解説/Review

分布関数による経済メカニズムの解明

水野 貴之 *1 *2

Analysis of Economic Mechanism Using Distribution Functions

Takayuki MIZUNO*1 *2

Abstract– One of the characteristic analytical methods in Econophysics is observation of distribution function of an economic phenomenon. By comparing the distribution function with a normal distribution, we can find the correlation structure hiding in the phenomenon through Lindeberg's central limit theorem and the generalized (stable distribution) central limit theorem. We introduce some examples of research which used the distribution function for the phenomena of financial markets, real estate markets, an online product market, and a retail market.

Keywords– normal distribution, exponential distribution, power-law distribution, central limit theorem, stable distribution

1. はじめに

経済現象を扱う学問分野には，経済学の範囲をどこまで詳細に述べるかによるが，マクロ経済学，ミクロ経済学，計量経済学，数理ファイナンス，金融工学，行動経済学，神経経済学や本特集で解説する経済物理学などがある．経済物理学を象徴する分析手法として，現象における分布関数の直接観測とスケーリング則が挙げられる．例えば，物質の臨界温度近傍における比熱や帯磁率の振る舞いや，バクテリアが作り出すコロニー，地震の規模や発生間隔における研究などで用いられる．これらの分析手法は，他の経済分野ではあまり用いられることがなく，統計物理学から派生した経済物理学の特徴である．本稿では，分布関数の直接観測による経済メカニズムの分析に焦点をあてて解説する．

2 節では，データ分析を行ううえで強い前提仮定を必要とする統計量や統計分析（検定）の短所を紹介する．3 節では，少ない前提仮定で現象の相関構造を抽出することが可能な分布関数を用いたデータ分析手法について述べる．このような解析が可能なのは，分布を直接観測できるほどのビックデータが昨今存在するからである．逆を返せば，十分なデータがない現象であれば，前提仮定を駆使して統計量や統計分析（検定）が活躍する．4 節では，金融市場，不動産市場，家電オンライン市場，小売市場を対象に行われた分布関数を用いた研究例を紹介する．5 節は全体のまとめである．

2. 統計量や統計分析（検定）の短所

経済物理学では，主に，めったに起きないが起きると経済に大きな影響を与える現象の解明をおこなっている．わかり易い例で言えば，金融市場のバブルや暴落などである．これらの現象が起きる確率は，1%以下であることが多く，有意水準 5%で判断する通常の統計検定では捉えることが難しい．そこで，経済物理学では分布関数の直接観測を行うことにより，分析対象の現象の背後に相関構造が存在するのかを明らかにする．

通常の統計量や統計分析（検定）には短所がある．相関構造の有無は，通常，データの情報を集約した統計量と統計検定を用いておこなわれる．そのときに問題になってくるのは，集約による情報の欠落や，検定が正しく機能するための対象データに関する前提条件である．

まず，前者について，分散を例に説明する．分散はデータのバラツキを示す最も基本的な統計量である．データ

*1 国立情報学研究所情報社会相関研究系 東京都千代田区一ツ橋 2-1-2
*2 総合研究大学院大学複合科学研究科情報学専攻 東京都千代田区一ツ橋 2-1-2
*1 National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo
*2 Graduate University for Advanced Studies (Sokendai), 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo

Received: 9 July 2013, 8 August 2013

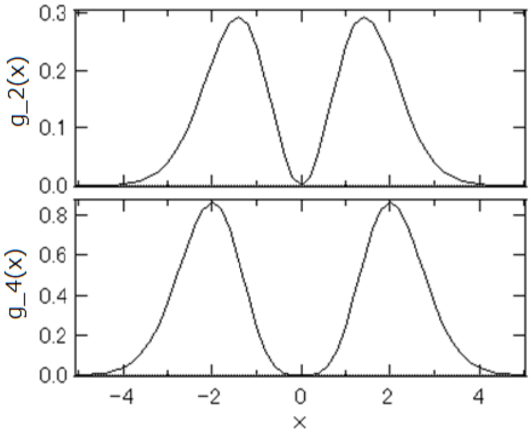

Fig. 1: $g_2(x)$ と $g_4(x)$ の分布

が従う確率密度関数を $f(x)$ とすると，分散は，

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x) dx$$

で定義される．この分散は，x の全領域の代表値として扱われることがあるが，それは誤りである．$f(x)$ が標準正規分布の場合の関数 $g_2(x) = x^2 f(x)$ を描いた図が **Fig. 1** である．この曲線 $g_2(x)$ の最大値は，およそ $x = \pm 1.5$ であり，分散にはこの辺りのデータの特徴が強く反映していることが分かる．一方，$x < -4$ や $x > 4$ の大きな値でのデータのバラツキは，ほとんど分散には反映されていない．もし，$x = \pm 4$ の特徴も捉えたければ，$g_4(x) = x^4 f(x)$ までモーメントを拡張する必要がある．**Fig. 1** の下図は $g_4(x)$ を表しており，$g_4(x = \pm 4)$ も若干の値を持っていることが確認できる．しかしながら，逆に $g_4(x)$ では $g_2(x)$ に比べて，$x = \pm 1.5$ の特徴は小さくなる．このように，データの情報を集約した統計量には，情報に対する偏りが存在してしまうために，分析対象の現象にあわせて統計量を正しく選択しなければいけないという短所がある．

