



HOKKAIDO UNIVERSITY

Title	高頻度金融データの統計的性質とそのマイノリティゲームによる再現性に関する基礎研究
Author(s)	日野, 光
Degree Grantor	北海道大学
Degree Name	学士
Issue Date	2009-03-25
Doc URL	https://hdl.handle.net/2115/43957
Type	bachelor thesis
File Information	hino2008.pdf



平成 20 年度 卒業論文

高頻度金融データの統計的性質とそのマイノリティゲーム
による再現性に関する基礎研究

情報工学コース 混沌系工学研究室
日野 光

目次

1	序論	1
2	金融データの性質	3
2.1	高頻度データの解析 (金融商品の価格変動幅)	3
2.1.1	データフォーマット	3
2.1.2	歪度 (skewness) と尖度 (kurtosis)	4
2.2	高頻度データの解析 (変動時間間隔)	5
2.2.1	ワイブル確率紙	5
2.2.2	ジニ係数	6
2.2.3	平均待ち時間	8
2.3	この節のまとめ	9
3	ミクロなモデルによるシミュレーション	9
3.1	システムの説明	9
3.2	マイノリティ・ゲームの価格変動について	12
4	変動間隔の揺らぎの生成	13
4.1	第一通過過程 (First Passage Process)	13
4.2	第一通過過程の幅 ϵ とマイノリティゲームの履歴数 M の関係	14
5	解析結果	15
5.1	価格変動幅の統計的性質	15
5.1.1	実際の金融データ	15
5.1.2	マイノリティ・ゲーム	15
5.2	変動時間間隔の統計的性質	16
5.2.1	考察	16
6	結論	17
7	Appendix	20
7.1	標準正規分布の歪度と尖度の証明	20
7.2	最小二乗法	21
7.3	平均待ち時間の導出	22
7.4	カルバック-ライブラ距離による第一通過過程の効果の測定: 実データ	23

1 序論

近年、数秒オーダーでの時間スケールにおける価格変動を記録した高頻度金融データの入手が可能となり、既存の経済理論、なかでも金融工学において、その数学的な取り扱いの容易さから用いられてきたいくつかの仮説/仮定が、これら高頻度データとの整合性の観点から見直されてきている。特に、ブラック・ショールズ方程式の導出において仮定される「金融商品の価格変動はブラウン運動に従う」という仮説、つまり「価格変動幅の分布が正規分布に従う」という前提は、実際のデータから観測される分布が「広い裾 (Fat tails)」を持つことが明らかにされることで修正の必要性が認識されるようになった。このような大きな変動が比較的大きな確率で生じる確率過程としてはレビ・フライト (Levy flight) などが知られており、この場合の価格変動幅の確率分布は裾が広いレビ分布となる。

しかし、このような金融商品の価格変動をより根源的に説明するためには、その商品の価格を決定する市場のメカニズムを深く理解することが不可欠である。例えば、市場には多くのトレーダが参加し、彼らは各自が商品の価格変動や政治情勢など様々な情報を考慮することで、自分の利得を可能な限り最大化する戦略を立て、金融商品を売り買いするが、そのようなトレーダ群の意思決定の結果として金融商品の価格変動が生じ、それを我々は観測するわけである。従って、このような社会システムは自ずとある種の「階層性」を持つことになる。つまり、「ミクロ」なレベルでのトレーダの判断から「マクロ」なレベルの金融商品の価格変動までの間には様々なスケールの経済活動が複雑に絡み合っており、このミクロとマクロの関連性を理解すること、言い方を変えれば、金融商品の価格変動(マクロ)をトレーダ群の意思決定(ミクロ)の結果として矛盾なく説明することは野心的課題ではあるが、既存の経済理論を「多体問題」「統計力学」の観点から一歩進めるためには重要なプロセスであり、そのような研究は既存の経済学に対し新しい視点を提供することが期待できる。

そこで本研究ではこの目標に対する第一歩として、高頻度金融データの価格変動時間間隔に着目し、この統計的性質を様々な観点から調べることで、金融工学における計算機シミュレーションで多くの場合に仮定されてきた「価格変動に関するポアソン過程」の前提の妥当性を検討する。ついで、この統計的性質をミクロなトレーダ群の意思決定の結果として再現するため、「マイノリティゲーム」と呼ばれる繰り返しゲームに対し、その第一通過過程を用いることで、得られる価格変動間隔の統計的性質を実データ解析による結果と比較する。これらの結果、従来の「ポアソン過程」の前提は必ずしも正しくなく、短い時間スケールの変動間隔と長い時間スケールの変動間隔の間にはある種の「クロスオーバー (crossover: 分布の切り変わり)」が生じ、分布は指数則から裾の広い分布へ変化することが実データ解析/マイノリティゲームによるシミュレーション双方で普遍的に現れることを明らかにする。

本論文の構成は以下の通りである。第2章では本研究で扱う金融データの説明

を行い、その価格変動幅および価格変動間隔に関する統計的性質を解析する。第3章ではミクロなモデルによるシミュレーションとして市場の履歴を伴うマイノリティゲームを導入し、続く第4章では第一通過過程を用いた価格変動間隔の生成方法について説明する。第5章では実データ解析、マイノリティゲームによる計算機シミュレーションの具体的な結果を示し、時間間隔の分布が指数則から裾の広い分布へ変化することが実データ解析/マイノリティゲームによるシミュレーション双方で普遍的に現れることを示す。第6章では全ての結果についての考察を行い、今後の研究計画についても言及する。

2 金融データの性質

2.1 高頻度データの解析 (金融商品の価格変動幅)

ブラックショールズの公式など初期の金融工学では価格変動の分布を正規分布と仮定しているが、金融商品に関する大規模なデータが入手可能である現在では、そのデータを直接的に解析することで、正規分布の仮定の妥当性を確認することができる。そこで、ここではBTP future(イタリアの国債で10年物固定利付債)の1997年1月から6月までのデータを用いて、その仮定の正当性を調べるために正規分布からの「ずれ」を以下で述べる統計量から計る。

2.1.1 データフォーマット

BTPのデータは以下の様なテキスト形式で表されている。BTPデータの変動のグラフは図1に表示する。図1ではデータの最初の時刻を0としている。

日月年 時分秒	価格
130397 080013	12530
130397 080015	12525
130397 080017	12525
130397 080019	12535
130397 080021	12525

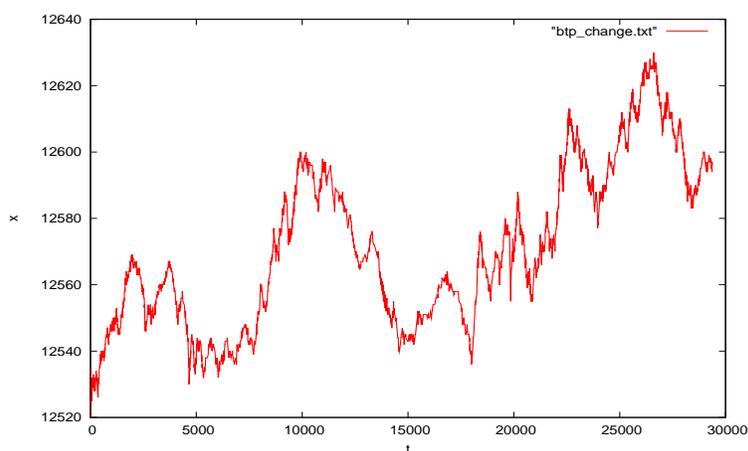


図 1: 金融商品の価格変動のグラフ

2.1.2 歪度 (skewness) と尖度 (kurtosis)

