dennou777.cls テスト文書

ひとみ

2019 年 4 月 16 日

この文書は、dennou777.cls のテスト文書である。

第 1 章　ベンフォードの法則

1.1　はじめに

等比数列 $\{2^n\}$ や、フィボナッチ数列 $\{F_n\}$ の、「一番上の桁」に出現しやすい数字は何でしょうか。

例: $2^5 = 32$ なら 3、$F_7 = 12$ なら 1。

1.2　証明

1.2.1　$\{2^n\}$ の場合

2^k の最上位の数字を d、m を整数とすると、

$$d \times 10^m \leq 2^k < (d+1) \times 10^m$$

常用対数を取ると、

$$\begin{gathered} m + \log[d] \leq k \log[2] < m + \log[d+1] \\ \frac{m + \log[d]}{\log[2]} \leq k < \frac{m + \log[d+1]}{\log[2]} \end{gathered}$$

$d = d_n$ のとき、不等式を満たす k の幅 x_n を求める。

$$\begin{aligned} x_n &= \frac{m + \log[d_n]}{\log[2]} - \frac{m + \log[d_n + 1]}{\log[2]} \\ &= \frac{1}{\log[2]} \cdot \log\left[\frac{d_n + 1}{d_n}\right] \end{aligned}$$

これより、x_n は m（桁数）によらない事がわかる。

後のために $\sum_{i=1}^{9} [x_i]$ を求めておく。

$$\begin{aligned} \sum_{i=1}^{9} [x_i] &= \frac{1}{\log[2]} \sum_{i=1}^{9} \log\left[\frac{d_i + 1}{d_i}\right] \\ &= \frac{1}{\log[2]} \cdot \log\left[\frac{2}{1} \cdot \frac{3}{2} \cdot \cdots \cdot \frac{10}{9}\right] \\ &= \frac{1}{\log[2]} \cdot \log[10] = \frac{1}{\log[2]} \end{aligned}$$

k は一様分布なので、$d = n$ となる確率 $\Pr[d = n]$ は、

$$\begin{aligned}\Pr[d = n] &= \frac{x_n}{\sum_{i=1}^{9}[x_i]} \\ &= \frac{1}{\log[2]} \cdot \log\left[\frac{n+1}{n}\right] \Bigg/ \frac{1}{\log[2]} \\ &= \log\left[\frac{d+1}{d}\right]\end{aligned}$$

1.2.2 $\{F_n\}$ の場合

$$\begin{cases} F_1 &= 1 \\ F_2 &= 1 \\ F_{n+2} &= F_{n+1} + F_n \end{cases}$$

フィボナッチ数列の漸化式から一般項を求めると、

$$F_n = \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right)$$

である。

ここで、$(1-\sqrt{5})/2 < 1$ であるから、n が大きいときには

$$F_n \approx \frac{\phi^n}{\sqrt{5}}, \qquad \phi = \frac{1+\sqrt{5}}{2} \text{（黄金比）}$$

となり、2^k のときと同様の議論によって、同じ確率が導かれる。

1.3 導いた式との比較

$n \leq 100$ の範囲で、導かれた式と、実際の頻度とを比較する（表 1.3.1 と図 1.3.1）。

1.4 Benford's law

今までの議論は自然界に出現する数値でも、経験則的に成り立つ。

Benford の法則

自然界に出現する数値で、最上位に n が現れる確率は

$$\log_{10}\left[\frac{n+1}{n}\right]$$

- n 進法表記の数値に拡張可能。
- 会計の分野で、不正を見つけるために使われる。

表 1.3.1　最上位の数字の頻度

数字	2^n	F_n
1	30	30
2	17	18
3	13	13
4	10	9
5	7	8
6	7	6
7	6	5
8	5	7
9	5	4

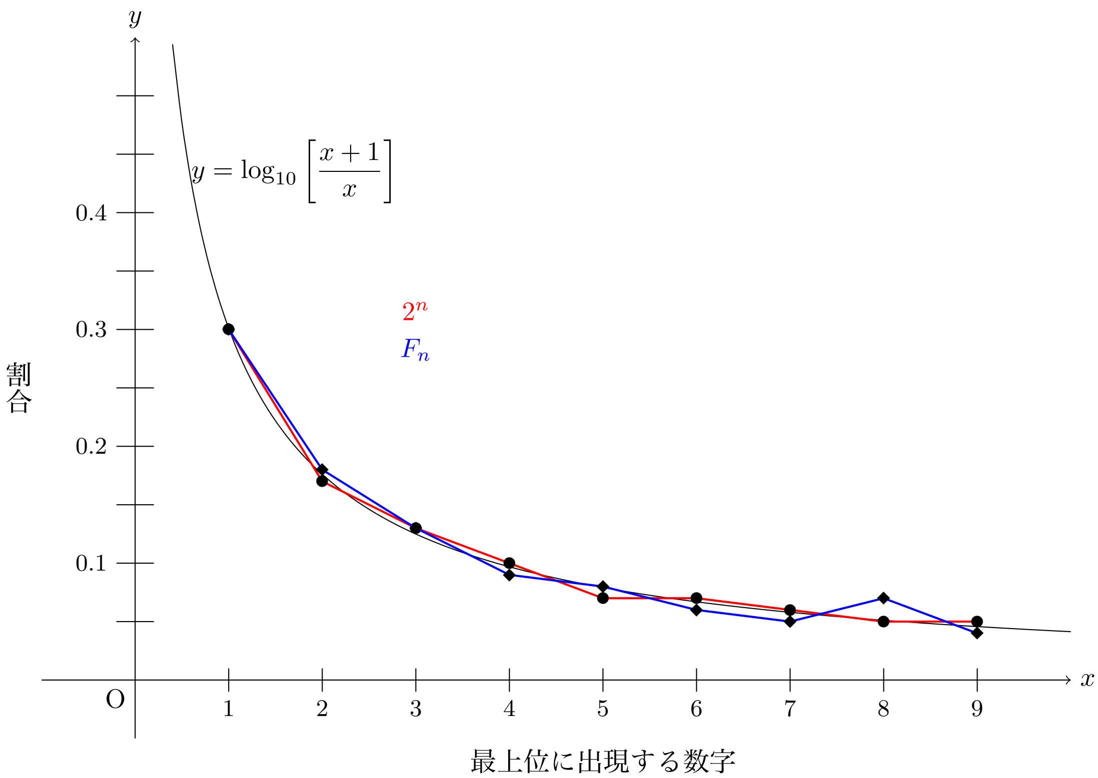

図 1.3.1　最上位の数字の頻度