次に後者の統計分析（検定）について述べる．統計検定や統計分析も，先の統計量と同じように，広く特徴を捉えるわけではなく，かなり限定した特徴を抽出する．ここでは相関構造の抽出でよく使われる相関関数を例にとって述べる．データ A とデータ B の関係を相関関数では，一次関数で回帰する．その時の傾きがゼロでなければ相関が存在するといい，ゼロであれば相関がないとする．つまり，データ A とデータ B に一次関数に従う関係がなく，二次関数に従うような関係があった場合，相関関数では相関構造を抽出することができない．従って，統計分析や統計検定も，分析対象の現象にあわせて正しく選択しなければいけないという短所がある．

3. 安定分布

現象に相関構造が存在するかしないかを明らかにするために，ありとあらゆる統計分析や統計検定を試すのは建設的ではない．そこで経済物理学では，少ない仮定で成り立つ安定分布の理論を応用して，何かしらの相関構造が存在すれば，それを検出することが可能な分布の分析をおこなう．

3.1 中心極限定理

安定分布における限定的な理論として中心極限定理がある．安定分布の理論では，多くの確率変数 x_i の和で定義された確率変数 S_n の分布を考える．中心極限定理では，これらの確率変数 x_i が独立で，分散が有限値 s^2 である同一分布関数に従うときに，確率変数 S_n が正規分布に従うことを証明している．従って，観測された現象が正規分布に従っていなければ，その現象における「確率変数 x_i に相関構造が存在する」 or 「x_i の分散が発散している」 or 「x_i の分布関数が i に依存して異なる」ことを示唆している．

3.2 リンドベルグの中心極限定理

リンドベルク条件 [1] による若干の仮定を導入することにより，「x_i の分布関数が i に依存して異なる」ことが原因で正規分布に従っていないという可能性をリジェクトすることができる．確率変数 x_i の平均値 μ_i と分散 s_i^2 が i に依存して有限である場合でも，リンドベルグ条件

$$\lim_{n\to\infty} \frac{1}{B_n^2} \sum_{i=1}^{n} \sum_{|x_i-\mu_i|\geq \tau B_n} (x_i - \mu_i)^2 f_i(x_i) = 0$$

を満たすときには，確率変数 S_n が正規分布になる．ここで，τ は $\tau > 0$ の任意の定数，$f_i(x_i)$ は x_i の確率密度関数，$B_n^2 = \sum_{i=1}^{n} \sigma_i^2$ である．

このリンドベルグ条件を噛み砕いて言うと，ある特定の確率変数 x_i が確率変数 S_n に大きな影響を与えないという状況を仮定している．例えば，ある $x_{i'}$ だけ極端に大きな値を持っていて，$S_n - x_{i'}$ が S_n と極めて異なる場合，リンドベルク条件は成立しない．

3.3 レヴィの安定分布

安定分布の理論 [2] は，正規分布に収束する中心極限定理を，他の分布に収束する場合まで拡張した極限定理である．「x_i の分散が発散している」場合の確率変数 S_n の分布の形状は，$|S_n|$ が十分に大きい領域において，確率密度関数のべき指数が 3 以下（累積分布におけるべき指数 α が 2 以下）のべき分布

$$P(|S_n|) \propto |S_n|^{-(\alpha+1)}, \quad 0 < \alpha \leq 2$$

に漸近することが分かっている．この $0<\alpha\leq 2$ の範囲をレヴィの安定領域という．この理論より，観測された現象が，この分布に従っていなければ「x_i の分散が発散している」可能性をリジェクトすることができる．

このようにして，観測された現象が従っている分布関数より,「確率変数 x_i に相関構造が存在する」ことを見出していく．

4．指数分布

ある現象において事象の発生がランダムであるのか，それとも相関構造が存在するのか，それを見分けるのに指数関数が利用される．もし，単位時間中に事象が平均 λ 回，過去の事象の発生状況に依存せずランダムに発生する場合，任意の時刻から始めて，その後に事象が最初に発生するまでにかかる時間間隔 t の分布は指数関数

$$P(t)=\lambda e^{-\lambda t}$$

に従う．従って，単位時間中の平均発生回数が過去の発生状況に依存して変化する場合には，時間間隔 t の分布は指数関数から外れる．例えば，部品の摩耗による機械の故障の場合には，古い機械ほど壊れやすいので時間とともに故障率 λ が大きくなる．商品の販売売り切りの場合には，定番商品になるほど打ち切られなくなり，時間とともに打ち切られ率 λ が小さくなる．これらのような場合，λ は

$$\lambda(t)=\frac{m}{\eta}\left(\frac{t}{\eta}\right)^{m-1}$$

で近似できることが多く，時間間隔 t の分布は拡張指数関数（ワイブル分布）

$$P(t)=\lambda(t)e^{-\left(\frac{t}{\eta}\right)^{m}}$$

になる．ここで，η は尺度パラメータと呼ばれる係数，m はワイブル係数である．機械の摩耗故障の場合は $m>1$ となり分布は指数関数に比べて上に凸，商品の販売打ち切りの場合には $m<1$ となり指数関数に比べて下に凸となる．

