アローヘッド市場における価格変動：Lévy 安定分布の可能性

田中 美栄子†

†明治大学総合数理学部 〒164-8525 東京都中野区中野 4-21-1 明治大学中野キャンパス 1011 号室

E-mail: †mieko@meiji.ac.jp

あらまし

東証において，アローヘッドシステムという 1 ミリ秒間隔で株式売買を行える仕組みが稼働を始めたのが 2010 年 1 月であり， 2015 年 9 月にはその第 2 ステップとして 0.5 ミリ秒間隔にまで短縮された．この高速取引によって価格変動の性質がどのように変わったのかを知るため，5 秒足にサンプリングされた 2013 年のデータから分布関数を求めたところ，1990 年頃の米国株価変動の解析結果により求められた α=1.4 の Lévy 分布と同じであるという結果を得た．更に 1 分足にサンプリングされた 2015 年のデータからも同様の結果を得た．

キーワード　アローヘッド市場，価格変動，Lévy の安定分布，統計分布関数

Price Fluctuation in the Arrowhead Market : Possibility of Lévy Stable Distribution

Mieko Tanaka-Yamawaki†

† Department of Mathematical Sciences Based on Modeling and Analysis, School of Interdisciplinary Mathematical Sciences, Meiji University, 1-21-4, Nakano-ku, Tokyo, 164-8525 Japan

E-mail: †mieko@meiji.ac.jp

Abstract

The arrowhead system is a trading system within a millisecond, which was launched in January 2010 at Tokyo Stock Exchange, Inc.(TSE) and was upgraded in September, 2015 to the level of half a millisecond. In order to investigate the nature of price fluctuation in the arrowhead market, we have analyzed price time series taken from TSE. The result is the statistical distribution can be identified as the Lévy distribution of index α =1.4, the same as the previous result reported in 1995 for the data of S&P500 index (1983-1989), and also for the Nikkei Index price in mid 1990's.

Keywords　Arrowhead Market, Price Fluctuation, Lévy's Stable Distribution, Statistical Distribution Function

1. はじめに

価格変動の統計性について 1900 年の Bachelier 論文以降，株価変動の統計分布は基本的に正規分布とされてきたが，1994 年の Nature 論文[1]では 1 分毎の S&P500 指数の実データの解析により，指数 1.4 の Lévy 分布に近いことが示された．その後，諸説あるものの fat-tail で narrow-neck な分布則に従うことには異論がない．

一方，2010 年 1 月 4 日以降、東京証券取引所は arrowhead と呼ぶ千分の 1 秒単位での取引を可能するシステムに移行した。更に 2015 年 9 月 24 日以降，2 千分の 1 秒単位での取引を可能にする arrowhead 2 となって、超高速化が進行し，海外から多くのファンドが参加してシステムトレードを行うようになった結果，従来の取引環境とは異なる投資状況が生じている．このようなシステム下で行われている市場の価格変動の性質もおそらく以前とは大きく異なるものと考えられる．

2. Arrowhead 市場の価格変動:データ(A)

そこで，我々は東証 Market Impact View で限定公開されていた時期（2013 年 4 月～12 月）の 100 銘柄に対する 5 秒足株価を入手し，Lévy 分布と一致する場合の指数とその安定性を調べた．使用するデータ は 1 銘柄につき 5 秒毎の全 640800 点のデータであり，今後これをデータ(A)呼ぶ．この各点が表す運動を酔歩とみなし，価格変動を確率変数として扱うことによりその分布関数の形を推定する．変数は価格ではなくその差分とし，更に銘柄毎の価格に依存しない変数として，対数収益（log-return）

$$Z_{\Delta t}(t) = log X(t+\Delta t) - log\, X(t) \qquad (1)$$

ⓒ 2016 Information Processing Society of Japan

を確率変数と見て，この確率分布を調べるのが普通である．

3. Lévy 安定分布のスケール不変性

Lévy 分布は$1 \le \alpha \le 2$に対し次式で定義される安定分布である．

$$P_{\alpha,\ \gamma\Delta t}(Z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{ikZ-\gamma\Delta t|k|^{\alpha}}dk \tag{2}$$

ここで Δt は時間間隔であり，最小単位 5 秒のとき $\Delta t=1$ と定義しておく.指数 α=1 に対してはローレンツ分布

$$P_{\alpha,\beta}(Z) = \frac{\beta}{\pi}\frac{1}{\beta^2+Z^2} \tag{3}$$

となり，指数 α=2 に対しては正規分布

$$P_{\alpha,\beta}(Z) = \frac{1}{2\sqrt{\pi\beta}}\exp(-\frac{Z^2}{4\beta}) \tag{4}$$

