Randomness and normality

alexander.shen@lirmm.fr, www.lirmm.fr/~ashen

equipe ESCAPE,
LIRMM, CNRS & University of Montpellier.
Supported by ANR RaCAF project.
Joint work with Alexander Kosachinskiy (HSE, Moscow)

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Individual random objects

Would you believe that a fair coin produced these sequences?

- 0000000000000000000000000000000000000000
- 0101010101010101010101010101010101010101
- 0111001101110011011100011000111111100111
- 0010010000111111011010101000100010000101

All 40-bit sequences are equally random (have the same probability 2^{-40} in a fair coin model), but some look more random than others

Not only a paradox, but a real problem for statisticians: can we reject the null hypothesis of a fair coin looking at the experimental data?

Simple normality

Borel (1909, Les probabilités dénombrables et leurs applications arithmétiques):

> *Un nombre simplement normal est donc caractérisé par le fait que,* $c_0, c_1, c_2, \ldots, c_8, c_9$ *désignant les nombres respectifs de fois que figurent les chiffres* $0, 1, 2, \ldots, 8, 9$ *parmi les* n *premières décimales, chacun des rapports:*
>
> $$\frac{c_0}{n}, \frac{c_1}{n}, \ldots, \frac{c_8}{n}, \frac{c_9}{n}$$
>
> *a pour limite* $\frac{1}{10}$ *lorsque* n *augmente indéfiniment.*

"In the limit every digit appears equally often"

`00000`... is not simply normal, but `010101`... is

Simple normality

Borel (1909, Les probabilités dénombrables et leurs applications arithmétiques):

> *Un nombre simplement normal est donc caractérisé par le fait que, $c_0, c_1, c_2, \ldots, c_8, c_9$ désignant les nombres respectifs de fois que figurent les chiffres $0, 1, 2, \ldots, 8, 9$ parmi les n premières décimales, chacun des rapports:*
>
> $$\frac{c_0}{n}, \frac{c_1}{n}, \ldots, \frac{c_8}{n}, \frac{c_9}{n}$$
>
> *a pour limite $\frac{1}{10}$ lorsque n augmente indéfiniment.*

"In the limit every digit appears equally often"

`00000`... is not simply normal, but `010101`... is

Simple normality

Borel (1909, Les probabilités dénombrables et leurs applications arithmétiques):

> *Un nombre simplement normal est donc caractérisé par le fait que, $c_0, c_1, c_2, \ldots, c_8, c_9$ désignant les nombres respectifs de fois que figurent les chiffres $0, 1, 2, \ldots, 8, 9$ parmi les n premières décimales, chacun des rapports:*
>
> $$\frac{c_0}{n}, \frac{c_1}{n}, \ldots, \frac{c_8}{n}, \frac{c_9}{n}$$
>
> *a pour limite $\frac{1}{10}$ lorsque n augmente indéfiniment.*

"In the limit every digit appears equally often"

`00000`... is not simply normal, but `010101`... is

Simple normality

Borel (1909, Les probabilités dénombrables et leurs applications arithmétiques):

> *Un nombre simplement normal est donc caractérisé par le fait que, $c_0, c_1, c_2, \ldots, c_8, c_9$ désignant les nombres respectifs de fois que figurent les chiffres $0, 1, 2, \ldots, 8, 9$ parmi les n premières décimales, chacun des rapports:*
>
> $$\frac{c_0}{n}, \frac{c_1}{n}, \ldots, \frac{c_8}{n}, \frac{c_9}{n}$$
>
> *a pour limite $\frac{1}{10}$ lorsque n augmente indéfiniment.*

"In the limit every digit appears equally often"

`00000`... is not simply normal, but `010101`... is

Simple normality

Borel (1909, Les probabilités dénombrables et leurs applications arithmétiques):

> *Un nombre simplement normal est donc caractérisé par le fait que, $c_0, c_1, c_2, \ldots, c_8, c_9$ désignant les nombres respectifs de fois que figurent les chiffres $0, 1, 2, \ldots, 8, 9$ parmi les n premières décimales, chacun des rapports:*
>
> $$\frac{c_0}{n}, \frac{c_1}{n}, \ldots, \frac{c_8}{n}, \frac{c_9}{n}$$
>
> *a pour limite $\frac{1}{10}$ lorsque n augmente indéfiniment.*

