

書籍販売市場における隠れた法則性

井庭 崇[†] 深見 嘉明^{††} 斉藤 優^{†††}

本論文の目的は、商品の売れ行きの背後にある隠れた法則性を探ることにある。書籍販売市場は、すべての商品が同じように売れるわけではなく一部の商品が爆発的に売れるというウィナー・テイク・オール市場になっている。本論文では、商品販売市場に潜む隠れた法則性を明らかにするために、日本全国における書籍販売の実データを用いて実証的に分析する。その結果、販売冊数と順位の関係がべき乗則に従っていることが明らかになった。また、ジャンルによっては、べき乗分布から乖離していることがわかり、ジャンルの特徴を理解できることが示唆された。

Hidden Law in the Book Sale Market

TAKASHI IBA,[†] YOSHIAKI FUKAMI^{††} and MASARU SAITOH^{†††}

In this paper, we explore the hidden law in the book sale market in Japan. The book sale market is known as "Winner-Take-All market" in which a very small number of the books are extremely sold although the rest of them are hardly sold. In this paper, we analyze the empirical data of bookstores across Japan, in order to reveal the hidden law. The results show that the relation between sale and rank follows power law. In addition, we observe the alienation between the empirical distribution and power law in some category.

1. はじめに

現在、日本に流通している書籍は12億冊以上といわれ、年間に発行される新刊タイトル数も7万7千点にのぼる。書籍は一般的な物財と異なり、商品の価値・効用が人によって多様であり、また流通する商品種類の絶対数が大きいという特徴がある。このような状況において、顧客の多様なニーズに応えるため、大規模書店では数十万タイトルを常備しているというが、それらの書籍が同じように売れるわけではない。書籍販売市場は、一部の商品が爆発的に売れるという「ウィナー・テイク・オール市場」¹⁾ だといわれている。このような経験則を実証するために、本研究では、日本全国における書籍販売の実データを解析し、商品販売市場に潜む隠れた法則性を明らかにする。

2. 書籍全体の分析

書籍販売市場の隠れた法則性を明らかにするために、

まず最初に書籍全体の販売データを分析することにした。本論文では、全国に分布する2,000書店以上のPOS(販売時点情報管理)システムの実データを用いる。対象期間は2006年5月の1ヶ月とする。その期間内に購入された書籍のタイトル数は、合計で300,000タイトルを超えている。

販売冊数と順位の関係进行分析するために、横軸に順位をとり、縦軸に販売冊数の割合(そのタイトルの販売冊数を書籍全体の販売冊数で割ったもの)をとってグラフ化したものが図1である。図中の+印が実データをプロットした点である。このグラフをみると、順位の上位にあるものは非常によく売れ、それ以外の大多数はほとんど売れていないということがわかる*。

これを両対数グラフにプロットしたものが図2である。両対数グラフでは、縦軸も横軸も対数軸となっており、それぞれ目盛りが1上がるごとに、10倍ずつ値が増加していく。この図から、書籍の販売冊数と順位との関係がべき乗分布になっている(両対数グラフ上で直線的になっている)ことがわかる**。

[†] 慶應義塾大学 総合政策学部

Faculty of Policy Management, Keio Univ.

^{††} 慶應義塾大学 政策・メディア研究科

Graduate School of Media and Governance, Keio Univ.

^{†††} 日本出版販売株式会社 システム部

Nippon Shuppan Hanbai, Inc.

* 図1のようなべき乗分布のグラフにおいて、右側に伸びた裾野は「ロングテール」(long tail)と呼ばれることがある。特にインターネット書籍販売市場の特徴に関する議論で取り上げられて話題となっている²⁾。

** べき乗分布では、正規分布の場合と異なり、平均や分散が事実上

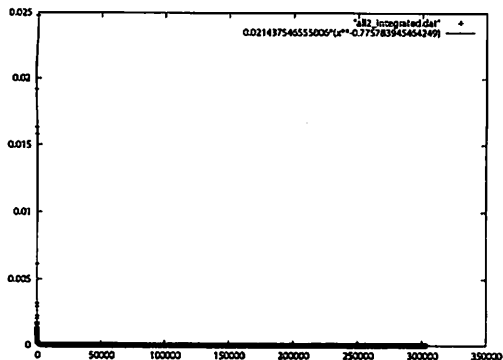


図1 書籍全体の販売冊数(割合)と順位の関係(線形グラフ)

図2に描かれたグラフの直線は、 $S(r) = \alpha r^{-\beta}$ で表される近似式であり、この図では $\alpha = 0.0214, \beta = 0.7758$ となっている(区間 $1 \leq r \leq 10000$ において決定係数が0.9965)。最上位の販売冊数(割合)は、第1位『ハリー・ポッターと謎のプリンス』が0.0247、第2位『ダ・ヴィンチ・コード(上)』が0.0191、第3位『ダ・ヴィンチ・コード(中)』が0.0163、第4位『ダ・ヴィンチ・コード(下)』が0.0158、第5位『国家の品格』が0.0061となっている。このグラフの右端にある最下位では0.0000000582289と、かなり少数の販売冊数になっている。