ここでは、金融商品の価格変動についての分布の対称性や尖度、すその厚みの性質などを調べる。歪度は確率分布の対称性を測る量である。原点まわりの2次および3次のモーメント μ_2 及び μ_3 を使って

$$s = \frac{\mu_3}{\mu_2^{\frac{3}{2}}}$$

によって定義される。対称分布では $s=0$ である。 m 次のモーメントは確率変数 x ($|x| \leq \infty$) の分布を $P(x)$ として

$$\mu_m = \int_{-\infty}^{\infty} x^m P(x) dx$$

で定義される。

一方、尖度は単峰分布の確率密度関数の形がモードの近傍で尖っているのがあるいは平らかという視点から正規分布からのずれを評価する量である。その尖度は分布の原点まわりの2次及び4次のモーメント μ_2 及び μ_4 を使って

$$k = \frac{\mu_4}{\mu_2^2}$$

によって定義される。標準正規分布ではパラメータに関係なく3になる。

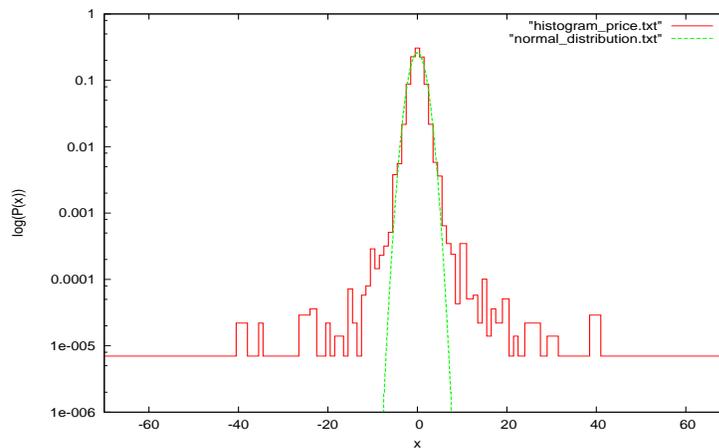


図 2: 正規分布と比べた金融商品の価格変動幅のグラフ (skewness=0.087201 kurtosis=23.752031)

図 2 は BTP の価格変動幅の分布と平均、分散が同じ正規分布を y 軸を対数スケールにしてプロットしたものである。図から見てわかるように、実際の金額のデータはブラック・ショールズ公式などで用いられている正規分布ではなく、すそが広い分布になっている。つまり、価格が大きく変動する確率は予想以上に大きいことが分かる。このことから、数学的取り扱いの容易さから用いられ、検証無

しに用いられてきた価格変動間隔の正規分布の仮定は見直す必要があることがわかる。従って、市場や金融商品の価格変動を数理モデルを用いて表現する際には、得られる結果を実データと照らし合わせて検証することが不可欠である。よって、本研究では数理モデルの構築と、その実データによる検証の双方を重視する。

2.2 高頻度データの解析 (変動時間間隔)

ここまで、金融商品の価格変動幅についての統計的な性質を調べてきたが、ここからは価格変動の変動間隔に着目する。価格の変動間隔を予想する際、最も素朴には変動が「ランダム」に発生するという仮定を用いるであろう。事実、多くの金融市場のシミュレーションでは変動をポアソン仮定に従って生成するものが多い。価格変動がポアソン過程に従う場合、計測を始めてからはじめの変動が生じるまでの変動時間間隔 t の従う分布はポアソン分布 $P_k(t) = (\lambda t)^k e^{-\lambda t} / k!$ で $k = 0$ とおいた

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

ような指数分布に従うことになる。しかし、前節で変動幅の正規分布仮定が正しくなかったように、変動間隔の従う「点過程」がポアソン過程に従うという仮定が正しいか否かを確認することは、市場をモデル化する際に重要となる。従って、ここでは実際の金融データに関し、そのポアソン過程からの「ずれ」を以下の統計量から計る。

2.2.1 ワイブル確率紙

価格の変動間隔がポアソン分布であるか否かを調べるためワイブル確率紙の方法を用いる。確率紙とは分布関数を普通目盛りのグラフに書くと曲線になるところを、横軸と縦軸のスケールを適切に取り直して直線であてはめ、その傾きなどから、分布の形状についての情報を引き出す古典的手法である。

我々が良く知っている例として、分布のべき則を確認する時、我々はべき則が $P(t) = t^{-\alpha}$ なので $\underbrace{\log P(t)}_Y = -\alpha \underbrace{\log t}_X$ となり、両対数プロットで直線に乗るか否かを調べた。

また、指数則を確認する場合は $P(t) = e^{-\beta t}$ なので $\underbrace{\log P(t)}_Y = -\beta \underbrace{t}_X$ である

から、片対数プロットで直線に乗るか否かを調べた。これらを一般化するものとして、以下のワイブル確率紙の方法が用いられる。

ワイブル分布は

$$P_{m,a}(t) = \frac{mt^{m-1}}{a} \exp\left(\frac{-t^m}{a}\right) \quad (1)$$

のように表され、指数分布を一般化したものである。 $m = 1$ であれば指数分布になり、従ってその変動間隔分布を持つ確率過程はポアソン過程になる。 a は分布のスケールを決定するパラメータである。ワイブル確率紙の横軸は $\log(t)$ 、縦軸はワイブル分布の式 (1) の累積分布

$$W(t) = 1 - \int_0^t P_{m,a}(t) dt = 1 - \exp\left(\frac{-t^m}{a}\right)$$

の $\log(-\log W(t))$ である。従ってデータをこのワイブル確率紙にプロットし、それが直線に乗れば問題の分布はワイブル則に従うことが分かり、このときの直線の傾きが $m=1$ であれば、その確率過程はポアソン過程であることが分かる。

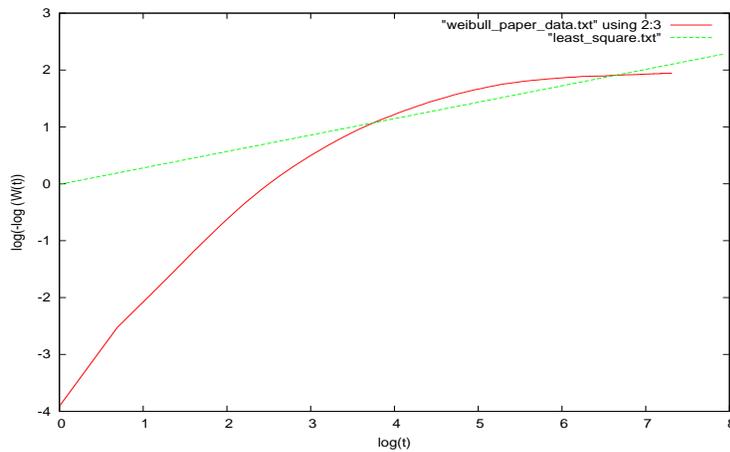


図 3: ここで扱った実データのワイブル確率紙 ($m=0.289009$)

図 3 はワイブル確率紙を用いてデータの分布をプロットしたものと、データの分布に対して最小二乗法を用いて、最もフィットする直線を求めたものである。最小二乗法に関しては 7 節 Appendix で記述する。