5．様々な経済現象の分布

5 節では，2 節と 3 節の分布を使った分析を，金融市場，不動産市場，家電オンライン市場，小売市場に適用させた例を紹介する．

5.1 金融市場

経済物理学において，一番初めに注目された経済現象が金融市場の大きな価格の揺らぎである．1995 年に Nature の誌面で R. N. Mantegna と H. E. Stanley によって，ニューヨーク証券取引所の株価変動の分布の特徴が報告された [3]．当時，分布が観測可能な経済現象に関するビックデータというと，金融市場のデータぐらいしか存在しなく，経済物理学ではあらゆる研究者がこぞって分析をおこなった．ここでは，水野らが明らかにした外国為替レートの変動の例を紹介する [4]．

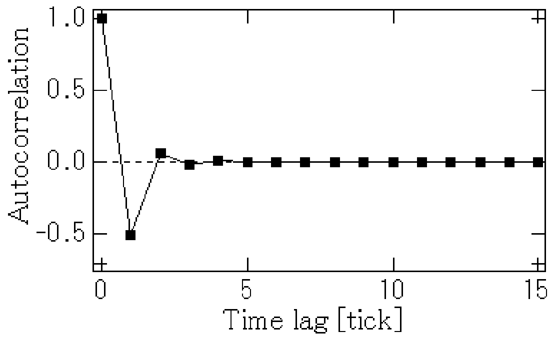

Fig. 2: 円ドルレート変動の自己相関関数

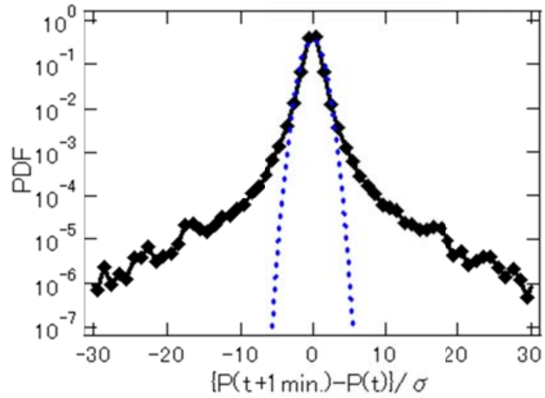

Fig. 3: 円ドルレート変動の確率密度関数

Fig. 2 は，円ドルレートにおける 1 取引単位の価格変動の自己相関関数を表している．取引ごとに時間が 1 つ進む尺度を tick といい，2000 年頃の円ドルレートでは 1 tick は約 7 秒であった．**Fig. 2** から自己相関関数は 2 tick 以降，ほとんどゼロであることが分かる．すなわち，ある時刻のレート変動は，2 tick 以上過去のレート変動と線形の相関を持たない．この結果を受けて，しばしば，金融市場の価格変動はランダムであると仮定されることがある．しかしながら，この結果は，2 節で説明したように，あらゆる相関構造が存在しないことを示唆しているのではなく，あくまで「線形の相関構造」が存在しないことを示しているに過ぎない．従って，ランダムの仮定を用いる場合には注意が必要である．

変動に関する相関構造の存在を，レート変動の分布を観測することにより確認する．**Fig. 3** は，1 分間のレート差の確率密度関数である．横軸はレート差を，それ自

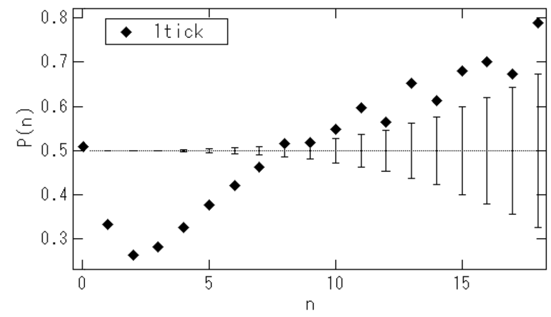

Fig. 4: 円ドルレート変動の符号の相関構造

身の標準偏差で規格化してある．図中の破線は，標準正規分布を表している．つまり，明らかに，レート差の分布は正規分布に従っておらず，正規分布では非常に稀にしか起こり得ない 10σ を越える大きな変動がしばしば起きている．これは，1 週間のレート差に広げても同様であることが報告されている．では,「レート変動の分散が発散していない」ことから確認する．円ドルレート差 Δp の確率密度関数は

$$P(|\Delta p|) \propto |\Delta p|^{-(3+1)} \quad \text{for} \quad |\Delta p| \geq 3\sigma$$

