Algorithmic randomness test for a class of measures

Péter Gács

Computer Science Department
Boston University

Coverfest

Martin-Löf's theory of randomness

(As presented by Levin). Let X be the space Σ^* of finite strings, or the space $\Sigma^{\mathbb{N}}$ of infinite strings. Let μ be a probability measure over X. A test

$$f_\mu(x)$$

quantifies the nonrandomness of outcome $x \in X$ with respect to μ. In Martin-Löf's theory, measure μ is assumed to be "computable" and fixed. Required:

- $\int f_\mu(x)\mu(dx) \leqslant 1$. (The measure of "non-random" objects is small.)
- f is lower semicomputable in x. (Sooner or later we will recognize non-randomness.)

Test t is universal if $\forall f\ \exists c > 0\ \forall x\ f_\mu(x) < c \cdot t_\mu(x)$.

Theorem

There is a universal test $\mathbf{t}_\mu(x)$.

Martin-Löf's theory of randomness

(As presented by Levin). Let X be the space Σ^* of finite strings, or the space $\Sigma^{\mathbb{N}}$ of infinite strings. Let μ be a probability measure over X. A test

$$f_\mu(x)$$

quantifies the nonrandomness of outcome $x \in X$ with respect to μ. In Martin-Löf's theory, measure μ is assumed to be "computable" and fixed. Required:

- $\int f_\mu(x)\mu(dx) \leqslant 1$. (The measure of "non-random" objects is small.)
- f is lower semicomputable in x. (Sooner or later we will recognize non-randomness.)

Test t is universal if $\forall f\ \exists c > 0\ \forall x\ f_\mu(x) < c \cdot t_\mu(x)$.

Theorem

There is a universal test $\mathbf{t}_\mu(x)$.

Martin-Löf's theory of randomness

(As presented by Levin). Let X be the space Σ^* of finite strings, or the space $\Sigma^{\mathbb{N}}$ of infinite strings. Let μ be a probability measure over X. A test

$$f_\mu(x)$$

quantifies the nonrandomness of outcome $x \in X$ with respect to μ. In Martin-Löf's theory, measure μ is assumed to be "computable" and fixed. Required:

- $\int f_\mu(x)\mu(dx) \leqslant 1$. (The measure of "non-random" objects is small.)
- f is lower semicomputable in x. (Sooner or later we will recognize non-randomness.)

Test t is universal if $\forall f\ \exists c > 0\ \forall x\ f_\mu(x) < c \cdot t_\mu(x)$.

Theorem

There is a universal test $\tilde{t}_\mu(x)$.

Martin-Löf's theory of randomness

(As presented by Levin). Let X be the space Σ^* of finite strings, or the space $\Sigma^{\mathbb{N}}$ of infinite strings. Let μ be a probability measure over X. A test

$$f_\mu(x)$$

quantifies the nonrandomness of outcome $x \in X$ with respect to μ. In Martin-Löf's theory, measure μ is assumed to be "computable" and fixed. Required:

- $\int f_\mu(x)\mu(dx) \leqslant 1$. (The measure of "non-random" objects is small.)
- f is lower semicomputable in x. (Sooner or later we will recognize non-randomness.)

Test t is universal if $\forall f\ \exists c > 0\ \forall x\ f_\mu(x) < c \cdot t_\mu(x)$.

Theorem

There is a universal test $\mathbf{t}_\mu(x)$.

Martin-Löf's theory of randomness

(As presented by Levin). Let X be the space Σ^* of finite strings, or the space $\Sigma^{\mathbb{N}}$ of infinite strings. Let μ be a probability measure over X. A test

$$f_\mu(x)$$

quantifies the nonrandomness of outcome $x \in X$ with respect to μ. In Martin-Löf's theory, measure μ is assumed to be "computable" and fixed. Required:

- $\int f_\mu(x)\mu(dx) \leqslant 1$. (The measure of "non-random" objects is small.)
- f is lower semicomputable in x. (Sooner or later we will recognize non-randomness.)

Test t is universal if $\forall f\ \exists c > 0\ \forall x\ f_\mu(x) < c \cdot t_\mu(x)$.

Theorem

There is a universal test $\mathbf{t}_\mu(x)$.

Martin-Löf's theory of randomness

(As presented by Levin). Let X be the space Σ^* of finite strings, or the space $\Sigma^{\mathbb{N}}$ of infinite strings. Let μ be a probability measure over X. A test

$$f_\mu(x)$$

quantifies the nonrandomness of outcome $x \in X$ with respect to μ. In Martin-Löf's theory, measure μ is assumed to be "computable" and fixed. Required:

- $\int f_\mu(x)\mu(dx) \leqslant 1$. (The measure of "non-random" objects is small.)
- f is lower semicomputable in x. (Sooner or later we will recognize non-randomness.)

Test t is universal if $\forall f\ \exists c > 0\ \forall x\ f_\mu(x) < c \cdot t_\mu(x)$.

Theorem

There is a universal test $\tilde{\mathbf{t}}_\mu(x)$.

Martin-Löf's theory of randomness

(As presented by Levin). Let X be the space Σ^* of finite strings, or the space $\Sigma^{\mathbb{N}}$ of infinite strings. Let μ be a probability measure over X. A test

$$f_\mu(x)$$

quantifies the nonrandomness of outcome $x \in X$ with respect to μ. In Martin-Löf's theory, measure μ is assumed to be "computable" and fixed. Required:

- $\int f_\mu(x)\mu(dx) \leqslant 1$. (The measure of "non-random" objects is small.)
- f is lower semicomputable in x. (Sooner or later we will recognize non-randomness.)

Test t is universal if $\forall f\ \exists c > 0\ \forall x\, f_\mu(x) < c \cdot t_\mu(x)$.

Theorem

There is a universal test $\tilde{\mathbf{t}}_\mu(x)$.

Test in terms of complexity

I assume familiarity with description (Kolmogorov) complexity. Let $X = \Sigma^*$. For $x \in X$, denote the complexity (the prefix version) of x by

$$H(x)$$

(same as $K(x)$ in Li-Vitányi). Let $\tilde{\mathbf{d}}_\mu(x) = \log \tilde{\mathbf{t}}_\mu(x)$, called the deficiency of randomness of x with respect to μ.

Theorem

The following holds, for constants c_μ: Over the set of finite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} -\log \mu(x) - H(x) + c_\mu.$$

Over the set of infinite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} \sup_n -\log \mu(x_{\le n}) - H(x_{\le n}) + c_\mu.$$

Test in terms of complexity

I assume familiarity with description (Kolmogorov) complexity. Let $X = \Sigma^*$. For $x \in X$, denote the complexity (the prefix version) of x by

$$H(x)$$

(same as $K(x)$ in Li-Vitányi). Let $\tilde{\mathbf{d}}_\mu(x) = \log \tilde{\mathbf{t}}_\mu(x)$, called the deficiency of randomness of x with respect to μ.