2.2.2 ジニ係数

主に社会における所得分配の格差を測る指標として、ジニ係数が知られている。この係数は $[0,1]$ 間の実数値をとり、0 に近いほど平等になり、1 に近いほど格差が大きくなる。また、所得のばらつきがランダムなポアソン分布に従う場合、そのジニ係数は 0.5 となる。これを、ここでは時間の変動間隔の「格差」を測る目的で使う。ジニ係数を解析的に求めるには、まずレートの変動間隔の分布からローレンツ曲線を求めなければならない。そのため分布を $P(t)$ とし、その累積分布と規格化された累積分布

$$X(r) = \int_0^r P(t)dt$$

$$Y(r) = \frac{\int_0^r tP(t)dt}{\int_0^\infty tP(t)dt}$$

を計算し、 r を媒介変数として X と Y の関係をプロットする。このとき描かれる曲線がローレンツ曲線である。これは $[0,1]$ 間で下に凸の曲線であり、完全平等の場合には $P(t) = \delta(t - t_*)$ を代入し

$$X(r) = \int_0^r \delta(t - t_*)dt = \Theta(r - t_*) \quad (2)$$

$$Y(r) = \frac{\int_0^r t\delta(t - t_*)dt}{\int_0^\infty t\delta(t - t_*)dt} = \frac{t_*\Theta(r - t_*)}{t_*} = \Theta(r - t_*) \quad (3)$$

となるから、 $Y = X$ がローレンツ曲線となる ($\Theta(\dots)$ は階段関数)。これを「完全平等線」と呼ぶことにする。ワイブル分布(1)のローレンツ曲線は簡単な計算の結果

$$Y = Q\left(\frac{1}{m} + 1, -\log(1 - X)\right)$$

となる。ここに、 $Q(z, x)$ は不完全ガンマ関数であり

$$Q(z, x) = \int_0^x t^{z-1}e^{-t}dt$$

で定義される。ジニ係数はこのローレンツ曲線と完全平等線との間の面積を2倍したものとして算出される。従って、ローレンツ曲線自体が完全平等線の場合にはゼロ。その対極のある時間でのみ変動が起こった場合、(2)(3)の結果に $t_* = r = \infty$ を代入して、 $Y = X = 1$ がローレンツ曲線となり、このときの面積は $Y = X$ と X 軸、 Y 軸とで囲まれる三角形の面積 $1/2$ であり、従ってジニ係数はその2倍の1となる。

具体的にジニ係数は

$$G = 2 \int_0^1 (X - Y)dX = 2 \int_0^\infty (X(r) - Y(r))\frac{dX}{dr}dr$$

と書けるので、具体的にワイブル分布に対して計算すると

$$G = 1 - \left(\frac{1}{2}\right)^m$$

が得られる。よって $m = 1$ としてレートの変動がランダムなポアソン過程に従う場合、ジニ係数は 0.5 となる。

実データがある場合、数値的にジニ係数を求めることもできる。まず各 Δt が $t_1 \leq t_2 \leq \dots \leq t_N$ となるようにデータをソートする。次にレート変動間隔の累積分布

$$X_i = \frac{i}{N}$$

と規格化された累積変動間隔

$$Y_i = \frac{\sum_{r=1}^i \Delta t_r}{\sum_{r=1}^N \Delta t_r} = \frac{\sum_{r=1}^i \Delta t_r}{N\mu}$$

$$\mu = \frac{1}{N} \sum_{i=0}^N \Delta t_i$$

を計算し、 $dX_i = (i + 1)/N - i/N = 1/N$ に注意すると

$$G = 2 \sum_{i=1}^N N(X_i - Y_i)dX = \frac{1}{N^2\mu} \sum_{i=1}^N (2i - N - 1)\Delta t_i$$

となり、この G が求めるジニ係数となる。この式を用いると、この実データでのジニ係数は 0.859 となった。これから、このデータのレート変動はランダムに起こっているのではなく、何らかの時間相関を持ち、飛びぬけて大きな間隔、あるいは小さな間隔が所々で現れていることが分かる。

2.2.3 平均待ち時間

ここまでの解析でジニ係数より我々がここで扱っている実データの変動間隔 Δt は一定ではなく、ある確率分布 $P(\Delta t)$ で特徴付けられる揺らぎを持っていることが分かった。ここからは、ある時刻にレートをチェックしてから、実際にレートが変動するまでの平均待ち時間を考える。時間間隔を Δt とすると平均待ち時間 $\langle s \rangle$ は以下の式で求まる。詳しい導出は Appendix で述べる。

$$\langle s \rangle = \frac{\langle (\Delta t)^2 \rangle}{2 \langle \Delta t \rangle}$$

$\langle \dots \rangle$ は時間間隔 Δt の分布 $P(\Delta t)$ についての平均、 $\langle \dots \rangle = \int_0^\infty p(\Delta t)(\dots)d\Delta t$ である。この実験では $\langle \Delta t \rangle = 46.2764$ $\langle s \rangle = 20059.3$ となった。レートが変動する間隔よりも、待っている時間のほうが長いのは、一見すると矛盾しているように思える。このことを Inspection Paradox(観測上のパラドクス)と呼ぶ。しかし、これはジニ係数から分かるように、レートの変動間隔に大きな偏りがあるため、極端に長い変動間隔、極端に短い変動間隔が現れ、前者のタイミングでチェッ

クするチャンスは、後者のものと比べて、圧倒的に多くなることが予想できる。このようなタイミングを含めて平均すると、結果として待ち時間のほうが大きくなることがあり得るわけである。

2.3 この節のまとめ

価格変動のモデルは正規分布ではなく、より裾の広い分布をしている。つまり、正規分布で仮定されているよりも多く大きな価格変動（暴落や暴騰）が無視できない頻度で起こりやすいことが分かった。価格変動のモデルとしてより近いものとして、価格変動のボラティリティが時間的に変動するような ARCH(Auto Regressive Conditional Heteroskedasticity) 過程や GARCH(Generalized ARCH) 過程が知られている。一方、レートの変動はランダムに起こるポアソン過程ではなく、何らかの相関を持っていることが分かった。これから変動の間隔は飛びぬけて大きな間隔や逆に小さな間隔があることが分かった。レートの変動間隔のモデルとしてより近いものとして、Mittag-Leffler 関数などが知られている。

3 ミクロなモデルによるシミュレーション

そもそも金融商品の価格の変動は、膨大な数のトレーダの売り買いの意思決定の結果起こっている。また、トレーダは開示されている為替レートの情報に基づいて売り買いを決定しているので、為替レートの変動というマクロな情報とトレーダの行動というミクロな情報は互いに影響していると考えられる。そこで、本節ではゲーム理論の中でマイノリティ・ゲームと呼ばれる繰り返しゲームに着目し、トレーダの行動が金融商品価格の過去 M ステップに遡る上下情報を共有し、それに基づき、自らの行動を決定する状況をシミュレーションすることで、ミクロなトレーダの動きがマクロな情報における確率変数の統計的性質にどのように影響するのか調べる。具体的には前節までで得られた実データの変動幅、及び変動間隔に関する統計的な性質が、マイノリティ・ゲームによって再現できるか否かを確かめる。

3.1 システムの説明

ここでのマイノリティ・ゲームとは N 人(奇数)が「売り」「買い」のどちらかの判断を行って少数派のグループに属したものが勝つというものである。各トレーダ i ($i=1, \dots, N$) は各ラウンド l で、入札価格 $b_i(l)$ を決める。ここで入札価格 $b_i(l)$ は ± 1 を (+1: 「売り」 -1: 「買い」) の 2 値をとるものとする。つまり、ここでの「入札」の解釈は「 $b_i(l)$ が正の値を取るならば、ラウンド l でトレーダ i は商品を

価格 b_i で買いたいと考えている」ということである。このとき、入札価格 $A(l)$ を

$$A(l) = \frac{1}{\sqrt{N}} \sum_{i=1}^N b_i(l) \quad (4)$$

とする。 $1/\sqrt{N}$ のファクタは $A(l)$ が N の大きさに影響されないようにオーダが1となるために付けた。 $A(l)$ が正の値を持つときは市場はそのラウンドで「売り超過」にあり、負の値を持つときは「買い超過」にあることを意味する。今の場合、 N を奇数に選んだ関係で必ず市場は「売り超過」か「買い超過」になる。各トレーダはラウンド l での意思決定に際し、過去の M ラウンドまで定義される市場の履歴ベクトル λ を用いる。

$$\lambda(l, A) = \begin{bmatrix} \text{sgn}[A(l-1)] \\ \text{sgn}[A(l-2)] \\ \dots \\ \text{sgn}[A(l-M)] \end{bmatrix}$$

