

統計分布の形状による分析*

——大学とロングテール・パレート法則を事例に——

西津 伸一郎

【要旨】

統計分布は有意性の検定など統計的検定に用いられるが一般的であるが、本稿は統計分布の形状自体から多くの示唆が得られることを示すのが目的である。

事例として用いたのが、大学の偏差値と志願者数、パレートの法則そしてロングテールの法則である。パレートの法則に対しロングテールの法則は、一見するとパレートの法則に対するアンチテーゼのように見える。ところが、この2つの法則はどちらも同一のパレート分布とよばれる統計分布から導かれている。なぜ異なる結論に至るのかを明らかにする。

【キーワード】 ロングテール, クリス・アンダーソン, パレートの法則
パレート分布, べき乗分布, 80:20の法則

【目次】

1. はじめに
2. 統計分布の形状
 - 2-1. 正規分布
 - 2-2. 一様分布

* 本研究は、立正大学経済研究所より研究助成を受けた。

- 2-3. パレート分布
- 2-4. その他の分布
- 3. 統計分布の形状による分析
 - 3-1. 大学の偏差値と志願者
 - 3-2. パレートの法則
 - 3-3. パレート分布の検証
 - 3-4. ロングテールの法則
 - 3-4-1. ロングテールの概要
 - 3-4-2. 販売サイドのロングテール
 - 3-4-3. 生産サイドのロングテール
 - 3-4-4. マイナー世界の需要喚起
- 4. おわりに

1. はじめに

統計分布は有意性の検定など統計的検定に用いられるが一般的である。本稿は、その方法論ではなく、統計分布の形状自体から多くの示唆が得られることを示すのが目的である。事例として用いたのが、大学の偏差値と志願者数、パレートの法則そしてロングテールの法則である。

大学については、受験生の多くが大学予備校の公表する偏差値を指標として受験している。この偏差値と入学試験の志願者数の関係を、その統計分布の形状から明らかにする。

パレートの法則は、イタリアの経済学者アルフレッド・パレートによって 1897 年に発見された、120 年以上前のことである。所得や富の分布は、平等ではなく偏りがあるというものである。約 20 パーセントの人々が約 80 パーセントの富を得ているという事実を発見した。このパレートの法則は、その後「80:20 の法則」として経済学の分野ではなく、経営学の分野において企業の経営戦略の手法として地位を確立してきた。例えば 80 パーセントの収入は、20 パーセントの生産物から得られると言われるように、「コア」となる売れ筋商品が存在する。このコア

となる 20 パーセントの商品を見出し、ここに経営資源を集中することで、より効率的に利益が得られるというものである。

一方で、2006 年にクリス・アンダーソンによって指摘されたロングテールの法則がある。これはパレートの法則においてコア以外の捨象されてきた商品を重視することが、利益の源泉になると主張した。これは一見すると、20 世紀においては主流であったパレートの法則から導かれる 80:20 の法則に対するアンチテーゼのように見える。ところが、この 2 つの法則はどちらも同一のパレート分布と呼ばれる統計分布から導かれている。なぜ異なる結論に至るのかを明らかにする。これは経済学の視点から一言で表現すると、費用構造の問題である。そしてその要因の 1 つは、IT 技術の進化と低廉化である。

2. 分布の形状

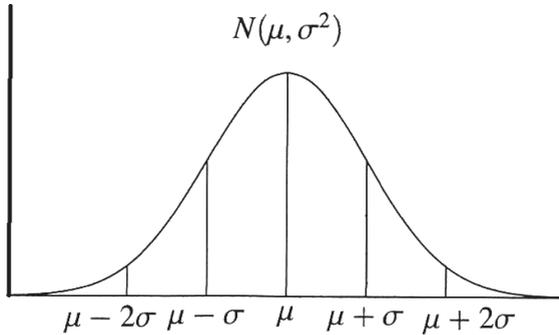
最初に本稿に登場する統計分布を紹介する。本稿では数学的な性質は利用しないので、ここでは数学的な説明は省略している。

2-1. 正規分布

代表的な分布である。一般に統計的検定において用いられ、我々が“分布”という名称を聞いたときに真っ先に思い浮かぶのがこの分布である。例えば全国の中学 3 年生の男女別の身長などがこの釣鐘型の分布をしている。平均を中心に左右対称の形状である。高身長や低身長の割合は少ない。また 4 メートルを超える異様な高身長もない。生物的に制約があるからである。他にも受験者数の多い模擬試験の得点分布などもこの形状をしている。図表 2-1 に正規分布（連続型）の形状を示す。

なお、本稿では離散型の場合でも図示するときには、議論に問題はないので、連続型で表示することにする。後で他の分布との比較を行うときに必要になるので、正規分布の基礎的部分について整理しておく。

標本のある区分に分け、その区分に入る標本数をカウントする。一般には度数と呼んでいる。図の縦軸は度数、横軸は区分を表している。図は連続型だが、こ



図表 2-1 正規分布

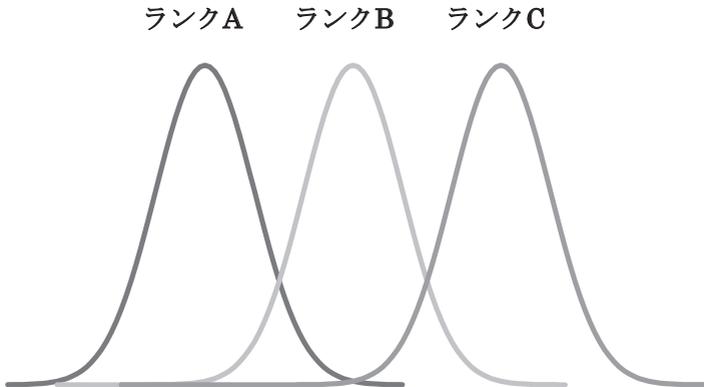
の場合には離散型の分布になる。平均は μ 、標準偏差は σ である。このとき $\mu \pm \sigma$ の範囲には、約 68.3% の標本が入る。 $\mu \pm 2\sigma$ の範囲には 95.5%、 $\mu \pm 3\sigma$ の範囲には 99.7% が入る。つまり平均周りが一番多く、平均から離れるにつれて、緩やかに少なくなっていく。