Theorem

The following holds, for constants c_μ: Over the set of finite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} -\log \mu(x) - H(x) + c_\mu.$$

Over the set of infinite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} \sup_n -\log \mu(x_{\le n}) - H(x_{\le n}) + c_\mu.$$

Test in terms of complexity

I assume familiarity with description (Kolmogorov) complexity. Let $X = \Sigma^*$. For $x \in X$, denote the complexity (the prefix version) of x by

$$H(x)$$

(same as $K(x)$ in Li-Vitányi). Let $\tilde{\mathbf{d}}_\mu(x) = \log \tilde{\mathbf{t}}_\mu(x)$, called the deficiency of randomness of x with respect to μ.

Theorem

The following holds, for constants c_μ: Over the set of finite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} -\log \mu(x) - H(x) + c_\mu.$$

Over the set of infinite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} \sup_n -\log \mu(x_{\le n}) - H(x_{\le n}) + c_\mu.$$

Test in terms of complexity

I assume familiarity with description (Kolmogorov) complexity. Let $X = \Sigma^*$. For $x \in X$, denote the complexity (the prefix version) of x by

$$H(x)$$

(same as $K(x)$ in Li-Vitányi). Let $\tilde{\mathbf{d}}_\mu(x) = \log \tilde{\mathbf{t}}_\mu(x)$, called the deficiency of randomness of x with respect to μ.

Theorem

The following holds, for constants c_μ: Over the set of finite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} -\log \mu(x) - H(x) + c_\mu.$$

Over the set of infinite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} \sup_n -\log \mu(x_{\le n}) - H(x_{\le n}) + c_\mu.$$

Test in terms of complexity

I assume familiarity with description (Kolmogorov) complexity. Let $X = \Sigma^*$. For $x \in X$, denote the complexity (the prefix version) of x by

$$H(x)$$

(same as $K(x)$ in Li-Vitányi). Let $\tilde{\mathbf{d}}_\mu(x) = \log \tilde{\mathbf{t}}_\mu(x)$, called the deficiency of randomness of x with respect to μ.

Theorem

The following holds, for constants c_μ: Over the set of finite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} -\log \mu(x) - H(x) + c_\mu.$$

Over the set of infinite strings,

$$\tilde{\mathbf{d}}_\mu(x) \stackrel{+}{=} \sup_n -\log \mu(x_{\leqslant n}) - H(x_{\leqslant n}) + c_\mu.$$

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").

New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").

New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").

New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").
New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").

New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").

New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").

New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Arbitrary measures

Restriction to computable measures is unnatural (it is particularly baffling to probabilists). How to extend the definition to arbitrary measures? Idea: just use (over $X = \Sigma^*$):

$$-\log \mu(x) - H(x).$$

Alas, this test has undesireable properties (does not "conserve randomness").
New idea (following early work of Levin): test $f_\mu(x)$:

- $\int f_\mu(x)\mu(dx) \leqslant 1$.
- f is lower semicomputable in the pair (μ, x).

What does this mean? If we mean that μ is defined by an infinite string S with $f_S(x)$, (lower semi)computable from (S, x) then different descriptions of the same S may give different tests.

Other idea: equip the space of measures with a computability structure, so that one can talk about (lower semi)computability in μ itself, independent of its the particular description. In other words, the dependence on μ must be extensional. Levin has done this for infinite binary sequences.

This approach is attractive, but we also leads to some unexpected results (neutral measure).

Other idea: equip the space of measures with a computability structure, so that one can talk about (lower semi)computability in μ itself, independent of its the particular description. In other words, the dependence on μ must be extensional. Levin has done this for infinite binary sequences.
This approach is attractive, but we also leads to some unexpected results (neutral measure).

Constructive topology

Computability extended: instead of only about random strings, to speak of random real numbers, even about a random path of the Brownian motion (non-compact space). (For the special case of Brownian motion the concept has been worked out already by Asarin.)

Here, I will only work with metric spaces.

Computable metric space

$\mathbf{X} = (X, d, D, \alpha)$.

- A distance function d over X.
- A fixed countable dense set $D \subseteq X$ (so, $\mathbf{X}$ is separable).
- An enumeration α of D.

Condition: $d(x, y)$ is computable for $x, y \in D$.

Computable metric space

$\mathbf{X} = (X, d, D, \alpha)$.

- A distance function d over X.
- A fixed countable dense set $D \subseteq X$ (so, $\mathbf{X}$ is separable).
- An enumeration α of D.

Condition: $d(x, y)$ is computable for $x, y \in D$.

Computable metric space

$\mathbf{X} = (X, d, D, \alpha)$.

- A distance function d over X.
- A fixed countable dense set $D \subseteq X$ (so, $\mathbf{X}$ is separable).
- An enumeration α of D.

Condition: $d(x,y)$ is computable for $x, y \in D$.

Computable metric space

$\mathbf{X} = (X, d, D, \alpha)$.

- A distance function d over X.
- A fixed countable dense set $D \subseteq X$ (so, $\mathbf{X}$ is separable).
- An enumeration α of D.

Condition: $d(x,y)$ is computable for $x,y \in D$.

Computable metric space

$\mathbf{X} = (X, d, D, \alpha)$.

- A distance function d over X.
- A fixed countable dense set $D \subseteq X$ (so, $\mathbf{X}$ is separable).
- An enumeration α of D.

Condition: $d(x,y)$ is computable for $x,y \in D$.

Computable metric space

$\mathbf{X} = (X, d, D, \alpha)$.

- A distance function d over X.
- A fixed countable dense set $D \subseteq X$ (so, $\mathbf{X}$ is separable).
- An enumeration α of D.

Condition: $d(x, y)$ is computable for $x, y \in D$.

Continuity and computability

Some concepts of topology, and their constructive versions:

Basis of open balls: balls with center in the dense set D and rational radius.

Open set: a union of basis elements.

R.e. open set: a union of a r.e. set of basis elements.

Continuity and computability

Some concepts of topology, and their constructive versions:

Basis of open balls: balls with center in the dense set D and rational radius.

Open set: a union of basis elements.

R.e. open set: a union of a r.e. set of basis elements.

Continuity and computability

Some concepts of topology, and their constructive versions:

Basis of open balls: balls with center in the dense set D and rational radius.

Open set: a union of basis elements.

R.e. open set: a union of a r.e. set of basis elements.

Continuity and computability

Some concepts of topology, and their constructive versions:

Basis of open balls: balls with center in the dense set D and rational radius.

Open set: a union of basis elements.

R.e. open set: a union of a r.e. set of basis elements.

Let $f : X \to Y$ between metric spaces.

Continuous: $f^{-1}(V)$ is open for all basis elements $V \subseteq Y$.

