

研究員 の眼

ジップの法則

法則性をもとに、将来の紛争発生を予測する !?

保険研究部 主席研究員 篠原 拓也
(03)3512-1823 tshino@nli-research.co.jp

メディアでは、ランキングをよく目にする。いま売れている本のランキング、インターネットのニュースアクセスランキング、ハードなスポーツランキング、人気温泉地ランキングなど、いろいろなものがランキング形式で紹介されている。

ランキングの定め方も、数量をもとにしたものだけでなく、アンケート、人気投票などさまざま。何かの数量をもとにしたランキングの場合、順位とその数量の間に、ジップの法則といわれる関係がみられることがある。本稿では、ジップの法則について、みていこう。

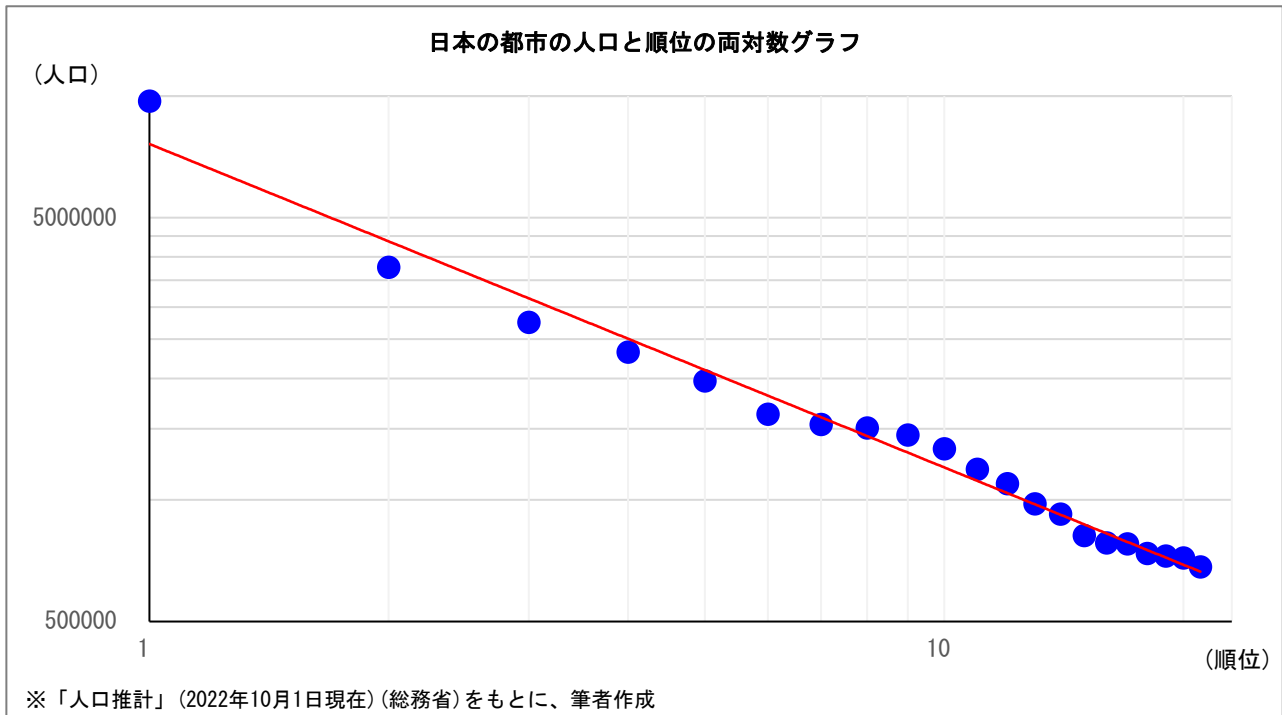
◇ n 番目に大きな都市の人口は、最大都市の人口の $1/n$ に比例する

まず、ジップの法則の内容を見ていこう。ある数量をもとに、ランキングをとることにする。このとき、第2位のものの数量は、第1位ものの数量の $1/2$ に比例する。第3位のものの数量は、第1位ものの数量の $1/3$ に比例する。……第 n 位のものの数量は、第1位ものの数量の $1/n$ に比例する。これが、法則の内容だ。