他にも正規分布をしているものがある。電子部品は高度な技術と製造装置により、同一特性を持つ個体が作られるように思えるが、この性能にもばらつきが存在する。例えば 半導体の特性の一つに直流増幅率 (hFE) がある。これも同一の生産ラインで作られた同じロットの個体にもばらつきがある。また製造ロットにより直流増幅率の平均値にもばらつきが発生する。

半導体には直流増幅率の大きさによりランクが存在するが、例えばランク名称を ABC とすると、A: 200~350, B: 300~450, C: 400~550 のように各ランクの上端と下端の値はオーバーラップしている。あるロットのランクが A になるか B や C になるかは製造時に決定される。ばらつきが大きいので、同一のマスクから作られる半導体の名称を複数取得したメーカーもあった。形式上、別の半導体としたのである。

それぞれのランクで直流増幅率の値は正規分布をしている。ランクごとに平均値に近いものが多数を占め、高い値のものも低い値のものも少なくなる。これを図示したのが図表 2-2 である。

パソコン等に使用される CPU (中央演算装置) の動作可能クロックも、製造さ

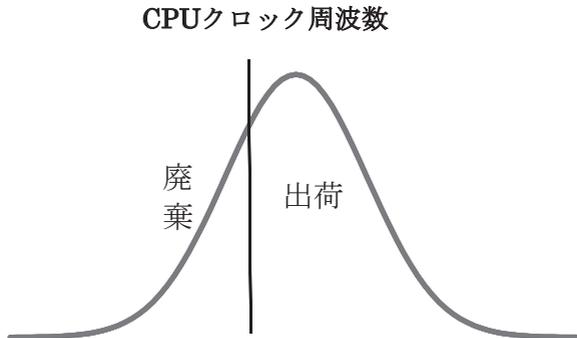


図表 2-2 半導体の直流増幅率 hFE の分布

れたものは同様に正規分布をしているはずである。これは公表されたことはないと思われるが、クロックが高くなるほど個体数は減少するはずである。CPU 製造メーカーは、動作クロックの下限を保証して販売しているから、むやみに動作クロックの範囲を広げると販売上支障が出る。例えば保証動作クロックを 1.6GHz ~ 3.2GHz のものを保証動作クロック 1.6GHz として販売すれば、PC 製造企業や個人は最大と最小では 2 倍の違いのある CPU を手にすることになる。PC 製造企業は安全のためクロックを 1.6GHz として PC を設計するであろう。中には 3.2GHz で動作するものがあるにも関わらずである。一方で個人の手作りパソコンでは、運のよい人は 2 倍の性能を手に入れることができる。

CPU 製造メーカーは、これを防ぐために保証動作クロックを細かく設定する。そして動作クロックの高いものほど数量が減ってくることに合わせて、価格も高く設定できるのである。これは需要側も高クロックの CPU を望むから、需要側も高価格でも購入する人がいる。需要曲線に沿った価格設定ができるということである。

一方で、動作保証をしているということは、製造した CPU をすべて出荷できないのではないと思われる。図表 2-3 のように、一定のクロック値のもの以上を出荷する。場合によっては約半数が廃棄される可能性がある。クロック下位の

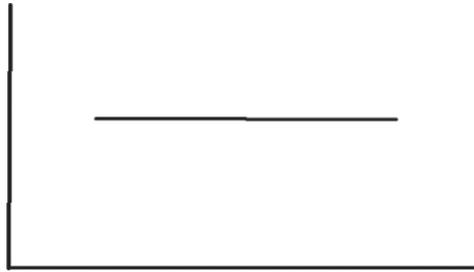


ものを出荷しても数量が少ないものを、低価格で販売せざるを得なくなる。また同一 CPU で性能に幅がありすぎるのも問題である。かつてのように 1 社独占ではないが、需要曲線に沿った利益を上げるには、むしろ半数を廃棄したほうが合理的であると推論できる。CPU 以外の半導体も例えばメモリーなどは「相性」といわれる不具合も知られている。メモリー製造メーカーによっては、特定の企業の PC ではうまく動作しないなどの現象である。これはメモリーの動作速度にばらつきがあるからである。それぞれの設計に余裕がなく、組み合わせる機器によって動作限界が運悪く来てしまうことが原因である。

以上の事例のように、生物に限らず物理的なものまで正規分布をしているのである。

2-2. 一様分布

ある一定の範囲で、確率密度が同一の分布である。例としてはサイコロの目の 1~6 それぞれの出る確率がある。この場合には離散型一様分布と呼ばれる。統計分布を意識しないで議論する人は、この分布の形状を意識せずに用いていると言える。一般的な社会現象で、図表 2-4 のような分布の形状のものは思い浮かばないが、本文中に登場するため掲載した。