Computable: $f^{-1}(V)$ is r.e. open, uniformly in the enumerated basis elements V.

Lower semicomputable: a constructive version of "lower semicontinuity": the set

$$\{ (x, r) : f(x) > r \}$$

is a r.e. open subset of $X \times \mathbb{Q}$.

Computable point: $x \in \mathbf{X}$: if the constant function $0 \mapsto x$ is.

Effective compactness: If for every k one can compute a covering of X by basic balls of radius $\leqslant 1/k$.

Let $f : X \to Y$ between metric spaces.

Continuous: $f^{-1}(V)$ is open for all basis elements $V \subseteq Y$.

Computable: $f^{-1}(V)$ is r.e. open, uniformly in the enumerated basis elements V.

Lower semicomputable: a constructive version of "lower semicontinuity": the set

$$\{\, (x, r) : f(x) > r \,\}$$

is a r.e. open subset of $X \times \mathbb{Q}$.

Computable point: $x \in \mathbf{X}$: if the constant function $0 \mapsto x$ is.

Effective compactness: If for every k one can compute a covering of X by basic balls of radius $\leqslant 1/k$.

Let $f : X \to Y$ between metric spaces.

Continuous: $f^{-1}(V)$ is open for all basis elements $V \subseteq Y$.

Computable: $f^{-1}(V)$ is r.e. open, uniformly in the enumerated basis elements V.

Lower semicomputable: a constructive version of “lower semicontinuity”: the set

$$\{ (x, r) : f(x) > r \}$$

is a r.e. open subset of $X \times \mathbb{Q}$.

Computable point: $x \in \mathbf{X}$: if the constant function $0 \mapsto x$ is.

Effective compactness: If for every k one can compute a covering of X by basic balls of radius $\leqslant 1/k$.

Let $f : X \to Y$ between metric spaces.

Continuous: $f^{-1}(V)$ is open for all basis elements $V \subseteq Y$.

Computable: $f^{-1}(V)$ is r.e. open, uniformly in the enumerated basis elements V.

Lower semicomputable: a constructive version of "lower semicontinuity": the set

$$\{\, (x, r) : f(x) > r \,\}$$

is a r.e. open subset of $X \times \mathbb{Q}$.

Computable point: $x \in \mathbf{X}$: if the constant function $0 \mapsto x$ is.

Effective compactness: If for every k one can compute a covering of X by basic balls of radius $\leqslant 1/k$.

Let $f : X \to Y$ between metric spaces.

Continuous: $f^{-1}(V)$ is open for all basis elements $V \subseteq Y$.

Computable: $f^{-1}(V)$ is r.e. open, uniformly in the enumerated basis elements V.

Lower semicomputable: a constructive version of "lower semicontinuity": the set

$$\{\, (x, r) : f(x) > r \,\}$$

is a r.e. open subset of $X \times \mathbb{Q}$.

Computable point: $x \in \mathbf{X}$: if the constant function $0 \mapsto x$ is.

Effective compactness: If for every k one can compute a covering of X by basic balls of radius $\leqslant 1/k$.

Let $f : X \to Y$ between metric spaces.

Continuous: $f^{-1}(V)$ is open for all basis elements $V \subseteq Y$.

Computable: $f^{-1}(V)$ is r.e. open, uniformly in the enumerated basis elements V.

Lower semicomputable: a constructive version of "lower semicontinuity": the set

$$\{ (x, r) : f(x) > r \}$$

is a r.e. open subset of $X \times \mathbb{Q}$.

Computable point: $x \in \mathbf{X}$: if the constant function $0 \mapsto x$ is.

Effective compactness: If for every k one can compute a covering of X by basic balls of radius $\leqslant 1/k$.

The following observation is useful, for a computable metric space X:

Proposition

Let $f :\subseteq X \to \mathbb{R}_+$ be a lower semicomputable function. Then it can be extended to a total lower semicomputable function $g : X \to \mathbb{R}_+$.

Topology of measures

We will always require $\mathbf{X}$ to be a complete computable metric space.

Weak convergence: $\mu_i \to \mu$ if $\mu_i f \to \mu f$ for all bounded continuous functions f. Can be metrized using, for example, the Prokhorov distance.

Dense set of measures: finite rational combinations of measures of form δ_x for $x \in D$.

This turns the set of probability measures into a computable complete metric space $\mathbf{M}(\mathbf{X})$.

Topology of measures

We will always require $\mathbf{X}$ to be a complete computable metric space.

Weak convergence: $\mu_i \to \mu$ if $\mu_i f \to \mu f$ for all bounded continuous functions f. Can be metrized using, for example, the Prokhorov distance.

Dense set of measures: finite rational combinations of measures of form δ_x for $x \in D$.

This turns the set of probability measures into a computable complete metric space $\mathbf{M}(\mathbf{X})$.

Topology of measures

We will always require **X** to be a complete computable metric space.

Weak convergence: $\mu_i \to \mu$ if $\mu_i f \to \mu f$ for all bounded continuous functions f. Can be metrized using, for example, the Prokhorov distance.

Dense set of measures: finite rational combinations of measures of form δ_x for $x \in D$.

This turns the set of probability measures into a computable complete metric space $\mathbf{M}(\mathbf{X})$.

Topology of measures

We will always require **X** to be a complete computable metric space.

Weak convergence: $\mu_i \to \mu$ if $\mu_i f \to \mu f$ for all bounded continuous functions f. Can be metrized using, for example, the Prokhorov distance.

Dense set of measures: finite rational combinations of measures of form δ_x for $x \in D$.

This turns the set of probability measures into a computable complete metric space $\mathbf{M}(\mathbf{X})$.

Topology of measures

We will always require $\mathbf{X}$ to be a complete computable metric space.

Weak convergence: $\mu_i \to \mu$ if $\mu_i f \to \mu f$ for all bounded continuous functions f. Can be metrized using, for example, the Prokhorov distance.

Dense set of measures: finite rational combinations of measures of form δ_x for $x \in D$.

This turns the set of probability measures into a computable complete metric space $\mathbf{M}(\mathbf{X})$.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_f such that for all x we have

$$f_\mu(x) < c_f \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Measure-dependent tests

Definition (μ-test)

A μ-test is a (possibly partial) function $f_\nu(x)$ with

- $f_\nu(x)$ is lower semicomputable in (ν, x).
- $\int f_\mu(x)\mu(dx) \leqslant 1$.

It is a uniform test if it is a ν-test for each ν.

Theorem (Hoyrup, Rojas)

There is a universal uniform test $\mathbf{t}_\mu(x)$: for all μ and each μ-test $f_\mu(x)$ there is a constant c_t such that for all x we have

$$f_\mu(x) < c_t \mathbf{t}_\mu(x).$$

Note that the universal uniform test $\mathbf{t}_\mu(x)$ dominates even the μ-tests for fixed μ.