直線から乖離している部分は、実データでは規模が有限であるために生じるカットオフ領域だと考えられる¹⁰⁾。なお、このグラフでは第2位と第3位が直線上から外れているが、これはこのタイトルが上・中・下巻に分けて販売されていることの影響である。上巻を購入する顧客の多くが、中巻や下巻も購入することになるため、ほぼ横ばいの販売冊数となるのである。

同じデータを用いて、別の角度からの分布を描くこともできる。ある冊数以上売れたタイトルがいくつあるのかをプロットしたのが図3である。このグラフは、販売冊数を横軸にとり、その冊数以上売れたタイトルの数(割合)を縦軸にとって、両対数グラフとして描いたものである。この累積分布のグラフにおいても、べき乗分布になっていることがわかる。

ゼロや無限大になってしまう。自然科学や社会科学の様々な領域において「べき乗分布」が発見されることから、新しい原理として注目を集めている³⁾。べき乗分布になる現象には、単語の出現頻度と順位の関係(ジップの法則)、地震の規模と順位の関係(グーテンベルク=リヒター則)、岩石を破砕した際に生じる破片のサイズと順位⁴⁾、自然界や社会におけるネットワークのノードの次数とその順位の関係⁵⁾、都市人口の規模と順位⁶⁾、個人や企業の所得分布⁷⁾⁸⁾、価格変動の規模と頻度の関係⁹⁾などがある。

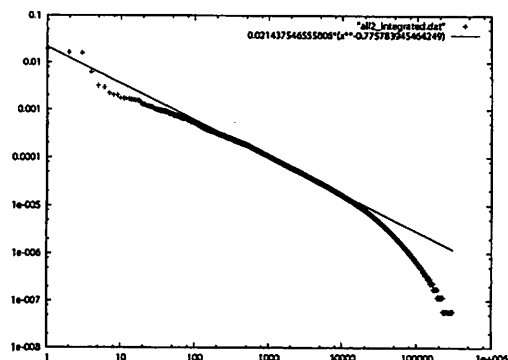


図2 書籍全体の販売冊数(割合)と順位の関係

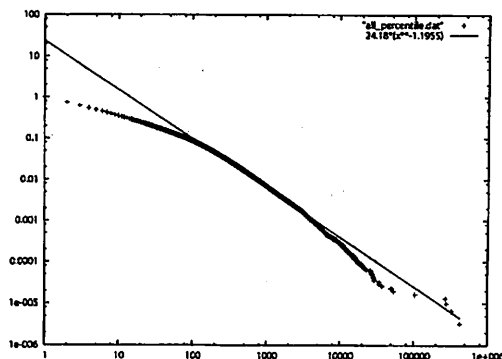


図3 書籍全体の販売冊数分布

3. ジャンル別の分析

書籍のジャンルによって、その分布に違いはあるのだろうか。以下では、ジャンル別に販売冊数と順位の関係性を分析していくことにする。ここでは、広く用いられている日本図書コードの分類コード(Cコード)にもとづいてジャンルを分類して分析を行う*。

3.1 文学

内容の大分類が「文学」である書籍をプロットすると図4のようになる。このグラフを見ると、書籍の販売冊数(割合)と順位との関係がべき乗分布になっていることがわかる(グラフ上の直線は $\alpha = 0.0546, \beta = 0.8752$ であり、区間 $1 \leq r \leq 5000$ において決定係数0.9865)。最上位の販売冊数(割合)は、第1位『ハ

* この分類コードには、「販売対象」、「発行形態」、「内容」のコードがある。販売対象コードには、「一般」や「学参II:高校生対象」などの分類がある。発行形態コードには、「単行本」や「新書版」などがある。内容コードは、大分類として「社会科学」、「自然科学」、「工学・工業」、「芸術・生活」、「文学」などがある。それぞれの大分類には、さらに細かい中分類がある。例えば「自然科学」には、中分類として「物理学」、「化学」、「生物学」などがある。

