素数の偏りと深リーマン予想
（前編）

小山信也

挿絵： 藤津麻里

今号と次号の2回にわたり，整数論で長年にわたり未解決問題とされてきた「チェビシェフの偏り」に対して得た進展を報告する.「チェビシェフの偏り」は19世紀にチェビシェフが指摘した問題であり,「4で割って3余る素数」が「4で割って1余る素数」よりも多めに存在するかにみえる現象を指す．よく知られているように,「ディリクレの算術級数定理」によってこれらは同数であることが証明されている.「チェビシェフの偏り」はこの事実に矛盾するかのような現象である．この奇妙な「偏り」はどのように定式化され，それが発生する原因は何なのか，これまで未解明であった.

このたび，東京工業大学名誉教授である黒川信重氏[7]によって2012年頃から提唱されてきた「深リーマン予想」を用いることにより，チェビシェフの偏りの定式化と証明に成功した．その結果，外見上は「不自然な偏り」に見えるこの現象は，実は，素数全体がバランスを保ちながら分布するために必要な「自然な現象」であることが判明した．そのバランスとは「素因数分解の一意性」すなわち「オイラー積」である．オイラー積の究極的な収束性を表す「深リーマン予想」を踏まえると,「チェビシェフの偏り」の自然な定義が得られ，その存在を証明でき，存在理由も明確になる.

今回，前編では，問題の端緒から本研究以前になされていた先行研究までを紹介し，次号の後編において，解明に至るアイディアと深リーマン予想について解説する.

1　チェビシェフの手紙

1853年，チェビシェフは知人への手紙[1]に「素数に関する新たな発見」を記した．それは,「4で割って3余る素数」と「4で割って1余る素数」の分布の比較から判明した奇妙な「偏り」であった．以下に，チェビシェフが記した2つの例を紹介する．なお，手紙[1]には「新しい結果（nouveau resultat）」と記されているものの証明は未記載であり，現在に至るまで証明は知られておらず未解決問題とされているため，以下では「予想」と呼ぶ.

予想１

次の級数は，$c \to 0$ で $+\infty$ に発散する.

$$e^{-3c} - e^{-5c} + e^{-7c} + e^{-11c} - e^{-13c} - e^{-17c} + e^{-19c} + e^{-23c} + \cdots \quad (1)$$

予想２

$f(x)$ が $\lim_{x\to\infty} x^{\frac{1}{2}} f(x) \neq 0$ を満たし，正で単調減少であるとき，次の級数は発散する.

$$f(3) - f(5) + f(7) + f(11) - f(13) - f(17) + f(19) + f(23) + \cdots \quad (2)$$

式(1)の指数や式(2)の変数に現れている数列

$$3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23, \cdots$$

[1]手紙の原文はブリティッシュコロンビア大学のサイトで入手可能である.
https://personal.math.ubc.ca/~gerg/teaching/592-Fall2018/papers/1853.Chebyshev.pdf

は，第 n 項が「小さい方から数えて n 番目の奇素数 p」であり，符号は法 4 の原始指標

$$\chi_{-4}(p) = \begin{cases} 1 & (p \equiv 1 \pmod 4) \\ -1 & (p \equiv 3 \pmod 4) \end{cases}$$

である．すなわち，予想 1 は

$$\lim_{c \to 0} \sum_{p:\ \text{奇素数}} \chi_{-4}(p) e^{-pc} = \infty$$

と言い換えられるし，予想 2 は「級数

$$\sum_{p:\ \text{奇素数}} \chi_{-4}(p) f(p)$$

が収束するためには，$\lim_{x \to \infty} x^{\frac{1}{2}} f(x) = 0$ が必要」と同値である．

これら 2 つの予想の意味するところは何だろうか．まず，予想 1 は，式 (1) を

$$\lim_{x \to \infty} \left(\sum_{\substack{p \le x \\ p \equiv 3 \pmod 4}} e^{-pc} - \sum_{\substack{p \le x \\ p \equiv 1 \pmod 4}} e^{-pc} \right)$$

