq 1.$$ 

This implies that the relation also holds in the limit, thus $u_h$ is a sumtest for $P$.

For the second claim of the lemma, note that we can choose an approximation $e_s$ (from above) of $e$ such that $e_s$ has computably bounded support and $\sum_s P_s(x)e_s(x) \leq 1$ for all $s$. By universality of $m(\cdot|\cdot)$, this implies that there exist a $c$ such that $P_s(x)e_s(x) \leq c \cdot m(x|s)$, and this $c$ does not depend on $x$ and $s$. We can wait for the first stage in the approximation of $m(x|s)$ for which this equation becomes true and let this stage define the function $h(x, s)$. This implies $e_s(x) \leq c \cdot m_{h(x, s)}(x|s)/P_s(x)$ for all $x$ and $s$, thus $e(x) \leq c \cdot u_h(x)$. $h(x, s)$ is defined for all $x$ and $s$, thus $h$ is computable.

Let $K^0(x)$ be the Kolmogorov complexity of $x$ on a machine with an oracle for to the Halting problem that is optimal.

---

2 This simplified proof was suggested by Alexander Shen.
Corollary 2. If $e$ is an upper semicomputable sumtest for a universal semimeasure, then $\log e(x) \leq K(x) - K^0(x) + O(1)$. (The constants implicit in the $O(\cdot)$ notation depend on $e$.)

Proof. It suffices to show the corollary for $e = u_h$. Let $m^0'(x) = 2^{-K^0(x)}$. We use [4, Theorem 2.1], which states

$$m^0'(x) = \Theta \left( \lim \inf_t m(x|t) \right).$$

By definition of $u_h$:

$$u_h \leq \lim \inf_t \frac{m_h(x,t|x|t)}{m_t(x)}.$$

For all but finitely many $t$, the denominator exceeds $m(x)/2$, and by the theorem mentioned above, $\lim \inf m(x|t)$ in the numerator is $O(m^0'(x))$. By choice of $m^0'(x)$, it equals $2^{-K^0(x)}$. Thus

$$\leq \lim \inf_t \frac{m(x|t)}{m(x)/2} \leq O \left( \frac{m^0'(x)}{m(x)} \right) = O \left( \frac{2^{-K^0(x)}}{2^{-K(x)}} \right).$$

\[ \square \]

3 Upper bound for upper semicomputable tests for $m$

Theorem 3. If $e$ is an upper semicomputable sumtest for $m$, then

$$e(x) \leq O \left( |x| (\log |x|)^2 \right).$$

Proof. By Theorem 1, it suffices to show the theorem for $u_h$ for all computable $h$. Note that $\sum_x 2^{-2|x|} = \sum_n 2^{-n} = 1$, thus for each universal semimeasure $m$ there exist $c > 0$ such that $m(x) > e2^{-2|x|}$. Assume that $m^0(\cdot)$ is an approximation from below of such $m(\cdot)$ such that $m_1(x) \geq c2^{-2|x|}$.

The idea is as follows. We consider some times $t_1, t_2, \ldots$ on which $m(t)$ is large. (For any number $n$ let $m(n)$ be the universal semimeasure for the string containing $n$ zeros.) This implies that $m_h(x,t_i|x|t_i)$ is not much above $m_{t_{i+1}}(x)$ if $t_{i+1}$ is sufficiently above $t_i$. From the definition of $u_h(x)$ with $t = t_i$ it follows that if $u_h(x)$ is large, then also $m_{t_{i+1}}(x)/m_{t_i}(x)$ must be large. On the other hand, $m(x)/m_1(x)$ is bounded, thus the first ratio can only be large for few $t_i$. On the other hand, our construction implies that for large $u_h(x)$ the first ratio must be large for many $t_i$.

We show the following claim

For all computable $h$ there exist a series of numbers $t_1, t_2, t_3, \ldots$ and a constant $c > 0$ such that for all $i \geq |x|$ either

$$\frac{m_{t_{i+1}}(x)}{m_{t_i}(x)} \geq 2 \quad \text{or} \quad u_h(x) < 2ci(\log i)^2.$$
Let us first show how this implies the theorem. The definition of a semi-measure implies $m(x) \leq 1$, thus $\frac{m(x)}{m_1(x)} \leq O \left(2^{2|x|}\right)$ by assumption on $m_1(i)$, and hence

$$\frac{m_{t_1}(x)}{m_1(x)} \cdots \frac{m_{t_{|x|-1}}(x)}{m_{t_{|x|-2}}(x)} \leq \frac{m(x)}{m_1(x)} \leq O \left(2^{2|x|}\right).$$

For large $x$, at most $2|x| + O(1) < 3|x|$ elements in the series $t_{|x|}, t_{|x|+1}, \ldots, t_{|x|-1}$ can satisfy the left condition. Thus, some element does not satisfy the condition and hence

$$u_h(x) < 2c \left(4|x|\right) \left(\log(4|x|)\right)^2.$$

This implies the theorem.