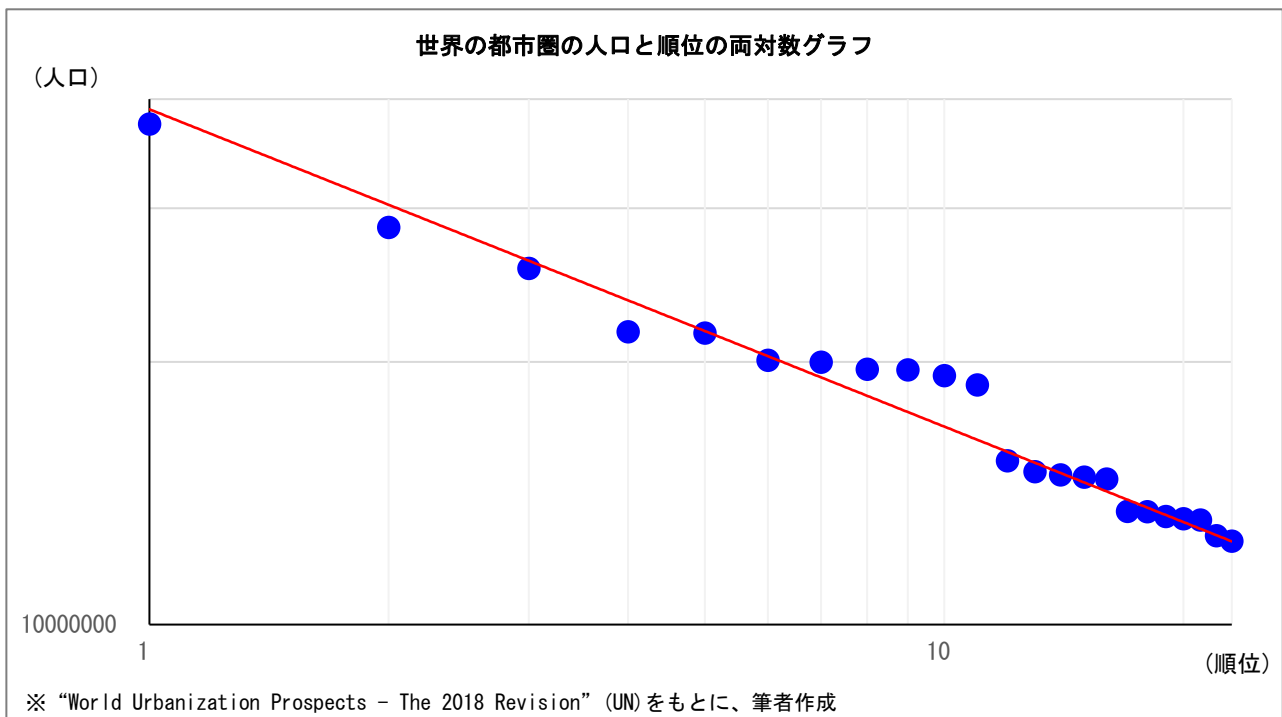
この法則は、経験則(経験の集積から得られた法則)だ。もちろん、どんなランキングにも成り立つというわけではない。もし成り立つ場合でも、数学や物理などの法則と違って、厳密に $1/2$ 、 $1/3$ 、…、 $1/n$ に比例するというのではなく、それらのべき乗に比例することが一般的だ。だが、順位とその数量の間に、簡単な関係が見出されるというのは、なかなか興味深い。

この法則が成り立つかどうかをグラフで見るときは、横軸に順位の対数、縦軸に数量の対数をとる“両対数グラフ”で表すとわかりやすい。法則が成り立つときには、右肩下がりの直線となるからだ。

例えば、都市の人口のランキングで、この法則は成り立つといわれる。実際に、日本の都市の人口をもとにランキングをとり、両対数グラフで表してみると次の通りとなった。実績データ(青色の丸)に対して、近似線(赤線)が右肩下がりの直線として描かれている。