履歴ベクトル λ の各成分はそのラウンドでシステム全体が「売り超過」か「買い超過」なのかを ± 1 で示している。履歴ベクトル λ は各成分が ± 1 の2値をとるので 2^M 通りある。例えば λ が以下のものであったとする。

$$\lambda(l, A) = \begin{bmatrix} +1 \\ +1 \\ \dots \\ +1 \end{bmatrix}$$

これは過去 M ラウンドで市場が常に「売り超過」であったことを意味する。この時、あるトレーダは「次も「売り超過」だろうから自分は買おう」と考えるかもしれないし、また別のトレーダは「次のラウンドあたりには「買い超過に転じるはず」と状況を裏読みして売るかもしれない。このように過去の履歴に対する各トレーダの行動はトレーダの戦略に依存する。ここで、提示された履歴ベクトル λ に対して各トレーダ i は2通りの戦略を取れるとする。

$$R^i_{(l,A)} = (R_{1, (l,A)}, R_{2, (l,A)})$$

このベクトルの成分値 $R_{a, (l,A)}$ は ± 1 をとり、トレーダが「売る」か「買う」かに対応することになる。ただし、戦略の値は確率 $1/2$ で ± 1 (「売り」か「買い」) を割り振る。つまり、ある履歴に対してあるトレーダの戦略が戦略1, 戦略2ともに「買い」でも構わないとする。

各トレーダはゲームの各ラウンド l で次の更新式に従って戦略 $a(=1,2)$ を選択した場

合の利得 $p_{ia}(l)$ を更新し、利得を最大化するような \tilde{a}_i を求めて、最適戦略 $R_{\tilde{a}_i(l), (l,A)}^i$ を決定する。係数 $\tilde{\eta}$ は更新の重みで、ここでは簡単のため $\tilde{\eta} = 1$ とする。

$$p_{ia}(l+1) = p_{ia}(l) - \frac{\tilde{\eta}}{\sqrt{N}} R_{\tilde{a}_i(l), (l,A)}^i A(l) \quad (5)$$

$$\tilde{a}_i(l) = \arg \max_a [p_{ia}(l)] \quad (6)$$

$$b_i(l) = R_{\tilde{a}_i(l), (l,A)}^i \quad (7)$$

ここで、式 (7) より $A(l)$ は

$$A(l) = \frac{1}{\sqrt{N}} \sum_i^N b_i(l) = \sum_i R_{\tilde{a}_i(l), (l,A)}^i \quad (8)$$

である。従って、この $A(l)$ を金融商品の価格 X の変動幅と考え、 X の確率過程を

$$X(l+1) = X(l) + A(l) \quad (9)$$

で記述する。この式から生成した x の動きを図 4 にプロットする。マクロな量である総入札価格 $A(l)$ の式 (8) にはミクロな情報である各トレーダの利得が内在しており、また、トレーダの戦略を決める利得の式 (5) には、総入札価格 $A(l)$ が含まれている。これより、我々が調べている市場の履歴をともなったマイノリティ・ゲームのシステムはミクロな情報とマクロな情報が無相関ではなく、互いに関係していることが分かる。

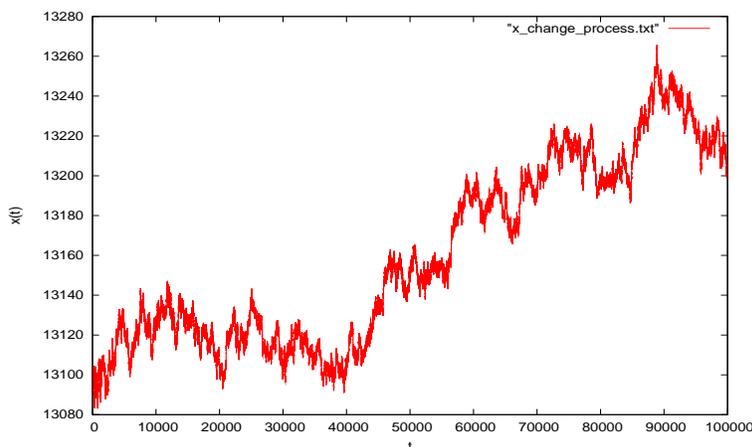


図 4: $A(l)$ を変動幅とした金融商品の (9) 式による確率過程

このマイノリティ・ゲーム以外にも、各トレーダ i ごとに異なり、かつ時間と共に異なる売値 $s_i(t)$, 買値 $b_i(t)$ を持ったトレーダが $b_i(t) \geq s_j(t)$ を満たしたときに取引が行われるようなエージェントモデルなどが知られている。この図 4 と BTPfuture の値動きの図 1 を比べると、定性的には類似の確率過程にしたがっているように見える。

3.2 マイノリティ・ゲームの価格変動について

ここではマイノリティ・ゲームの基本的な性質として、価格変動に着目したボラティリティと履歴数の関係について述べる。ボラティリティとは、ある金融商品の不安定性を示す量であり、ボラティリティが大きいほど価格の揺らぎが大きくなる。これをマイノリティゲームでは以下のように定義する。

$$\sigma^2 = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{l=1}^L A(l)^2 - \left\{ \frac{1}{L} \sum_{l=1}^L A(l) \right\}^2$$

ボラティリティと履歴数の関係を図 5 に示す。図 5 より明らかに遡る履歴数が大きい

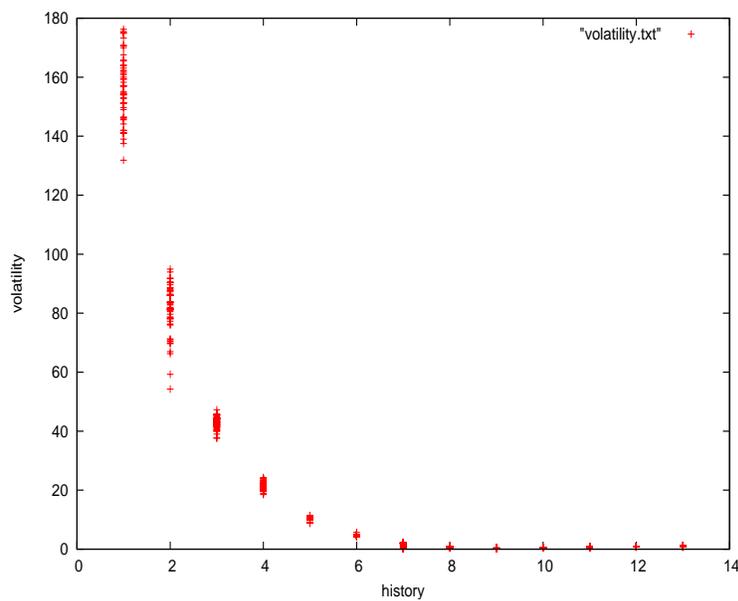


図 5: 履歴数とボラティリティの関係

い場合の方がボラティリティは小さい。同じ M の値で異なるのは、情報ベクトルや初期値などの「データ」に対応している。従って、遡る履歴数が少ない場合にはデータによるばらつきも多い。また、ボラティリティを最小化するような最適な履歴が存在することが分かる。