Class tests

Another way to avoid the problem of uncomputable measures is to test whether an object is random with respect to any measure in a whole (natural) class $\mathscr{C}$. Say, whether there is a $0 < p < 1$ such that

$$x = x_1 x_2 \dots$$

is random with respect to the Bernoulli measure with parameter p.

Definition

$f(x)$ is a class test for class $\mathscr{C}$ of measures if

- It is lower semicomputable in x.
- $\int f(x)\mu(dx) \leqslant 1$ for all $\mu \in \mathscr{C}$.

Class tests

Another way to avoid the problem of uncomputable measures is to test whether an object is random with respect to any measure in a whole (natural) class $\mathscr{C}$. Say, whether there is a $0 < p < 1$ such that

$$x = x_1 x_2 \dots$$

is random with respect to the Bernoulli measure with parameter p.

Definition

$f(x)$ is a class test for class $\mathscr{C}$ of measures if

- It is lower semicomputable in x.
- $\int f(x)\mu(dx) \leqslant 1$ for all $\mu \in \mathscr{C}$.

Class tests

Another way to avoid the problem of uncomputable measures is to test whether an object is random with respect to any measure in a whole (natural) class $\mathscr{C}$. Say, whether there is a $0 < p < 1$ such that

$$x = x_1 x_2 \dots$$

is random with respect to the Bernoulli measure with parameter p.

Definition

$f(x)$ is a class test for class $\mathscr{C}$ of measures if

- It is lower semicomputable in x.
- $\int f(x)\mu(dx) \leqslant 1$ for all $\mu \in \mathscr{C}$.

Minimization

Given the universal uniform test $\mathbf{t}_\mu(x)$, there is a natural candidate for a class test:

$$\mathbf{t}_{\mathscr{C}}(x) = \min_{\mu \in \mathscr{C}} \mathbf{t}_\mu(x).$$

Theorem

Assume that the class $\mathscr{C}$ is effectively compact. Then $\mathbf{t}_{\mathscr{C}}(x)$ is a class test and it is universal (dominates all other class tests for $\mathscr{C}$).

Levin has noted the existence of class tests long ago. Examples of natural effectively compact classes: Bernoulli, m-state Markov, stationary.

Minimization

Given the universal uniform test $\mathbf{t}_\mu(x)$, there is a natural candidate for a class test:

$$\mathbf{t}_{\mathscr{C}}(x) = \min_{\mu \in \mathscr{C}} \mathbf{t}_\mu(x).$$

Theorem

Assume that the class $\mathscr{C}$ is effectively compact. Then $\mathbf{t}_{\mathscr{C}}(x)$ is a class test and it is universal (dominates all other class tests for $\mathscr{C}$).

Levin has noted the existence of class tests long ago. Examples of natural effectively compact classes: Bernoulli, m-state Markov, stationary.

Minimization

Given the universal uniform test $\mathbf{t}_\mu(x)$, there is a natural candidate for a class test:

$$\mathbf{t}_{\mathscr{C}}(x) = \min_{\mu \in \mathscr{C}} \mathbf{t}_\mu(x).$$

Theorem

Assume that the class $\mathscr{C}$ is effectively compact. Then $\mathbf{t}_{\mathscr{C}}(x)$ is a class test and it is universal (dominates all other class tests for $\mathscr{C}$).

Levin has noted the existence of class tests long ago. Examples of natural effectively compact classes: Bernoulli, m-state Markov, stationary.

Minimization

Given the universal uniform test $\mathbf{t}_\mu(x)$, there is a natural candidate for a class test:

$$\mathbf{t}_{\mathscr{C}}(x) = \min_{\mu \in \mathscr{C}} \mathbf{t}_\mu(x).$$

Theorem

Assume that the class $\mathscr{C}$ is effectively compact. Then $\mathbf{t}_{\mathscr{C}}(x)$ is a class test and it is universal (dominates all other class tests for $\mathscr{C}$).

Levin has noted the existence of class tests long ago. Examples of natural effectively compact classes: Bernoulli, m-state Markov, stationary.

Martin-Löf's approach

Martin-Löf also defined Bernoulli tests. We present them in the integral-constraint version. Denote

$$\mathbb{B}(n,k) = \{\, x \in \mathbb{B}^n : \sum_i x(i) = k \,\}.$$

Definition

A combinatorial Bernoulli test is a function $f : \mathbb{B}^* \to \overline{\mathbb{R}}_+$ with the following constraints:

- It is lower semicomputable.
- It is monotonic with respect to the prefix relation.
- For all $0 \le k \le n$ we have $\binom{n}{k}^{-1} \sum_{x \in \mathbb{B}(n,k)} f(x) \le 1$.

Martin-Löf's approach

Martin-Löf also defined Bernoulli tests. We present them in the integral-constraint version. Denote

$$\mathbb{B}(n,k) = \{\, x \in \mathbb{B}^n : \sum_i x(i) = k \,\}.$$

Definition

A combinatorial Bernoulli test is a function $f : \mathbb{B}^* \to \overline{\mathbb{R}}_+$ with the following constraints:

- It is lower semicomputable.
- It is monotonic with respect to the prefix relation.
- For all $0 \leqslant k \leqslant n$ we have $\binom{n}{k}^{-1} \sum_{x \in \mathbb{B}(n,k)} f(x) \leqslant 1$.

Martin-Löf's approach

Martin-Löf also defined Bernoulli tests. We present them in the integral-constraint version. Denote

$$\mathbb{B}(n,k) = \{ x \in \mathbb{B}^n : \sum_i x(i) = k \}.$$

Definition

A combinatorial Bernoulli test is a function $f : \mathbb{B}^* \to \overline{\mathbb{R}}_+$ with the following constraints:

- It is lower semicomputable.
- It is monotonic with respect to the prefix relation.
- For all $0 \leqslant k \leqslant n$ we have $\binom{n}{k}^{-1} \sum_{x \in \mathbb{B}(n,k)} f(x) \leqslant 1$.

Martin-Löf's approach

Martin-Löf also defined Bernoulli tests. We present them in the integral-constraint version. Denote

$$\mathbb{B}(n,k) = \{\, x \in \mathbb{B}^n : \sum_i x(i) = k \,\}.$$

Definition

A combinatorial Bernoulli test is a function $f : \mathbb{B}^* \to \overline{\mathbb{R}}_+$ with the following constraints:

- It is lower semicomputable.
- It is monotonic with respect to the prefix relation.
- For all $0 \leqslant k \leqslant n$ we have $\binom{n}{k}^{-1} \sum_{x \in \mathbb{B}(n,k)} f(x) \leqslant 1$.