と書き換えれば,「4 で割って 3 余る素数」にわたる和と「4 で割って 1 余る素数」にわたる和の比較とみなせる．$c \to 0$ と $x \to \infty$ は極限操作の順序交換ができないため単純には論じられないが，仮に $c = 0$ を代入すれば，上の式は

$$\lim_{x \to \infty} (\pi(x, 4, 3) - \pi(x, 4, 1)) \tag{3}$$

となる．ここで，$\pi(x, q, a)$ は「q で割って a 余るような x 以下の素数の個数」を表す記号である．もし，極限 (3) が ∞ となれば,「4 で割って 3 余る素数」が「4 で割って 1 余る素数」よりも無数に多く存在することになる．実際には，チェビシェフ以前の 1837 年に「ディリクレの算術級数定理」が証明されており，それら 2 組の素数は同数であることが知られていた．算術級数定理を法 4 の場合に記すと

$$\lim_{x \to \infty} \frac{\pi(x, 4, 3)}{\pi(x, 4, 1)} = 1 \tag{4}$$

となる．(3) と (4) を比べると，(3) は (4) よりも精密な主張である．このことを，後年 1896 年に証明された素数定理

$$\lim_{x \to \infty} \frac{\pi(x)}{\mathrm{Li}(x)} = 1 \tag{5}$$

を用いて説明しよう．ここで，$\pi(x)$ は「x 以下の素数の個数」であり，$\mathrm{Li}(x)$ は対数積分関数

$$\mathrm{Li}(x) = \int_2^x \frac{dt}{\log t}$$

で，$x \to \infty$ における主な挙動は $x/\log x$ である．素数定理 (5) より，

$$\pi(x) = \mathrm{Li}(x) + (\text{誤差項})$$

となり,「誤差項」の部分は $\mathrm{Li}(x)$ よりも小さい．算術級数定理 (4) より,「4 で割って 3 余る素数」と「4 で割って 1 余る素数」は素数全体の半分ずつを占めるので,

$$\begin{aligned} \pi(x, 4, 3) &= \frac{1}{2}\mathrm{Li}(x) + (\text{誤差項}) \\ \pi(x, 4, 1) &= \frac{1}{2}\mathrm{Li}(x) + (\text{誤差項}) \end{aligned} \tag{6}$$

となる．この 2 式の「誤差項」は $\mathrm{Li}(x)$ より小さければ何でも良い．その大きさは (4) の正否に影響がないし，2 つの誤差項が互いに同程度の大きさである必要もない．つまり，(4) は主要項 $\frac{1}{2}\mathrm{Li}(x)$ のみを規定しており，それ未満の項のことは何も主張していない．これに対し，(3) は全く異なる．(6) を (3) に代入すれば，引き算によって $\frac{1}{2}\mathrm{Li}(x)$ が打ち消しあうので,

$$(3) = \lim_{x \to \infty} (\text{誤差項どうしの差})$$

となる．すなわち，(3) の挙動は素数定理の誤差項の大きさや，それらの内訳が「3 余る素数」と「1 余る素数」にどのように配分されるかといった問題に深くかかわっている．そして，これらの問題は，(4) のあずかり知らないことなのである．

ここまで，説明のために仮に $c = 0$ として式の意味を考えてきたが，実際には極限操作の順序交換が成り立たず，極限 (3) は存在しない上

に，差 $\pi(x,4,3)-\pi(x,4,1)$ は，正になったり負になったりを無限回繰り返すことが証明されている（第 3 節「リトルウッドの定理」）．この事実は，ディリクレの算術級数定理を踏まえれば，直感的に理解しやすいだろう．$\pi(x,4,3)$ と $\pi(x,4,1)$ は漸近的に等しいのだから，一方的な大小関係は無く，両者が「抜きつ抜かれつ」を永遠に繰り返す光景は，自然に想像できる．したがって，式 (1) を考える上で，$c=0$ を直に代入することは，厳密には意味を持たない．

では，式 (1) の意味をどう考えたら良いだろうか．それは，チェビシェフが素数に「重み」を付けることで，存在しない極限に解釈を与えたものと捉えられる．すなわち，従来の $\pi(x,q,a)$ は，どの素数も一様に「1 個」と数えていたが，これを「大きな素数を軽く」数えるように変えると真実が見えてくる．これがチェビシェフの洞察であった．式 (1) の各項 e^{-pc} は，$c>0$ である限り，p の増大に伴い指数関数的に減少する．したがって，収束性が良くなり，通常の個数関数に対して存在しなかった (3) に相当する極限が存在したり，∞ や $-\infty$ に確定したりする可能性がある．単に「無限回符号が変わるから収束しない」といって話を終わるのではなく，無理やりにでも収束させることによって「3 余る素数」と「1 余る素数」の違いを見極めようとしているのである．式 (1) は，このような「重みづけ」を施して素数を数えた「重み付き個数」の差とみなせる．チェビシェフは，この「重み付き個数」において，「3 余る素数」が「1 余る素数」よりも（無数に）多く存在することを観察したのである．