以下では、前章で導入したいいくつかの統計量を介して両者を定量的に比較し、マイノリティゲームを用いてBTPfutureの値動きが再現できるか否かを調べることにする。

4 変動間隔の揺らぎの生成

前節で導入したマイノリティゲームに基づく金融商品の価格変動についての確率モデル

$$X(l+1) = X(l) + A(l) \quad (10)$$

では、一定時間間隔おきに変動が生じ、その意味において変動間隔に揺らぎは存在しない。しかし、2章で見たように、実データにおいてはその種の揺らぎは存在した。よって、ここでは第一通過過程を用いて人工的に変動間隔揺らぎを生成し、得られるデータの統計的性質を調べて行くことにする。

4.1 第一通過過程 (First Passage Process)

第一通過過程を用いたレート提示方法とは、図6のように頻繁に変動する市場レートに対し、 2ϵ の幅を設定し市場レートがこの幅を超えた場合のみ、その時点でのレートを提示する手法である。実際に我々が目にしているレートも、プロのトレーダたちの取引によって決まる「市場レート」があり、そこから一日単位、あるいは一時間単位と区切って表示されている。しかし昨今、一般的に普及しているインターネットを介した金融取引では、例えば大きな事件があったりして、レートが乱高下するようなことがあった時、一日や一時間単位では顧客は乱高下のタイミングを利用することができなくなってしまう。逆に数秒毎に区切ると、レートは平時では極微小にしか動いていないので、顧客にとってあまり有益とはいえない。そこで、間引かれた情報によって作られるレートが市場レートの性質をある程度反映しつつ、情報として有益なものとしてソニー銀行などで第一通過過程を採用している。

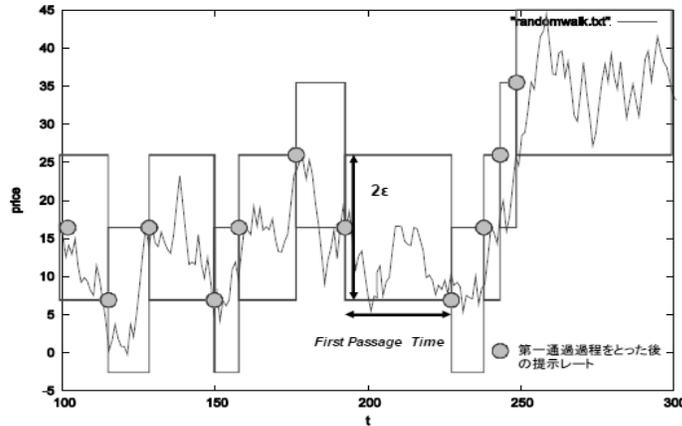


図 6: 第一通過過程を用いたとき

この第一通過過程を用いると、明らかに細かな価格変動はレートに反映されず、変動間隔は長くなり、逆に大きな変動間隔は直ちにレートに反映し、従って、変動間隔は短くなる。その結果として、変動間隔 Δt は揺らぐことになり、その分布 $P(\Delta t)$ は非自明な関数形を持つことが予想される。

4.2 第一通過過程の幅 ϵ とマイノリティゲームの履歴数 M の関係

時間の分布には第一通過過程の幅 ϵ とマイノリティゲームのシステムで用いた履歴入の参照するラウンド数（履歴数） M が関係している。図 7 では時間の分布をジニ係数で評価している。

第一通過過程の幅 ϵ が増加すると、通過しにくくなりジニ係数が増えるのは明らかだと思いが、履歴入が参照するラウンド数 M が増加するとジニ係数が増えるのは 3.2 節で述べた関係がある。つまり、履歴数が小さいとボラティリティが高くなり第一通過過程のフィルタを通過しやすくなる。フィルタを通過するデータが多いと時間の揺らぎは小さくなり、ジニ係数は小さくなるのである。また、本実験ではボラティリティが履歴数が $M=9$ の時に最小になった。その時、ジニ係数は最大となっていることから、履歴数とジニ係数に関係があることが言える。なお、第一通過過程の幅 ϵ が 4 より大きいとき、履歴数が大きいと通過するデータが存在しなくなったので図では表示していない。

次章では、第一通過過程を介して修正された価格変動間隔揺らぎをもつマイノリティ・ゲームが BTPfuture の値動きを再現できるか否かを調べることにする。

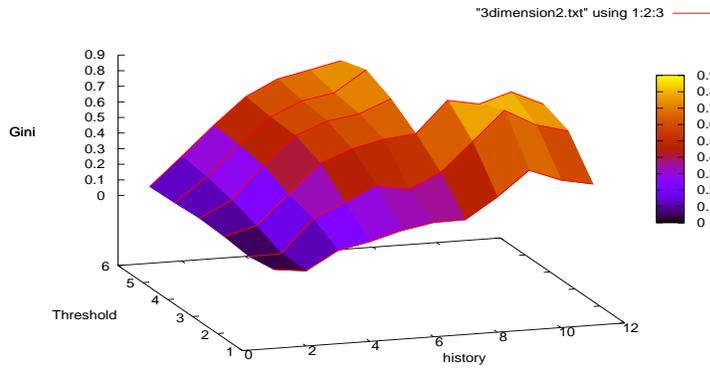


図 7: ϵ と履歴数とジニ係数の関係

5 解析結果

ここからは、マイノリティ・ゲームによる BTPfuture の値動きの再現性を調べていく。本章は本論文の中心部分である。

5.1 価格変動幅の統計的性質

5.1.1 実際の金融データ

レート分布については 2.1 節で述べた図 2 のようになる。第一通過過程を用いると、0 から幅 2ϵ の分だけデータがカットオフされる。

5.1.2 マイノリティ・ゲーム

結果は図 8 である。実際の金融データと同様、第一通過過程を用いると、中央から幅 2ϵ の分だけデータがカットオフされる。履歴数 M に依存して分布が変わることが分かる。この理由は 3.2 節で述べたように履歴数とボラティリティに相関があるからである。図 2 と比べると履歴数 M が高いと分布が近づくように思える。しかし、実データのボラティリティを計算すると 40076 と非常に高い数値となり、履歴数 M が小さい方が近いことになるので矛盾する。これは今回用いた実験データでは変動するまでの時間が長くかかった場合、価格が急激に変動しやすい傾向があり、それによってボラティリティが高くなっていると考えられる。しかし、統計的に調べてはいないので今後の課題となる。

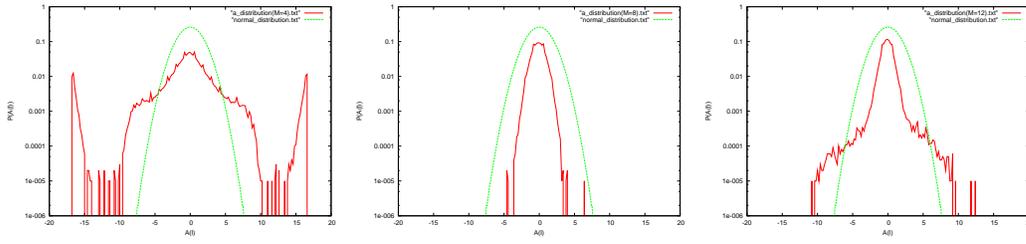


図 8: マイノリティゲームのレートの分布と正規分布、左から $M = 4, 8, 12$ で行い kurtosis は 9.16, 3.12, 19.8 となった

5.2 変動時間間隔の統計的性質

結果は図 9, 10 である。実データ、マイノリティ・ゲーム共に $\epsilon = 2$ でマイノリティ・ゲームは履歴数 $M = 9$ とした。変動時間間隔 t が $t \leq 100$ 以下では指数分布のような振る舞いをするが、2.2.2 節で述べたように、飛びぬけて大きな間隔が存在し、 t の大きい所では他の分布になっていると考えられる。このような性質を第一通過過程を用いたマイノリティ・ゲームによって再現することが出来た。