Martin-Löf's approach

Martin-Löf also defined Bernoulli tests. We present them in the integral-constraint version. Denote

$$\mathbb{B}(n,k) = \{\, x \in \mathbb{B}^n : \sum_i x(i) = k \,\}.$$

Definition

A combinatorial Bernoulli test is a function $f : \mathbb{B}^* \to \overline{\mathbb{R}}_+$ with the following constraints:

- It is lower semicomputable.
- It is monotonic with respect to the prefix relation.
- For all $0 \leqslant k \leqslant n$ we have $\binom{n}{k}^{-1} \sum_{x \in \mathbb{B}(n,k)} f(x) \leqslant 1$.

Martin-Löf's approach

Martin-Löf also defined Bernoulli tests. We present them in the integral-constraint version. Denote

$$\mathbb{B}(n,k) = \{\, x \in \mathbb{B}^n : \sum_i x(i) = k \,\}.$$

Definition

A combinatorial Bernoulli test is a function $f : \mathbb{B}^* \to \overline{\mathbb{R}}_+$ with the following constraints:

- It is lower semicomputable.
- It is monotonic with respect to the prefix relation.
- For all $0 \leqslant k \leqslant n$ we have $\binom{n}{k}^{-1} \sum_{x \in \mathbb{B}(n,k)} f(x) \leqslant 1$.

Standard methods show:

Proposition

There is a universal combinatorial Bernoulli test.

Fix a universal combinatorial Bernoulli test $b(x)$ and extend it to infinite sequences ξ by $b(\xi) = \sup_n b(\xi^{\leqslant n})$.

Standard methods show:

Proposition

There is a universal combinatorial Bernoulli test.

Fix a universal combinatorial Bernoulli test $b(x)$ and extend it to infinite sequences ξ by $b(\xi) = \sup_n b(\xi^{\leqslant n})$.

Standard methods show:

Proposition

There is a universal combinatorial Bernoulli test.

Fix a universal combinatorial Bernoulli test $b(x)$ and extend it to infinite sequences ξ by $b(\xi) = \sup_n b(\xi^{\leqslant n})$.

Let $\mathbf{t}_{\mathscr{B}}(\xi)$ be a universal class test for Bernoulli measures, for infinite sequences.

Theorem

We have $b(\xi) \stackrel{*}{=} \mathbf{t}_{\mathscr{B}}(\xi)$.

In words: a sequence is nonrandom with respect to all Bernoulli measures if and only if it is rejected by a universal combinatorial Bernoulli test; moreover, even the degree of nonrandomness for random sequences, defined in the two ways, is the same.

Let $\mathbf{t}_{\mathscr{B}}(\xi)$ be a universal class test for Bernoulli measures, for infinite sequences.

Theorem

We have $b(\xi) \stackrel{*}{=} \mathbf{t}_{\mathscr{B}}(\xi)$.

In words: a sequence is nonrandom with respect to all Bernoulli measures if and only if it is rejected by a universal combinatorial Bernoulli test; moreover, even the degree of nonrandomness for random sequences, defined in the two ways, is the same.

Orthogonality

Bernoulli sequences

Let $0 < p < 1$, and $x = x_1 x_2 \dots$ an infinite sequence, with

$$S_n(x) = x_1 + \cdots + x_n.$$

Then x is Bernoulli random with respect to p if

a. x is random with respect to the class $\mathcal{B}$ of Bernoulli measures.
b. $S_n(x)/n \to p$.

Requirement b (as convergence in general) can be replaced with a stronger one, say:

- $|S_n(x)/n - p| < n^{-1/3}$ for all but finitely many n.

We want to generalize this observation.

Orthogonality

Bernoulli sequences

Let $0 < p < 1$, and $x = x_1 x_2 \dots$ an infinite sequence, with

$$S_n(x) = x_1 + \cdots + x_n.$$

Then x is Bernoulli random with respect to p if

a. x is random with respect to the class $\mathscr{B}$ of Bernoulli measures.
b. $S_n(x)/n \to p$.

Requirement b (as convergence in general) can be replaced with a stronger one, say:

- $|S_n(x)/n - p| < n^{-1/3}$ for all but finitely many n.

We want to generalize this observation.

Orthogonality

Bernoulli sequences

Let $0 < p < 1$, and $x = x_1 x_2 \dots$ an infinite sequence, with

$$S_n(x) = x_1 + \cdots + x_n.$$

Then x is Bernoulli random with respect to p if

- (a) x is random with respect to the class $\mathscr{B}$ of Bernoulli measures.
- (b) $S_n(x)/n \to p$.

Requirement (b) (as convergence in general) can be replaced with a stronger one, say:

- $|S_n(x)/n - p| < n^{-1/3}$ for all but finitely many n.

We want to generalize this observation.

Orthogonality
Bernoulli sequences

Let $0 < p < 1$, and $x = x_1 x_2 \ldots$ an infinite sequence, with

$$S_n(x) = x_1 + \cdots + x_n.$$

Then x is Bernoulli random with respect to p if

a. x is random with respect to the class $\mathscr{B}$ of Bernoulli measures.
b. $S_n(x)/n \to p$.

Requirement b (as convergence in general) can be replaced with a stronger one, say:

- $|S_n(x)/n - p| < n^{-1/3}$ for all but finitely many n.

We want to generalize this observation.

Orthogonality

Bernoulli sequences

Let $0 < p < 1$, and $x = x_1x_2\dots$ an infinite sequence, with

$$S_n(x) = x_1 + \cdots + x_n.$$

Then x is Bernoulli random with respect to p if

- (a) x is random with respect to the class $\mathscr{B}$ of Bernoulli measures.
- (b) $S_n(x)/n \to p$.

Requirement (b) (as convergence in general) can be replaced with a stronger one, say:

- $|S_n(x)/n - p| < n^{-1/3}$ for all but finitely many n.

We want to generalize this observation.

Separating test

Definition

For a class $\mathscr{C}$ of measures over a computable metric space $\mathbf{X} = (X, d, D, \alpha)$, a lower semicomputable function $s : X \times \mathscr{C} \to \overline{\mathbb{R}}_+$ is a separating test for a subclass $\mathscr{C}' \subseteq \mathscr{C}$ if

- $s_\mu(\cdot)$ is a test for each $\mu \in \mathscr{C}$.
- if μ or $\mu' \in \mathscr{C}'$ and $\mu \neq \nu$ then $s_\mu(x) \vee s_\nu(x) = \infty$ for all $x \in X$.

We call an element x typical for $\mu \in \mathscr{C}'$ if $s_\mu(x) < \infty$.

A typical element determines uniquely the measure μ for which it is typical.

Separating test

Definition

For a class $\mathscr{C}$ of measures over a computable metric space $\mathbf{X} = (X, d, D, \alpha)$, a lower semicomputable function $s : X \times \mathscr{C} \to \overline{\mathbb{R}}_+$ is a separating test for a subclass $\mathscr{C}' \subseteq \mathscr{C}$ if

- $s_\mu(\cdot)$ is a test for each $\mu \in \mathscr{C}$.
- if μ or $\mu' \in \mathscr{C}'$ and $\mu \neq \nu$ then $s_\mu(x) \vee s_\nu(x) = \infty$ for all $x \in X$.