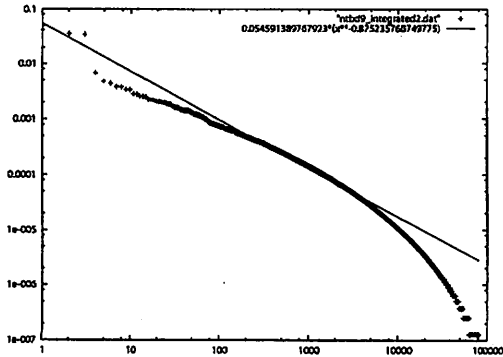


図4 文学における販売冊数(割合)と順位の関係

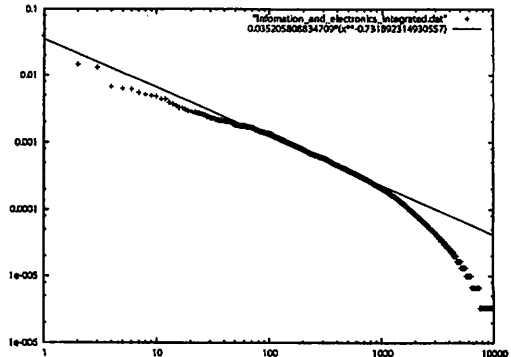


図6 情報・通信における販売冊数(割合)と順位の関係

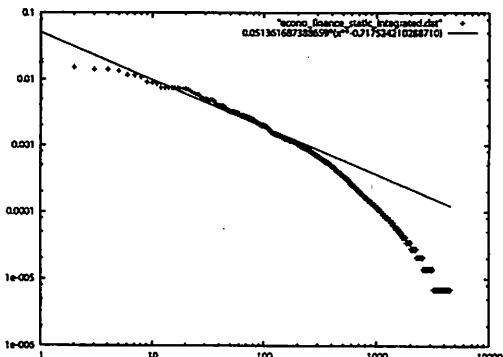


図5 経済・財政における販売冊数(割合)と順位の関係

リー・ポッターと謎のプリンス』が0.0535, 第2位『ダ・ヴィンチ・コード(上)』が0.0413, 第3位『ダ・ヴィンチ・コード(中)』が0.0353, 第4位『ダ・ヴィンチ・コード(下)』が0.0341, 第5位『東京タワー』が0.0068となっている。最上位の点が横ばいになっているのは、書籍全体のグラフと同じ理由による。

3.2 経済・財政・統計

内容の大分類が「社会科学」で中分類が「経済・財政・統計」の書籍のうち、販売対象が学習参考書であるものを除いてプロットすると図5のようになる。このグラフを見ると、書籍の販売冊数(割合)と順位との関係がべき乗分布になっていることがわかる(グラフ上の直線は $\alpha = 0.0514, \beta = 0.7175$ であり、区間 $1 \leq r \leq 200$ において決定係数0.9690)。

このグラフでは、第1位『一番売ってる株の雑誌ダイヤモンドザイが作った』が0.0217, 第2位『株式投資これだけはやってはいけない』が0.0178, 第3位『株価チャート練習帳』が0.015, 第4位『たった7日で株とチャートの達人になる!』が0.0141, 第5位『ヤバい経済学』が0.0140となっている。このジャン

ルの上位は、ほとんどが株や投資に関する書籍である。

3.3 情報・通信

内容の大分類「総記」の中分類「情報科学」、および大分類「工学・工業」の中分類「電子通信」に含まれる書籍のうち、販売対象が学習参考書であるものを除いてプロットすると図6のようになる。このグラフを見ると、書籍の販売冊数(割合)と順位との関係がべき乗分布になっていることがわかる(グラフ上の直線は $\alpha = 0.0778, \beta = 0.8903$ であり、区間 $1 \leq r \leq 1000$ において決定係数0.9822)。

最上位の販売冊数(割合)は、第1位『99・9%は仮説』が0.0420, 第2位『グーグル完全活用本』が0.0393, 第3位『グーグル明解検索術』が0.0144, 第4位『ヤフー・ジャパン完全活用本』が0.0130, 第5位『超図解ビギナーズエクセル』が0.0066となっている。このジャンルの上位は、ほとんどがコンピュータ・ソフトウェアの使い方に関する解説書である。

3.4 物理学

内容の大分類が「自然科学」の中分類「物理学」の書籍のうち、販売対象が学習参考書であるものを除いてプロットすると図7のようになる。このグラフを見ると、書籍の販売冊数(割合)と順位との関係がべき乗分布になっていることがわかる(グラフ上の直線は $\alpha = 0.3077, \beta = 1.1547$ であり、区間 $1 \leq r \leq 500$ において決定係数0.9681)。最上位の販売冊数(割合)は、第1位『新しい高校物理の教科書』が約0.1000, 第2位『人物で語る物理入門』が約0.0416, 第3位『人物で語る物理入門』が約0.0338となっている。これらは、どれも新書版書籍である。