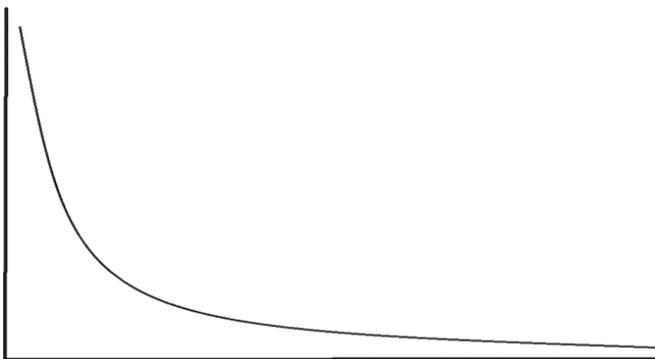


図表 2-4 一様分布

2-3. パレート分布

統計学では一般に「べき乗分布」と呼ばれる形状である。右側に向かってなだらかに裾を引く分布である。一般にパレート分布と呼ばれるものは、縦軸に所得を、横軸には個人あるいは世帯をとり所得の多い順に左からならべたものである。したがって曲線の下方面積を合計すれば、人々に配分された総所得ということになる。図表 2-5 で分かるように、左側の少数の人々が総所得の多くの割合を占めていることが分かる。

パレートは約 120 年前のイタリアにおいて、約 20 パーセントの人々が約 80 パーセントの富を得ているという事実を発見した。のちに“80:20 の法則”と呼ば



図表 2-5 パレート分布

れることになる比率である。また 1940 年代に J. M. ジュラン (J. M. Juran) がこの発見の功績を讃えて、この現象を“パレートの法則 (Pareto's Principle)”と呼んだ。

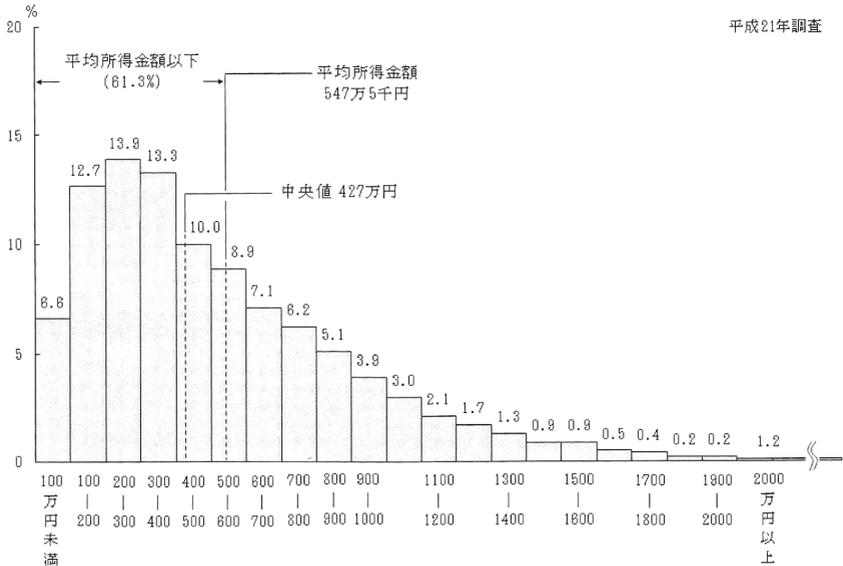
1992 年の国連開発プログラム報告書によると、世界の GDP の 82.7 パーセントが上位 20 パーセントの人々によって占められていることが示された。パレートの発見以来 100 年経過しても、世界レベルでも同様の現象が見られているのである。

パレート分布と形状が似たものがある。例えば信頼性分析に用いられるワイブル分布である。製造した製品の故障率を表すのに用いられたりする。縦軸は故障率、横軸は製品寿命(稼働時間)である。時間の経過とともに故障率が逓減する状態を表している。製品の保証期間を何年あるいは何時間にするかを決定するとき用いられる。当初高いのは、初期不良である。パレート分布と形状は同じだが、意味が異なることに注意しなくてはならない。これを明確にするため、経済現象の具体的な例を用いて説明する。

日本における所得の分布もパレート分布をしているはずである。ところが、平成 22 年の厚生労働白書に掲載された世帯別所得分布図は、図表 2-6 のようになっている。右側のなだらかに逓減する形状はパレート分布に似ているが、左側は歪んだ正規分布の形状である。歪度がプラスの値になる典型的な例である。所得金額に上限はなく下限はゼロの分布であるから、もともと偏りが発生する性質を持っている。しかも所得金額には生物のような制約はない。平均値は 547.5 万円、中央値は 427 万円です約 120 万円の差がある。中央値はちょうど真ん中の世帯で、半数以上の世帯が平均値より低い。

この分布図は、横軸は世帯所得の区分であり、1000 万円未満は 100 万円刻み、1000 万円以上は 200 万円刻みで 2000 万円までである。上は 2000 万円以上である。そして縦軸は各所得区分の世帯数を表示したものである。なお最近の厚生労働白書は、上は 1000 万円以上でまとまっており、上位所得層の分布が分からなくなっている。

ここでパレート分布は、先に述べたように縦軸に世帯所得を、横軸には個別の世帯をとり世帯所得の多い順に左からならべたものである。横軸には対象となる



図表 2-6 平成 21 年の日本の世帯別所得分布

出所：平成 22 年版厚生労働白書

すべての世帯が並ぶのである。つまり両者は次元がまったく異なるのである。世帯別所得分布が縦軸に世帯数、横軸に所得区分をとった図で仮に正規分布をしていても、パレート分布の次元で図を描くと、右下がりの形状になるのである。なお厚生労働省が公表しているデータでは、パレート分布を描くことはできない。

