How to tame randomness?

Closing lecture

Mateusz Topolewski

Zakopane 2024

Randomness – meaning

Randomness – meaning

Following Wikipedia...

Randomness refers to the apparent or actual lack of a definite pattern or predictability in information.

Random sequences

Generators

- Flipping a coin
- Rolling the dice
- Drawing from a box
- Electromagnetic noise
- Computer program
- ...

Is it random?

- 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1

Is it random?

- 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
- 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0

Is it random?

- 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
- 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0
- 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0

Is it random?

- 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
- 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0
- 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0
- 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1

Is it random?

- 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
- 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0
- 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0
- 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1
- 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0

Is it random?

- 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
- 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0
- 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0
- 0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1
- 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0
- 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0

How to recognize?

- Similar numbers of 0's and 1's

How to recognize?

- Similar numbers of 0's and 1's
  1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0
  1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0
  0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1

How to recognize?

- Similar numbers of 0's and 1's
- Similar numbers of 00's, 01's, 10's and 11's

How to recognize?

- Similar numbers of 0's and 1's
- Similar numbers of 00's, 01's, 10's and 11's
- Similar numbers of all 0-1 triples, quadruples, ...

How to recognize?

- Similar numbers of 0's and 1's
- Similar numbers of 00's, 01's, 10's and 11's
- Similar numbers of all 0-1 triples, quadruples, ...
- Series

Which is *more random*?

10111001110101110100010101011100110110001100101001
00111010000000111011100011011111100001101101110011
11010111000101111011100110010111011110011000000100
1111111001011001110000000101000100000111010011000001

0010010111000100111001001011100001111100100001 0001
01001001000100101001001001010010010010001001000100
01000111110011010100101010111001010101010100101001
01101001010100101010010100100101001010100101011001

Which is *more random*?

10111001110101110100010101011100110110001100101001
00111010000000111011100011011111100001101101110011
11010111000101111011100110010111011110011000000100
11111110010110011100000010100010000111010011000001

00100101110001001110010010111000011111001000010001
01001001000100101001001001010010010010001001000100
01000111110011010100101010111001010101010100101001
01101001010100101010010100100101001010100101011001

Long series

- Long series should appear quite often
- In a sequence of n coin flips, expected number of k heads is

$$\frac{n}{2^{k+2}}.$$

Large n

$n = 1000$		$n = 100\,000$	
series length	exp. frequency	series length	exp. frequency
1	125	1	12 500
2	62	2	6 250
3	31	3	3 125
4	15	4	1 562
5	8	5	781
6	4	6	390
7	2	7	195
8	1	8	97
9	0	9	49
10	0	10	24
11	0	11	12

Runs test

- (Abraham) Wald-(Jacob) Wolfowitz runs test – statistical test for randomness

Runs test

- (Abraham) Wald-(Jacob) Wolfowitz runs test – statistical test for randomness
- run = maximal series of equal elements
  $\mathbf{111}00\mathbf{11}0\mathbf{1}000000\mathbf{111} \rightarrow$ 7 runs

Runs test

- (Abraham) Wald-(Jacob) Wolfowitz runs test – statistical test for randomness
- run = maximal series of equal elements
  11100**11**0**1**000000**111** $\rightarrow$ 7 runs
- Number of runs should be *proper* (not too small, not too large)

Runs test

- (Abraham) Wald-(Jacob) Wolfowitz runs test – statistical test for randomness
- run = maximal series of equal elements
  11100**11**0**1**000000**111** → 7 runs
- Number of runs should be *proper* (not too small, not too large)
- Example no. 5 (primes) → 15 runs

Runs test

- (Abraham) Wald-(Jacob) Wolfowitz runs test – statistical test for randomness
- run = maximal series of equal elements
  11100**11**0**1**000000**111** → 7 runs
- Number of runs should be *proper* (not too small, not too large)
- Example no. 5 (primes) → 15 runs
- Example no. 6 → 14 runs

Example – iPod random playlist

Example – iPod random playlist

We changed the mechanism to a less random one to make it seem more random.

Steve Jobs

Statistical data

Consider any statistical dataset, e.g.

- financial data of some company,
- population of world countries,
- areas of deserts

and focus on the first digits of the numbers.

Afghanistan **6**47497
Albania **2**7399
Algeria **2**381730
Andorra **4**51
Angola **1**246693
...

Statistical data

Consider any statistical dataset, e.g.

- financial data of some company,
- population of world countries,
- areas of deserts

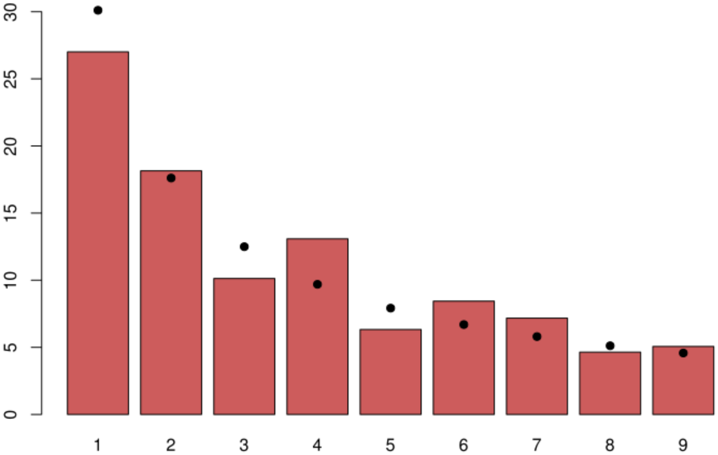
and focus on the first digits of the numbers.

Afghanistan **6**47497
Albania **2**7399
Algeria **2**381730
Andorra **4**51
Angola **1**246693
...

Question

Do each of the digits 1–9 appear in the first place equally often?

Population (July 2010)



Source: https://upload.wikimedia.org/wikipedia/commons/1/13/Benfords_law_illustrated_by _world%27s_countries_population.svg

Benford's law

A set of numbers is said to satisfy Benford's law if the leading digit d occurs with probability

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right).$$

Benford's law

A set of numbers is said to satisfy Benford's law if the leading digit d occurs with probability

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right).$$

d	1	2	3	4	5	6	7	8	9
$P(d)$ (in %)	30.1	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6

Benford's law

A set of numbers is said to satisfy Benford's law if the leading digit d occurs with probability

$$P(d) = \log_{10}\left(1 + \frac{1}{d}\right).$$

d	1	2	3	4	5	6	7	8	9
$P(d)$ (in %)	30.1	17.6	12.5	9.7	7.9	6.7	5.8	5.1	4.6

Most statistical data satisfy Benford's law!

Remember

Some things seem to be non-random, but they are.
Other things seem to be random, but they are not.

Mateusz Topolewski