先頭の数字が 1 になるのは三割だが、それは桁をまたいで広がるデータだけである

— 2^n（$n = 1$ から 10000）の先頭の数字は 1 が 0.301000 で、$\log_{10}(1 + 1/d)$ との差は九つの数字すべてで 0.000092 以内 — 分かれ目は、データが桁をまたいで広がるかどうかである — [論文612]

山岸 勇輝 (Yuuki YAMAGISHI)
Yamagishi Complexity Torus Lab, Japan｜山岸勇輝ラボ
2026年9月

— From thinking to syncing. —

要旨

数の一覧で、先頭の数字（いちばん上の桁）が 1 から 9 のどれになるかを数えると、1 が約三割、9 が五パーセント足らずになることが多い（ベンフォードの法則）。それがどんなデータで成り立ち、どんなデータで成り立たないかを数える。新しい定理も法則も主張しない。 第一に、2 の冪は、ベンフォードの法則にぴったり従う——2^n を $n = 1$ から 10000 まで正確な整数で計算して先頭の数字を数えると、1 が 3010 回（0.301000）、9 が 458 回で、$\log_{10}(1 + 1/d)$ との差は九つの数字すべてで **0.000092** 以内だった。フィボナッチ数（初めの 5000 個）でも差は 0.000347 以内だった（第2節）。 第二に、全部を同じ数倍しても、分布は変わらない——2^n をすべて 3 倍すると、先頭が 1 の割合はやはり 0.301000 で、差の最大は 0.000108 だった（第3節）。 第三に、桁をまたがないデータは従わない——1 以上 10 未満の一様な数では先頭が 1 の割合は 0.110675（1/9 に近い）で、それを 2 倍すると 0.554595 に変わる。平均 170 cm の身長ではほぼ全部（0.999985）が 1 である（第4節）。 第四に、これが本稿の芯である。分離子は、データが桁をまたいで広がるかどうかである——0 から 1 の一様な数を 20 個掛けた積は 17 桁あまりに広がり、先頭が 1 の割合は 0.300080 だった。桁の目盛り（対数）で見て一様に広がるデータは、何倍しても分布が変わらず、$\log_{10}(1 + 1/d)$ に従う。1 桁に収まる一様な数や、0.170 桁しか広がらない身長は従わない（第4節）。 キーワード：ベンフォードの法則、先頭の数字、尺度不変性、対数の一様分布、桁の広がり。

本稿の射程（射程注記）：新しい定理も法則も主張しない——先頭の数字が偏ること（ニューカム、1881／ベンフォード、1938）、何倍しても変わらない分布がこの法則に限られること（ピンカム、1961）、いろいろな分布から取り混ぜた数がこの法則に近づくこと（ヒル、1995）、$n\log_{10}2$ の小数部分が一様に散ること（ワイル、1916）は、いずれも既知である。応用を扱わない——帳簿の不正の検出などには立ち入らない。なぜ多くの実データが従うかは論じない——従うデータと従わないデータを、桁の広がりで分けて数えるだけである。既刊との関係：論文117 は、スケール不変性が指数を四通りに選ぶことを分けた——本稿は、尺度を変えても変わらない分布が先頭の数字では $\log_{10}(1+1/d)$ になることを、数で確かめる。加えたのは、2^n の先頭の数字を正確な整数で 10000 個数えたこと、3 倍しても変わらないこと、一様な数では 2 倍で変わることを並べたこと、分離子を「桁をまたいで広がるか」に置き、一様・身長・積の広がりを桁で測ったことである。

1. 記法と問い

設定：正の数 x の先頭の数字を $D(x)\in\{1,\dots,9\}$ とする。ベンフォードの法則は、先頭が d になる割合を

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

とする。$d=1$ から 9 で
$0.301030, 0.176091, 0.124939, 0.096910, 0.079181, 0.066947, 0.057992, 0.051153, 0.045757$ で、和は 1 である。

問い：どんな数の並びがこの割合に従うか。従わないのはどんな場合か。何が分けるか。

2. 2 の冪

計算 2.1：2^n を $n=1$ から 10000 まで正確な整数で計算し、先頭の数字を数えた。

先頭の数字	回数	割合	$\log_{10}(1+1/d)$	差
1	3010	0.301000	0.301030	−0.000030
2	1761	0.176100	0.176091	+0.000009
3	1249	0.124900	0.124939	−0.000039
4	970	0.097000	0.096910	+0.000090
5	791	0.079100	0.079181	−0.000081
6	670	0.067000	0.066947	+0.000053
7	579	0.057900	0.057992	−0.000092
8	512	0.051200	0.051153	+0.000047
9	458	0.045800	0.045757	+0.000043

差の最大は **0.000092** だった。フィボナッチ数の初めの 5000 個でも、先頭が 1 の割合は 0.301000、差の最大は 0.000347 だった。

> **観察 2.2**
>
> 2^n の先頭が d になるのは、$n\log_{10}2$ の小数部分が $\log_{10}d$ と $\log_{10}(d+1)$ の間に落ちるときである。$\log_{10}2$ は分数で書けないので、n を増やすとその小数部分は 0 から 1 に一様に散る——だから先頭が d になる割合は、区間の幅 $\log_{10}(d+1)-\log_{10}d=\log_{10}(1+1/d)$ になる。