となるが,一般の α に対しては式(2)から数値積分により求めることができる．

安定分布に従う確率変数は重ね合わせた結果も同じ分布になるという性質を持つ．初期値 0 から 5 秒後に最初の価格変動 x が起き，更に 5 秒後にはもう一度同じ確率分布 f に従う価格変動 y が起きて $z=x+y$ を得る時，z の確率分布は x と y の確率分布の積となるが，同一の z の値を取る多くの x と y の重ね合わせとして

$$f_{\Delta t=2}(z) = \int_0^z f_{\Delta t=1}(x) f_{\Delta t=1}(z-x)dx \tag{5}$$

の畳み込みで表される．$F(Z)$のフーリエ変換で表すと

$$F_{\Delta t=2}(k) = (F_{\Delta t=1}(k))^2 \tag{6}$$

となり，Δt というパラメータは時間が経つにつれて同一の統計分布に従う確率変数が畳み込まれて行く度に増加するが，指数 α は変化しないので同じ指数の分布が保たれる．異なる時間間隔の分布同士はスケール変換で繋がり，

$$P(Z(t)) = c^{-\frac{1}{\alpha}}\ P_{\alpha,\ \Delta t}(c^{\frac{1}{\alpha}}Z(t)) \tag{7}$$

によって,単一の分布に重なる.この性質を利用して，データ解析により指数 α の値を推定することができる．

4. アローヘッドデータ(A)と Lévy 分布

データ(A)の最小単位である 5 秒を $\Delta t=1$ として，$\Delta t=3, 12, 60, 120$（それぞれ,15 秒,60 秒,300 秒,600 秒に対応する）とした疎視化を行い,100 銘柄の平均を取った値の出現頻度分布を式(7)に従ってスケール変換し，描画したものを Fig.1 に示す．指数 α を 1<α<2 の範囲で 0.1 刻みに比較した結果,最も良く一致したのは α=1.4 であった．単一株価分布は Z=0 の付近で α=1 のローレンツ分布に近い株価も存在し,株価毎に異なるようにも見える．

また,疎視化の程度を上げてゆくにつれデータ数が減少することにも注意が必要である.表 1 に示すように,$\Delta t=3$ で元のデータの 3 分の 1 の 21,360 点にまで減少し,1 分の時間間隔に対応する $\Delta t=12$ で 53,400 点にまで減少してしまい，Δt が 1 分を超えると，イベント数の少ない両端では分布の精度が悪化する[2]．

表 1　時間スケールの変化によるデータ点の減少

Δt	時間間隔(s)	データ点
1	5	640800
3	15	213600
6	30	106800
12	60	53400

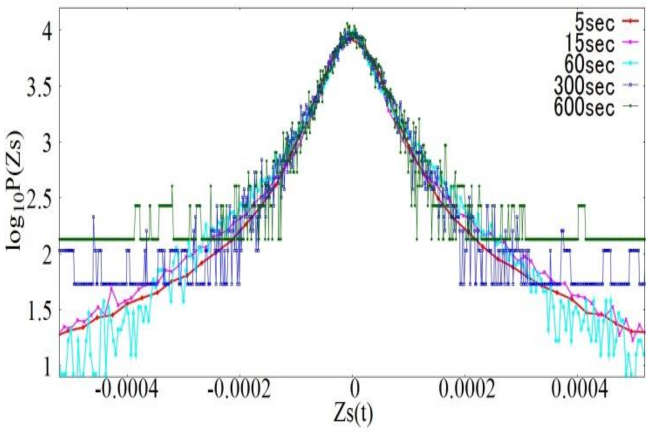

Fig. 1 スケール変換後の確率分布と$\alpha = 1.4$の Lévy 分布

Lévy 分布を特徴づける指数 α を更に正確に求めるには式(7)から得られる P(0)の Δt 依存性

$$\log\left(P_{\gamma\Delta t}(0)\right) = -\frac{1}{\alpha}\log(\Delta t) + \log C \tag{8}$$

を利用することができる．P(0)を縦軸に，Δt を横軸にした両対数グラフを Fig.2 に示す.直線部分($\Delta t = 1$~12)の領域の傾きから α=1.41（$\Delta t=24$ まで含めると 1.38)となり Fig.1 の示す α=1.4 に近い値となっている．