We now construct a sequence $t_1, t_2, t_3, \ldots$ satisfying the conditions of the claim. This construction depends on a parameter $c$, which will be chosen later. Let $t_1 = 1$. For $i \geq 1$, $t_{i+1}$ is given by the first stage in the approximation of $m(\cdot)$ such that

$$\frac{m_{h(x,t_i)}(x|t_i)}{i(\log i)^2} \leq c \cdot m(x),$$

(1)

for all $x$ of length at most $i$.

We first argue why for appropriate $c$, such a stage $t_{i+1}$ exist, i.e. why (1) holds for all $x$. Note that the sequence is recursively enumerated uniformly in $c$, thus $m(c)/(i(\log i)^2) \leq O(m(i)m(c)) \leq O(m(t_i))$. On the other side

$$m_{h(x,t_i)}(x|t_i) m(t_i) \leq O(m(x)),$$

thus for some $c'$ independent of $c, i$ and $x$:

$$\frac{m_{h(x,t_i)}(x|t_i) m(c)}{i(\log i)^2} \leq c' \cdot m(x).$$

and (1) is satisfied if $c \geq c'/m(c)$. This relation holds if we choose $c$ to be a large power of two (indeed $m(2^a) \leq \alpha/2^2$ for some $\alpha > 0$, thus choose $l$ such that $2^a \geq c'^2/\alpha$).

It remains to show the claim. Assume that the right condition is not satisfied and choose $t = t_i$ in the definition of $u_h$:

$$2ci(\log i)^2 \leq u_h(x) \leq \frac{m_{h(x,t_i)}(x|t_i)}{m_{t_i}(x)} m_{t_{i+1}}(x) \frac{m_{t_{i+1}}(x)}{m_{t_i}(x)} \leq c(\log i)^2 \frac{m_{t_{i+1}}(x)}{m_{t_i}(x)}.$$

This implies the left condition. 

4. Construction of large upper semicomputable tests

For each pair $(f, g)$ of computable functions, an upper semicomputable function $e_{f,g}$ is constructed. Afterwards, it is shown that $\sum_x m(x)e_{f,g}(x) \leq O(1)$ for appropriate $f$. Finally, we construct $g$ such that $e_{f,g}$ equals $|x|/|\log |x||^2$ for infinitely many $x$. Before constructing $e_{f,g}$, we show a technical lemma.
Lemma 4. There exists a sequence of numbers $t_1, t_2, \ldots$ such that for all $k$

$$\sum_{x} \left\{ m(x) : \frac{m(x)}{m_{t_k}(x)} \geq 2 \right\} \leq 2^{-k+1}$$

and such that for some computable function $f$ and large $k$

$$m_{f(t_k)}(x) \leq 2^{-k}/k^2.$$ 

Proof. The proof of the lemma is closely related to the proof that strings with high Kolmogorov complexity of Kolmogorov complexity given the string are rare [6] (see [2] for more on the technique). The construction of $t_k$ uses the function $k_t$, which is in turn defined using an approximation from below for the famous number $\Omega$ (see [5]):

$$\Omega = \sum_{x} m(x) \quad \text{and} \quad \Omega_t = \sum_{|x| \leq t} m_t(x).$$

For each $t$ let $k_t$ be the position of the leftmost bit of $\Omega_t$ (in binary) that differs from $\Omega_{t-1}$. Note that $k_t$ tends to infinity for increasing $t$. Let

$$t_k = \max \{ t : k_t \leq k \},$$

i.e. the largest $t$ for which there is a change in the first $k$ bits of $\Omega_t$. Clearly,

$$\sum_{x} m(x) - \sum_{|x| \leq t_k} m_{t_k}(x) = \Omega - \Omega_{t_k} \leq 2^{-k}.$$ 

and this inequality implies the first inequality of the lemma.