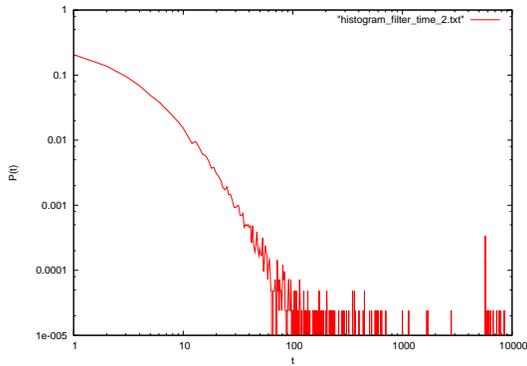


図 9: 実データの時間分布

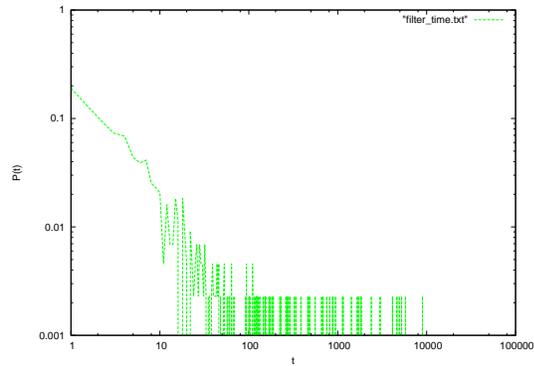


図 10: マイノリティゲームの時間分布

5.2.1 考察

実データ、マイノリティゲーム共に t が大きい所では他の分布に切り替わると考えられる。そこで他の分布に切り替わる点を t_* としたとき、 $t \geq t_*$ での分布 $P(t)$ が何かを考える。候補として

$$P(t) = t^{-\alpha} (\alpha \geq 1) \quad (11)$$

$$P(t) = \begin{cases} t^{-\alpha} & (\alpha < 1) \\ 0 & (t_c \leq t) \end{cases} \quad (12)$$

のようなベキ分布が考えられる。ベキ分布になると考える理由はある確率分布 $P(t) = \beta e^{-\alpha t}$ が $\alpha \rightarrow 0$ となる時、指数ではなく $t^{-\alpha}$ とベキで表せるからである。これを臨界緩和という。 t_c はカットオフである。仮に (11) 式が正しいとすると $P(t)$ は

$$P(t) = \begin{cases} Ae^{-\alpha t} (0 \leq t \leq t_*) \\ t^{-\gamma} (t_* \leq t) \end{cases} \quad (13)$$

となり、ジニ係数を $G(\alpha, \beta)$ のように α, β の関数として表すことができ、実データのジニ係数と比較できるようになる。しかし、(11) 式が正しいかどうかの定量的判断には統計的検定が必要になる。

このような分布の切り替わりが発生する理由として時間変動間隔の短い領域と長い領域でトレーダの意識・意思決定に違いが生じていると考えられる。短い時間間隔でレートが変動する領域では、トレーダ達も競って取引を行い、それが成立する確率も高くなるので、ますます頻繁に取引が成立し、価格が細かく変動する。従って、クロスオーバの点よりも短い時間スケールでは、このような取引に関する「正のフィードバック効果」があると考えられる。結果、極端に短い間隔も高頻度で現れて、指数、あるいはそれよりも速い減衰をするような分布で記述される。一方、長い間隔で取引が起こる領域では、なかなか価格が確定しないので、トレーダ達も自分の取引にある種の疑心を持ち始め、売る、買うのアクションを躊躇し始める。すると、その「躊躇」という行動が商品の価格にも反映し、なかなか価格が上下動しないことから、それを見てトレーディングするトレーダ達の行動もますます鈍くなる。従って、切り替わりよりも長い時間スケールでは、このような「負のフィードバック」が生じ、従って、相対的に長い間隔が高頻度で出ることにより、分布が広い裾を持ち始め、これが冪分布に従う。つまり、負のフィードバック効果により、取引間隔が長く長く伸びていくことになると考えられる。以上は実データに関する考察だが、第一通過過程を用いたマイノリティ・ゲームでは、情報を間引くことによって、時間の揺らぎを生成しているので、切り替わりが起こるメカニズムは異なるかもしれないが、同様のメカニズムが働いている可能性もある。

6 結論

本論文では、第一通過過程を介した金融データの統計的な性質を調べ、マイノリティゲームを用いてその一部を再現した。特に変動時間間隔においては、従来、

検証なしに仮定されていたポアソン過程ではなく、時間が大きい所では指数分布ではない他の分布になるような切り替わりがあることを実データで確認し、マイノリティゲームによって再現することが出来た。一方、価格変動に関しては第一通過過程が分布にどのように影響しているかは確認できなかった。変動時間間隔の揺らぎが内在してなかったことや、エージェントが価格変動を上がったか下がったかの2値で捉え、価格変動の大きさに考慮しなかった点など、実際の市場や人の行動をどのようにモデル化するかが課題となる。

マイノリティゲームの統計的な性質の再現には以上のような問題があったが、基本的な性質を再現するモデルとしては有効であったと考える。また、実データでは変動時間間隔の大きな所などの発生しにくく、データ数が少ない所の分析が難しいがシミュレーションではいくらかでも行えるので、データを蓄積し分析することができ、シミュレーションの意義があると考え。今後取り組むべき課題については以下のようなものが挙げられる。

- マイノリティゲームによる価格更新式への変動間隔の導入
具体的には価格 $X(l)$ の状態更新を

$$X(l + \Delta l) = X(l) + A(l) \quad (14)$$

で与え、変動間隔 Δl の分布 $P(\Delta l)$ を実データ解析から見積もることで連続時間ランダムウォーク (Continuous-Time Random Walk: CTRW) を構成する。

- BTP future の $X(l)-X(l+1)$ プロットから示唆される情報を用いたマクロな発展式と、そこで得られる知見のマイノリティゲームへのフィードバック
BTP future の連続する時刻間の価格の散布図, 図 11 から

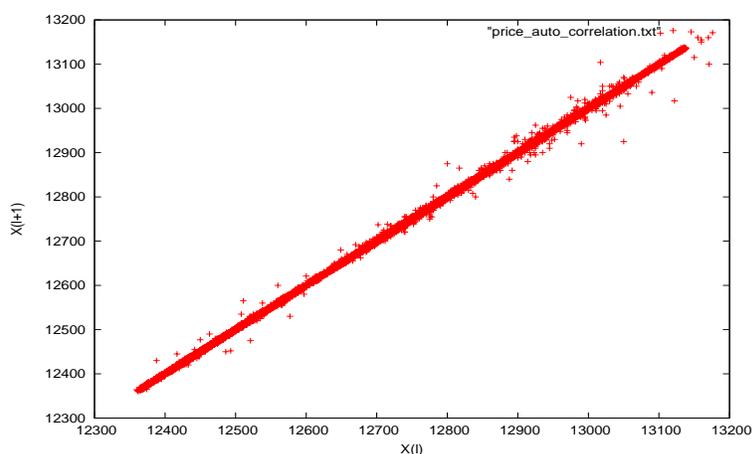


図 11: $X(l)-X(l+1)$ プロット

価格変動は第 1 近似として次のマクロな更新式に従うように見える。

$$X(l+1) = aX(l) + \eta(l) \quad (15)$$

ここに a は定数, $\eta(l)$ は白色ノイズ $\langle \eta(l)\eta(l') \rangle = \delta_{l,l'}$ である。従って, 価格変動に関し, このマクロな更新式をミクロなレベルで再現するようにマイノリティゲームに修正を加えていく。