仮に完全平等（全世帯が同一所得）の場合、パレート分布の次元で図を描くと 2-2 で説明した一様分布の形状になる。つまり同一所得の世帯が左から右にずらっと並ぶのである。正規分布の形式の図だと、平均にすべてが集まり、標準偏差はゼロになる。このように一般には正常な世界を“正規分布の世界”，偏りのある世界を“パレート分布の世界”と表現することが多いが、図を描く上での次元が異なることに注意する必要がある。軸の説明はあるが、このことに触れた文献は見ることがない。次元の異なる分布の形状は、単純に比較はできないのである。先に述べたワイブル分布は、パレート分布と同様の形状だが、ここでの正規分布の形

式と同一の次元である。

パレート分布は描く次元によっては、必ずしも正規分布と異なる世界ではないのである。パレート分布は、偏りを表現するあるいは強調するのに適した形なのである。

2-4. その他の分布

今回は取り上げていないが、他にも代表的な統計分布が多数ある。1 時間に特定の交差点を通過する車両台数を表すポワソン分布、これをもとに交通渋滞のシミュレーションを行うこともある。結果が成功か失敗のいずれかであるベルヌーイ試行を独立に n 回行った時の成功回数を確率変数とする離散型確率分布である二項分布、 n を十分に大きくとると正規分布で近似できる。その他、対数正規分布、ガンマ分布、スチューデントの t 分布、カイ二乗分布など他にも多くの分布が存在する。

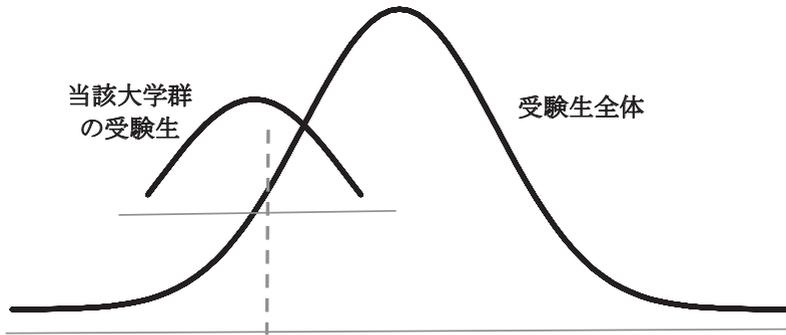
3. 統計分布の形状による分析

3-1. 大学の偏差値と志願者

分布の形状を意識しないで、議論した場合に誤った結論を得る例として大学を取り上げる。ここで偏差値が年々下がり続けている大学があったとしよう。

18 歳人口が減少する中では、さほど特異なことではない。それに伴って志願者数も減り続けている。

ここで、この大学の志願者が減り続けることには必然性があるのである。大学の偏差値は大学予備校が公表している。一般に模擬試験でその偏差値とった人が、実際の入学試験において合格者 50 パーセント、不合格者 50 パーセントとなった値が用いられている。偏差値 55 の受験生が偏差値 50 の大学を受験すれば、必ず合格するかと言えば、そうではない。以前に予備校のデータを調べたところ、合格者も不合格者もほぼ正規分布をしていた。また先に述べた正規分布の説明における半導体の直流増幅率のランクごとの分布(図表 2-2)と同様に合格者と不合格者の分布はオーバーラップしていた。したがって、偏差値 45 の受験生が偏差値



図表 3-1 受験生の偏差値分布

50の大学に合格し、先述の偏差値55の受験生が不合格になることも割合は低いものの起きてしまうのである。公表された50パーセント合格率の偏差値とは、そのような性質のものである。模擬試験の平均点をとると、偏差値は50になる。

図表3-1のように、大規模模擬試験の結果である受験生の偏差値の分布は、正規分布である。先に見たように、平均±標準偏差の範囲に、70パーセント弱の受験生が入る。図表3-1には大きい正規分布と小さい正規分布が描かれているが、大きいほうの正規分布が受験生全体の分布である、

点線の位置が当該大学とすると、その大学の受験生は、その周辺にやはり正規分布に近い形で分布している。図表3-1の小さいほうの正規分布である。受験生数は大きい正規分布で表されているので、点線に位置すると受験生はかなり少なくなっていることが分かる。偏差値が下がると受験生が減少することには必然性があるのである。偏差値が下がるほど受験生は減るのである。偏差値が平均±標準偏差の範囲から外れると、この大学と同じ偏差値の大学群には、全体の多くても10数パーセントの受験生しか志願してこなくなる。その少ない受験生減少への対策として、この大学が教育内容の改革・充実をアピールすることを打ち出したとしよう。この対策は、受験生の偏差値分布が先にみた一様分布をしている場合には有効かもしれない。しかし現実には正規分布をしているのである。分布から見える偏差値と受験生の関係はまぎれもない事実なのである。

この正規分布の反対側に位置する偏差値上位の大学は、同様の視点から見ると、

受験生は減少しなくてはならない。ところが、現実には偏差値上位の大学は多くの志願者を集めている。この理由は「3-3 パレートの検証」で説明する。

3-2. パレートの法則

今から 120 年以上前に提唱されたパレートの法則は、その後 80:20 の法則として経済学ではなく、経営学の分野でその地位を確立した。一般によく言われていることは、

- (1) 80 パーセントの収入は、20 パーセントの顧客から得られる。
 - (2) 80 パーセントの収入は、20 パーセントの生産物から得られる。
 - (3) 80 パーセントのクレームは、20 パーセントの顧客からである。
- などである。

