The BFP (Benford-Fibonacci-Perez) method validates the consistency of COVID-19 epidemiological data in France and Italy

Jean-Claude Perez, PhD Maths § Computer Science Bordeaux University, RETIRED interdisciplinary researcher (IBM Emeritus, IBM European Research Center on Artificial Intelligence Montpellier), Bordeaux metropole, France, jeanclaudeperez2@gmail.com

Abstract

The Benford method can be used to detect manipulation of epidemiological or trial data during the validation of new drugs. We extend here the Benford method after having detected particular properties for the Fibonacci values 1, 2, 3, 5 and 8 of the first decimal of 10 runs of official epidemiological data published in France and Italy (positive cases, intensive care, and deaths) for the periods of March 1 to May 30, 2020 and 2021, each with 91 raw data. This new method – called “BFP” for Benford-Fibonacci-Perez - is positive in all 10 cases (i.e. 910 values) with an average of favorable cases close to 80%, which, in our opinion, would validate the reliability of these basic data.

Introduction

On the one hand, there is Benford's law (http://www.fusioninvesting.com/2009/11/benfords-law-and-fibonacci-numbers/) which stipulates that the majority of series of measurements more or less linked to natural or biological phenomena are confirmed, if they are now, to this law which is defined as follows:

In (http://www.fusioninvesting.com/2009/11/benfords-law-and-fibonacci-numbers/) we note:

« Benford’s law, also called the first-digit law, states that in lists of numbers from many real-life sources of data, the leading digit is distributed in a specific, non-uniform way. According to this law, the first digit is 1 almost one third of the time, and larger digits occur as the leading digit with lower and lower frequency, to the point where 9 as a first digit occurs less than one time in twenty.

This counter-intuitive result has been found to apply to a wide variety of data sets, like electricity bills, street addresses, stock prices, population numbers, death rates, lengths of rivers, physical and biological (which are very common in nature).

It is named after physicist Franck Benford, who stated it in 1938, although it had been previously stated by Simon Newcomb in 1881. Particularly, in epidemiology and health drugs trials, this law permits to validate accuracy and réalité of basic data ».

This law is used in various areas like stock exchange, social phenomena, epidemiology etc... (Sarkar, 2018).
This can therefore help detect fraud in scientific publications as well as unintentional errors in these datasets.

Often, we present the Fibonacci sequence as an example of a distribution obeying my Benford law fairly well.

On the other hand, there is, precisely, this Fibonacci law:

Well known in natural forms: nautilus spiral, sunflower flowers, pineapple, palm trees or pine cones, Fibonacci numbers also control the relative proportions of TCAG nucleotides in DNA: we had already demonstrated this 30 years ago (Perez, 1991), (Perez, 1997).

More recently, we have shown that these same Fibonacci proportions of the genome of the mitochondria, the energy source of the human cell, are deteriorated by mutations associated with various cancers (Perez, 2017). We also demonstrate how these same Fibonacci proportions of DNA make it possible to distinguish a genome of a real bacterium from its attempt at a synthetic chimera (Perez, 2019).

In the field of SARS-CoV2, its mRNA vaccines, and its multiple variants, we have demonstrated since the start of the COVID-19 pandemic how these Fibonacci numbers offered a new angle for the analysis of mRNA sequences and mutations of SARS-CoV2: a biomathematic point of view of the genome (Perez, 2020), (Perez & Montagnier, 2020), mRNA vaccines or variants (Perez, 2021a), or the last Indian variant "Delta" B.1.617.2 (Perez, 2021b).

The paradox which is at the source of our method:

On the one hand, Benford's law is often illustrated by its "good correlation" when applied to the Fibonacci sequence, which everyone knows is at the root of many forms of nature.

On the other hand, when we observe this same histogram, taken as proof of Benford's law by the primes, I note, on the contrary, that the (Fibonacci) numbers 1 2 3 5 and 8 differ in this histogram other numbers 4 6 7 and 9 (see Figures 2 § 3, and Table 1).

It is this observation which will be at the root of our method, then illustrated by this article.
Figure 2 - percentages of Benford's law over the first 200 Fibonacci numbers.