同様に、国連の報告書に掲載されている世界の都市圏の人口をもとにランキングをとり、両対数グラフで表してみると次の通りとなった。日本の都市と同様に、近似線が右肩下がりの直線として描かれている。(なお、都市圏の定義は国ごとに異なるため、その点の注意が必要だ。)



◇ 単語の出現頻度、科学誌の執筆数、生物学の種の数など、様々なもので、ジップの法則が成り立つ

ジップの法則は、アメリカの言語学者ジョージ・キングスリー・ジップ氏が1942年の論文で示した内容によるとされる。戦後、この法則は、多くの分野の研究者を巻き込みながら、議論されてきた。

その中でも有名なものが、1978年にノーベル経済学賞を受賞した、ハーバート・サイモン氏による議論だ。経済学のみならず、政治学、認知心理学、経営学、情報科学など多くの分野を究めた同氏は、1955年の論文のなかで、文中での単語の出現頻度、科学誌における著者別の執筆数、都市の人口、収入の分布、生物学における種数を例に、この法則が幅揃い分野の経験データに成り立つことを述べている。

特に、単語の出現頻度については、言語科学分野の他の研究者から寄せられた批判に対して、鋭い反駁(はんぱく)を展開している。

ただ、どうしてジップの法則が成り立つのか、という点については、分野横断的な説明はまだ得られていない。各分野の研究者が、その解明に向けて、取り組んでいる状況と言われている。

◇ ジップの法則は紛争発生の予見にも用いることができる !?

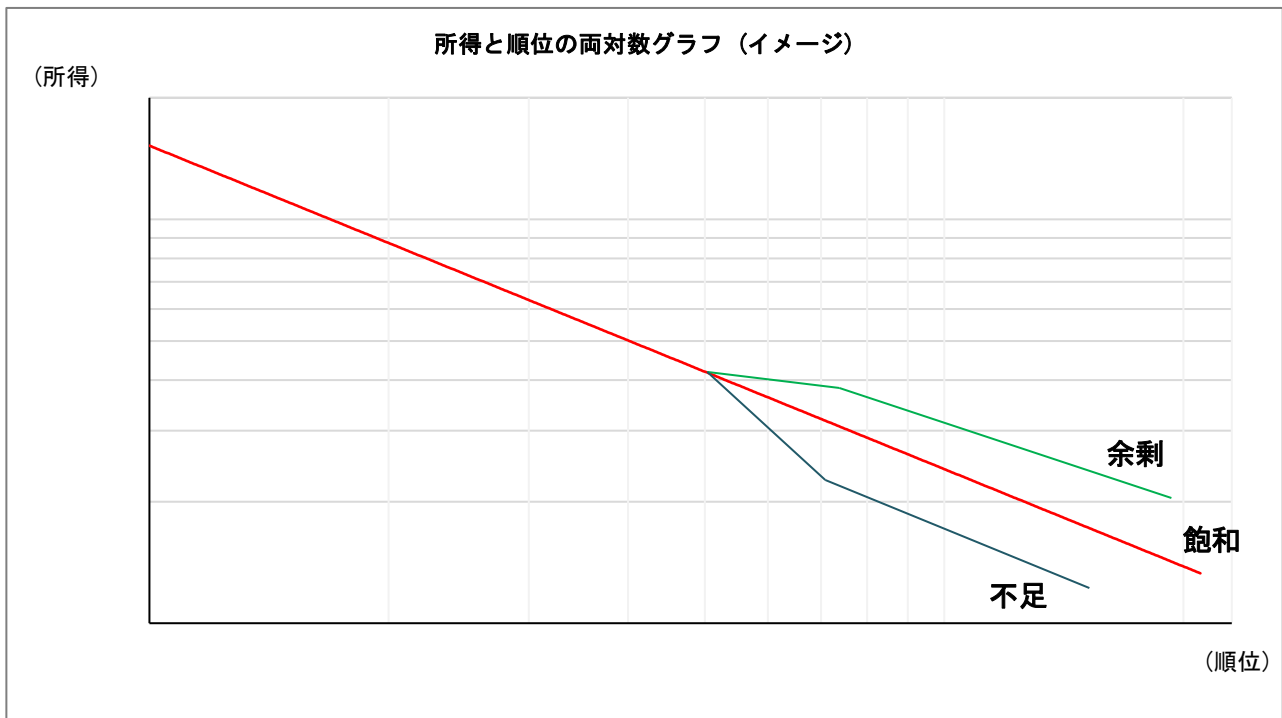
それでは、ジップ氏自身の研究はどうだったのか。彼は、言語学者として、単語の出現頻度の研究を進めた。1922年出版の、アイルランドの作家ジェイムズ・ジョイスによる長編小説「ユリシーズ」について、使用されている26万語以上の出現頻度を調べて、この法則を確かめている。