に従っている．このべき指数の値はレヴィの安定領域の外側にあり，分散が発散していないことが分かる．次に,「レート変動分布の時間依存性」であるが，レート差のべき指数は時期に依存することなく，ほぼ $\alpha = 3$ を維持していることが確認されている．つまり，時間により非定常に分布の形状が大きく変化して，正規分布からのズレを生み出している訳ではない．これらの結果は，レート変動に，線形以外の相関構造が存在することを示唆している．

相関構造の抽出をおこなう．**Fig. 3** のレート差の確率密度関数は 3σ 以上が正規分布より高い確率を持っている．さらに注意深く観測すると，$\sigma = 0$ 付近も正規分布より高い確率を持っていることも読み取れる．つまり，レートが極端に大きく動く場合と全く動かない場合が共存している．これらの特徴をとらえるために，レートが下降方向 $(-)$ に動いたあと，上昇方向 $(+)$ に n 回連続して動いた条件下で，次も上昇方向 $(+)$ に動く確率 $P(+|+\ldots+-)$ を調査する．この結果は **Fig. 4** において横軸に連続回数 n を，縦軸に確率 $P(+|+\ldots+-)$ をとることにより示される．**Fig. 4** のエラーバーは，サンプル数 n に依存する標準誤差を表し，$0.5 \pm 1/(2\sqrt{n})$ で与えられる．レートの連続上昇回数が $n \leq 8$ 回では，確率が $P(+|+\ldots+-) \leq 0.5$ である．つまり，レートは過去の変動を打ち消す方向へ動きやすい．他方，レートの連続上昇回数が $n{>}8$ 回では，確率が $P(+|+\ldots+-){>}0.5$ である．つまり，8 回以上連続して上昇するとトレンドが形成され，レートのカスケード的な上昇が起きる．これらの結果は下降方向 $(-)$ の場合でも同じであることが確認されている．このようなレート変動の複雑な相関構造が，レート差分布に正規分布からの乖離を生み出す．

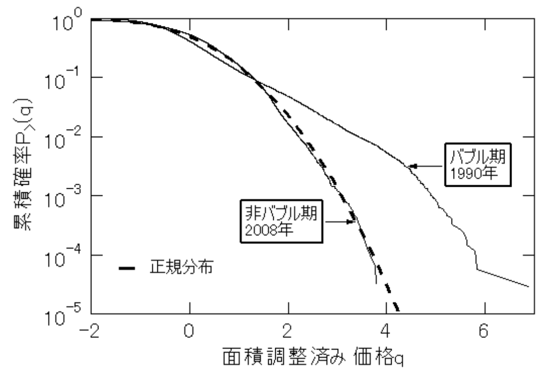

Fig. 5: 面積調整済み価格の分布

5.2 不動産市場

分布を用いた不動産バブルの検出に関する大西らの研究について紹介する [5]．バブルとは，資産価格（この場合，不動産の価格）が実体経済（不動産が生み出す価値）の成長以上のペースで過剰投機により高騰し続ける経済状態のことである．実体経済を測定することは要因が多すぎて難しいために，バブルの検出は一筋縄ではいかない．

首都圏都心部のマンションの物件の価格に注目する．はじめに間取り等の物件属性の違いによる価格の違いを調整しておく．属性の違いは主に面積で調整することができる．これは，面積が広い物件ほど，床暖房が完備されたりと，その他の多くの価格決定要因を面積の影響に置き換えることができるからである．物件の面積 s と価格 p の関係は，指数関数

$$p(s) = e^{0.013s+q}$$

に従っている [5]．ここで q は面積で調整できなかった価格，別の言葉でいえば，面積による価格の影響を取り除いた物件価格である．この q を面積調整済み価格と呼ぶことにする．

物件の価格は，その物件の持つ様々な属性の価格を足しあわせた合計と考えることができる．3 節のリンドベルグの中心極限定理により，各属性の価格が物件の価格に極端に大きな影響を与えないという状況下であれば，物件の価格の分布は正規分布になる．そこで，バブル崩壊後の 2008 年の面積調整済み価格 q の累積確率分布 $P_{>}(q)$ を **Fig. 5** で示す．累積確率分布とは価格 q 以上の価格が存在する確率 $P_{>}(q)$ をプロットした分布である．2008 年の分布は，図中の破線で示された正規分布と一致していることが分かる．