“In the limit every digit appears equally often”

`00000`... is not simply normal, but `010101`... is

Normality

- Stronger condition: $00, 01, 10, 11$ appear equally often, and the same is true for k-bit blocks for $k = 3, 4, 5, \ldots$
- `01010101`... is not normal
- definition for infinite sequences only
- two versions of the definition:

01 11 00 11 01 11 00 11 01 00 01 aligned

0 1 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 1 0 0 non-aligned

- Equivalent if required for all k (Borel, without proof)
- Not so obvious proofs (1940s): Pillai, Niven, Zuckerman, Maxfield,...

Normality

- Stronger condition: $00, 01, 10, 11$ appear equally often, and the same is true for k-bit blocks for $k = 3, 4, 5, \ldots$
- 01010101... is not normal
- definition for infinite sequences only
- two versions of the definition:

01 11 00 11 01 11 00 11 01 00 01 aligned

0 1 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 1 0 0 non-aligned

- Equivalent if required for all k (Borel, without proof)
- Not so obvious proofs (1940s): Pillai, Niven, Zuckerman, Maxfield,...

Normality

- Stronger condition: 00, 01, 10, 11 appear equally often, and the same is true for k-bit blocks for $k = 3, 4, 5, \ldots$
- `01010101`... is not normal
- definition for infinite sequences only
- two versions of the definition:

01 11 00 11 01 11 00 11 01 00 01 aligned

0 1 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 1 0 0 non-aligned

- Equivalent if required for all k (Borel, without proof)
- Not so obvious proofs (1940s): Pillai, Niven, Zuckerman, Maxfield,...

Normality

- Stronger condition: $00, 01, 10, 11$ appear equally often, and the same is true for k-bit blocks for $k = 3, 4, 5, \ldots$
- `01010101...` is not normal
- definition for infinite sequences only
- two versions of the definition:

01 11 00 11 01 11 00 11 01 00 01 aligned

0 1 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 1 0 0 non-aligned

- Equivalent if required for all k (Borel, without proof)
- Not so obvious proofs (1940s): Pillai, Niven, Zuckerman, Maxfield,...

Normality

- Stronger condition: $00, 01, 10, 11$ appear equally often, and the same is true for k-bit blocks for $k = 3, 4, 5, \ldots$
- `01010101`... is not normal
- definition for infinite sequences only
- two versions of the definition:

01 11 00 11 01 11 00 11 01 00 01 aligned

0 1 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 1 0 0 non-aligned

- Equivalent if required for all k (Borel, without proof)
- Not so obvious proofs (1940s): Pillai, Niven, Zuckerman, Maxfield,...

Normality

- Stronger condition: 00, 01, 10, 11 appear equally often, and the same is true for k-bit blocks for $k = 3, 4, 5, \dots$
- `01010101`... is not normal
- definition for infinite sequences only
- two versions of the definition:

01 11 00 11 01 11 00 11 01 00 01 aligned

0 1 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 1 0 0 non-aligned

- Equivalent if required for all k (Borel, without proof)
- Not so obvious proofs (1940s): Pillai, Niven, Zuckerman, Maxfield,...

Normality

- Stronger condition: $00, 01, 10, 11$ appear equally often, and the same is true for k-bit blocks for $k = 3, 4, 5, \ldots$
- `01010101`... is not normal
- definition for infinite sequences only
- two versions of the definition:

01 11 00 11 01 11 00 11 01 00 01 aligned

0 1 1 1 0 0 1 1 0 1 1 1 0 0 1 1 0 1 0 0 non-aligned

- Equivalent if required for all k (Borel, without proof)
- Not so obvious proofs (1940s): Pillai, Niven, Zuckerman, Maxfield,...

Normality $\neq$ randomness

- `0 1 10 11 100 101 110 111 1000 1001`... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]00100100001111110110101010001000100000101\ldots$ conjectured to be normal
- = example #4
- normality is “weak randomness”
- what else?

Normality $\neq$ randomness

- 0 1 10 11 100 101 110 111 1000 1001... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]001001000011111101101010100010001000010 1\ldots$ conjectured to be normal
- = example #4
- normality is “weak randomness”
- what else?