なお，後年の研究（ハーディ・リトルウッド [2]，ランダウ [4],[5]）により，予想 1 はディリクレ L 関数 $L(s,\chi_{-4})$ に対するリーマン予想と同値であることが証明された．よく知られているように，ディリクレ L 関数のリーマン予想は「ディリクレ素数定理（算術級数定理）の誤差項」の精密化とも同値であるから，先ほど $c=0$ の場合に言及した「(3) が (4) よりも誤差項に関してより精密な主張である」という点は，$c\to 0$ でも成り立つことがわかる．

次に，予想 2 について見てみよう．予想 1 は，指数関数的に減少する例を用いたが，減少の度合いはどこまで緩やかにできるだろうか．たとえば，$f(x)=1/x^{\alpha}$ $(\alpha>0)$ とおいて考えてみよう．すると，予想 2 は，級数

$$\sum_{p:\ \text{奇素数}} \frac{\chi_{-4}(p)}{p^{\alpha}}$$

の収束を論じていることになる．この級数の収束は，ディリクレ L 関数 $L(s,\chi_{-4})$ のオイラー積

$$L(s,\chi_{-4})=\prod_{p:\ \text{奇素数}}(1-\chi_{-4}(p)p^{-s})^{-1}$$

の $s=\alpha$ における収束と同値である．予想 2 は，この収束のために，$\lim_{x\to\infty} x^{\frac{1}{2}-\alpha}=0$，すなわち，$\alpha>1/2$ が必要であることを主張している．ここで得た境界 1/2 は，後編で解説する「深リーマン予想」にも登場する．深リーマン予想は，チェビシェフが触れなかった「境界 $\alpha=1/2$ における挙動」にまで踏み込んでおり，それを用いると「偏り」が解明されるのである．

過去に「チェビシェフの偏り」を解説した多くの文献で，予想 1 のみが言及されることが多かった．これに対して予想 2 はほとんど無視されてきたと言える．予想 2 を詳しく論じた文献を私は見たことがない．しかし，チェビシェフは手紙を書いた段階で，すでに境界 1/2 を意識しており，それを予想 2 に記した．ある意味で，彼は深リーマン予想に近づきつつあったのかもしれない．

2 チェビシェフの偏り

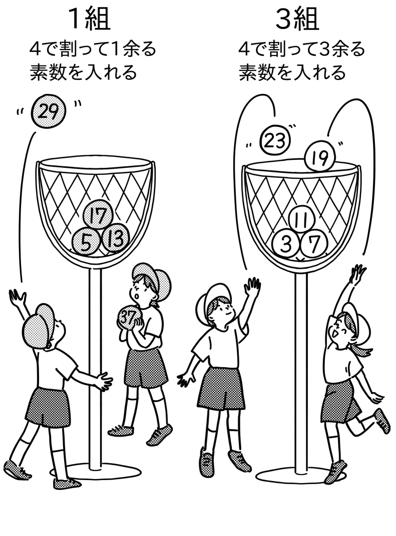
予想 1 では，「4 で割って 3 余る素数」が「4 で割って 1 余る素数」よりも多く存在するように見える現象を，素数に重みを付けて考えたが，実際の素数の個数はどうなのか，確認してみよう．

目を見張るのは，不等式

$$\pi(x,4,3) \geq \pi(x,4,1) \tag{7}$$

が成り立つような x が，圧倒的に多いことである．$\pi(x,q,a)$ の値が変わるのは x が素数のときのみであるから，素数 x についてのみ不等式を考察すれば十分である．すると，$x < 26861$ なるすべての素数 x に対して不等式 (7) は成り立ち，この間で不等式 (7) の等号が成り立つ素数 x は，たったの 4 個しかない．そして，素数 $x = 26861$ において，初めて逆側の不等号「$<$」が成立するが，すぐ次の素数 $x = 26863$ で再び等号が成り立ち，その次の素数 $x = 26879$ で再び不等号「$>$」が成り立つ．そして，その後はしばらく不等号の向きが変わらない．次に不等号「$<$」が成り立つのは，なんと，$x = 616841$ のときである．61 万過ぎまで圧倒的に「3 余る組」が優勢なのである．

