

Power's Law とランク付け

2005MM070 杉浦加奈

指導教員：尾崎俊治

1 はじめに

今日、わたしたちの周りには多くの数字で溢れている。都市の人口、所得、アンケート結果などあらゆるものに順位付けがされている。ランク付けされたものの並び方や数字の大きさには法則性を持つものも少なくない。人の行動の下で起こったものにさえ法則性が適応されているものも存在する。本論文では、そのような法則性について研究、考察していこうと思う。

2 Power's Law とは

Zipf's Law, Pareto's Law, Benford's Law など観測量がパラメータのべき乗に比例するという法則である。

3 Zipf's Law について

サイズが k 番目に大きい要素が 1 番大きい要素の $1/k$ に比例するという経験則である。アメリカの言語学者ジョージ・キングズリー・ジップが提唱した言語学上の法則である。ジップは英単語の出現率を調べるために、ジェームズ・ジョイスの「ユリシーズ」に含まれる単語 (260,430 個) と、いくつかの新聞記事の語彙数 (43,989 個) を比べた。すると、頻出する単語は次のような順位になった ([1] 参照)。

- (1) the (全体の 10%)
 - (2) of (全体の 5%)
 - (3) and (全体の 3.3%)
- 単語の出現率 (%) = 10 / 順位

その他に以下の現象は Zipf's Law に適応している

- ・ ウェブページへのアクセス頻度
- ・ 都市の人口
- ・ 上位 3% の人々の収入
- ・ 音楽における音符の使用頻度
- ・ 細胞内での遺伝子の発現量
- ・ 地震の規模
- ・ 固体が割れたときの破片の大きさ

4 オイラー定数と Zipf's Law

オイラー定数 γ は (1) 式で表わされる。

$$\gamma = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \log n\right) \quad (1)$$

$$\gamma = 0.5772156649\dots ([3] \text{ 参照})$$

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} \approx \log n + \gamma \quad (2)$$

Zipf's Law による値の総和は

$$S_n = m \sum_{k=1}^n \frac{1}{k} \quad (m: 1 \text{ 番大きい要素}) \quad (3)$$

$$S_n = m \left(\sum_{k=1}^n \frac{1}{k}\right) \approx m(\log n + \gamma) \quad (4)$$

となりオイラー定数を用いて表すことが出来る。

5 ゴルフの賞金に関する Zipf's Law

2008 年 10 月 16 日 (水) ~ 10 月 19 日 (日) に行われた日本オープンゴルフ選手権競技での第 3 ラウンド進出選手 (第 2 ラウンド 7 2 ホール終了時に 60 位以内の選手) の順位に対する賞金に関して Zipf's Law が成り立っている。

実際の獲得賞金は優勝の選手以外の同率順位の場合の賞金は、その人数分の下位までの賞金の平均額となる。

また、アマチュアが取得順位に該当した場合は次位のプロフェッショナルに振りあてられる ([3] 参照)。

5.1 Zipf's 分布

x 軸に順位 y 軸に賞金とするジップ分布は以下のようになる。

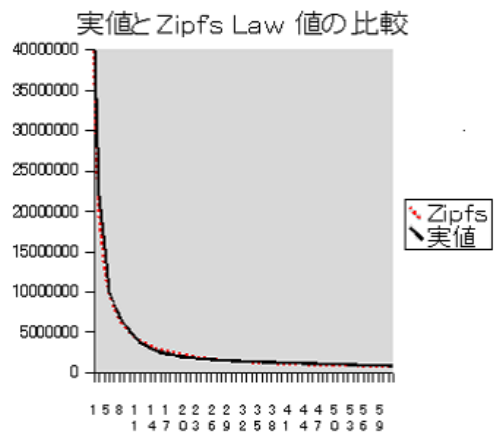


図 1 Zipf's Law 予測値と実際の値の Zipf's 分布

図 1 のグラフでは Zipf's Law の予測値とほぼ同じ曲線を描いていることがわかる。よってゴルフの賞金は Zipf's Law に従っている。

6 Pareto's Law

経済において、全体の数値の大部分は、全体を構成するうちの一部の要素が生み出しているという法則である。この法則は多数ある項目の中で「重要なものはごくわずかしかな」という基本理念で成り立っている。商品の品質管理の分野で活用されている例で、重点的に改良す

べきものを重要度のあるものから順番に10項目あげる。そして、その最上位の2つの項目だけを改良する。すると、効果は、全体の80%を改良したのと同等の結果が期待できるというものである ([1] 参照)。