Normality $\neq$ randomness

- `0 1 10 11 100 101 110 111 1000 1001`... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]0010010000111111011010101000100010000101\ldots$ conjectured to be normal
- = example #4
- normality is “weak randomness”
- what else?

Normality $\neq$ randomness

- 0 1 10 11 100 101 110 111 1000 1001... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]0010010000111111011010101000100010000101\ldots$ conjectured to be normal
- = example #4
- normality is "weak randomness"
- what else?

Normality ≠ randomness

- `0 1 10 11 100 101 110 111 1000 1001`... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]0010010000111111011010101000100010000101\ldots$ conjectured to be normal
- = example #4
- normality is "weak randomness"
- what else?

Normality $\neq$ randomness

- `0 1 10 11 100 101 110 111 1000 1001`... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]00100100001111110110101010001000100001 01\ldots$ conjectured to be normal
- = example #4
- normality is “weak randomness”
- what else?

Normality $\neq$ randomness

- `0 1 10 11 100 101 110 111 1000 1001`... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]0010010000111111011010101000100010000101\ldots$ conjectured to be normal
- = example #4
- normality is "weak randomness"
- what else?

Normality $\neq$ randomness

- 0 1 10 11 100 101 110 111 1000 1001... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]00100100001111110110101010001000100001 01\ldots$ conjectured to be normal
- = example #4
- normality is “weak randomness”
- what else?

Normality $\neq$ randomness

- 0 1 10 11 100 101 110 111 1000 1001... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]001001000011111101101010100010001000010 1\ldots$ conjectured to be normal
- = example #4
- normality is “weak randomness”
- what else?

Normality $\neq$ randomness

- 0 1 10 11 100 101 110 111 1000 1001... is normal (Champernowne)
- the same is true if we use only composite numbers (Champernowne)
- or prime numbers (Copeland, Erdös)
- or perfect squares (Besicovitch)
- $\pi = [11.]001001000011111101101010100010001000010 1\ldots$ conjectured to be normal
- = example #4
- normality is “weak randomness”
- what else?

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\dots000$ not random: short description: "n zeros"
- the same for $010101\dots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: “n zeros”
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- “compressed size” (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "*n* zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: "n zeros"
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- "compressed size" (no decompression)
- randomness: Kolmogorov complexity close to length

Randomness as incompressibility

- individual random sequences: plausible as outcomes of coin tossing
- (classical) probability theory: no idea
- Kolmogorov, Levin, Chaitin,...: randomness = incompressibility
- $000\ldots000$ not random: short description: “n zeros”
- the same for $010101\ldots$ and for π in binary
- ...and for all computable sequences
- algorithmic information theory: description of x = a program that produces x; random = no short descriptions
- Kolmogorov complexity: minimal length of a description
- “compressed size” (no decompression)
- randomness: Kolmogorov complexity close to length

Normality as weak incompressibility

- randomness $\Longleftrightarrow$ incompressibility
- normality: weak randomness
- normality: weak incompressibility (Agafonov, Schnorr, Becher, Heiber,...)

Theorem (informal)

A sequence is normal if and only if it is incompressible with finite memory

Proof (informal).

$\Leftarrow$ if different blocks appear with different frequencies, then Shannon–Fano code can be used; it uses finite memory

$\Rightarrow$ with finite memory decompression is local: N-bit blocks for large N are decompressed almost independently, and most of them are incompressible (and all blocks have the same frequency) □

Normality as weak incompressibility

- randomness $\Leftrightarrow$ incompressibility
- normality: weak randomness
- normality: weak incompressibility (Agafonov, Schnorr, Becher, Heiber,...)

Theorem (informal)

A sequence is normal if and only if it is incompressible with finite memory

Proof (informal).

$\Leftarrow$ if different blocks appear with different frequencies, then Shannon–Fano code can be used; it uses finite memory

$\Rightarrow$ with finite memory decompression is local: N-bit blocks for large N are decompressed almost independently, and most of them are incompressible (and all blocks have the same frequency) □

Normality as weak incompressibility

- randomness $\Leftrightarrow$ incompressibility
- normality: weak randomness
- normality: weak incompressibility (Agafonov, Schnorr, Becher, Heiber,...)

Theorem (informal)

A sequence is normal if and only if it is incompressible with finite memory

Proof (informal).

