岡山理科大学紀要 第42号 A pp. 7-11(2006)

ベキ乗の先頭桁の数字について

高嶋　恵三　・　小谷　真美*
岡山理科大学理学部応用数学科
* 岡山理科大学理学研究科修士課程応用数学専攻

(2006年10月2日受付、2006年11月6日受理)

はじめに

ベキ乗 a^n の 10 進法表示における先頭桁の数字について考える. 例えば, $a = 2$ の場合

$2^1 = \underline{2}$, $2^2 = \underline{4}$, $2^3 = \underline{8}$, $2^4 = \underline{1}6$, $2^5 = \underline{3}2$, $2^6 = \underline{6}4$, $2^7 = \underline{1}28$, $2^8 = \underline{2}56$, $2^9 = \underline{5}12$, $2^{10} = \underline{1}024$, $2^{11} = \underline{2}048$, $2^{12} = \underline{4}096$, $2^{13} = \underline{8}192, \ldots$

エルゴード理論でよく知られているように, 先頭桁の数字には 1 から 9 までの数字が出現することが知られている. この事実は $a = 3, 4, \ldots, 9$ でも同様であり, いずれの場合にも, 数字 k の出現頻度は

$$\log_{10}(k+1) - \log_{10} k\,, \qquad k = 1, 2, \ldots, 9,$$

であることが Weyl の補題から分かる.

本報告では, $a = 2, 3, 4, \ldots, 9$ に対して, $n = 1$ から $n = 10{,}000{,}000$ まで計算した結果について検討する. 特に $a = 7$ の場合, 各数字の出現頻度の漸近的挙動についてこれまでの予想と大きく異なる結果が示されることについて報告する.

エルゴード理論からの準備

ここではまず, Berger[1] の教科書に基づき, エルゴード理論の基礎概念について簡単にまとめる.

X を距離空間とし, B_X は X のボレル集合族, μ は (X, B_X) 上の確率測度とし, $T : X \to X$, 連続写像, を考える.

定義　T が (X, B_X, μ) 上の**保測変換**であるとは, $T\mu = \mu$ を満たすことである. ここで $T\mu(A) = \mu(T^{-1}(A))$, $A \in B_X$, である.

定義　保測変換 T が **エルゴード的 ergodic** であるとは, 任意の $A \in B_X$ が T-不変, すなわち $T^{-1}(A) = A$, なら $\mu(A) = 0$ または $\mu(A) = 1$, となることである.

定義　保測変換 T が **混合的 mixing** であるとは

$$\lim_{n\to\infty} \mu(T^{-n}(A) \cap B) = \mu(A)\mu(B)\,, \quad \forall A, B \in B_X\,.$$

定義（無理数回転） 複素平面上の単位円 S^1 を考える. S^1 上の変換, **無理数回転**, R_θ を

$$R_\theta : \begin{cases} S^1 & \to & S^1 \\ z & \mapsto & e^{2\pi i\theta} z \end{cases}$$

で定義する. ここで $0 \le \theta < 1$ である.

R_θ は θ が無理数であるときに限り, エルゴード的であることが知られている. また, R_θ は（ θ が有理数, 無理数に関係なく）混合的ではない, ことも知られている.

保測変換 T について次の有名な２つのエルゴード定理が知られている：

定理 T を (X, B_X, μ) 上の保測変換とし, f を (X, B_X, μ) 上の可積分関数とする. この時

(1) **(Birkhoff Ergodic Theorem)**
(X, B_X, μ) 上の可積分関数 f^* が存在して,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) = f^*(x) , \quad \mu\text{-almost all } x .$$

(2) **(von Neumann Ergodic Theorem)**
(X, B_X, μ) 上の可積分関数 f^* が存在して,

$$\lim_{n\to\infty} \int_X \left| \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)) - f^*(x) \right| \mu(dx) = 0 .$$

これらの結果は独立確率変数列に対して, **大数の法則** として知られている結果の拡張である.