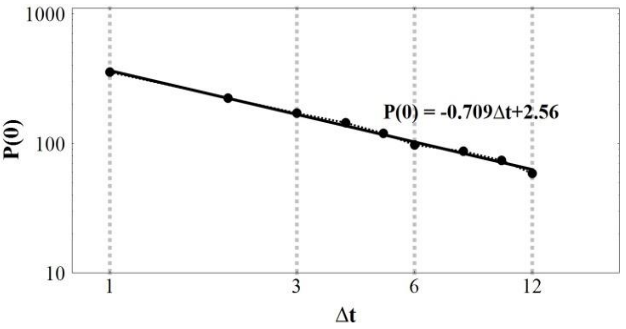

Fig.2　原点回帰率 P(0)の Δt 依存性による指数 α

arrowhead 移行後の株価はそれ以前のものと大きな違いがあるかも知れないとの予想と異なり，2013 年の 5 秒足価格は $\alpha=1.4$ の Lévy 分布に非常によく重なるという結果が得られた．文献[1]にある，1983～1989 の S&P500(ニューヨーク市場)の $\Delta t=1$ 分~1000 分の従う分布と同じであるという結論が得られた．つまり **arrowhead** システムに移行する前と後で，本質的に変わらない価格変動の性質を持つことがわかった．但しここで用いたのは東証 Market Impact View で限定公開されていた時期(2013 年 4 月～12 月)の 100 銘柄という特殊なデータの平均値に対する結果であり,結果をどれほど一般化できるかは疑問である.データ長が短いため，分布の両端については信頼度が薄いと考えられる．

また,文献[1]の結果は著者等により後に否定されている事情もあり,更なる検証が必要である．

加えて，そもそも理論的な問題として，分布全体が Lévy 分布であれば分散が無限大となり，現実的ではないため，分布の両端では Lévy 分布から外れていると考えるのが妥当である．また，時間スケール Δt の大きく異なる価格変動が全て単一の統計分布でフィットできるほどに単純な構造をしているとも考えにくい．

そこで別のデータを用いて独立の解析を行う．google finance の頁からダウンロードした 2015 年 6 月 16 日から 11 月 4 日までの東証株価 440 銘柄の 1 分足価格変動データ(時系列長 29,386)である．今後これをデータ(B)と呼ぶ．

5. アローヘッドデータ(B)と Lévy 分布

データ(B)の最小スケールである 1 分を $\Delta t=1$ として価格変動の確率分布 p(x)を求め，様々の α 値を持つ Lévy 分布 q(x)と比較して KL 距離

$$\mathbf{D(p||q)} = \sum_{\mathbf{x}} \mathbf{p(x)} \log(\mathbf{p(x)}/\mathbf{q(x)}) \qquad (9)$$

の最小となる α を選ぶ．価格として

(1)440 銘柄の平均価格
(2)個別株価 3 銘柄（関西電力,日産自動車,東芝）

を選び，分布のヒストグラムを求める．この際,横軸である対数収益 Z(t)の範囲[-0.01:0.01]の分割数を，中心部分に出現頻度 0 の枠が出ない範囲で最大となるよう選択する．440 銘柄の平均価格に対しては 2000 分割,個別銘柄に対しては 100 分割が選ばれる．このような条件下で，Lévy 分布の二つの指数 α と β を，$1<\alpha<2$ の範囲内で 0.05 刻み，$10^{-7}<\beta<10^{-4}$ の範囲内で 10^{-7} 刻みの中から最良値を選択した．結果は表 2 に示すように東証 440 銘柄の平均価格に対しては α=1.4 が最も良い近似であり，3 銘柄の個別株価に対しては α=1.55～1.65 という値となった．個別株の場合に大きな α 値を取るが，同時に KL 距離が大きいので，平均株価に比較して信用度は落ちる．

表 2　東証(2015)1 分足価格変動と Lévy 分布の一致度

株式銘柄	α	β	KL 距離
440 銘柄平均	1.40	5.4×10^{-6}	0.039
関西電力	1.55	10.0×10^{-6}	0.286
日産自動車	1.65	3.9×10^{-6}	0.423
東芝	1.55	8.8×10^{-6}	0.156

表 2 の 1 行目にある 440 銘柄平均値の場合を例として，分割数の違いによる指数の最適値と KL 距離との関連を調べる．個別株価に対するのと同じ分割数 100 を選ぶと KL 距離は 0.095 となるが,分割数を 500, 1000, と増加させるに従い，KL 距離が減少して行くが，分割数を 2000 に取ると逆に増加する.これ以上分割を細かくしても近似は悪化する.最適 α 値は 1.40 で安定するため，この値を最良値と考えることができる．この様子を表 4 にまとめた．