$\Leftarrow$ if different blocks appear with different frequencies, then Shannon–Fano code can be used; it uses finite memory

$\Rightarrow$ with finite memory decompression is local: N-bit blocks for large N are decompressed almost independently, and most of them are incompressible (and all blocks have the same frequency) □

Normality as weak incompressibility

- randomness $\Leftrightarrow$ incompressibility
- normality: weak randomness
- normality: weak incompressibility (Agafonov, Schnorr, Becher, Heiber,...)

Theorem (informal)

A sequence is normal if and only if it is incompressible with finite memory

Proof (informal).

$\Leftarrow$ if different blocks appear with different frequencies, then Shannon–Fano code can be used; it uses finite memory

$\Rightarrow$ with finite memory decompression is local: N-bit blocks for large N are decompressed almost independently, and most of them are incompressible (and all blocks have the same frequency) □

Normality as weak incompressibility

- randomness $\Longleftrightarrow$ incompressibility
- normality: weak randomness
- normality: weak incompressibility (Agafonov, Schnorr, Becher, Heiber,...)

Theorem (informal)

A sequence is normal if and only if it is incompressible with finite memory

Proof (informal).

$\Leftarrow$ if different blocks appear with different frequencies, then Shannon–Fano code can be used; it uses finite memory

$\Rightarrow$ with finite memory decompression is local: N-bit blocks for large N are decompressed almost independently, and most of them are incompressible (and all blocks have the same frequency) □

Normality as weak incompressibility

- randomness $\Leftrightarrow$ incompressibility
- normality: weak randomness
- normality: weak incompressibility (Agafonov, Schnorr, Becher, Heiber,...)

Theorem (informal)

A sequence is normal if and only if it is incompressible with finite memory

Proof (informal).

$\Leftarrow$ if different blocks appear with different frequencies, then Shannon–Fano code can be used; it uses finite memory
$\Rightarrow$ with finite memory decompression is local: N-bit blocks for large N are decompressed almost independently, and most of them are incompressible (and all blocks have the same frequency) □

Automatic complexity

- description mode: relation $D(x, y)$ on binary strings
- $D(p, x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| : D(p, x)\}$: complexity with respect to D
- not all modes are useful: if $D(x, y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0, 1, \varepsilon\} \times \{0, 1, \varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p, x) obtained in this way

Automatic complexity

- description mode: relation $D(x, y)$ on binary strings
- $D(p, x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| : D(p, x)\}$: complexity with respect to D
- not all modes are useful: if $D(x, y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0, 1, \varepsilon\} \times \{0, 1, \varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p, x) obtained in this way

Automatic complexity

- description mode: relation $D(x,y)$ on binary strings
- $D(p,x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| \colon D(p,x)\}$: complexity with respect to D
- not all modes are useful: if $D(x,y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0,1,\varepsilon\} \times \{0,1,\varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p,x) obtained in this way

Automatic complexity

- description mode: relation $D(x,y)$ on binary strings
- $D(p,x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| : D(p,x)\}$: complexity with respect to D
- not all modes are useful: if $D(x,y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0,1,\varepsilon\} \times \{0,1,\varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p,x) obtained in this way

Automatic complexity

- description mode: relation $D(x,y)$ on binary strings
- $D(p,x)$ reads “p is a description of x”
- $C_D(x) = \min\{|p| : D(p,x)\}$: complexity with respect to D
- not all modes are useful: if $D(x,y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0,1,\varepsilon\} \times \{0,1,\varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p,x) obtained in this way

Automatic complexity

- description mode: relation $D(x, y)$ on binary strings
- $D(p, x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| : D(p, x)\}$: complexity with respect to D
- not all modes are useful: if $D(x, y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0, 1, \varepsilon\} \times \{0, 1, \varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p, x) obtained in this way

Automatic complexity

- description mode: relation $D(x,y)$ on binary strings
- $D(p,x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| : D(p,x)\}$: complexity with respect to D
- not all modes are useful: if $D(x,y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0,1,\varepsilon\} \times \{0,1,\varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p,x) obtained in this way

Automatic complexity

- description mode: relation $D(x, y)$ on binary strings
- $D(p, x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| : D(p, x)\}$: complexity with respect to D
- not all modes are useful: if $D(x, y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0, 1, \varepsilon\} \times \{0, 1, \varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p, x) obtained in this way