For the second, we show that there exist a computable $f$ such that for all $t$:

$$m_{f(t)}(t) \leq O\left(2^{-k_t}/k_t^2\right).$$

Indeed, given the first $k_t$ bits $y$ of $\Omega_t$ we can compute $t$ (by waiting until the first stage $s$ such that $y$ is a prefix of $\Omega_s$). This implies $m(y) \leq O(m(t))$. Note that $2^{-|x|}/|x|^2 \leq o(m(z))$ for all $z$, thus

$$2^{-k_t}/k_t^2 \leq m(t)$$

for large $t$. $k_t$ is computable from $t$ and we can wait until the current approximation of $m(t)$ is large enough to satisfy the equation. Let this stage be $f(t)$. Note that it is defined for all $t$, and hence $f$ is computable and satisfies the inequality at the start of the paragraph. \qed

Let $e_{f,g}(x)$ be equal to $|x|/(\log |x|)^5$ if for all $t$ either

$$m_{f(t)}(t) \leq \frac{6 \log |x|}{|x|} \quad \text{or} \quad \frac{m_{g(x,t)}(x)}{m_t(x)} \geq 2,$$

otherwise let $e_{f,g}(x) = 1.$

7
Proposition 5. For some \( c > 0 \), some computable \( f \) and for all computable \( g \) the function \( c \cdot e_{f,g} \) is an upper semicomputable sumtest for \( m \).

Proof of Proposition 5. \( e_{f,g} \) is upper semicomputable. Indeed, for at most finitely many \( x \) the relation \( |x|/(\log |x|)^5 \geq 1 \) is false; for all other \( x \), let the test be equal this value until we find a \( t \) that does not satisfy the conditions. It remains to construct \( f \) such that
\[
\sum_x e_{f,g}(x)m(x) \leq O(1).
\]

For some \( x \) of length \( n \), let \( k = \log n - 3 \log n - 3 \), and suppose that \( t_k \) and \( f \) satisfy the conditions of Lemma 4. One can calculate that
\[
m_f(t_k) > 2^{-k} \geq (6 \log n)/n.
\]

The set of all \( x \) of length \( n \) such that \( e_{f,g}(x) > 1 \) has measure at most
\[
2^{-k+1} \leq O\left(\frac{(\log n)^3}{n}\right).
\]

We are now ready to bound the sum at the beginning of the proof. It is sufficient to consider \( x \) such that \( e_{f,g}(x) > 1 \).
\[
\sum_x \{e_{f,g}(x)m(x) : e_{f,g}(x) > 1\} \leq \sum_n \frac{n}{(\log n)^3} \sum_{|x|=n} \{m(x) : e_{f,g}(x) > 1\},
\]
by the above bound on the measure of the sets in the right sum, this is at most
\[
\leq O(1) \sum_n \frac{n}{(\log n)^3} \frac{(\log n)^3}{n} \leq O(1) \sum_n \frac{1}{(\log n)^2} \leq O(1).
\]

Theorem 6. There exist an upper semicomputable sumtest for \( m \) that exceeds
\[
\Omega \left( |x|/(\log |x|)^5 \right)
\]
for some \( x \) of all large lengths.

Proof. By Proposition 5, it suffices to construct a computable \( g \) and an \( x \) of each large length such that \( e_{f,g}(x) = |x|/(\log |x|)^5 \). Our construction works for any function \( f \).

Fix a large \( n \). By construction of \( e_{f,g} \), the function exceeds one on some \( x \) of length \( n \), only if \( \log(1/m_f(x)) \) increases each time \( m_f(t)(t) \) is large. Let \( t_1, t_2, \ldots \) be the set of all \( t \) such that \( m_f(t)(t) < 6(\log n)/n \). Clearly, there are less than \( n/(6 \log n) \) such \( t \) and the sequence can be enumerated uniformly in \( n \).

To construct \( x \) we maintain a lexicographically ordered list that initially contains all strings of length \( n \). At stages \( t_i \) we remove all \( x \) from the list for which
\[
\log(1/m_f(x)) \leq n - 6i \log n.
\]

Let \( x \) be the first string in the list that is never removed. Such \( x \) exist for large \( n \), because in total there are less than \( 2^{n-6 \log n+1} \) strings removed.