- マイノリティゲームにおける複数コミュニティの導入とコミュニティ内での情報共有効果

本論文で扱ったマイノリティゲームでは, 各トレーダは市場履歴のみからマクロな情報を得て, 自らの状態を更新したが, システムに複数のコミュニティを導入し, そのコミュニティ内のトレーダ間の情報のやり取りを許すことで, ゲームのマクロな振る舞いに違いがでる可能性がある。各トレーダの各コミュニティへの分割指針としては, 例えば, 複雑ネットワークが形成されるようにするなどが考えられる。

- 第一通過過程フィルタの相互情報量による特徴づけ: 情報理論的アプローチ
第一通過過程を経たデータが, 元の高頻度データの持つ情報をどの程度「保存」しているのかを定量化するため, 高頻度な市場データの時間間隔列を「入力」: $\mathbf{X} = \{\Delta l_i | i = 1, \dots, N\}$, フィルタを介した後のデータ (顧客提示データ) の時間間隔列を「出力」: $\mathbf{Y} = \{\Delta t_i | i = 1, \dots, N\}$ に見立てた場合の相互情報量:

$$I(\mathbf{X} : \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X}) \quad (16)$$

を窓幅 ϵ の関数として計算することで, 第一通過過程によって間引かれたデータの損失した情報量を定量化する。ここに, $H(\mathbf{Y}), H(\mathbf{Y}|\mathbf{X})$ はそれぞれエントロピー, 条件付きエントロピーである。もし, 上記の意味での入力列と出力列がともに i.i.d. (independent identical distributed) ならば

$$P(\mathbf{X}) = P(\Delta l_1, \dots, \Delta l_N) = P(\Delta l_1) \cdots P(\Delta l_N) = \{P(\Delta l)\}^N \quad (17)$$

$$P(\mathbf{Y}) = P(\Delta t_1, \dots, \Delta t_N) = P(\Delta t_1) \cdots P(\Delta t_N) = \{P(\Delta t)\}^N \quad (18)$$

と簡略化され, 単位データあたりの出力情報のエントロピー $H(\mathbf{Y})$ は

$$\frac{H(\mathbf{Y})}{N} = - \sum_{\Delta t} P(\Delta t) \log P(\Delta t) = - \int_0^\infty P(\Delta t) \log P(\Delta t) d\Delta t \quad (19)$$

として評価できる。また, 入力データにおける変動間隔のフィルタによる「収縮」と「拡大」が各間隔で独立であるとする

$$P(\mathbf{Y}|\mathbf{X}) = P(\Delta t_1|\Delta l_1) \cdots P(\Delta t_N|\Delta l_N) = \{P(\Delta t|\Delta l)\}^N \quad (20)$$

であるから、単位データあたりの条件付きエントロピー $H(Y|X)$ は

$$\begin{aligned} \frac{H(Y|X)}{N} &= - \sum_{\Delta l} \sum_{\Delta t} P(\Delta t|\Delta l) \log P(\Delta t|\Delta l) \\ &= - \int_0^\infty d\Delta t \int_0^\infty d\Delta l P(\Delta t|\Delta l) \log P(\Delta t|\Delta l) \quad (21) \end{aligned}$$

で与えられる。従って、あとは実データから $P(\Delta t)$, $P(\Delta l)$, $P(\Delta t|\Delta l)$ を同定することで、所望の相互情報量を計算することができる。

- マイノリティゲームから導かれる資産分布の解析
 マイノリティゲームにおいて、自らの意思決定の結果、マイノリティグループに属したトレーダには +1 点、マジョリティグループに属したトレーダには 0 点を与える場合、十分なラウンドを終えた時点でトレーダ i の資産 c_i はトレーダごとに異なるはずであり、この分布 $P(c)$ がどのような性質を持つのかは興味深い。このような「富の分布」は昔から研究者の興味をかきたてているが、その分布には普遍的にパレート則: $P(c) \sim c^{-\gamma}$ が成り立つことが知られている。このマイノリティゲームに参加しているトレーダのコミュニティにおける資産分布がパレート則に従うか否かを調べることは「富がいかに配分されるか」という社会科学的観点からも興味深い。

今後は上記の課題を段階的に解決し、市場のマクロな性質をミクロの法則から導き出すための基礎理論を構築していきたい。

7 Appendix

7.1 標準正規分布の歪度と尖度の証明

確率密度 $f(x)$ をもつ連続分布 F に対して、 $t = 0$ の近傍で

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

が存在するとき、この関数を分布 F のモーメント母関数という。この関数は $t = 0$ の近傍で何回でも微分でき、 F の k 次の原点まわりのモーメントが

$$\mu'_k = M^{(k)}(0)$$

で与えられる。また、異なる分布が同じモーメント母関数をもつことはない。正規分布 (μ, σ^2) のモーメント母関数は $M(t) = \exp(\sigma^2 \frac{t^2}{2} + \mu t)$ であるから $M(t)$ の t

についての微分を $M'(t)$ と定義すると 4 次までのモーメントは

$$\begin{aligned} M'(t) &= (\sigma^2 t + \mu) \exp(\sigma^2 \frac{t^2}{2} + \mu t) \\ M''(t) &= (\sigma^2 t + \mu)^2 \exp(\sigma^2 \frac{t^2}{2} + \mu t) + \sigma^2 \exp(\sigma^2 \frac{t^2}{2} + \mu t) \\ M'''(t) &= (\sigma^2 t + \mu)^3 \exp(\sigma^2 \frac{t^2}{2} + \mu t) + 3\sigma^2 (\sigma^2 t + \mu) \exp(\sigma^2 \frac{t^2}{2} + \mu t) \\ M''''(t) &= (\sigma^2 t + \mu)^4 \exp(\sigma^2 \frac{t^2}{2} + \mu t) + 6\sigma^2 (\sigma^2 t + \mu)^2 \exp(\sigma^2 \frac{t^2}{2} + \mu t) + 3\sigma^4 \exp(\sigma^2 \frac{t^2}{2} + \mu t) \end{aligned}$$

となる。これから $t = 0, \mu = 0$ とおくと

$$\begin{aligned} s &= \frac{M'''(0)}{M''(0)^{\frac{3}{2}}} = 0 \\ k &= \frac{M''''(0)}{M''(0)^2} = \frac{3\sigma^4}{\sigma^4} = 3 \end{aligned}$$

よって標準正規分布では $s=0, k=3$ となることが分かる。

7.2 最小二乗法

n 個のデータ $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ が得られたとする。最もフィットする直線を $y = ax + b$ とおくと、各データとの距離は $d_i = y_i - (ax_i + b)$ となる。最小二乗法では $S = \sum_{i=1}^N d_i^2$ を最小とするような a, b を求める。S は a と b に関する二次関数と考えられるので

$$\begin{aligned} \frac{\partial S}{\partial a} &= 0 \\ \frac{\partial S}{\partial b} &= 0 \end{aligned}$$

として、 a, b について解くと

$$\begin{aligned} b &= \frac{\sum_{i=1}^N y_i - a \sum_{i=1}^N x_i}{N} \\ &= \frac{\sum_{i=1}^N x_i^2 \sum_{i=1}^N y_i - \sum_{i=1}^N x_i y_i \sum_{i=1}^N x_i}{n \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} \\ a &= \frac{n \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{n \sum_{i=1}^N x_i^2 - (\sum_{i=1}^N x_i)^2} \end{aligned}$$