Automatic complexity

- description mode: relation $D(x,y)$ on binary strings
- $D(p,x)$ reads "p is a description of x"
- $C_D(x) = \min\{|p| \colon D(p,x)\}$: complexity with respect to D
- not all modes are useful: if $D(x,y) \equiv \texttt{true}$, then $C_D(x) \equiv 0$

Automatic description mode:

- every p is a description of $O(1)$ strings
- the relation D corresponds to an automaton of a special type

directed graphs; edges are labeled by $\{0,1,\varepsilon\} \times \{0,1,\varepsilon\}$; going along a path, we collect first letters into p, second letters into x; D consists of all pairs (p,x) obtained in this way

Automatic complexity version

Theorem

An infinite binary sequence $a_1a_2\ldots a_n\ldots$ is normal if and only if for every automatic description mode D we have

$$\liminf \frac{C_D(a_1a_2\ldots a_n)}{n} \geqslant 1$$

"Normality: no way to compress significantly all prefixes of the sequence if only automatic description modes are allowed"

Automatic complexity version

Theorem

An infinite binary sequence $a_1 a_2 \ldots a_n \ldots$ is normal if and only if for every automatic description mode D we have

$$\liminf \frac{C_D(a_1 a_2 \ldots a_n)}{n} \geqslant 1$$

"Normality: no way to compress significantly all prefixes of the sequence if only automatic description modes are allowed"

Automatic complexity version

Theorem

An infinite binary sequence $a_1 a_2 \ldots a_n \ldots$ is normal if and only if for every automatic description mode D we have

$$\liminf \frac{C_D(a_1 a_2 \ldots a_n)}{n} \geqslant 1$$

"Normality: no way to compress significantly all prefixes of the sequence if only automatic description modes are allowed"

Local complexity version

A function $K(x)$ on strings with non-negative integer values is called a local complexity measure if

- $K(xy) \geqslant K(x) + K(y)$ [locality]
- the number of strings x such that $K(x) \leqslant n$ is $O(2^n)$ [calibration]

Theorem

An infinite binary sequence $a_1 a_2 \ldots a_n \ldots$ is normal if and only if for every local complexity function K we have

$$\liminf \frac{K(a_1 a_2 \ldots a_n)}{n} \geqslant 1$$

Local complexity version

A function $K(x)$ on strings with non-negative integer values is called a local complexity measure if

- $K(xy) \geqslant K(x) + K(y)$ [locality]
- the number of strings x such that $K(x) \leqslant n$ is $O(2^n)$ [calibration]

Theorem

An infinite binary sequence $a_1 a_2 \ldots a_n \ldots$ is normal if and only if for every local complexity function K we have

$$\liminf \frac{K(a_1 a_2 \ldots a_n)}{n} \geqslant 1$$

Local complexity version

A function $K(x)$ on strings with non-negative integer values is called a local complexity measure if

- $K(xy) \geqslant K(x) + K(y)$ [locality]
- the number of strings x such that $K(x) \leqslant n$ is $O(2^n)$ [calibration]

Theorem

An infinite binary sequence $a_1 a_2 \ldots a_n \ldots$ is normal if and only if for every local complexity function K we have

$$\liminf \frac{K(a_1 a_2 \ldots a_n)}{n} \geqslant 1$$

Local complexity version

A function $K(x)$ on strings with non-negative integer values is called a local complexity measure if

- $K(xy) \geqslant K(x) + K(y)$ [locality]
- the number of strings x such that $K(x) \leqslant n$ is $O(2^n)$ [calibration]

Theorem

An infinite binary sequence $a_1 a_2 \dots a_n \dots$ is normal if and only if for every local complexity function K we have

$$\liminf \frac{K(a_1 a_2 \dots a_n)}{n} \geqslant 1$$

Classical results as corollaries

- equivalence between non-aligned and aligned definitions of normality
- Wall: real number remains normal when multiplied/divided by an integer
- Champernowne, Copeland, Erdös, Besicovitch examples of normal numbers (sufficient conditions for the concatenation $B_1 B_2 \ldots$ of blocks: average block complexity close to average block length)
- Agafonov, Schnorr: finite state selection rule preserves normality
- Piatetski-Shapiro: if block frequencies are bounded by $O(2^{-k})$ for blocks of length k, then the sequence is normal