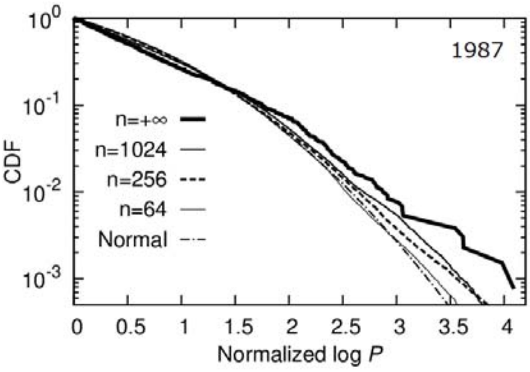

Fig. 6: 立地面積調整済み価格の分布

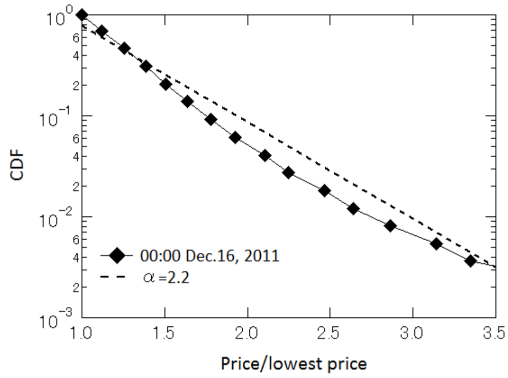

Fig. 7: 価格のバラツキの分布

一方，バブル期の 1990 年の面積調整済み価格 q の累積確率分布 $P_>(q)$ は **Fig. 5** で示されるように正規分布には従っていない．つまり，バブル期では面積では調整できない価格決定に影響を与える際立った属性が存在する．それは「物件の地域」である．非バブル期では首都圏都心部で地域差による価格のばらつきは小さく，分布を正規分布から乖離させる影響の尺度で考えれば，都市部全域を同一地区とみなすことができていた．しかしバブル期には，同一とみなせる地域の範囲が，バブルが進むに連れてどんどん狭くなっていく．ここで，物件 i の面積調整済み価格 q_i を，さらにその物件 i の近隣 n 物件の平均価格 $\overline{q_{i,n}}$ で割ることにより，立地と面積で調整した価格

$$q'_{i,n} = \frac{q_i}{\overline{q_{i,n}}}$$

を導入する．**Fig. 6** は，バブル期の 1987 年の立地面積調整済み価格 $q'_{i,n}$ の累積確率分布 $P_>(q'_n)$ を $n=\infty$（首都圏都心部全域），$n=1024,256,64$ について示した図である．隣接 64 物件のとき，立地面積調整済み価格は正規分布に従うことが分かる．これはおよそ 1.8 km 圏内を同一の地域とみなすことができること意味している．

バブル期の不動産に対する投機は特定の地域に集中し，その結果として，非バブル期において立地として同じ価値を持っていた物件間で大きな差が生まれたことが，分布の分析により分かった．面積調整価格の分布の正規分布からの乖離を測定することにより，アメリカなど他国の不動産バブルも検出が可能であることが確認されている．

5.3 家電オンライン市場

同時刻における同一製品の価格のバラツキに関する分布の研究を紹介する．同一製品が違う価格で売られていた場合，消費者は安い方の店で買うはずであり，高い店では売れない．従って，商売を成り立たせるために，自然と，どの店舗も同じ価格を提示するという考え方が「一物一価の法則」である．しかしながら，実際に一物一価の法則が観測される例は極めて少なく，何故，成り立たないかを調べることにより，複雑な経済メカニズムの一端を解明することができる．

経済学において，小売業界で一物一価の法則が成り立たない理由は，サーチコストとメニューコストだと言われている．サーチコストとは，消費者がどの店が安いかを探したり，購入店舗までの移動にかかる金額のことである．メニューコストとは，店舗が値札を書き換えるのにかかる金額のことである．これらのコストがあるから，例えば福岡と東京の店で価格が異なることがありえると説明される．しかし，最近，これらのコストが限りなくゼロに近いインターネットの価格比較サイトにおいても，価格の大きなバラツキが存在することが指摘されている．この価格のバラツキのメカニズムについて分布を用いて解明した水野らの研究を紹介する [6].

Fig. 7 は，価格比較サイト「価格.com」における 2011 年 12 月 16 日午前 0 時 0 分時点に存在する 20 店舗以上が取り扱っている全ての製品について，各店舗の各製品の価格を各製品のその時点の最安値で割った値の累積確率分布を表している．この分布から，最安値の 1.5 倍以上の価格を提示する店舗が全体の 20%程度存在することが分かる．

