

POS データにおける売上と商品数のべき的相関の観測

藤本 祥二[†] 石川 温[†] 水野 貴之[‡]金沢学院大学[†] 国立情報学研究所[‡]

1. はじめに

近年、世界各国企業の網羅的な財務データが入手可能になってきたことにより詳細な分析が可能になってきた。コンピュータの処理速度の向上やインターネットの普及に伴い大量のデジタルデータが記録され利用しやすい状態で管理されている。このような大規模データ（いわゆるビッグデータ）を統計的に分析することによって、普遍的に観測される法則が幾つも発見されてきている。そのような法則を体系的に結び付けることにより、経済を理解するのが本研究の大きな目的である。

これまでの分析対象となる基本単位は企業であり、企業の売上・従業員数・資産あるいは利益などの企業の規模を表す量の分布や、それらの成長率分布などが研究されてきた。これらの量は高額域（大規模域）では Pareto 則と呼ばれるべき分布に従い、中額域（中規模域）では対数正規分布に従うことが知られている。企業を基本単位とした企業規模量に関する法則およびそれらの関係については、多くの研究が存在する。しかし一方で、企業が生み出している製品を基本単位とする研究はあまり多くない。

このような状況下で、Stanley のグループは企業の生産する商品の種類数と売上との相関を記述するモデルや、商品の種類数の成長から企業規模量の成長率を導出するモデルを提案し、モデルから導かれる結果を実データにより確認した[1]。しかし、彼らが検証に用いたデータベースはヨーロッパと北アメリカの製薬会社の製品（医薬品）に関するものであった。薬の販売や価格付けには様々な規制や制約があることより、彼らの研究のみでは典型的な企業の製品による検証とは言い難かった。これに対して坂井・渡辺らは、日本のスーパーマーケット数百社の POS (Point of Sales) データベース『日経スーパーマーケット POS データ』を用いて、Stanley グループのモデルの検証を行った。スーパーマーケットが扱う商品は、医薬品に比べて標準的な企業の製品であり、販売や価格付けに関する規制や制約が比較的少ないと考えられるからである。その結果、Stanley らが観測した企業規模量の成長率は確認できるものの、“企業の各製品が他の製品と独立に成長する”という彼らのモ

デルの前提は、実データでは成立していないことを示した [2]。

このような状況のもと、我々は先行研究 [3] において、坂井・渡辺らが用いたのと同じデータベースにより、標準的な企業の製品でも Stanley のグループが報告した各企業の生産する商品の種類数と売上の相関が観測されることを確認した。また、それら商品を販売する各スーパーマーケットが扱う商品の種類数と売上にも、同様の相関があることを確認した。そして、それらの相関関係は 2004 年から約 10 年間に変化しないことより、これらの性質が普遍的なものであると結論付けた。

本研究では、2001 年から 2012 年の 12 年間のアメリカのスーパーマーケットの POS データベース[4]を用いて、我々の先行研究と同様の分析を行い、上記の相関関係およびその時間的変化を確認する。

2. 企業別分析

我々が日本のスーパーマーケットで扱われる日用品の POS データで観測した企業の製品の種類数と売上の相関が、アメリカのスーパーマーケットの POS データでも確認できるかを調査した。図 1 は 2012 年のデータで企業数は 4,849 である。図 1 の散布図に最適直線を引くと次式のようなになる：

$$S \propto N^\alpha. \quad (1)$$

この性質は興味深いものであり、企業の売上は製品の種類数に単純に比例するのではなく、製品の種類が多くなるとべき的に売上が大きくなることを示している。

図 1 と同様の測定を 2001 年より毎年行った結果が図 2 である。図 2 は、収集する POS データ数の増加により、企業数が増加するにも関わらず、べき的相関関係を特徴づける指数 α はほとんど変化せず、約 2.0 の値であることを示している。

先行研究で行った日本のスーパーマーケットの POS データを分析した場合も同様の性質が観測され、約 10 年間で α の値はほぼ一定で、およそ 1.5 であった。本稿の分析結果であるアメリカのスーパーマーケットでは、それより少し大きい値が観測されている。