Table 1 - percentages of Benford's law over the first 500 Fibonacci numbers.

| d | % théorique | % observé |
|---|-------------|-----------|
| 1 | 30.100      | 30.130    |
| 2 | 17.600      | 17.560    |
| 3 | 12.490      | 12.570    |
| 4 | 09.691      | 09.381    |
| 5 | 07.918      | 07.984    |
| 6 | 06.694      | 06.586    |
| 7 | 05.799      | 05.788    |
| 8 | 05.115      | 05.389    |
| 9 | 04.575      | 04.391    |
Benford Law applied to the 500 first Fibonacci Numbers

Evidence of a specificity for Fibonacci digits 1 2 3 5 8

Figure 3 - percentages of Benford's law over the first 500 Fibonacci numbers.

What about the “BFP” method running on the firsts Fibonacci numbers?

Table 2 – 2 clusters partition of the 34 firsts Fibonacci numbers and BFP digits (Benford-Fibonacci-Perez).

| Fibonacci BFP digit | 1 1 |
|---------------------|-----|
| 1 1                 |     |
| 2 2                 |     |
| 3 3                 |     |
| 5 5                 |     |
| 8 8                 |     |
| 13 1                |     |
| 21 2                |     |
| 34 3                |     |
| 55 5                |     |
| 89 1                |     |
| 144 1               |     |
| 233 2               |     |
| 377 3               |     |
| 610 6               |     |
| 987 9               |     |
| 1597 1              |     |
| 2584 2              |     |
| 4181 4              |     |
It seems that our “BFP” law is all the more clear that the Fibonacci numbers are small here 27 on the first 34 = 79.41%.

Methods and datas

Fibonacci numbers:
0 1 1 2 3 5 8 13 21 34 55 89 144 233 377 610 987 1597 2584 4181 6765 10946 17711 28657 46368 75025 121393 196418 317811 514229 832040 1346269 2178309 3524578 5702887

For any whole number in the list, consider only its decimal with the highest weight decimal.

Example:
13 ==> 1
3398 ===> 3
4765 ===> 4

If the selected decimal digit belongs to fibonacci 1 2 3 5 8 do +1
Otherwise 4 6 7 9 0 do +0

We then calculate the% of positives / total.

Basic datas:

Main data sources from:
for France,
https://www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/
and for Italy:

https://www.sciencedirect.com/science/article/pii/S2352340920304200

Table 3 - Italy: from 1 March to 30 May 2020 and 2021.

| Positive cases | Death | Intensive care |
|----------------|-------|---------------|
| 2020 | 2021 | 2020 | 2021 | 2020 | 2021 |
| 566 | 13114 | 12 | 246 | 140 | 2289 |
| 342 | 17083 | 11 | 343 | 166 | 2327 |
| 466 | 20884 | 27 | 347 | 229 | 2411 |
| 587 | 22865 | 28 | 339 | 295 | 2475 |
| 769 | 24036 | 41 | 297 | 351 | 2525 |
| 778 | 23641 | 49 | 307 | 462 | 2571 |
| 1247 | 20765 | 36 | 207 | 567 | 2700 |
| 1492 | 13902 | 133 | 318 | 650 | 2700 |
| 1797 | 19749 | 97 | 376 | 733 | 2756 |
| 977 | 22409 | 168 | 332 | 877 | 2827 |
| 2313 | 25673 | 196 | 373 | 1028 | 2859 |
| 2651 | 26824 | 189 | 380 | 1153 | 2914 |
| 2547 | 26062 | 250 | 317 | 1328 | 2982 |
| 3497 | 21315 | 175 | 264 | 1518 | 3082 |
| 3590 | 15267 | 368 | 354 | 1672 | 3157 |
| 3233 | 20396 | 349 | 502 | 1851 | 3256 |
| 3526 | 23059 | 345 | 431 | 2060 | 3317 |
| 4207 | 24935 | 475 | 423 | 2257 | 3333 |
| 5322 | 25735 | 427 | 386 | 2498 | 3364 |
| 5986 | 23832 | 627 | 401 | 2655 | 3387 |
| 6557 | 20159 | 793 | 300 | 2857 | 3448 |
| 5560 | 13846 | 651 | 386 | 3009 | 3510 |
| 4789 | 18765 | 601 | 551 | 3204 | 3546 |
| 5249 | 21267 | 743 | 460 | 3390 | 3588 |
| 5210 | 23798 | 683 | 460 | 3489 | 3620 |
| 6203 | 23987 | 712 | 457 | 3612 | 3628 |
| 5909 | 23839 | 919 | 380 | 3732 | 3635 |
| 5974 | 19611 | 889 | 297 | 3856 | 3679 |
| 5217 | 12916 | 756 | 417 | 3906 | 3721 |
| 4050 | 16017 | 812 | 529 | 3981 | 3716 |
| 4053 | 23904 | 837 | 467 | 4023 | 3710 |
| 4782 | 23649 | 727 | 501 | 4035 | 3681 |
| 4668 | 21932 | 760 | 481 | 4053 | 3704 |
| 4585 | 21261 | 766 | 376 | 4068 | 3714 |
| 4805 | 18025 | 681 | 326 | 3994 | 3703 |
| 4316 | 10680 | 525 | 296 | 3977 | 3737 |
| 3599 | 7767 | 636 | 421 | 3898 | 3743 |
| 3039 | 13708 | 604 | 627 | 3792 | 3683 |
| 3836 | 17221 | 542 | 487 | 3693 | 3663 |
| 4204 | 18938 | 610 | 718 | 3605 | 3603 |
| 3951 | 17567 | 570 | 344 | 3497 | 3588 |
| 4694 | 15746 | 619 | 331 | 3381 | 3585 |
| 4092 | 9789 | 431 | 358 | 3343 | 3593 |
| 3153 | 13447 | 566 | 476 | 3260 | 3526 |
| 2972 | 16168 | 602 | 469 | 3186 | 3490 |
| 2667 | 16974 | 578 | 380 | 3079 | 3417 |
| 3786 | 15943 | 525 | 429 | 2936 | 3366 |
| 3493 | 15370 | 575 | 310 | 2812 | 3340 |
| 3491 | 12694 | 482 | 251 | 2733 | 3311 |
| 3047 | 8864 | 433 | 316 | 2635 | 3244 |
| 2256 | 10274 | 454 | 390 | 2573 | 3151 |
Table 4 - France: from 1 March to 30 May 2020 and 2021.