最安値に比べて価格の高い店舗にて，商品が売れているのかを調査した結果が **Fig. 8** である．図の (◆) は，横軸に各製品における店舗の価格順位をとり，縦軸にその価格順位の店舗のレジに行くためのボタンがクリックされる確率をプロットした．1 位（最安値）が選ばれる確率は約 13%，10 位が選ばれる確率は約 3.3%と，25 位まで，およそ指数関数

$$P(r) \propto e^{-0.122r}$$

に従って減衰していく．ここで r は店舗の価格順位である．このように商品購入時に価格の高い店舗も選ばれている．

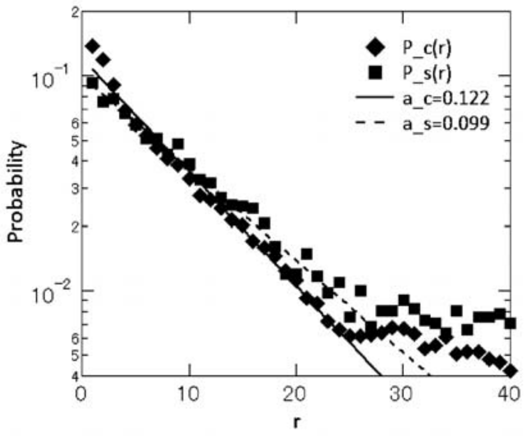

Fig. 8: 順位に対する購入確率と参入確率

店舗が選ばれる確率が，何故に順位に対して指数関数に従うのかについて考える．4 節から指数関数がポアソン的な事象の発生により生まれることが分かる．そこで，次のような単純な仮説を導入する．消費者は店舗に対する好みを持っている．例えば，ある人はクレジットカードで商品が購入可能な店舗を好み，またある人は延長保証が可能な店舗を好む．消費者の店舗に対する好みは多岐にわたっていると想定し，各店舗が好まれる確率は店舗に依らず θ で与えられていると仮定する．消費者は商品を購入する店舗を決める意思決定の過程として，まず好みの店舗をいくつか選んで，次にそれらの中から一番安い店舗で購入すると仮定する．この仮説のもとでは，順位 r の店舗で商品が購入される確率は，1 位から $r-1$ 位までの店舗が好みではなく，r 位が好みの店舗である確率によって与えられるので，

$$P(r) = \theta(1-\theta)^{r-1}$$

となり，まさに先程観測した指数関数が現れる．この指数関数を **Fig. 8** に当てはめると $\theta = 1 - e^{-0.122} \approx 0.115$ である．この仮説での関数型と実際の関数型の一致は，このような消費者の購買行動が存在する可能性を示唆している．

次に，店舗の価格設定に目を向ける．**Fig. 8** の (■) は各店舗が市場に参入及び再参入したときに提示した価格の順位の分布を表している．店舗は他の店舗の価格を知っているのにもかかわらず，最安値を提示するとは限らないことが読み取れる．また，この分布は，消費者が店舗のレジに行くためのボタンがクリックされる確率 (◆) の分布とほぼ一致していることがわかる．この結果は，店舗は消費者が高い価格でさえクリックすることを認識しており，その消費者の需要に合わせて高値を設定していると考えられる．つまり，サーチコストとメニューコストがゼロに近いインターネットの価格比較サイトでさえ，消費者の店舗に対する好みによって同一製品で大きな価格のバラツキが生まれる．

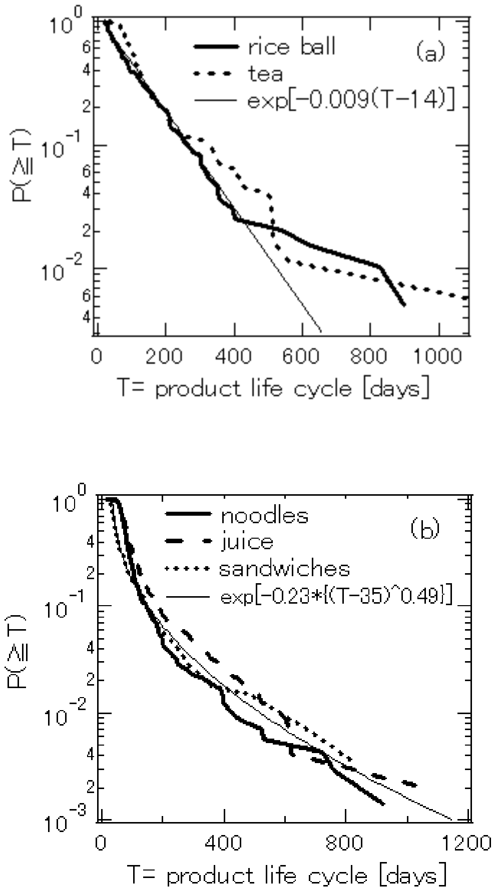

Fig. 9: 販売寿命の分布

5.4 小売市場

コンビニエンスストアにおける商品の販売寿命と消費者の支払いに関する分布についての研究を紹介する．はじめに，水野らの商品の販売寿命についての研究を紹介する [7]．コンビニエンスストアは売り場の面積が限られることから，商品の販売サイクルは非常に短い．例えば，三ヶ月間で約 110 種類のおにぎりが販売され，発売開始から 60 日以上生き残るものは僅か 30%しか存在しない．このように非常に短いことから，工場での生産管理を行ううえで，販売寿命の予測が大きな課題となっている．