By construction
\[
\log (1/m_f(x)) > n - 6i \log n.
\]
Now we argue, why there is a computable $g$ such that
\[
\log \left( 1/m_{g(x,t)}(x) \right) \leq n - 6i \log n - 2 \log n + O(1)
\]
(2)
for all $i$ such that $t_i$ exists; and by construction of $e_{f,g}$ this is sufficient for the theorem. After stage $t_i$, we remove at most
\[
\exp(n - 6(i + 1) \log n) + \exp(n - 6(i + 2) \log n) + \ldots < \exp(n + 1 - 6(i + 1) \log n)
\]
strings. Let $P(y|i, n)$ be the uniform distribution over the first $\exp(n + 1 - 6(i + 1) \log n)$ strings that remain in the lexicographically ordered list. Note that we have $i < n / (6 \log n)$, thus this list is never empty, and hence contains $x$. By universality there exist a $c$ such that for all strings $y$
\[
c \cdot m(y) \geq P(y|i, n)/(i^2 n^2).
\]
Now we construct the function $g(y, t)$. From $i$ and $n$ we can compute the largest $i$ such that $t_i \leq i$ if such $i$ exist (for $t < t_i$ the value of $g(y, t)$ may be anything). From $t_i$ and $i$ we can compute the lexicographically ordered list at stage $t_i$ and evaluate $P(z|i, n)$ for all $z$ in the list. Then we wait for the stage until the equation above is satisfied for all $z$ in the list. Let this stage be $g(y, t)$. This implies
\[
c \cdot m_{g(x,t)}(x) \geq P(x|i, n)/(i^2 n^2) \geq \exp(-n - 1 + (6i + 1) \log n) / n^4,
\]
and this implies (2).
\[\square\]

The next corollary follows from the proof above.

**Corollary 7.** There exist no test that is universal in the set of upper semicomputable sumtests for $m$.

**Proof.** We show that for every test $e$, we can construct $f$, $g$ and infinitely many $x$, such that $e(x) \leq O(1)$ and $e_{f,g}(x) = |x|/(\log |x|)^9$. Remind the construction of $u_h$ and by Theorem 1 there exist a computable $h$ be such that $e(x) \leq O(u_h(x))$. Let $m(t) \cdot 2^{17}$ be an approximation of $m(\cdot)$ from below such that $m(x) \geq \Omega (2^{-|x|^1})$. Now, for every $n$, we follow the construction of $x$ from the proof above with the following modification: we do not start from a list of all $x$ of length $n$, but from all $x$ of length $n$ such that
\[
m_{h(x,1)}(x|1) \leq 2^{-n+1}.
\]
There are less than $2^{n-1}$ strings with $m(x|1) > 2^{n-1}$, thus the list contains at least $2^n - 2^{n-1} = 2^{n-1}$ strings, and this is sufficient for the proof above. Let $t = 1$ in the definition of $u_h$. This implies
\[
u_h(x) \leq \frac{m_{h(x,1)}(x|1)}{m_1(x)} \leq \frac{2^{-n+1}}{\Omega (2^{-n})} \leq O(1),
\]
thus $e(x) \leq O(1)$. On the other hand, we can follow the proof above to construct $g$ and $x$ such that $e_{f,g}(x) = n/(\log n)^8$. (The function $f$ is still obtained from Lemma 4. On the other hand, the function $g$ might be larger than in the proof above, and depends on $h$. Indeed, it equals the stage on which $m(x)$ increases sufficiently above $2^{-n+1}$, and this time is related to $h(x,1)$.) \[\square\]
References

[1] B. Bauwens. Computability in statistical hypotheses testing, and characterizations of independence and directed influences in time series using Kolmogorov complexity. PhD thesis, Ugent, May 2010.

[2] B. Bauwens. m-sophistication. In Proceedings of the 6th Conference on Computability in Europe (CiE), July 2010.

[3] B. Bauwens and S. Terwijn. Notes on sum-tests and independence tests. Theory of Computing Systems, 48:247–268, 2011. open access.

[4] L. Bienvenu, A. Muchnik, A. Shen, and N. Vereshchagin. Limit complexities revisited. Theory of Computing Systems, 47(3):720–736, 2010.

[5] G.J. Chaitin. A theory of program size formally identical to information theory. J. Assoc. Comput. Mach., 22(3):329–340, 1975.

[6] P. Gács. Lecture notes on descriptional complexity and randomness. http://www.cs.bu.edu/faculty/gacs/papers/ait-notes.pdf, 1988–2011.

[7] M. Li and P.M.B. Vitányi. An Introduction to Kolmogorov Complexity and Its Applications. Springer-Verlag, New York, 2008.

[8] R.E. Walpole, R.H. Myers, S.L. Myers, and K.Ye. Probability & Statistics for Engineers & Scientists. Pearson Education, New Jersey, 2007.

[9] G A Young and R L Smith. Essentials of statistical inference. Cambridge series in statistical and probabilistic mathematics. Cambridge Univ. Press, Cambridge, 2005.