つまり、コアとなる 20 パーセントのものに経営資源を集中することで、より効率的に利益が得られるというものである。なお、この 80 パーセントと 20 パーセントは、おおよその数字であり、これより大きかったり小さかったりする。80 パーセントが 70 パーセントになることもある。要は偏りがあることを示す数値である。(1)~(3) のような事実を前にすれば、コアでない 80 パーセントのものについては、捨象するのが合理的ということになる。あるいは残しても経営資源を多くは投入しないということである。IT 時代が到来する前の 20 世紀までは、この考え方が主流であった。パレートの法則については、後で述べるロングテールの法則の説明と重複する部分が多いので、ここでは詳細には触れないことにする。ロングテールの法則は、パレートの法則あるいは 80:20 の法則のアンチテーゼかと言われることがある。どちらもパレート分布を前提とした議論である。結論から言うと、どちらも正しいのである。パレートの法則あるいは 80:20 の法則は、何らかの制約がある場合に正しく、ロングテールの法則は、その制約が外れた場合に成立するのである。

3-3. パレート分布の検証

パレート分布の事例として私立大学の志願者数を取り上げる。国公立大学は、入試の機会が私立大学のように自由度がないため除外した。

データは、週刊東洋経済臨時増刊『本当に強い大学 2019』（2019年6月5日発行）に掲載された2019年志願者ランキングを利用した。増加傾向あるいは減少傾向とは言えない。旺文社教育情報センターのデータによると、2019年の私立大学志願者数は延べ383万人である。また4(6)年制私立大学受験生数は68.2万人である。なお、総志願者数はセンター試験利用入試を含む一般入試を対象としている。

この掲載誌にもパレートの法則がみられる。国公立を含むと全体で大学は764校ある。それにも関わらず、掲載大学は100校である。第1位の大学が15万人を超えているのに対して100位の大学は8593人(公立大学)である。

私立大学に限ってみると、掲載された95校の累計は308万4068人である。全体の15.6パーセントの大学で、総志願者数383万人の80.5パーセントを集めているのである。残りの約85パーセントの大学が残りの19.5パーセント、人数にして75万人を分け合っているのである。これは延べ人数である。旺文社教育情報センターのデータによると平均併願回数は5.58回である。単純計算だが実際の受験生数は $75/5.58=13.44$ (万人)、掲載のない513大学の1校当たりの受験生数は、 $13.44 \text{ 万人} / 513 = 262$ 人となる。大学の規模にもよるが、実に厳しい数字である。大学自体も掲載されることを望まないところもあるかもしれない。ちなみに上位27大学、比率にして僅か4.4パーセントの大学で、総志願者数の50パーセントを超えている。

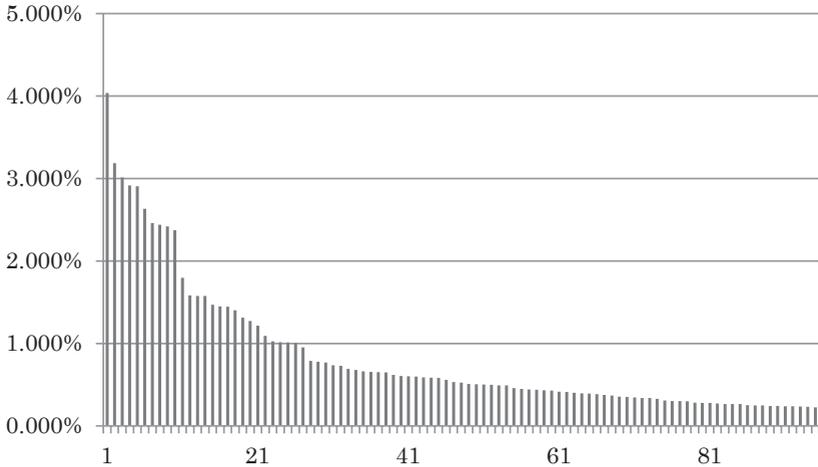
紙面制約から100校を掲載しているのであろうが、国公立を含む764校全体の13パーセントである100校で大学全体の傾向を掴むには十分なのである。これと同様に、多くのことがベスト100あるいはベスト50のような形で掲載される。これから漏れたものは、従来の紙メディアでは情報として登場してこない。つまりコアではないものは捨象されるのである。図表3-2と図表3-3に、志願者数の数値表と志願者数を比率化した図を示す。大学名は省略した。

図表3-3を見ると上位校の部分に多少滑らかさに欠ける箇所がある。30番目あたりの大学からはなだらかに減少していく。べき乗分布としてどの程度の当て嵌まりを示すのか、最小二乗法で推計してみた。べき乗分布は一般に次のように表される。