S^1 上の無理数回転に対して, 対応する結果は次の **Weyl の補題**として知られている定理である.

定理 S^1 上の任意の Riemann 積分可能な関数 f と無理数 θ に対して

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(R_\theta^k(z)) = \int_{S^1} f(z)\, \lambda(dz) , \quad \forall z \in S^1 .$$

但し, λ は S^1 上の通常の測度（弧長の自然な拡張）で, $\lambda(S^1) = 1$, を満たすものとする.

この Weyl の補題より, 特に $f(z) = I_A(z)$ ($A \subset S^1$ は弧であり $I_A(z)$ は A の定義関数) とすると,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_A(R_\theta^k(z)) = \frac{|A|}{2\pi} , \quad \forall z \in S^1 .$$

ここで $|A|$ は A の弧長である.

べき乗について

a^n が 10 進法で ℓ 桁の数であり, その先頭桁の数字が k となるのは, 常用対数をとって考えることにより

$$\ell - 1 + \log_{10} k \le n \log_{10} a < \ell - 1 + \log_{10}(k+1) ,$$

が成り立つ場合に限ることがわかる．この不等式の整数部分を無視して小数部分のみを考え，S^1 の上で $n\ \log_{10} a$ の挙動を考える．

R^1 から S^1 への自然な写像 $\pi : R^1 \to S^1$,

$$\pi(x) = e^{2\pi i x} \ , x \in R^1 \ ,$$

を導入して考えると，a^n の先頭桁が k となるのは，

$$R_\theta^n\,(1) \in \pi([\log_{10} k, \log_{10}(k+1))) \ , \quad k = 1, 2, \ldots, 9 \ , \ \theta = \log_{10} a \ ,$$

となる場合に限ることがわかる．

$\log_{10} a$ ($a = 2, 3, \ldots, 9$) は無理数であるので，上記の Weyl の補題が適用でき，先頭桁の数字の出現頻度漸近挙動に関して，次が成り立つことがわかる：

命題　$a = 2, 3, \ldots, 9$ に対して，$n \to \infty$ のとき，

$$\frac{a^\ell \text{ の先頭桁が } k\text{，となる } \ell\,(1 \le \ell \le n) \text{ の個数}}{n} \to \log_{10}(k+1) - \log_{10} k \ ,$$

但し，$k = 1, 2, \ldots, 9$.

計算機実験の結果について

独立確率変数列の場合，よく知られているように大数の法則における極限への収束の速さは，$O(\frac{1}{\sqrt{n}})$ のオーダーである．一方，無理数回転の場合，混合的でなく単にエルゴード的であるので収束の速さについては一般的な結果は知られていないが，混合的な場合より速い，と予想される．

そこで出現頻度の極限分布への収束の速さを見るために，観測分布と極限分布との差を数理統計学で適合度の検定として知られている，χ^2 検定

$$\chi_0^2 = \sum_{k=1}^{9} \frac{(f_k - e_k)^2}{e_k} \ ,$$

を用いて調べてみた．ここで f_k は数字 k の観測度数，e_k は期待度数，$n \times (\log_{10}(k+1) - \log_{10} k)$ である．この場合の自由度は 8 (9 - 1) である．

$a = 2, 3, 4, 5, 6, 8, 9$ の場合は計算機実験の結果は，上記の予想に沿ったものである．例えば，$a = 2$ については

表 1．　$a = 2$ の場合，n' は 100,000 単位．

n'	1	2	3	4	5	6	7	8	9
χ_0^2	0.0017	0.0010	0.0004	0.0013	0.0017	0.0003	0.0005	0.0007	0.0005
10	20	30	40	50	60	70	80	90	100
0.0002	0.0003	0.0007	0.0002	0.0001	0.0001	0.0002	0.0001	0.0002	0.0003

上記の表には載せていないが，実際の計算は n に対して 1,000 単位で行っている．この計算結果から，自由度 8 に対して n が極めて大きいので，n が 10,000 程度ですでに十分極限分布に近づいていると考えてよ