We call an element x typical for $\mu \in \mathscr{C}'$ if $s_\mu(x) < \infty$.

A typical element determines uniquely the measure μ for which it is typical.

Separating test

Definition

For a class $\mathscr{C}$ of measures over a computable metric space $\mathbf{X} = (X, d, D, \alpha)$, a lower semicomputable function $s : X \times \mathscr{C} \to \overline{\mathbb{R}}_+$ is a separating test for a subclass $\mathscr{C}' \subseteq \mathscr{C}$ if

- $s_\mu(\cdot)$ is a test for each $\mu \in \mathscr{C}$.
- if μ or $\mu' \in \mathscr{C}'$ and $\mu \neq \nu$ then $s_\mu(x) \vee s_\nu(x) = \infty$ for all $x \in X$.

We call an element x typical for $\mu \in \mathscr{C}'$ if $s_\mu(x) < \infty$.

A typical element determines uniquely the measure μ for which it is typical.

Separating test

Definition

For a class $\mathscr{C}$ of measures over a computable metric space $\mathbf{X} = (X, d, D, \alpha)$, a lower semicomputable function $s : X \times \mathscr{C} \to \overline{\mathbb{R}}_+$ is a separating test for a subclass $\mathscr{C}' \subseteq \mathscr{C}$ if

- $s_\mu(\cdot)$ is a test for each $\mu \in \mathscr{C}$.
- if μ or $\mu' \in \mathscr{C}'$ and $\mu \neq \nu$ then $s_\mu(x) \vee s_\nu(x) = \infty$ for all $x \in X$.

We call an element x typical for $\mu \in \mathscr{C}'$ if $s_\mu(x) < \infty$.

A typical element determines uniquely the measure μ for which it is typical.

Separating test

Definition

For a class $\mathscr{C}$ of measures over a computable metric space $\mathbf{X} = (X, d, D, \alpha)$, a lower semicomputable function $s : X \times \mathscr{C} \to \overline{\mathbb{R}}_+$ is a separating test for a subclass $\mathscr{C}' \subseteq \mathscr{C}$ if

- $s_\mu(\cdot)$ is a test for each $\mu \in \mathscr{C}$.
- if μ or $\mu' \in \mathscr{C}'$ and $\mu \neq \nu$ then $s_\mu(x) \vee s_\nu(x) = \infty$ for all $x \in X$.

We call an element x typical for $\mu \in \mathscr{C}'$ if $s_\mu(x) < \infty$.

A typical element determines uniquely the measure μ for which it is typical.

Separating test

Definition

For a class $\mathscr{C}$ of measures over a computable metric space $\mathbf{X} = (X, d, D, \alpha)$, a lower semicomputable function $s : X \times \mathscr{C} \to \overline{\mathbb{R}}_+$ is a separating test for a subclass $\mathscr{C}' \subseteq \mathscr{C}$ if

- $s_\mu(\cdot)$ is a test for each $\mu \in \mathscr{C}$.
- if μ or $\mu' \in \mathscr{C}'$ and $\mu \neq \nu$ then $s_\mu(x) \vee s_\nu(x) = \infty$ for all $x \in X$.

We call an element x typical for $\mu \in \mathscr{C}'$ if $s_\mu(x) < \infty$.

A typical element determines uniquely the measure μ for which it is typical.

Separating test

Definition

For a class $\mathscr{C}$ of measures over a computable metric space $\mathbf{X} = (X, d, D, \alpha)$, a lower semicomputable function $s : X \times \mathscr{C} \to \overline{\mathbb{R}}_+$ is a separating test for a subclass $\mathscr{C}' \subseteq \mathscr{C}$ if

- $s_\mu(\cdot)$ is a test for each $\mu \in \mathscr{C}$.
- if μ or $\mu' \in \mathscr{C}'$ and $\mu \neq \nu$ then $s_\mu(x) \vee s_\nu(x) = \infty$ for all $x \in X$.

We call an element x typical for $\mu \in \mathscr{C}'$ if $s_\mu(x) < \infty$.

A typical element determines uniquely the measure μ for which it is typical.

Example

For a binary sequence x, and for $p \in [0,1]$ the function

$$s_p(x) = s_{B_p}(x) = c \cdot \sup\{\, k : |S_{2^k}(x) - 2^k p| > 2^{0.6k} \,\}$$

is a separation test for the Bernoulli class $\mathscr{B}$, for an appropriately chosen constant c.
In creating this test we exploited the existence of a computable convergence speed in the law of large numbers.

A generalization of the Bernoulli example:

Example

$\mathscr{C}$ is the class of m-state stationary Markov chains, $\mathscr{C}'$ the class of ergodic chains.

Example

For a binary sequence x, and for $p \in [0,1]$ the function

$$s_p(x) = s_{B_p}(x) = c \cdot \sup\{\, k : |S_{2^k}(x) - 2^k p| > 2^{0.6k} \,\}$$

is a separation test for the Bernoulli class $\mathscr{B}$, for an appropriately chosen constant c.

In creating this test we exploited the existence of a computable convergence speed in the law of large numbers.

A generalization of the Bernoulli example:

Example

$\mathscr{C}$ is the class of m-state stationary Markov chains, $\mathscr{C}'$ the class of ergodic chains.

When a separating test exists, it helps structuring randomness tests:

Theorem

Let $\mathscr{C}$ be an effectively compact class of measures, let $\mathbf{t}_\mu(x)$ be the universal uniform test and let $\mathbf{t}_{\mathscr{C}}(x)$ be a universal class test for $\mathscr{C}$. Assume that $s_\mu(x)$ is a separating test for $\mathscr{C}' \subseteq \mathscr{C}$. Then we have the representation

$$\mathbf{t}_\mu(x) \stackrel{*}{=} \mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$$

for all $\mu \in \mathscr{C}'$.

The theorem separates the randomness test into two parts. One part tests randomness with respect to the class $\mathscr{C}$, the other typicality with respect to the measure μ.

When a separating test exists, it helps structuring randomness tests:

Theorem

Let $\mathscr{C}$ be an effectively compact class of measures, let $\mathbf{t}_\mu(x)$ be the universal uniform test and let $\mathbf{t}_\mathscr{C}(x)$ be a universal class test for $\mathscr{C}$. Assume that $s_\mu(x)$ is a separating test for $\mathscr{C}' \subseteq \mathscr{C}$. Then we have the representation

$$\mathbf{t}_\mu(x) \stackrel{*}{=} \mathbf{t}_\mathscr{C}(x) \vee s_\mu(x)$$

for all $\mu \in \mathscr{C}'$.