| 順位 | 志願者数 | 順位 | 志願者数 | 順位 | 志願者数 | 順位 | 志願者数 |
|----|---------|----|--------|----|--------|----|-----------|
| 1 | 154,672 | 31 | 28,221 | 61 | 15,950 | 91 | 9,158 |
| 2 | 122,010 | 32 | 27,916 | 62 | 15,794 | 92 | 9,067 |
| 3 | 115,447 | 33 | 26,474 | 63 | 15,443 | 93 | 8,979 |
| 4 | 111,755 | 34 | 26,027 | 64 | 15,131 | 94 | 8,922 |
| 5 | 111,338 | 35 | 25,307 | 65 | 15,016 | 95 | 8,649 |
| 6 | 100,853 | 36 | 25,034 | 66 | 14,696 | 合計 | 3,084,068 |
| 7 | 94,198 | 37 | 24,964 | 67 | 14,328 | | |
| 8 | 93,452 | 38 | 24,802 | 68 | 14,158 | | |
| 9 | 92,686 | 39 | 23,720 | 69 | 13,612 | | |
| 10 | 90,876 | 40 | 23,238 | 70 | 13,462 | | |
| 11 | 68,796 | 41 | 23,075 | 71 | 13,265 | | |
| 12 | 60,593 | 42 | 22,891 | 72 | 12,993 | | |
| 13 | 60,404 | 43 | 22,562 | 73 | 12,951 | | |
| 14 | 60,360 | 44 | 22,471 | 74 | 12,560 | | |
| 15 | 56,201 | 45 | 22,289 | 75 | 11,772 | | |
| 16 | 55,444 | 46 | 21,335 | 76 | 11,532 | | |
| 17 | 55,350 | 47 | 20,334 | 77 | 11,520 | | |
| 18 | 53,751 | 48 | 20,101 | 78 | 11,401 | | |
| 19 | 50,281 | 49 | 19,444 | 79 | 10,728 | | |
| 20 | 48,715 | 50 | 19,316 | 80 | 10,680 | | |
| 21 | 46,505 | 51 | 19,230 | 81 | 10,669 | | |
| 22 | 41,875 | 52 | 19,143 | 82 | 10,424 | | |
| 23 | 39,293 | 53 | 18,870 | 83 | 10,114 | | |
| 24 | 38,845 | 54 | 18,783 | 84 | 10,090 | | |
| 25 | 38,826 | 55 | 17,602 | 85 | 10,086 | | |
| 26 | 38,512 | 56 | 17,161 | 86 | 9,584 | | |
| 27 | 36,490 | 57 | 16,874 | 87 | 9,482 | | |
| 28 | 30,237 | 58 | 16,791 | 88 | 9,457 | | |
| 29 | 29,790 | 59 | 16,548 | 89 | 9,230 | | |
| 30 | 29,471 | 60 | 16,397 | 90 | 9,219 | | |

図表 3-2 2019 年私立大学入学試験志願者数 (上位 95 校)

出所: 週刊東洋経済臨時増刊『本当に強い大学 2019』(2019 年 6 月 5 日)



図表 3-3 2019 年私立大学入学試験志願者数比率 (上位 95 校)

出所: 週刊東洋経済臨時増刊『本当に強い大学 2019』(2019 年 6 月 5 日) を用いて筆者が算出した。

$$P = AX^{-n}$$

P は一般には確率だが、今回の場合は 1 大学の志願者数の全体に占める割合である。 X はランクを示す指標だが、今回は単純に順位を使用した、これを対数変換し、次式を単純最小二乗法で推計すると、

$$\ln P = \ln A - n \ln X$$

推計結果は、 $\ln A = -2.25523$ (-28.8), $-n = -0.78987$ (-37.3), 自由度修正済決定係数 = 0.936, () 内は t 値である。元の式に戻すと、

$$P = 0.10485 X^{-0.78987}$$

である。

ここからは、3-1 の偏差値上位大学の志願者数がなぜ多いのかという回答になる。先に見たように上位 15.6 パーセントの大学で 80.5 パーセントの志願者を集めていた。志願者が特定の大学に集中する傾向があるのである。つまりパレート分布になっているのである。やはり 80:20 に近い。この傾向は 2019 年に突然起きた訳ではない。大学の順位に変動はあるものの、長期に渡って観察される事実

である。偏差値は学部ごとに異なるので、大学全体でまとめた偏差値というのは存在しない。精査した訳ではないが、志願者が集中した大学は、少なくとも受験生全体の正規分布において、その右側に位置した大学である。

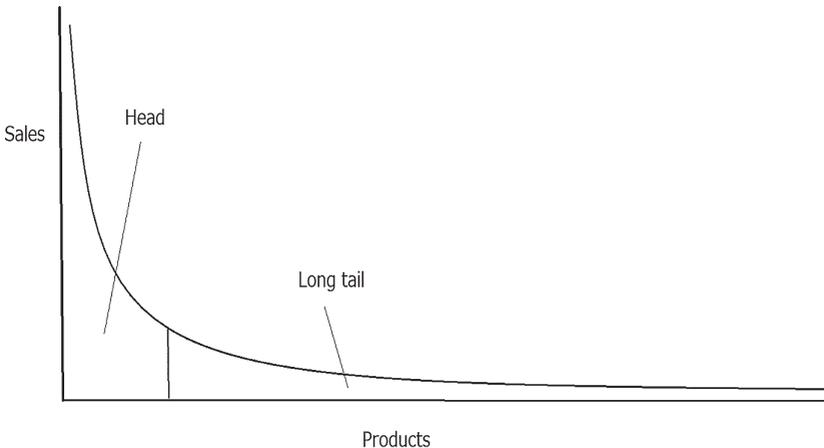
3-4. ロングテールの法則

3-4-1. ロングテールの概要

ロングテールは、クリス・アンダーソン [1] によって 2006 年に提唱されたものである。従来のパレートの法則あるいは 80:20 の法則が、コアとなる商品に集中することを主張したのに対し、アンダーソンは一見して真逆の主張をしたのである。図表 3-4 のように、縦軸に商品の売り上げをとり、横軸は売り上げの大きさ順に商品を並べたもので、典型的なパレートの分布になる。当時アンダーソンは雑誌ワイヤード (Wired Magazine) の編集長をしていたので、その著作では主にインターネットにおける楽曲のダウンロードを事例として取り上げている。縦軸は販売額ではなく、ダウンロード数である。しかし自動車 1 台とチョコレート 1 枚を同じ「1」のカウントとして同一視はできない。アンダーソンが取り上げた楽曲のダウンロードは、単価が同一なのでダウンロード数は、そのまま売上販売額になるのである。