| Positive cases | Death |
|----------------|-------|
| 2020 2021 2020 2021 | 0 114 0 75 |
| 43 20412 0 114 |
| 23 20453 0 75 |
| 48 19786 1 410 |
| 34 21912 1 322 |
| 73 13157 0 278 |
| 138 2364 3 405 |
| 179 29327 2 196 |
| 103 23466 1 127 |
| 410 23706 9 358 |
| 286 23945 11 356 |
| 371 26255 3 264 |
| 497 17026 15 265 |
| 586 4135 13 290 |
Results and Discussion

BASIC RESULTS:

Here is the expected result on the public data covid19 in France and in Italy between March 1 and May 30 in 2020 and in 2021. Then, having the RELIABILITY of these basic data, we will illustrate an example of application: bravais correlations pearson in France (data smoothed over 7 sliding days) on time between positive test and death.

Synthetic results: Test "BFP" method to validate SARS-CoV2 epidemiologic data.

| Italy:                           |
|---------------------------------|
| Positive cases 2020: 65 / 91 = 71.4% |
| Death 2020: 53 / 91 = 58.2%      |
| Positive cases 2021: 80 / 91 = 87.9% |
| Death 2021: 71 / 91 = 78.02%    |
| Intensive care 2020: 73 / 91 = 80.2% |
| Intensive care 2021: 91 / 91 = 100% |

| France:                          |
|---------------------------------|
| Positive cases 2020: 63 / 91 = 69.2% |
| Death 2020: 83 / 91 = 91.2%       |
| Positive cases 2021: 65 / 91 = 71.4% |
| Death 2021: 81 / 91 = 89%         |
Average 725 / 910 = 79.67% for 10 batches with 91 cases each, then a total of 910 cases.

It seems that “BFP” law is all the more clear that the Fibonacci numbers are small here 27 on the first 34 = 79.41%.
We notice that everything is > in 2021 than in 2020.

2020:
64 + 53 + 73 + 63 + 65 = 318/455 = 69.89%

2021:
80 + 71 + 91 + 83 + 81 = 406/455 = 89.23%

How to explain?
It may be because the 2021 values are > the 2021 values.
So the method would prefer larger values?

COMPARING WITH RANDOM VALUES:

The results obtained here, that is to say nearly 80% success for 910 real values cumulating 10 races of 91 values each coming from epidemiological measurements in France and Italy, are they GREATER than what would be produced by CHANCE? To answer this question, we performed 100 random batches, each simulating 910 representative random values, for a total of 91,000 random tests. Here are the results:
While the number of successes of real cases is 725 favorable cases (first significant number = 1,2,3,5 or 8), the 100 batches produce an average number of successes of 667.96 with random values between 641 and 697 (Figure 4).
Figure 4 – Comparing the 910 real France§Italy results with 100 RANDOM RUNS, each simulating 910 random values between 1 and max value France§Italy (i.e. [1, 53843]).