The theorem separates the randomness test into two parts. One part tests randomness with respect to the class $\mathscr{C}$, the other typicality with respect to the measure μ.

When a separating test exists, it helps structuring randomness tests:

Theorem

Let $\mathscr{C}$ be an effectively compact class of measures, let $\mathbf{t}_\mu(x)$ be the universal uniform test and let $\mathbf{t}_{\mathscr{C}}(x)$ be a universal class test for $\mathscr{C}$. Assume that $s_\mu(x)$ is a separating test for $\mathscr{C}' \subseteq \mathscr{C}$. Then we have the representation

$$\mathbf{t}_\mu(x) \overset{*}{=} \mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$$

for all $\mu \in \mathscr{C}'$.

The theorem separates the randomness test into two parts. One part tests randomness with respect to the class $\mathscr{C}$, the other typicality with respect to the measure μ.

In the Bernoulli example,

- Part $\mathbf{t}_{\mathscr{B}}(x)$ checks "Bernoulliness", that is independence. It encompasses all the irregularity criteria.
- Part $s_p(x)$ checks (crudely) for the law of large numbers: whether relative frequency converges (fast) to p.

If the independence of the sequence is taken for granted, we may assume that the class test is satisfied. What remains is typicality testing, similar to ordinary statistical parameter testing.

Remark

Separation is the only requirement of the test $s_\mu(x)$, otherwise, for example in the Bernoulli test case, no matter how crude the convergence criterion expressed by $s_\mu(x)$, the maximum $\mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$ is always (essentially) the same universal test.

In the Bernoulli example,

- Part $\mathbf{t}_{\mathscr{B}}(x)$ checks "Bernoulliness", that is independence. It encompasses all the irregularity criteria.
- Part $s_p(x)$ checks (crudely) for the law of large numbers: whether relative frequency converges (fast) to p.

If the independence of the sequence is taken for granted, we may assume that the class test is satisfied. What remains is typicality testing, similar to ordinary statistical parameter testing.

Remark

Separation is the only requirement of the test $s_\mu(x)$, otherwise, for example in the Bernoulli test case, no matter how crude the convergence criterion expressed by $s_\mu(x)$, the maximum $\mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$ is always (essentially) the same universal test.

In the Bernoulli example,

- Part $\mathbf{t}_{\mathscr{B}}(x)$ checks “Bernoulliness”, that is independence. It encompasses all the irregularity criteria.
- Part $s_p(x)$ checks (crudely) for the law of large numbers: whether relative frequency converges (fast) to p.

If the independence of the sequence is taken for granted, we may assume that the class test is satisfied. What remains is typicality testing, similar to ordinary statistical parameter testing.

Remark

Separation is the only requirement of the test $s_\mu(x)$, otherwise, for example in the Bernoulli test case, no matter how crude the convergence criterion expressed by $s_\mu(x)$, the maximum $\mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$ is always (essentially) the same universal test.

In the Bernoulli example,

- Part $\mathbf{t}_{\mathscr{B}}(x)$ checks "Bernoulliness", that is independence. It encompasses all the irregularity criteria.
- Part $s_p(x)$ checks (crudely) for the law of large numbers: whether relative frequency converges (fast) to p.

If the independence of the sequence is taken for granted, we may assume that the class test is satisfied. What remains is typicality testing, similar to ordinary statistical parameter testing.

Remark

Separation is the only requirement of the test $s_\mu(x)$, otherwise, for example in the Bernoulli test case, no matter how crude the convergence criterion expressed by $s_\mu(x)$, the maximum $\mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$ is always (essentially) the same universal test.

In the Bernoulli example,

- Part $\mathbf{t}_{\mathscr{B}}(x)$ checks "Bernoulliness", that is independence. It encompasses all the irregularity criteria.
- Part $s_p(x)$ checks (crudely) for the law of large numbers: whether relative frequency converges (fast) to p.

If the independence of the sequence is taken for granted, we may assume that the class test is satisfied. What remains is typicality testing, similar to ordinary statistical parameter testing.

Remark

Separation is the only requirement of the test $s_\mu(x)$, otherwise, for example in the Bernoulli test case, no matter how crude the convergence criterion expressed by $s_\mu(x)$, the maximum $\mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$ is always (essentially) the same universal test.

In the Bernoulli example,

- Part $\mathbf{t}_{\mathscr{B}}(x)$ checks "Bernoulliness", that is independence. It encompasses all the irregularity criteria.
- Part $s_p(x)$ checks (crudely) for the law of large numbers: whether relative frequency converges (fast) to p.

If the independence of the sequence is taken for granted, we may assume that the class test is satisfied. What remains is typicality testing, similar to ordinary statistical parameter testing.

Remark

Separation is the only requirement of the test $s_\mu(x)$, otherwise, for example in the Bernoulli test case, no matter how crude the convergence criterion expressed by $s_\mu(x)$, the maximum $\mathbf{t}_{\mathscr{C}}(x) \vee s_\mu(x)$ is always (essentially) the same universal test.

Stationary processes

Much more complicated case: arbitrary stationary processes (say 0-1 valued).

The separating test that we gave for the Bernoulli measures (and can be given for the above simple Markov chains) makes use of a known speed of convergence in the law of large numbers.

V'yugin proved that for arbitrary stationary processes, no recursive speed of convergence can be guaranteed in the Ergodic Theorem (which is the appropriate generalization of the law of large numbers).

Question

What can be said about the convergence speed for ergodic processes, at least in the countable-state Markov case (such is V'yugin's example).

Stationary processes

Much more complicated case: arbitrary stationary processes (say 0-1 valued).

The separating test that we gave for the Bernoulli measures (and can be given for the above simple Markov chains) makes use of a known speed of convergence in the law of large numbers.

V'yugin proved that for arbitrary stationary processes, no recursive speed of convergence can be guaranteed in the Ergodic Theorem (which is the appropriate generalization of the law of large numbers).

Question

What can be said about the convergence speed for ergodic processes, at least in the countable-state Markov case (such is V'yugin's example).

Stationary processes

Much more complicated case: arbitrary stationary processes (say 0-1 valued).

The separating test that we gave for the Bernoulli measures (and can be given for the above simple Markov chains) makes use of a known speed of convergence in the law of large numbers.

V'yugin proved that for arbitrary stationary processes, no recursive speed of convergence can be guaranteed in the Ergodic Theorem (which is the appropriate generalization of the law of large numbers).

Question

What can be said about the convergence speed for ergodic processes, at least in the countable-state Markov case (such is V'yugin's example).