図表 3-4 においてアンダーソンは、分布の形状を恐竜の姿に例えている。図の左側つまり 80:20 の法則におけるコアとなる 20 パーセントの部分に恐竜の頭であるヘッド (Head) と呼び、それ以外の右側の部分を恐竜の長い尻尾としてロングテール (Long Tale) と呼んだ。従来は捨象されてきた部分である。アンダーソンは、IT 時代においては、このロングテールが重要であると主張したのである。クリス・アンダーソン [1] の副題は、『「売れない商品」を宝の山に変える新戦略』である。80:20 の法則あるいはパレートの法則が支配する時代においては、捨象されてきた部分が「宝の山」であるという。一見すると、80:20 の法則に対するアンチテーゼと思えるが、そうではないのである。どちらも同じパレートの分布を前提とした理論である。

同一の分布を前提にして、一見逆に思える結論が導かれるのはなぜだろうか。これはインターネット時代の到来が大きな理由である。また経済学の視点から言



図表 3-4 パレート分布のロングテール

えば、単に生産（販売）費用の問題と捉えることができる。ロングテールの部分は、従来は捨象する対象であった。

クリス・アンダーソン [1] において、楽曲のダウンロードを例に説明をしている。図表 3-4 のヘッドの部分はダウンロード数の多い曲、つまりヒット曲である。一方でロングテールの部分は、あまりダウンロードの多くない曲である。

これらの楽曲が CD で販売されると、CD ショップはヘッドの部分の曲の CD を店舗に置き、ロングテールの CD は置かない。展示コストが掛からなければ、いくらでも置けるが、一定面積の店舗を借りれば、まず賃貸料がかかる、また店舗面積の制約から置ける枚数も限られる。めったに売れない CD では、展示コストのほうが上回る。実店舗ではヘッドの部分の販売するのが合理的なのである。

それに対してインターネットのダウンロードでは、IT 技術の進化・低廉化によりロングテールの楽曲を数百万曲も展示することができる。実店舗の展示のように多額の費用はかからない。オンライン音楽配信サービスのラプソディのケースでは、第 1 位のヒット曲の 1 か月のダウンロード数が約 18 万回であった。ずっと進んで第 2 万 5000 位では約 650 回である。第 1 位のわずか 0.4 パーセントである。これが第 5 万位では約 300 回、第 10 万位では約 120 回であった。これは

ずっと続いていくのであるが、決してゼロにはならないと言う。しかし 1 つ 1 つを見ると少額であり、ほんとうにこれらが販売に貢献するとは思えない回数である。ところが、このロングテールの楽曲は、その数が圧倒的に多いのである。1 つ 1 つは少額でも、それに楽曲数を掛け合わせた金額は、非常に大きなものになるのである。例えば第 10 万位から第 80 万位の楽曲を合計した 1 か月のダウンロード数は約 1600 万回である。この部分だけで売上全体の 15 パーセントを超えるという。この順位では、通常の CD ショップで見つけることは困難である。つまり従来なら完全に捨棄されてきたものが、インターネット販売では大きな割合を占めるのである。なかにはロングテール部分の売上が 50 パーセントを占めたという企業も現れた。

3-4-2. 販売サイドのロングテール

販売側と生産側では事情が異なる。最初に販売側の視点でみよう。80:20 の法則が支配的な時代には、実店舗で販売していた。商品を販売するためには、店舗で商品を展示しなければならない。展示には費用がかかる。また店舗にすべての商品を並べるということは、店舗の広さからも制約がある。

そうすると、あまり売れない商品は置けないということになる。仮に売れても展示の費用すら賄えない商品は置かないのが合理的である。そうすると、80:20 の法則が支配的になるのである。コアとなる売れ筋商品を展示するのである。このときロングテールの部分は捨象することになる。ところが、実店舗の場合には、「コア」を自分たちで見つけなければならない。

コアとなる商品が常に一定であれば、問題はない。ところがコアの商品は変動するのである。一定ではないコアを見つけるのは非常に困難である。例えばコンビニなどでは 1990 年代から POS システムによって、在庫管理・売れ筋商品の開拓を行ってきている。

一方でインターネットによる仮想店舗での販売には、商品の展示スペースには制約はない。費用面からみてもコンピュータのハードウェアの価格は劇的に下がっている。1980 年代のようにミニコンの 1 ギガバイトのハードディスクの価格が 1 億円という時代には、費用面からも制約がないとは言えない。このような状況で

は、むしろ実店舗での販売が有利であったと思われる。この時代にはコンピュータは、高速処理が必要なことにだけ使われる特別な存在であった。

現在のインターネットにおける仮想店舗の販売は、仮想店舗のホームページを作成するための固定費用が中心である。販売に伴う変動費用あるいは限界費用はほとんどかからない。その固定費用が劇的に低下したのである。インターネットの仮想店舗販売ではさらに大きな利点がある。

展示スペースに費用的にも物理的にも制約がないことは、商品数を多数展示できることになる。実店舗の時代には、「コア」となる商品を自分で見出す必要があったが、インターネット時代には「コア」を需要者が決めてくれる。販売サイドはすべての商品の展示をするだけでよいのである。

一つの例としてインターネットにおけるオークションを取り上げる。一般にオークションは中古品で1品だけの場合が多い。同一品の大量販売とは状況が異なるが、インターネット販売の特徴が極めてよく表れている。オークションの出品者（個人ではなく事業者）は、一般に非常に多くの商品を扱っている。インターネット・オークションで比較的良好に見られることがある。その商品説明の中で、「素人のため、この商品の知識はまったくありません。ご質問いただいても回答できません。画像をよくご覧になってご入札をお願いします。」というものである。そして開始価格を1円に設定して始まったオークションが、最終的にはなんと27万円という価格で落札されたのである。需要者が価格を決める典型的な例である。需要者は安価に購入したいと思っているが、競争者が値段を吊り上げるのである。需要者の効用と価格が一致したところで落札価格が決まったのである。この事業者が仮に実店舗で販売したら、1円ではないにせよ、かなりの安価で販売することになったかもしれない。自分では気が付かないまま大きな損失を出すことになる。このようにインターネットにおける仮想店舗販売は、個別の商品知識がなくても販売が可能であり、最低限の経営知識で運営ができるのである。