3.5 化学

内容の大分類が「自然科学」の中分類「化学」の書籍のうち、販売対象が学習参考書であるものを除いてプロットすると図8のようになる。このグラフを見る

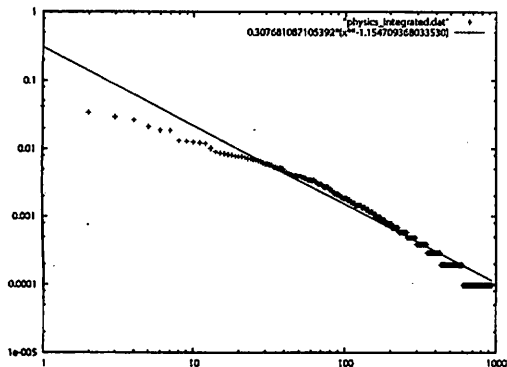


図7 物理学における販売冊数(割合)と順位の関係

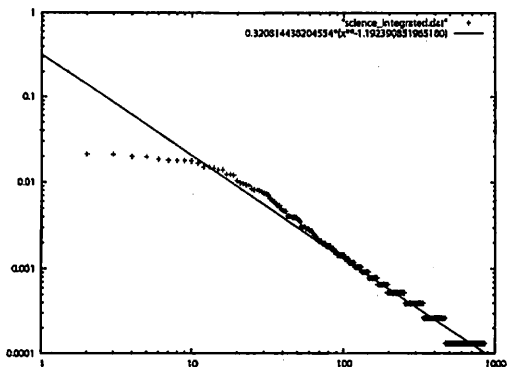


図8 化学における販売冊数(割合)と順位の関係

と、書籍の販売冊数(割合)と順位との関係が、約10位以降でべき乗分布になっていることがわかる(グラフ上の直線は $\alpha = 0.3208, \beta = 1.1924$ であり、全区間 $1 \leq r \leq 857$ において決定係数 0.9732)。

最上位の販売冊数(割合)は、第1位『新しい高校化学の教科書』が0.1238、第2位『化学の不思議がわかる本』が0.0404、第3位『アトキンス物理化学(上)』が0.0212、第4位『入門 機器分析化学』が0.0210、第5位『アトキンス物理化学(下)』が0.0199となっている。このグラフでは約10位までが横ばいになっているが、これは化学の上位に大学の教科書・参考書となる書籍が続いているためであろう。

この傾向を、物理学の場合と比較してみると、興味深いことがわかる。物理学においても、第8位『熟・統計力学』や第12位『解析力学』などの書籍が上位に存在するものの、物理学の上位の多くは、相対性理論や量子力学などを扱った新書版や単行本である。つまり、物理学では、相対性理論や量子力学などの解説書が一般に広く読まれているため、その分布がべき乗分布に近づいているということがわかる。

4. おわりに

本論文では、書籍販売市場における販売冊数と順位の関係にべき乗分布が観察されることを実証的に示した。言語、都市、地震などにおいて知られているこの法則性が、商品市場においても観察できるということを示した点で、画期的な報告であると思われる。このようなべき乗分布がどのようなメカニズムで生成されるのかについてはまだ明らかにされていないが、私たちは「自己組織的臨界現象」¹¹⁾が関係していると考えている。今後は、実データを用いた詳細な分析を進めるとともに、このような分布が生じるメカニズムのモデル化にも取り組んでいきたい。その意味で、本論文は、商品市場の新しい「見えざる手」を理解するための第一歩と位置づけることができるだろう。

謝 辞

本研究は、日本出版販売株式会社と慶應義塾大学SFCの共同研究として行われたものである。日本出版販売株式会社システム部の藤重雅継さん、早津昌輝さん、松島崇さん、SFC研究所訪問研究員の高部陽平さんに感謝の意を述べたい。また研究の遂行にあたり、高安秀樹さんに貴重なアドバイスをいただいた。ありがとうございました。

参 考 文 献

- 1) Frank, R. H. and Cook, P. J.: *The Winner-Take-All Society*, The Free Press.
- 2) Anderson, C.: *The Long Tail*, Hyperion (2006).
- 3) 公文俊平: 情報社会学序説: ラストモダンの時代を生きる, NTT出版(2004).
- 4) 高安秀樹: 経済物理学の発見, 光文社新書(2004).
- 5) Barabási, A.-L.: *LINKED: The New Science of Networks*, Perseus Book Group, Perseus Book Group.
- 6) Simon, H.: On a Class of Skew Distribution Functions, *Biometrika* (1995).
- 7) Pareto, V.: *Cours d'economie politique*, Macmillan (1897).
- 8) 高安秀樹, 高安美佐子: 経済・情報・生命の臨界ゆらぎ, ダイヤモンド社(2000).
- 9) 高安秀樹, 高安美佐子: エコノフィジックス: 市場に潜む物理法則, 日本経済新聞社(2001).
- 10) Watts, D. J.: *Six Degrees: The Science of a Connected Age*, W. W. Norton & company.
- 11) Bak, P. and Chen, K.: Self-organized Criticality, *Scientific American*, Vol. 264, pp. 46-53.