Fig. 9(a) はおにぎりと 500 ml のペットボトルのお茶の販売寿命 τ の累積確率分布である．14 日から 400 日程度まで指数関数に従い減衰している．

$$P(\tau) \propto e^{-0.009(\tau-14)}, \quad 14 \leq \tau \leq 400$$

一方，**Fig. 9(b)** はカップ麺と 500 ml のペットボトルのジュース，サンドイッチの販売寿命 τ の累積確率分布で

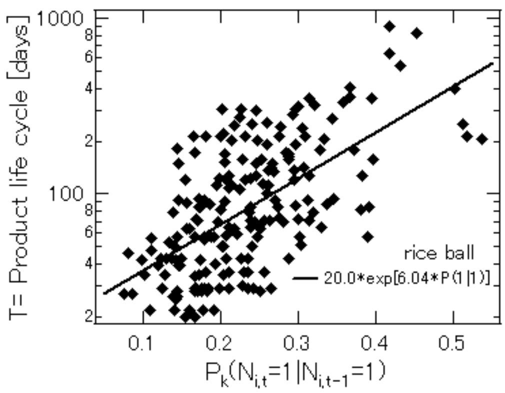

Fig. 10: リピート率と販売寿命の相関

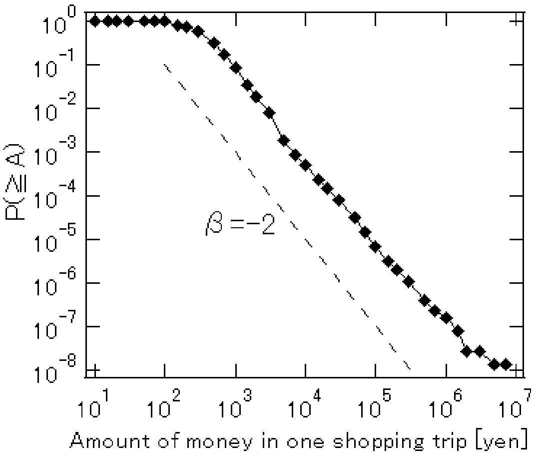

Fig. 11: コンビニでの支払金額の分布

ある．これらは約1ヶ月（35 日）以上に関して拡張指数関数に従い減衰している．

$$P(\tau) \propto e^{-0.23(\tau-35)^{0.49}}, \quad \tau \geq 35$$

この結果は4節の指数関数の議論より，カップ麺と 500 ml のペットボトルのジュース，サンドイッチでは長く市場に生き残るほど，販売が打ち切られる確率が下がっていくことが分かる．従って，過去の販売日数から残り販売寿命の予測が可能である．一方，おにぎりや 500 ml のペットボトルのお茶は，過去の販売日数と残りの販売寿命に相関はなく，過去に依存せず打ち切られる確率は一定である．つまり，過去の販売日数による残りの販売寿命の予測には意味が無い．

では，おにぎりや 500 ml のペットボトルのお茶について，残りの販売寿命を予測するには，どのようにすれば良いのであろうか．実は，それぞれの商品に対するリピーターの割合と販売寿命に相関があることが分かっている．各おにぎりについて，前回購入した「おにぎり」と同じ「おにぎり」を購入する割合を調べ，それをリピート率と定義する．**Fig. 10** は，各おにぎりについて，横軸にリピート率を縦軸に販売寿命をプロットした図である．リピート率が高いほど，販売寿命が長いことが読み取れ，残りの販売寿命の予測に使えることが分かる．この図では一次関数による近似が用いられているが，どのような関数が最適であるのかは，今後の課題である．

リピート率は，店舗が保有する顧客の電子マネーの履歴のデータを用いれば，マイナーな商品でも発売後3日程度あれば算出することが可能である．つまり，十分に工場の生産管理に利用できる．

一昔前のバブルの頃，東京の大手百貨店は1日に 1,000 万円以上購入した大口顧客が何組みいたかを数えるだけで，その日の全体の売上が予測できたと言われている．何故，そのようなことが可能だったのか．そこには消費者の支払い分布に潜む法則が関係する．次に，その法則について調べた水野らの研究を紹介する [8, 9].

当時，水野らは，コンビニエンスストアにおける顧客が一回の買い物で支払う金額に注目し，客単価を上げるにはどのようなキャンペーンが有効であるか模索していた．**Fig. 11** は，一回の買い物で支払う金額 A の累積確率分布である．客の 10 人に 1 人が 1000 円以上を，1000 人に 1 人が 1 万円以上を，なんと，1000 万人に 1 人が 100 万円以上を使っていることが分かる．しかも，これらの確率が同じ関数で連続的に繋がっているのである．この分布は金融市場と同じべき関数

$$P_{>}(A) \propto A^{-2}, \quad A \geq 700\text{ 円}$$

に従っている．このべき指数 2 は，立地や消費者の性別，年齢に依らず常に観測される．つまり，このべき法則が成り立つため，高額の支払者の金額と人数さえ分かれば，この関数に従って少額の支払者の人数を推定することができる．