最近になり書店の廃業や閉店が報道されることが度々ある。書店が1軒もない地方都市の報道もあった。これもアマゾンに代表されるインターネット販売の影響である。インターネットの売り場面積は、実店舗に比較すれば無限と言ってもよい。選択肢が多いのである。また一方で出版不況と呼ばれる現象もある。情報提

供を中心とする雑誌などは、インターネットの大量の情報の前には、紙媒体として存続するのはなかなか困難である。

CD や DVD など元々がデジタル信号のものは、さらにこの動きが強く働く。メディアを介さず、直接に需要者がインターネット上で購入することが可能である。メディアの場合には発生する在庫切れもない。また紙媒体の書籍のように輸送コストも発生しない。そして必然的に安価で提供される。IT 時代のメリットを販売サイドは十分に享受できるのである。急激に成長しているインターネットにおける販売だが、小売業における販売額比率はまだ 10 パーセントに満たない。その経済全体への影響は、始まったばかりとも言える。

企業サイドの視点ばかりで述べたが、需要者も若年層を中心として、この新しいシステムを受け入れている。今後の IT 技術の変化とともに、これからも変化が続いていくと思われる。

3-4-3. 生産サイドのロングテール

生産サイドは、販売サイドとは事情が異なる。「コア」となる商品を自分で見出し設計・生産しなければならない。先にも述べたように、需要してもらえないコアとなる商品は常に一定ではない。それを予測し開発しなければならない。図表 3-4 のヘッドの部分となるものの開発である。生産サイドにとっては、ロングテールは入りたくない領域である。生産サイドでは、販売サイドのようにロングテールにおけるメリットは出てこない。

3-4-4. マイナー世界の需要喚起

例としてオーディオマニアを取り上げてみよう。マニアックなオーディオマニアは、例えば売上ベスト 10 のような製品から選択し購入するだろうか。これらの製品は一般受けするものばかりで、マニアが購入することはまず起こらない。例えば、このマニアが 5000 人に 1 人の割合で存在し、年に 2 回購入することを考えてみましょう。人口 5 万人の都市では、10 人の顧客が存在し、年に 20 回の取引機会がある。この都市に店舗を構えたマニアックな商品を揃えたオーディオ店は早晚倒産することになる。つまり一定の人口規模の都市でしか成り立たな

いのである。人口1000万人の都市であれば、顧客は2000人、年間取引回数も4000回になる。つまり大都市でしか実店舗販売は成立しないのである。通信販売という方法もあるが、紙媒体の情報提供の時代には、写真1枚か2枚という状況であり、商品についての詳細は分からない。また雑誌に掲載された広告では、更新は通常1か月に1回である。どうしても実物を見てからということになる。また紙媒体の雑誌では、費用対効果を考えると、広告はベスト10やベスト20を中心したものになってしまう。マニアックな商品はここでも捨象されることになる。

ところが、インターネットの仮想店舗販売では、低廉なコストで大量の画像の提供・情報の提供が自由に行える。更新も随時行える。また顧客についての地域性はない。実物を見ることはできないが、大量の画像などで補えるのである。つまり一般には捨象されてきたマイナーな商品が選択肢として登場するのである。マイナー世界の需要を喚起することができるのである。ロングテールの登場である。

これをパレート分布で表現すると、その形状の変化である。従来は登場しなかったロングテール部分が、需要者の選択肢に入ったのである。フラット化というのは極論としても、需要が喚起されたことで、より緩やかな形状に変化する可能性がある。個々に見ると小さいロングテールがまとまることによって、より重要性がより増大するのである。

4. おわりに

世の中の多くのものが正規分布をしていることから、いくつかの現象の説明ができたと思っている。同様の視点の研究がないか探してみたが、このような視点の分析はなかった。もしかすると見落としたかもしれないが、分析の視点は筆者のオリジナルである。

ロングテールの法則は、パレートの法則のアンチテーゼではなく、パレートの法則と両立し得るものであることを示した。ロングテールの法則は、インターネット技術の進化・IT機器の低廉化がなければ、生じなかった。これは経済学的に見れば、費用の問題であることが明確にできたと思っている。コッチ [2] は、人間

行動にパレートの法則を適用した。人間行動には時間という大きな制約がある。経済学における機会費用の視点である。本稿では構成上この視点の内容は省くことにした。

なぜパレート分布のように偏った現象が生まれるのかについては触れなかった。野口 [3] ではマタイ効果について述べられている。有名になった研究者が、より優秀な研究成果をあげることができたかを示すプロセスの説明である。これはパレート分布の偏りのごく一部の説明である。しかし、正規分布にせよパレート分布にせよ、それによって表される現象が、長期に渡って観察されている。そのため、なぜ分布がそうなるかを考えずに、統計分布を前提とした分析を行っても良いと考えたのである。

【参考文献】

- [1] クリス・アンダーソン (2006) 『ロングテール』 (篠森ゆりこ 訳) 早川書房
- [2] リチャード・コッチ (2011) 『新版 人生を変える 80 対 20 の法則』 (仁平和夫・高遠裕子 訳) CCC メディアハウス
- [3] 野口悠紀雄 (2015) 『「超」集中法』 講談社現代新書
- [4] マーク・ブキャナン (2009) 『歴史は「べき乗則」で動く』 (水谷淳 訳) 早川書房
- [5] 奥野正寛・池田信夫 (2001) 『情報化と経済システムの転換』 東洋経済新報社