Classical results as corollaries

- equivalence between non-aligned and aligned definitions of normality
- Wall: real number remains normal when multiplied/divided by an integer
- Champernowne, Copeland, Erdös, Besicovitch examples of normal numbers (sufficient conditions for the concatenation $B_1 B_2 \ldots$ of blocks: average block complexity close to average block length)
- Agafonov, Schnorr: finite state selection rule preserves normality
- Piatetski-Shapiro: if block frequencies are bounded by $O(2^{-k})$ for blocks of length k, then the sequence is normal

Classical results as corollaries

- equivalence between non-aligned and aligned definitions of normality
- Wall: real number remains normal when multiplied/divided by an integer
- Champernowne, Copeland, Erdös, Besicovitch examples of normal numbers (sufficient conditions for the concatenation $B_1 B_2 \ldots$ of blocks: average block complexity close to average block length)
- Agafonov, Schnorr: finite state selection rule preserves normality
- Piatetski-Shapiro: if block frequencies are bounded by $O(2^{-k})$ for blocks of length k, then the sequence is normal

Classical results as corollaries

- equivalence between non-aligned and aligned definitions of normality
- Wall: real number remains normal when multiplied/divided by an integer
- Champernowne, Copeland, Erdös, Besicovitch examples of normal numbers (sufficient conditions for the concatenation $B_1 B_2 \ldots$ of blocks: average block complexity close to average block length)
- Agafonov, Schnorr: finite state selection rule preserves normality
- Piatetski-Shapiro: if block frequencies are bounded by $O(2^{-k})$ for blocks of length k, then the sequence is normal

Classical results as corollaries

- equivalence between non-aligned and aligned definitions of normality
- Wall: real number remains normal when multiplied/divided by an integer
- Champernowne, Copeland, Erdös, Besicovitch examples of normal numbers (sufficient conditions for the concatenation $B_1 B_2 \ldots$ of blocks: average block complexity close to average block length)
- Agafonov, Schnorr: finite state selection rule preserves normality
- Piatetski-Shapiro: if block frequencies are bounded by $O(2^{-k})$ for blocks of length k, then the sequence is normal

Classical results as corollaries

- equivalence between non-aligned and aligned definitions of normality
- Wall: real number remains normal when multiplied/divided by an integer
- Champernowne, Copeland, Erdös, Besicovitch examples of normal numbers (sufficient conditions for the concatenation $B_1B_2\ldots$ of blocks: average block complexity close to average block length)
- Agafonov, Schnorr: finite state selection rule preserves normality
- Piatetski-Shapiro: if block frequencies are bounded by $O\left(2^{-k}\right)$ for blocks of length k, then the sequence is normal

Finite state dimension

- Hausdorff dimension of a subset of $[0,1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \ldots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \ldots a_n)/n$

Finite state dimension

- Hausdorff dimension of a subset of $[0, 1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \ldots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \ldots a_n)/n$

Finite state dimension

- Hausdorff dimension of a subset of $[0,1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \ldots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \ldots a_n)/n$

Finite state dimension

- Hausdorff dimension of a subset of $[0, 1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \ldots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \ldots a_n)/n$

Finite state dimension

- Hausdorff dimension of a subset of $[0,1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \dots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \dots a_n)/n$

Finite state dimension

- Hausdorff dimension of a subset of $[0,1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \ldots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \ldots a_n)/n$

Finite state dimension

- Hausdorff dimension of a subset of $[0,1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \ldots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \ldots a_n)/n$

Finite state dimension

- Hausdorff dimension of a subset of $[0,1]$
- effective Hausdorff dimension
- is maximum of the dimension of individual points
- defined as $\liminf C(a_1 \ldots a_n)/n$
- finite state version of Hausdorff dimension (Dai, Lathrop, Lutz, Mayordomo)
- characterized as aligned/non-aligned limit entropy
- or $\inf_D \liminf C_D(a_1 \ldots a_n)/n$

More information and references

`https://arxiv.org/pdf/1701.09060.pdf`

(last version, 2019, see also FCT 2017 and 2019 papers)

Randomness discussed in a movie:

`https://www.youtube.com/embed/3YHHHEg3ioc?start=181&end=359`