そして、彼も言語科学分野にとどまらず、様々な分野で、この法則が成り立つことを述べている。特に、所得分布をもとに、国家の統一や分裂について論じている。その流れはこうだ。

まず、所得のランキングをもとに、その分布を両対数グラフで表す。データが右肩下がりの直線上に分布していれば、飽和(saturation)の状態だという。そして、もしある所得層で、右肩下がりの直線の上側に膨らんでいる部分があれば、余剰(surfeit)の状態。逆に、直線の下側にへこんでいる部分があれば、不足(deficient)の状態だという。

そのうえで、欧米の国々について、過去の所得分布を振り返っている。特に、ドイツの所得分布について、1920~30年代の過去のデータを確認しており、1932年と1934年の曲線で、下位中産階級の所得層で不足、それより下の所得層で余剰が生じていることを示した。これは、下位中産階級の所得層の人々が所得を減らしたためと分析している。

この余剰の層は、1910年代末のドイツ革命で組織化された所得層であり、地元の商店主、職人、貧しい専門職など、一般の人々を含んでいたとしている。そして、さらにその下には、グラフにはあらわれない大規模な失業者のグループがいたという。



ジップ氏は、著書の中で、「実際、所得分布におけるこの種の湾曲の出現は、革命前の状態の兆候である可能性が高く、これらの湾曲が消失しないことは、国家社会経済にとって不吉なことかもしれない」(※)と述べている。(ジップ氏は書籍の中で言及してはいないが) これらの余剰の層や失業者の層が、1933年のナチス台頭の遠因であったことが示唆される。

(※) 原文では以下の通り。

“Indeed the emergence of bends of this sort in an income-distribution may well be symptomatic of a pre-revolutionary condition, and the failure of these bends to disappear may be ominous for our national social-economy.”

(“National Unity and Disunity - The Nation As a Bio-Social Organism” George Kingsley Zipf (The Principia Press, 1941)の Chapter Five より)

少し、見方を変えてみよう。所得分布がジップの法則から乖離していることを確認することにより、社会のなかで格差が高じていることを把握する。そして、そのことが将来の紛争発生の予見にも用いる、といった活用方法が考えられるかもしれない。

もちろん、所得分布を紛争発生に直接結びつけて予見を行うというのは、いささか短絡的ともいえるだろう。ただ、所得分布の歪みを、紛争事態の端緒として捉えることには、アーリーウォーニングとしてそれなりに意味があるということもできよう。

ジップの法則は、いまから 80 年以上も前に示された経験則だ。現代では当たり前のスマートフォン、パソコン、インターネット、クラウド、AI はおろか、電卓すら開発されていなかった時代に、人類の叡智を駆使して、限られたデータやツールをもとに発見された法則といえるだろう。

しかし、その内容は、データ処理や AI ツールが整った現代でも色あせていない。この経験則は、いまでも、データを読み解くうえでの参考ツールとして、活用の道が考えられるだろう。

(参考文献)

「直感を裏切る数学 — 『思い込み』にだまされない数学的思考法」神永正博(ブルーバックス B-1888, 講談社, 2014 年)

「人口推計」(2022 年 10 月 1 日現在)(総務省)

“World Urbanization Prospects - The 2018 Revision” (UN)

“On a class of skew distribution functions” Herbert A. Simon (Carnegie Institute of Technology, 1955)

“National Unity and Disunity - The Nation As a Bio-Social Organism” George Kingsley Zipf (The Principia Press, 1941)