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0,1]$, μ_1 uniform over $[0,1/2]$.
- μ_2 uniform over $[1/2,1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0, 1]$, μ_1 uniform over $[0, 1/2]$.
- μ_2 uniform over $[1/2, 1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0, 1]$, μ_1 uniform over $[0, 1/2]$.
- μ_2 uniform over $[1/2, 1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \lessdot c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0,1]$, μ_1 uniform over $[0,1/2]$,
- μ_2 uniform over $[1/2,1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \lessdot p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0,1]$, μ_1 uniform over $[0,1/2]$,
- μ_2 uniform over $[1/2,1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0,1]$, μ_1 uniform over $[0,1/2]$,
- μ_2 uniform over $[1/2,1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \lesssim c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0, 1]$, μ_1 uniform over $[0, 1/2]$,
- μ_2 uniform over $[1/2, 1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \lesssim p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0, 1]$, μ_1 uniform over $[0, 1/2]$,
- μ_2 uniform over $[1/2, 1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0,1]$, μ_1 uniform over $[0,1/2]$,
- μ_2 uniform over $[1/2,1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Desireable test properties

Randomness with respect to computable measures has certain—intuitively meaningful—monotonicity:

- $\mu \leqslant c\nu \Rightarrow$ if x is random with respect to μ it is random with respect to ν.

This property does not survive for the uniform test.

Example

Let

- μ_0 uniform over $[0, 1]$, μ_1 uniform over $[0, 1/2]$,
- μ_2 uniform over $[1/2, 1]$.

With $p < 1/2$ be random with respect to μ_0, let $\nu_1 = p\mu_1 + (1-p)\mu_2$. Then p is not random with respect to ν_1, though but $\mu_0 \leqslant p^{-1}\nu_1$.

There is still a question whether all the good properties can be preserved in an appropriate definition.

Neutral measure

If S is a sequence describing μ, and a test $t_S(x)$ is computed as a function of the sequence S (intensional, or "via oracle S" as opposed to our extensional definition) then one can compute from S an object x with $\mu(\{x\}) = 0$, so there is a nonrandom object for μ.

But if we require (as we did) tests to be intensional in μ then this is not true anymore:

Theorem (Levin's neutral measure)

If $\mathbf{X}$ is compact then there is a measure M with the property that for all x, $\mathbf{t}_M(x) \leqslant 1$.

Levin called this measure "apriori probability", I use the more neutral term neutral measure.

For our test definition, this measure will not have nice computability properties.

Neutral measure

If S is a sequence describing μ, and a test $t_S(x)$ is computed as a function of the sequence S (intensional, or "via oracle S" as opposed to our extensional definition) then one can compute from S an object x with $\mu(\{x\}) = 0$, so there is a nonrandom object for μ. But if we require (as we did) tests to be intensional in μ then this is not true anymore:

Theorem (Levin's neutral measure)

If $\mathbf{X}$ is compact then there is a measure M with the property that for all x, $\mathbf{t}_M(x) \leqslant 1$.

Levin called this measure "apriori probability", I use the more neutral term neutral measure.

For our test definition, this measure will not have nice computability properties.

Neutral measure

If S is a sequence describing μ, and a test $t_S(x)$ is computed as a function of the sequence S (intensional, or "via oracle S" as opposed to our extensional definition) then one can compute from S an object x with $\mu(\{x\}) = 0$, so there is a nonrandom object for μ. But if we require (as we did) tests to be intensional in μ then this is not true anymore:

Theorem (Levin's neutral measure)

If $\mathbf{X}$ is compact then there is a measure M with the property that for all x, $\mathbf{t}_M(x) \leqslant 1$.

Levin called this measure "apriori probability", I use the more neutral term neutral measure.

For our test definition, this measure will not have nice computability properties.

Neutral measure

If S is a sequence describing μ, and a test $t_S(x)$ is computed as a function of the sequence S (intensional, or "via oracle S" as opposed to our extensional definition) then one can compute from S an object x with $\mu(\{x\}) = 0$, so there is a nonrandom object for μ. But if we require (as we did) tests to be intensional in μ then this is not true anymore:

Theorem (Levin's neutral measure)

If $\mathbf{X}$ is compact then there is a measure M with the property that for all x, $\mathbf{t}_M(x) \lesssim 1$.

Levin called this measure "apriori probability", I use the more neutral term neutral measure.

For our test definition, this measure will not have nice computability properties.

Neutral measure

If S is a sequence describing μ, and a test $t_S(x)$ is computed as a function of the sequence S (intensional, or "via oracle S" as opposed to our extensional definition) then one can compute from S an object x with $\mu(\{x\}) = 0$, so there is a nonrandom object for μ. But if we require (as we did) tests to be intensional in μ then this is not true anymore:

Theorem (Levin's neutral measure)

If $\mathbf{X}$ is compact then there is a measure M with the property that for all x, $\mathbf{t}_M(x) \leqslant 1$.

Levin called this measure "apriori probability", I use the more neutral term neutral measure.
For our test definition, this measure will not have nice computability properties.

Randomness conservation

Theorem

Let $f : X \to Y$ be computable. Then

$$\mathbf{t}_{f^*\mu}(f(x)) \overset{*}{<} \mathbf{t}_\mu(x).$$

There is a more general theorem, for computable random transitions.

Relation to complexity

The extensional nature of the test makes it hard to relate it to description complexity, in the case of noncomputable measures. Here, more research is needed.

Information

Relative algorithmic entropy

$$H_\nu(x) = -\log \mathbf{t}_\nu(x)$$

is a generalization of complexity (algorithmic entropy). Indeed, generalizing to non-probability measures ν (example: the counting measure $\#$)

$$H_\#(x) \stackrel{+}{=} H(x).$$

This is in analogy to the definition of relative (information-theoretical) entropy of μ with respect to ν,

$$\mathscr{H}_\nu(\mu) = -\int \log \frac{d\mu}{d\nu} d\mu,$$

(which is the negative of the so-called Kullback distance). Special cases: $\nu = \#$ gives ordinary entropy. For $\nu =$ Lebesgue measure gives $-\int f(x) \log f(x) dx$.

Addition theorem

Let us generalize the well-known addition theorem

$$H(x,y) \stackrel{+}{=} H(x) + H(y \mid x,\ H(x)).$$

Theorem (General Addition)

$$H_{\mu\times\nu}(x,y) \stackrel{+}{=} H_{\mu}(x \mid \nu) + H_{\nu}(y \mid x,\ H_{\mu}(x \mid \nu),\ \mu).$$

The proof is somewhat subtle.

Question

Applications?

Addition theorem

Let us generalize the well-known addition theorem

$$H(x,y) \stackrel{+}{=} H(x) + H(y \mid x,\ H(x)).$$

Theorem (General Addition)

$$H_{\mu\times\nu}(x,y) \stackrel{+}{=} H_{\mu}(x \mid \nu) + H_{\nu}(y \mid x,\ H_{\mu}(x \mid \nu),\ \mu).$$

The proof is somewhat subtle.

Question

Applications?