EXAMPLE OF APPLICATION:

Table 5 – Comparing France Covid-19 March to May years 2020 and 2021, distances between positive case and death using Bravais-Pearson method on 7 days average splines values.

| TEST TO DEATH | 2020  | 2021  |
|---------------|-------|-------|
| 7days         | 7514  | 6625  |
| 8days         | 8063  | 7154  |
| 9days         | 8577  | 7647  |
| 10days        | 9067  | 7883  |
| 11days        | 9408  | 8042  |
| 12days        | 9658  | 8212  |
| 13days        | 9832  | 8364  |
| 14 DAYS       | 9836  | 8419  |
| 15days        | 9731  | 8458  |
| 16days        | 9530  | 8482  |
| 17days        | 9187  | 8506  |
| 18days        | 8795  | 8496  |
| 19days        | 8308  | 8422  |
| 20days        | 7708  | 8338  |
| 21days        | 7103  | 8195  |
Conclusions

Benford's law already makes it possible to validate or doubt the relevance, reliability and non-manipulation of batches of natural or medical data.
What we are proposing today is beyond this Benford law, it is a PARTITION of the first 9 digits (or 10 when, as here, there is also some null data) in 2 clusters: Fibonacci cluster (1 2 3 5 8) and non-Fibonacci cluster (0 4 6 7 9).
We suggest that the Fibonacci numbers cluster are all the more in the majority the more the data set is reliable and real.
This constitutes a breakthrough in the analysis of natural, social and medical data. This method and the prospects that it should now be consolidated and deepened.

Finally, we have demonstrated by 91,000 random values draws that the "BFP" law applied to the 910 COVID-19 epidemiological values of France and Italy studied here produces results which cannot result from mere chance.

Acknowledgements

My thanks:
To Dr Gérard Maudrux ([https://blog-gerard.maudrux.fr/](https://blog-gerard.maudrux.fr/) who suggested to me) compare these epidemiological data for France and Italy between 2020 and 2021.

To Marc Niaufre (ex MICROSOFT international product manager for Excel) which introduced me to the Benford method and its subtle links with the Fibonacci sequence.

Thanks also for fructuous discussions about this article to Megawaty Tan (A private researcher based in South Sumatera, Indonesia), Robert Freeman M D. (author of "Nature's secret nutrient, golden ratio biomimicry, for PEAK health, performance and longevity), Philippe Risby (initiator of "Learning to Survive") project in Portugal, Valère Lounnas, (Free lance researcher at CMBI European Molecular Biology Laboratory (EMBL) Heidelberg ), Dr Daniel Favre, independent researcher, Brent, Switzerland, Christian Marc, ( retired, MSEE-Dipl-Eng Physics, MBA (Beta Gamma Sigma, USA), Harvard HBS Alumn, General Director [https://www.caravanedelapaix.com/](https://www.caravanedelapaix.com/), and Ethirajan Govindarajan (adjunct Professor, Department of Cybernetics, School of Computer Science, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India, Director, PRC Global Technologies Inc., Ontario, Canada, President, Pentagram Research Centre Pvt. Ltd., Hyderabad, India) and Xavier Azalber, Director FRANCE-SOIR newspaper ([https://www.francesoir.fr/info-en-direct](https://www.francesoir.fr/info-en-direct)).

Particularly, this work is the result of multiple exchanges and advice, since the very beginning of the COVID-19 pandemic, for which I must thank Professor Luc Montagnier (Nobel prizewinner for his discovery of HIV, Fondation Luc Montagnier Quai Gustave-Ador 62 1207 Geneva, Switzerland).

References

(Perez, 1991), J.C. Perez (1991), "Chaos DNA and Neuro-computers: A Golden Link", in Speculations in Science and Technology vol. 14 no. 4, ISSN 0155-7785, January 1991

[https://www.researchgate.net/publication/258439719_JC_Perez_1991_Chaos_DNA_and_Neuro-computers_A_Golden_Link_in_Speculations_in_Science_and_Technologyvol_14_no_4_ISSN_0155-7785](https://www.researchgate.net/publication/258439719_JC_Perez_1991_Chaos_DNA_and_Neuro-computers_A_Golden_Link_in_Speculations_in_Science_and_Technologyvol_14_no_4_ISSN_0155-7785)

Jean-Claude PEREZ. (2020). WUHAN COVID-19 SYNTHETIC ORIGINS AND EVOLUTION. International Journal of Research - Granthaalayah, 8(2), 285–324. [http://doi.org/10.5281/zenodo.3724003](http://doi.org/10.5281/zenodo.3724003)