支払金額のべき分布が現れるメカニズムについて考察する．3 節の安定分布の議論から，コンビニエンスストアで売られている商品間での価格の分散が発散しているか，または，一回の買い物で連鎖的に商品を購入してしまう相関構造が存在しているかという可能性が考えられる．通常，コンビニエンスストアでは日常品のみが販売されていることから，前者の価格の発散が分散するほど大きいとは考えにくく，後者の消費者の購買行動に何かしらの相関構造があると考えるほうが自然である．

消費者の購買行動に関する相関構造を示す．5.1 節の金融市場の研究例のように，べき分布を示す現象には，しばしば，トレンド形成のような相関構造が存在する．**Fig. 12** は，1 回の買い物において $N-1$ 個の商品をカゴに入れた条件下で，N 個目の商品をカゴに入れる確率を表している．$N \leq 5$ 個までは確率が減少する一方，$N>5$

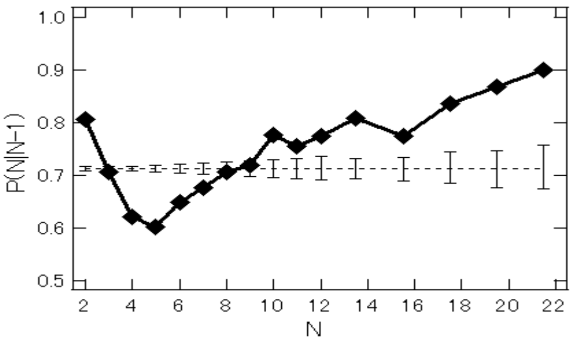

Fig. 12: 購買行動に関する相関構造

個以上では，N に比例して次の商品をカゴに入れる確率が上昇している．つまり，4 個以上の商品をカゴに入れると，他の商品を連鎖的に購入しやすくなり，この性質は支払金額の分布を正規分布から乖離させる．4 個は約 700 円に相当し，この金額を超えさせることが店舗経営において重要である．そのために，しばしば行われる 700 円を境にしたキャンペーンは，このようなメカニズムを経験的に利用したものであると推察される．

6. まとめ

本稿では，経済物理学の特徴的な分析手法の 1 つである分布関数の直接観測による相関構造の検出について，その背後にある安定分布の理論と，金融市場，不動産市場，家電オンライン市場，小売市場に関する研究例を紹介した．多くの経済現象で共通して，べき分布が観測される．現在，これらの研究例のように，安定分布の理論をもとにして各論的にべき分布が現れるメカニズムが解明され，それらの多くは経済主体間の何らかの相互作用が原因であることが分かってきた．近い将来，分布関数と相互作用の間の普遍的な関係が明らかになることが，経済物理学では期待されている．

謝辞: 本稿は日本学術振興会科研費若手研究（B）「ブーム学の基盤構築：経済主体間の創発メカニズムの解明」（課題番号：24710156）の助成のもと活動の一環として作成されたものである．

参考文献

[1] 蓑谷千凰彦: 統計分布ハンドブック, ISBN-10: 4254121547，朝倉書店，2003.

[2] R. N. Mantegna and H. E. Stanley, “Introduction to Econophysics,” ISBN-10: 0521039878, Cambridge University Press, 2007.

[3] R. N. Mantegna and H. E. Stanley, “Scaling behavior in the dynamicsofan economic index,” Nature, Vol.376, pp. 46-49, 1995.

[4] T. Mizuno, S. Kurihara, M. Takayasu, and H. Takayasu: “Analysis of high-resolution foreign exchange data of USD-JPY for 13 years,” Physica A, Vol.324, pp. 296-302, 2003.

[5] T. Ohnishi, T. Mizuno, C. Shimizu, and T. Watanabe: “Power laws in real estate prices during bubble periods,” Int. J. of Modern Physics: Conf. Series, Vol.16, pp. 61-81, 2012.

[6] T. Mizuno and T. Watanabe: “Why are product prices in online markets not converging?,” to appear in PLOS ONE.

[7] T. Mizuno and M. Takayasu: “The Statistical Relationship between Product Life Cycle and Repeat Purchase Behavior in Convenience Stores,” Progress of Theoretical Physics Supplement, Vol.179, pp. 71-79, 2009.

[8] T. Mizuno: “Power Law of Customers’ Expenditures in Convenience Stores,” J. of the Physical Society of Japan, Vol.77, 035001, 2008.

[9] T. Mizuno, M. Toriyama, T. Terano, and M. Takayasu: “Pareto law of the expenditure of a person in convenience stores,” Physica A, Vol.387, pp. 3931-3935, 2008.

水野 貴之

1977 年 12 月 15 日生．2005 年中央大学大学院理工学研究科物理学専攻博士後期課程修了．2013 年国立情報学研究所・総合研究大学院大学複合科学研究科准教授，現在に至る．経済物理学，データサイエンスなどの研究に従事．理学博士．日本物理学会，情報処理学会，日本経済学会などの正会員．