7.3 平均待ち時間の導出

レートが変動してから次に変動するまでの時間、つまり第一通過時間を τ とする。これは実データ解析からワイブル分布であることがわかっている。次に、直前のレート変動時刻から顧客がレートをチェックするまでの時間を t 、最後にレートをチェックしてから実際にレートが変動するまでの待ち時間を s とする。まず、待ち時間 s は

$$s = \tau - t$$

と表される。この平均値が平均待ち時間になる

そこで、 t の分布を $P_O(t)$ とし、待ち時間 s の分布を $\Omega(s)$ をこれと $P_{m,a}(\tau)$ を用いて表すと

$$\Omega(s) \propto \int_0^\infty d\tau \int_0^\infty dt Q(s|\tau, t) P_O(t) P_{m,a}(\tau)$$

となる。 $Q(s|\tau, t)$ は第一通過時間が τ である間隔に直前のレート変動から t だけ経過した時点で顧客がログインし、レートをチェックした条件下でその顧客が次のレート変動まで s だけ待つ確率であり、ここでは

$$Q(s|\tau, t) = \delta(s - r + \tau)$$

で与えられる。従って、待ち時間 s の分布は規格化定数を含めて

$$\Omega(s) = \frac{\int_0^\infty d\tau P_{m,a}(\tau) \int_0^\tau dt \delta(s - r + t) P_O(t)}{\int_0^\infty ds \int_0^\infty d\tau P_{m,a}(\tau) \int_0^\tau dt \delta(s - r + t) P_O(t)}$$

と書くことが出来る。顧客のログイン時間は一様： $P_O(t) = 1$ と考えるのが自然であろうから、これを考慮して待ち時間 s の平均を計算すると

$$\begin{aligned} \langle s \rangle &= \int_0^\infty ds s \Omega(s) = \frac{\int_0^\infty ds s \int_s^\infty d\tau P_{m,a}(\tau)}{\int_0^\infty ds \int_s^\infty d\tau P_{m,a}(\tau)} \\ &= \frac{\int_0^\infty \frac{d}{ds} \frac{s^2}{2} ds \int_s^\infty d\tau P_{m,a}(\tau)}{\int_0^\infty \frac{d}{ds} s ds \int_s^\infty d\tau P_{m,a}(\tau)} = \frac{(\frac{1}{2}) \int_0^\infty s^2 P_{m,a}(s) ds}{\int_0^\infty s P_{m,a}(s) ds} = \frac{E(\tau^2)}{2E(\tau)} \end{aligned}$$

が得られる。ここで第一通過時間の分布でのモーメントを

$$E(r^n) = \int_0^\infty ds s^n P_{m,a}(s)$$

と定義している。

7.4 カルバック-ライブラ距離による第一通過過程の効果の測定: 実データ

第一通過過程の変動間隔への効果を調べるため、第一通過過程を介さないデータからの分布と、第一通過過程を介した後の分布とのカルバック-ライブラ距離を計算してみる。カルバック-ライブラは2つの確率関数の差の尺度である。第一通過過程を介さない時間間隔と介して提示されるレート変動の時間間隔をそれぞれ確率関数 $f(\Delta t), g(\Delta t)$ とすると式は以下ようになる。

$$K(f||g) = \sum_{\Delta t} f(\Delta t) \log \frac{f(\Delta t)}{g(\Delta t)} \Delta t$$

結果は以下の図 12 である

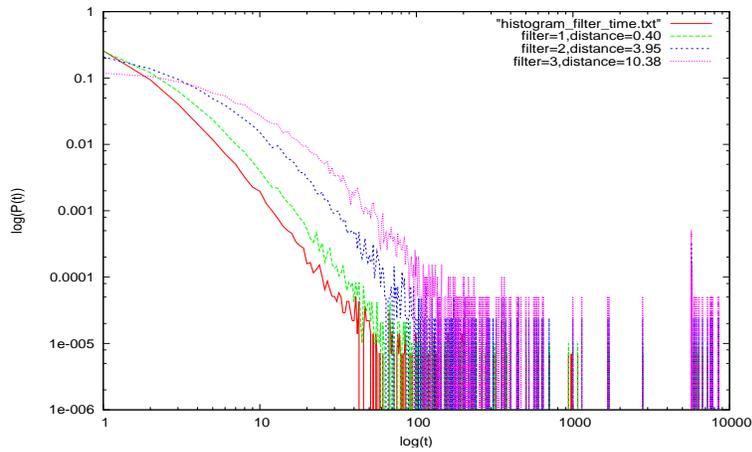


図 12: フィルターを介したときの時間の分布 ($\epsilon = 1, 2, 3$ でカルバックライブラ距離は元の分布と比べて 0.40, 3.95, 10.38 となった)

謝辞

指導教員である井上純一准教授に感謝します。また、本研究で用いた高頻度金融データを提供して下さったイタリア東ピアモンテ大学 Enrico Scalas 教授に感謝します。最後に、多くの励ましのお言葉を頂いた故和田充雄教授にこの場を借りて感謝します。

参考文献

- [1] Bouchaud, J.-P. and Potters, M. (2000): Theory of Financial Risk and Derivative Pricing, Cambridge University Press

- [2] Arthur, W.B. (1994): Inductive Reasoning and Bounded Rationality (The El Farol Problem), *Am. Econ. Rev.*, Vol.84, pp. 488-500
- [3] Challet, D. and Zhang, Y.-C.(1997): *Physica A*, Vol. 246, pp.407-418
- [4] Challet, D., Marsili, M. and Zhang, Yi.-C. (2005): *Minority Games*, Oxford University Press
- [5] Coolen, A.C.C.(2005): *The Mathematical Theory Of Minority Games: Statistical Mechanics Of Interacting Agents*, Oxford University Press
- [6] von Neumann, J. and Morgenstern, O. (1947): *Theory of Games and Economic Behavior*, Princeton University Press
- [7] 船木由喜彦 (2001): *エコノミックゲームセオリー — 協力ゲームの応用 —*, 臨時特別冊・数理科学, SGC ライブラリ 11, サイエンス社
- [8] 梶井厚志, 松井彰彦 (2000): *ミクロ経済学 — 戦略的アプローチ*, 日本評論社
- [9] Weibull, J.W. (1995): *Evolutionary Game Theory*, The MIT Press
- [10] Sazuka, N. (2007): On the gap between an empirical distribution and an exponential distribution of waiting times for price changes in a financial market, *Physica A*, Vol. 376, pp. 500-506
- [11] Sazuka, N. and Inoue, J. (2007): Fluctuations in time intervals of financial data from the view point of the Gini index *Physica A*, Vol. 383, pp. 49-53
- [12] Inoue, J. and Sazuka, N. (2007): Crossover between Levy and Gaussian regimes in first-passage processes, *Physical Review E* Vol. 76, 021111 (9 pages)
- [13] Inoue, J. and Sazuka, N. (2008): Queueing theoretical analysis of foreign currency exchange rates *Quantitative Finance*, 印刷中
- [14] Sazuka, N., Inoue, J. and Scalas, E. (2008): The distribution of first-passage times and durations in FOREX and future markets, Working paper, arXiv:0808.0372
- [15] Cavagna, A. (1998): Irrelevance of memory in the minority game, *Physica Review E*, Vol. 59, pp. R3783-R3786
- [16] Coolen, A.C.C. and Shayeghi, N. (2008): Generating functional Analysis of minority games with inner product strategy definitions, *Journal of Physics A: Mathematical and Theoretical*, Vol. 41, No.32, pp. 324005 (30pages)

- [17] 今野浩 (2000): 金融工学の挑戦, 中公新書
- [18] Ohira, T., Sazuka N., Maruo, K., Shimizu, T., Takayasu M. and Takayasu, H. (2002): Predicability of currency market exchange, Physica A, Vol. 308, pp. 368-374