い. また, $n = 100,000$ 以上では観測分布と極限分布の誤差は極めて小さい, といえることがわかる. 一方, $a = 7$ の場合に結果は上記の $a = 2,3,4,5,6,8,9$ などの場合と明らかに異なり, 以下のようになった:

表 2. $a = 7$ の場合, n' は 100,000 単位.

n'	1	2	3	4	5	6	7	8	9
χ_0^2	4.3383	6.9158	11.3697	16.8227	22.5452	28.4779	34.5438	40.6478	43.4946
10	20	30	40	50	60	70	80	90	100
46.4432	10.1113	4.5709	14.3735	0.0687	8.9961	2.2461	2.4188	5.9951	0.1107

なお、表 2 では分からないが, n を 1,000 単位で増やした場合の計算結果では、 χ_0^2 の値の最大値は $n = 1,165,000$ で $\chi_0^2 = 52.6637$, であり, $n = 2455000, 2456000$ では $\chi_0^2 = 0.0000$ となった. さらに, $n = 3744000$ では $\chi_0^2 = 17.1317$, $n = 4912000, 4913000$ では $\chi_0^2 = 0.0000$, などと, 約 1,100,000 程度の「周期」で増大と減少を繰り返していることがわかった.

ここで注意しなければならないことは, この計算結果は決して「先頭桁の数字の出現頻度が $\log_{10}(k+1) - \log_{10} k$ に漸近的に近づく」という, Weyl の補題に反するものではない, ということである. 上記の場合では n が, 自由度 8 に対して, 極めて大きな値であるので χ^2 検定はわずかな分布の違いを大きく見せている, という点に注意する必要がある.

このように,

- 1,000,000 程度の極めて長い間隔で増大・減少を繰り返す.
- χ_0^2 の値が 50 を超える,

という現象は知られてはいないようであり, 極めて不可思議な現象といえる. この現象に対する理論的説明はいまだ知られていないようであり, 今後の研究課題である.

参考文献

[1] Berger, A., : *Chaos and Chance*, Walter de Gruyter (2001)

[2] Weyl, H., : *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. **77**, 313 – 352 (1916)

On Leading Digits of Powers a^n

Keizo TAKASHIMA and Mami OTANI*

Department of Applied Mathematics, Faculty of Science,
**Graduate School of Science,*
Okayama University of Science,
1-1 Ridai-cho, Okayama 700-0005, Japan

(Received October 2, 2006; accepted November 6, 2006)

It is well-known that the leading digits of powers a^n, $a > 1$ is a natural number, are distributed over from 1 to 9, the asymptotic distribution is given as follows:

the leading digit equals to $k = 1, 2, \ldots, 9$ with probability $\log_{10}(k+1) - \log_{10} k$.

This is derived from the ergodic theorems, especially from Weyl's theorem on irrational rotations on the unit circle S^1, applied to $\log_{10} a$. Since the irrational rotation with $\log_{10} a$ is not *mixing* but merely *ergodic*, the empirical distributions of leading digits are expected to converge rather rapidly to the above asymptotic distribution.

We have checked the convergence of empirical distributions, using chi square test, up to $n = 10,000,000$ and for $a = 2, 3, \ldots, 9$. For all a, other than 7, the empirical distributions converge very quickly to the theoretic distribution, however, in case of $a = 7$, we observed very strange phenomena:

when n = 1,165,000, the value of chi square test equals to a huge value, 52.6637. When n = 2455000, 2456000, the value of chi square test decrease down to extraordinally small value, 0.0000. Values of the chi square test again go up to a maximal value 17.1317, when n = 3744000. After that, the chi square values converge to 0, with waving up and down in a period of about 1,100,000.

The theoretical explanations to the above phenomena is unknown and an open problem.

Bibliography

Berger, A., : *Chaos and Chance*, Walter de Gruyter (2001)
Weyl, H., : *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. **77**, 313 – 352 (1916)

Keywords: irrational rotations; Weyl's lemma; Birkoff's ergodic theorem.