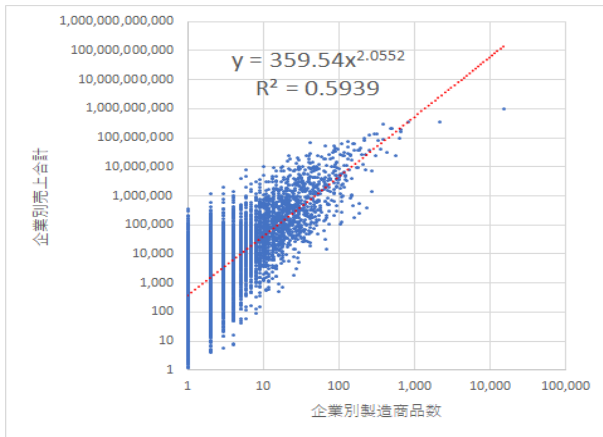


図1 2012年の企業別の製造商品の種類(N)と売上合計(S)の散布図。破線は最適直線： $S \propto N^\alpha$ 。

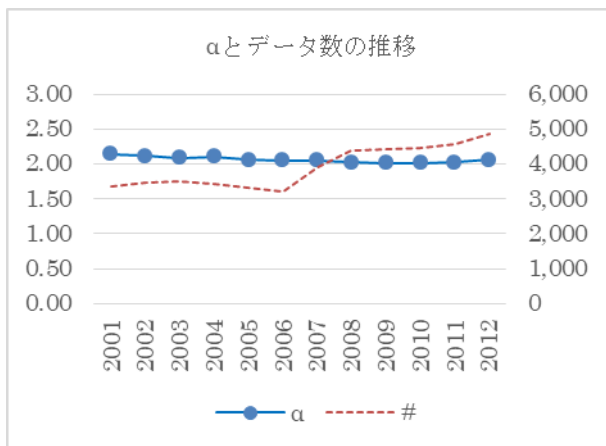


図2 2001年から2012年まで12年の α (●)と企業数(破線)の推移。

3. 店舗別分析

次に、各スーパーマーケットで扱う商品の種類数とそれらの売上合計の相関について調べる。企業別分析と同様に、それらを N, S と表すと、やはりのように N と S には強い相関が観測される。前節の企業別分析と同様に、散布図に最適直線を引くと次式ようになる：

$$S \propto N^\beta \quad (2)$$

企業別分析と同様にこの性質は興味深いものであり、スーパーマーケットの売上は扱う商品の種類数に単純に比例するのではなく、商品の種類が多くなるとベキ的に売上が大きくなることを示している。

企業別分析と同様に、2001年から2012年までのベキ的相関を特徴づける指数 β を観測すると、 β はほとんど変化せず、約 2.3 の値であり、年次に依らず次の関係が成り立っている：

$$\beta > \alpha \quad (3)$$

この大小関係は、日本のスーパーマーケットの分析でも確認されている。

4. まとめと今後の展望

日本とアメリカのスーパーマーケットのベキ相関の指数を比較すると、日本の場合は $\alpha \sim 1.5$ 、 $\beta \sim 2.0$ であったのに対して、アメリカの場合は $\alpha \sim 2.0$ 、 $\beta \sim 2.3$ と、 α と β ともに日本のベキ指数よりアメリカのベキ指数が大きい結果が得られた。これは、企業別分析および店舗別分析の両方において、アメリカの方がより商品数の増加により売上合計が増加する非線形性が強いことを意味している。

製品を製造企業別に集約する場合と、販売店舗別に集約する場合のいずれにも、製品の種類数と売上にベキ的相関関係が普遍的に観測されるという事は、企業を単位として経済を分析する研究手法は、企業の扱う製品を単位として記述するレベルの研究にも適用することができ、その結果を標準的な製品の実データで実証的に確認することが可能なことを示している。

今後の課題は、製品の集約法の違いにより、ベキ的相関の指数に違いが生じる理由、そして 1.5 や 2.0、あるいは 2.0 や 2.3 というベキ指数が現れる理由を明らかにすることである。この実証研究が進めば、従来の経営学とは違ったアプローチで企業の経営を分析・理解することが可能になると考えられる。

謝辞

本研究は JSPS 科研費 17K01277、18H05217、16H05904、国立情報学研究所、大林財団、大川情報通信基金の助成を受けています。

参考文献

- [1] Pammolli, F., Fu, D., Buldyrev, S. V., Riccaboni, M., Matia, K., Yamasaki, K., and Stanley, H. E.: *Eur. Phys. Lett. J. B* 57, 131 (2007)
- [2] Watanabe, T., Sakai, K.: *Eur. Phys. J. B* 76, 507 (2010)
- [3] 石川 温, 藤本 祥二, 水野 貴之, 渡辺 努: 人工知能学会: 経営課題に AI を! ビジネス・インフォマティクス研究会, 4th JSAI Special Interest Group on Business Informatics (SIG-BI) 2016, paper2.
- [4] IRi, <http://www.iriworldwide.com/en-US/>

Observation of Power-law Correlation between Sales and Product Volume in POS Data

†FUJIMOTO Shouji, ISHIKAWA Atsushi, Kanazawa Gakuin University

‡MIZUNO Takayuki, National Institute of Informatics