(Perez, 1997), Perez J.C, L'ADN décrypté, Ed. Marco Pietteur, ISBN : 2-87211-017-8, [https://www.editionsmarcopietteur.com/resurgence/91-adn-decrypte-9782872110179.html](https://www.editionsmarcopietteur.com/resurgence/91-adn-decrypte-9782872110179.html)

(Perez, 2009), Perez J.C, Codex biogenesis – Les 13 codes de l'ADN (French Edition) [Jean - Claude ... 2009); Language: French; ISBN -10: 2874340448; ISBN -13: 978-2874340444 [https://www.amazon.fr/Codex-Biogenesis-13-codes-ADN/dp/2874340448](https://www.amazon.fr/Codex-Biogenesis-13-codes-ADN/dp/2874340448)

(Perez, 2017), J.C Perez, 2017, Sapiens Mitochondrial DNA Genome Circular Long Range Numerical Meta Structures are Highly Correlated with Cancers and Genetic Diseases mtDNA Mutations January 2017 Journal of Cancer Science and Therapy 09(06) DOI: [10.4172/1948-5956.1000469](10.4172/1948-5956.1000469)

(Perez, 2018), Perez, J.C. Six Fractal Codes of Biological Life:perspectives in Exobiology, Cancers Basic Research and Artificial Intelligence Biomimetism Decisions
Making. Preprints 2018, 2018090139 (doi: 10.20944/preprints201809.0139.v1). Perez, J.C. Six Fractal Codes of Biological Life: perspectives in Exobiology, Cancers Basic Research and Artificial Intelligence Biomimetism Decisions Making. Preprints 2018, 2018090139 (doi: 10.20944/preprints201809.0139.v1).
https://www.preprints.org/manuscript/201809.0139/v1

(Perez, 2019), Perez, J. Epigenetics Theoretical Limits of Synthetic Genomes: The Cases of Artificials Caulobacter (C. eth-2.0), Mycoplasma Mycoides (JCVI-Syn 1.0, JCVI-Syn 3.0 and JCVI_3A), E-coli and YEAST chr XII. Preprints 2019, 2019070120 (doi: 10.20944/preprints201907.0120.v1).
https://www.preprints.org/manuscript/201907.0120/v1

Perez J.C, (2020). “WUHAN COVID-19 SYNTHETIC ORIGINS AND EVOLUTION.” International Journal of Research - Granthaalayah, 8(2), 285-324. https://doi.org/10.5281/zenodo.3724003

(Perez&Montagnier, 2020a), Perez, j.c, & Montagnier, L. (2020, April 25). COVID-19, SARS and Bats Coronaviruses Genomes unexpected Exogeneous RNA Sequences.
https://doi.org/10.31219/osf.io/d9e5g

(Perez%Montagnier, 2020b), Jean claude Perez, & Luc Montagnier. (2020). COVID-19, SARS AND BATS CORONAVIRUSES GENOMES PECULIAR HOMOLOGOUS RNA SEQUENCES. International Journal of Research - GRANTHAALAYAH ISSN (print): 2394-3629 July 2020, Vol 8(07), 217 – 263 DOI: https://doi.org/10.29121/granthaalayah.v8.i7.2020.678, Vol 8(07), 217 – 263 (Vol 8(07), 217 – 263), Vol 8(07), 217–Vol 8(07), 263. http://doi.org/10.5281/zenodo.3975578

(Perez, 2021) Perez, J. SARS-CoV2 Variants and Vaccines mRNA Spikes Fibonacci Numerical UA/CG Metastructures. Preprints 2021, 2021040034 (doi: 10.20944/preprints202104.0034.v5). Perez, J. SARS-CoV2 Variants and Vaccines mRNA Spikes Fibonacci Numerical UA/CG Metastructures. Preprints 2021, 2021040034 (doi: 10.20944/preprints202104.0034.v5).
https://www.preprints.org/manuscript/202104.0034/v5

Perez, DUF1220 Homo Sapiens and Neanderthal fractal periods architectures breakthrough(2017) SDRP Journal of Cellular and Molecular Physiology 1(1)p:25-49
http://dx.doi.org/10.25177/JCMP.1.1.4

(Sarkar, 2018), Tirthajyoti Sarkar, What is Benford law? Oct 2018, https://towardsdatascience.com/what-is-benfords-law-and-why-is-it-important-for-data-science-312cb8b61048