この情景は，運動会の玉入れに例えるとイメージしやすいかもしれない．1 組と 3 組の試合で，小さい素数から順に,「1 余る素数」は 1 組のかごに,「3 余る素数」は 3 組のかごに玉を入れていく．



ディリクレ素数定理（算術級数定理）によって，両者が最終的に同点になることはわかっているが，そこに至る過程では，3 組がリードする区間が圧倒的に長い．次ページの図は，300 万以下の代表的な x に対して，得点差をまとめたものである．ここからも，3 組の優勢が見てとれる．

ディリクレ素数定理（算術級数定理）で同数であることが証明されているにもかかわらず，実際の数値にこれだけ圧倒的な偏りが生ずる理由は，いったい何なのだろうか？これが，19 世紀以来の未解決問題「チェビシェフの偏り」である．

3　偏りの研究史

チェビシェフが指摘した「偏り」を研究するためには，まず「偏り」を定義する必要がある．前節で紹介したデータを見ると，1 組にほぼ勝ち目は無いと思えるので，すぐに思いつく素朴な予想は，以下のものだろう．

> ある X が存在して，任意の $x > X$ に対して次の不等式が成り立つ．
>
> $$\pi(x,4,3) > \pi(x,4,1)$$

しかし，1914 年にリトルウッドが次の定理を証明したことにより，この予想は否定された．

リトルウッドの定理 (1914)

> $x \to \infty$ において，差 $\pi(x,4,3) - \pi(x,4,1)$ は，無限回符号を変える．

したがって，1 組は弱いながらも，たまに勝つことを永遠に止めないことがわかる．そこで，次に思いつく自然な予想は，以下のものだろう．

ナポウスキー・テュランの予想 (1962)

> 不等式 $\pi(x,4,3) > \pi(x,4,1)$ を満たす x の自然密度は，$x \to \infty$ において 1 に収束する．すなわち，区間 $A(X)$ を，$\pi(x,4,3) > \pi(x,4,1)$ が成り立つ実数 $x \leq X$ の集合とおき，$\mathrm{vol}(A(X))$ を区間 $A(X)$ の長さとするとき，次式が成り立つ．
>
> $$\lim_{X\to\infty} \frac{\mathrm{vol}(A(X))}{X} = 1. \tag{8}$$

100 以下： 13 対 11・・・ 3 組が 2 点リード
1000 以下： 87 対 80・・・3 組が 7 点リード
1 万以下： 619 対 609・・・3 組が 10 点リード
10 万以下： 4808 対 4783・・・3 組が 25 点リード
100 万以下： 39322 対 39175・・・3 組が 147 点リード
200 万以下： 74516 対 74416・・・3 組が 100 点リード
300 万以下： 108532 対 108283・・・3 組が 249 点リード

この予想は，実数全体の中で「3 組が勝つ x が占める割合は 100%」という主張である．データからはこれも正しそうに思える予想だが，証明に向けた進展は全く得られないまま 30 年以上が経過した．そして 1995 年，この予想は以下の定理によって事実上，否定された．

カゾロフスキーの定理 (1995)

ディリクレ L 関数 $L(s,\chi)$ に対するリーマン予想を仮定すると，ナポウスキー・テュランの予想の極限値 (8) は，存在しない．

こうして「チェビシェフの偏り」の証明どころか定義すら不明のまま，長い年月が経過した．数値的なデータからは「3 組優勢」の状況が歴然としているのに，その圧倒的な偏りを数学的に定式化できなかったのである．

そんな状況下で登場したのが，ルビンスタインとサルナックである．彼らは，リーマン予想に加えて「線形独立予想」を仮定した．これは，L 関数の非自明零点の虚部が，有理数体 $\mathbb{Q}$ 上互いに線形独立であるという予想である．一般に，ゼータ関数や L 関数の非自明零点たちは超越的であり，互いに一切の線形関係が無いと考えられているため，この予想の真偽については賛否が無いと思われる．ただ,「非自明零点の虚部の線形独立性」を扱うこの予想は，実部を規定する「リーマン予想」の先にある予想であり，現状では証明の糸口すら不明で，解決の見込みは全くない．夢のまた夢である．

ルビンスタイン・サルナックの定理 [6]