3. 尺度を変える

計算 3.1：2^n をすべて 3 倍した数の先頭の数字を数えると、先頭が 1 の割合は 0.301000、差の最大は 0.000108 だった。

> **観察 3.2**
>
> 何倍しても、分布が変わらない。3 倍すると $n\log_{10}2$ の小数部分に $\log_{10}3$ が足されるだけで、一様に散っていることは変わらない——単位を変えても（円からドル、メートルからフィート）先頭の数字の割合が変わらないデータは、この割合に従うほかない。

4. 分離子

計算 4.1：四つの並び（それぞれ 200000 個）で、先頭が 1 の割合と、データが広がる桁数を比べた。

データ	先頭が 1 の割合	広がる桁数
1 以上 10 未満の一様な数	0.110675	1
それを 2 倍した数	0.554595	1
身長（平均 170 cm・標準偏差 7 cm）	0.999985	0.170
0 から 1 の一様な数を 20 個掛けた積	0.300080	17.44

一様な数の 0.110675 は $1/9=0.111111$ に、2 倍した 0.554595 は 10 から 20 が占める $10/18=0.555556$ に近い。積の割合と $\log_{10}(1+1/d)$ の差の最大は 0.001036 だった。

> **観察 4.2**
>
> これが本稿の芯である。分離子は、データが桁をまたいで広がるかどうかである。一様な数や身長は一つの桁の中に収まり、その桁のどこに集まるかで先頭の数字が決まる。2 倍すれば集まる場所が動き、割合も変わる——積は 17 桁あまりに広がり、桁の目盛りで見てほぼ一様になる。そうなると何倍しても分布が変わらず、先頭の数字は $\log_{10}(1+1/d)$ に従う。

> **観察 4.3**
>
> 論文117 と並べる。117 は、スケール不変性が指数を四通りに選ぶことを示した——先頭の数字の分布では、尺度を変えても変わらないことを要求すると、割合が $\log_{10}(1+1/d)$ に一つに決まる。3 倍しても変わらない 2^n と、2 倍で変わる一様な数が、その両側である。

> **観察 4.4**
>
> 言わないこと。帳簿の不正の検出などの応用は扱わない。なぜ多くの実データがこの法則に従うかの理論には立ち入らない——示したのは、2 の冪で割合が合うこと、尺度を変えても変わらないこと、桁をまたがないデータでは合わないことである。

5. 点呼と結び

1. 2^n で先頭が 1 は 3010 回（計算 2.1）：0.301000。
2. 九つの数字すべてで差 0.000092 以内（計算 2.1）：正確な整数で。
3. フィボナッチ数でも差 0.000347 以内（計算 2.1）：同じく従う。
4. 小数部分が一様に散るから（観察 2.2）：区間の幅が割合になる。
5. 3 倍しても 0.301000（計算 3.1）：尺度を変えても変わらない。
6. 一様な数は 0.110675（計算 4.1）：1/9 に近い。
7. それを 2 倍すると 0.554595（計算 4.1）：尺度で変わる。
8. 身長は 0.999985（計算 4.1）：0.170 桁しか広がらない。
9. 20 個の積は 0.300080（計算 4.1）：17 桁あまりに広がる。
10. 分離子は桁をまたいで広がるか（観察 4.2）：本稿の芯。

結び：数の一覧で先頭の数字を数えると、1 が約三割になることが多い。それは、桁をまたいで広がるデータだけである——2^n を 10000 個正確な整数で数えると、先頭の数字は九つとも $\log_{10}(1+1/d)$ から 0.000092 以内で、全部を 3 倍しても変わらなかった。$n\log_{10}2$ の小数部分が一様に散るからである。分離子は、データが桁をまたいで広がるかどうかである——1 から 10 の一様な数は 1/9 に近く、2 倍すると 10/18 に変わり、身長はほぼ全部が 1 だった。20 個の一様な数の積は 17 桁あまりに広がり、先頭が 1 の割合は 0.300080 だった。既刊との位置——論文117 のスケール不変性は、先頭の数字では割合を一つに決める。正直に言えば——応用にも、なぜ実データが従うかの理論にも立ち入っておらず、数え上げと広がりの比較だけである。

参考文献

1. Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*.
2. Hill, T. P. (1995). A statistical derivation of the significant-digit law. *Statistical Science*.
3. Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*.
4. Pinkham, R. S. (1961). On the distribution of first significant digits. *The Annals of Mathematical Statistics*.

5. Weyl, H. (1916). Über die Gleichverteilung von Zahlen mod. Eins. *Mathematische Annalen*.

6. 山岸勇輝 (2026). 『スケール不変性は、四通りに指数を選ぶ』[論文117]. Zenodo. 10.5281/zenodo.22005515

— From thinking to syncing. —

山岸 勇輝 (Yuuki YAMAGISHI)

© 2026 Yamagishi Complexity Torus Lab, Japan · Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0)