6.1 パレートの法則が用いられる例

- ・ビジネスにおいて売り上げの8割は全顧客の2割が生み出している。
売り上げを伸ばすには上の2割の顧客に的を絞ったサービスを行うことが効果的である
- ・所得税の8割は、課税対象者の2割が担っている。

6.2 ロングテール現象

ロングテール現象とは、売上金額の上位20%に入らない残りの80%に属する少数のアイテムの方が、上位20%で売れる数よりもはるかに多くなるような現象のことである。

これまでのビジネスでは、上の例のように20:80の法則で、上位20%のものに注力すればだいたいOKと考えられてきたが、最近では残りの80%の部分でビジネスをして収益出すというモデルも成立してきている。

今までのオフライン小売店では在庫の制限などでこの上位20%に当たる商品を多く揃えなければならず、その他80%は軽視されることが多かった。しかし、Amazon.comなどのオンライン小売店は在庫や物流にかかるコストが従来の小売店と比べて遥かに少ないので今まで見過ごされてきたこの80%をビジネス上に組み込むことが可能になった。そこからの売り上げを集積することにより新たなビジネスモデルを生まれた。

ロングテール・ビジネスの発展には「すべての商品が手に入るようにする」、「ほしい商品を見つける手助けをする」という要求を満たし、インターネットの普及が広まったことが大きな要因であると思う。PCだけでなく携帯端末の普及も影響している ([4] 参照)。また、アマゾンには「バーチャル在庫」を設け、選択肢を拡げることを可能にした。デジタルの在庫はコストが低く済む。そして注文や個人情報顧客が入力することにより人件費も抑えることができる。

7 Benford's Law

あるデータの中の数字の最大桁数は1が多いという法則である。数字 N が登場する確率は

$$\log_{10}\left(1 + \frac{1}{N}\right) \quad (5)$$

で表される。

- $N=1$ になるのは $\log_{10}(2) = 0.3010$
- $N=2$ になるのは $\log_{10}\left(\frac{3}{2}\right) = 0.1761$
- $N=3$ になるのは $\log_{10}\left(\frac{4}{3}\right) = 0.1249$
- $N=4$ になるのは $\log_{10}\left(\frac{5}{4}\right) = 0.0970$
- $N=5$ になるのは $\log_{10}\left(\frac{6}{5}\right) = 0.0791$
- $N=6$ になるのは $\log_{10}\left(\frac{7}{6}\right) = 0.0670$
- $N=7$ になるのは $\log_{10}\left(\frac{8}{7}\right) = 0.0579$
- $N=8$ になるのは $\log_{10}\left(\frac{9}{8}\right) = 0.0512$

$$N=9 \text{ になるのは } \log_{10}\left(\frac{10}{9}\right) = 0.0458$$

7.1 フィボナッチ数列における Benford's Law

フィボナッチ数列 (1,1,2,3,5,8,13,21,...)

表1 フィボナッチ数列における最大桁数登場回数

n	回数	割合
1	29	0.315
2	17	0.179
3	12	0.126
4	9	0.095
5	7	0.074
6	6	0.063
7	5	0.053
8	6	0.063
9	4	0.042
合計	95	

表1と Benford's Law に従う値はほぼ等しくフィボナッチ数列は自然発生の際の並びである。

Benford's Law を用いることによって数字が意図的に作られたものか、自然発生によるものかを判別することが出来る。必ずしも不正に作られた数字を当てることが出来るわけではないが、株の取引や、粉飾決算などの不正発覚などに深く関係していると思う。

8 おわりに

Power's Law が成り立つものには3つの特徴があると思う。1つ目は「不平等性」である。母数のうちの一部分が圧倒的多数を占拠する状況を生み出している。2つ目は「自己相似性」である。部分の性格が全体の性格と一致するのである。ゴルフの賞金に関する Zipf's Law では賞金の額と順位の対数の関係が常に傾き-1を推移している。3つ目は「無限大の分散」である。Power's Law が成り立つものは母数を増やすほど分散が無限大に発散する。Power's Law にはさまざまな現象をべき乗に成り立たせる不思議がある。べき乗則の分布はいくらエネルギーを入力してもどんどん圧縮されて無限に分散して定常状態となる。人間が考えて行動することまでも母数が大きくなればなるほどべき乗則に従う結果になると思う。これからも Power's Law に従う現象とそれを破り新たに変わってく現象などを調べ続けたいと思う。

9 参考文献

- [1] ウィキペディア <http://ja.wikipedia.org>
- [2] JGA 日本ゴルフ協会 <http://jga.or.jp>
- [3] ジョン・タバク (松浦 俊輔 訳):『数学と自然法則』。青土社、東京、2005。
- [4] ロングテールとは <http://johou.net/syoseki/longtail.htm>