ディリクレ L 関数 $L(s,\chi)$ に対するリーマン予想と線形独立予想を仮定すると，区間 $A(X)$ の対数密度は，$0.9959\cdots$ に収束する．すなわち，次式が成り立つ．

$$\lim_{X\to\infty}\frac{1}{\log X}\int_{t\in A(X)}\frac{dt}{t}=0.9959....$$

対数密度という名称は，仮に $A(X)$ が全区間だとすると，上の積分が

$$\int_{t\in A(X)}\frac{dt}{t}=\int_1^X\frac{dt}{t}=\log X$$

と対数を用いて表されることに由来する．対数密度は，対象を実数から自然数に代えても定義できる．彼らの結論は，自然数の密度として述べると，次のようになる．

$$\lim_{X\to\infty}\frac{1}{\log X}\sum_{n\in A(X)}\frac{1}{n}=0.9959....$$

左辺の $\sum_{n\in A(X)}\frac{1}{n}$ は，チェビシェフが考えた「重み付き個数」の類似とみなせる．なぜなら，自然数を一様に「1 個」ずつ数える代わりに，各自然数 n を「$\frac{1}{n}$ 個」と数えて「重み付き個数」を求めた式と似ているからである．ただ，素数や自然数の個数を直接数えているわけではなく,「x 以下で数えたときに 3 組リードとなるような x」の個数を考えている点が異なる．この点は，次節および後編で説明する「先行研究の欠陥」に関連する．

彼らの論文 [6] が発表されて以来，素数全体が，同じ大きさの 2 つの組に分かれている場合の「チェビシェフの偏り」の定義は，

一方の組がリードする区間の対数密度が，1/2 を超えること

とされるようなった．1/2 は，各組が素数全体の中に占める自然密度から来ている．

これを一般化すると，必ずしも同じ大きさでない複数の組に素数全体が分けられているときの「チェビシェフの偏り」は，次のように定義される．

特定の組の素数がリードする区間の対数密度が，その組が素数全体の中で占める自然密度よりも大きいこと．

一応の定義はできたものの，偏りが発生する理由は未解明のまま，30 年の歳月が流れた．この間，世界中の研究者によって，この定義に則った「チェビシェフの偏り」に関する研究論文が，数多く出版された．

4　先行研究の欠陥

論文 [6] によって得られた「チェビシェフの偏り」の定式化は，はたして正しいのだろうか．私は，以下の 3 つの理由から，この定義は不十分であると考える．

1. 「リードする区間の長さの割合」を用いても,「偏りの大きさ」は表現できない．
2. 仮定の「線形独立予想」が唐突で，不自然である．
3. 1/2 という境界値は，真実の $0.9959\cdots$ と，かけ離れている．

次号の後編では，これらの欠陥について解説した後，1〜3 のすべてを解消した「チェビシェフの偏り」の新たな定式化を提唱する．そして，実際に深リーマン予想の仮定下で，偏りの存在を証明する．

また，以上の着想は広範なゼータ関数・L 関数に拡張が可能である．したがって，様々な一般化により，これまで知られていなかった新たな偏りを発見でき，素数分布論への貢献を得るのである．

なお，本稿および次号の後編で報告する内容の一般向けの解説が，拙著 [8] にある．興味のある方は参照されたい．

参考文献

[1] P. L. Chebyshev: Oeuvres, vol. 1, p.697.

[2] G. H. Hardy and J. E. Littlewood: “Contributions to the theory of the Riemann zeta function and the theory of the distribution of primes” Acta Math. **41** (1917) 119-196.

[3] J. E. Littlewood: “Sur la distribution des nombres premiers” Comptes Rendus de l’Acad. Sci. Paris **158** (1914) 1869-1875.

[4] E. Landau: “Über einige ältere Vermutungen und Behauptungen in der Primzahltheorie I” Math. Z. **1** (1918) 1-24.

[5] E. Landau: “Über einige ältere Vermutungen und Behauptungen in der Primzahltheorie II” Math. Z. **1** (1918) 213-219.

[6] Rubinstein and P Sarnak: “Chebyshev’s bias” Exprimental Math. **3** (1994) 173-197.

[7] 黒川信重：「リーマン予想の探求 〜ABC から Z まで〜」技術評論社．2012 年

[8] 小山信也：「素数って偏ってるの？」技術評論社．2023 年