2015 年のアローヘッド 1 分足価格の実データから求めた価格の対数収益(z)に指数 $\alpha=1.4$ の Lévy 分布が最もよくフィットすると結論できる．この様子を Fig.3 に図示する．特に $|z|<0.01$ となる領域での誤差が非常に小さい様子を Fig.4 に示す．

表 4　平均株価(440 銘柄)に対する最適分割数

分割数	α	KL 距離
100	1.75	0.095
500	1.55	0.072
1000	1.40	0.027
2000	1.40	0.039

ⓒ 2016 Information Processing Society of Japan

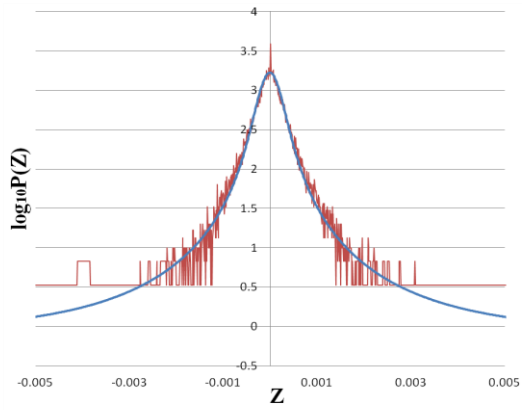

Fig.3 東証 440 銘柄の平均株価と Lévy 分布 α =1.4

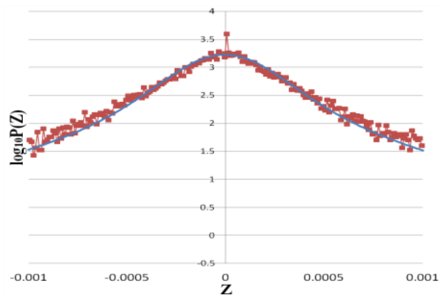

Fig.4 平均株価(Fig.3)の|z|<0.01 領域は高精度で一致

6. まとめと今後の課題

本研究は，2010 年 1 月より始まった東証アローヘッドシステムの開始後に株価変動の性質が以前と比べて如何に変化したのかを調べるため，先行研究[1]と同様の手法を用いて解析し比較した結果について報告するものである．

文献[1]の解析対象は 1984-1989 年のアメリカ市場の平均株価の一つである S&P500 であったが，そこで示されたのは価格変動の分布が正規分布ではなく，横軸の両端に大きな確率を持つ裾野の広い Fat-Tail 分布で，横軸が 0 の付近では α=1.4 のレヴィ分布に従うという事実であった．

この指数 α 決定には，式(7)に示す Lévy 分布の自己相似性を用いて，異なる Δt の分布関数を単一の分布に重ねることや，その対数を取った式(9)を用いて，分布の中心値 P(0)の時間スケール Δt 依存を利用して最小二乗法により直線の傾きの逆数として α を求める手法が使われた．

同様の分析を日本の株価指数 TOPIX(1995 年 1 月 4 日-12 月 28 日 1 分毎)約 6 万点に対して行った研究[4]においても α=1.4 の Lévy 分布が確認されていた．

本研究では文献[1]と同じ手法によりアローヘッドシステム移行後の価格変動を解析したが，結果は以前と同じ指数 α=1.4 のレヴィ分布が最適分布であるという結論を得た．但しこれは分布の中心部の narrow neck の部分についてであり，裾野部分の fat tail の挙動についてはデータ量が及ばず何とも言えない．

ビッグデータブームの振興により，豊富なデータが利用できるようになれば，この方面の研究に画期的な進歩が見込めると期待される．

謝辞

本稿は鳥取大学工学部知能情報工学科知識 A 研究室における研究内容に基づいたものです．使用したデータのうち 5 秒足のものは 2013 年度に卒業研究を行った山本貴範を中心とする計 3 名の協力により収集し，2015 年度に卒業研究を行った藤井猛の協力により解析を行った結果を 4 章に使用しました．1 分足のものは 2015 年に卒業研究を行った向谷和希の協力により収集でき,2015 年度修士論文を作成した吉井勝俊の計算協力を得たもので，この機会に謝意を表します．また，文献[4]を頂きました島田一平先生に深謝します。

文 献

[1] R.N.Mantegna and H.E. Stanley, Nature Vol. 376, pp46-49, 1995
[2] 田中美栄子,藤井猛,「アローヘッド市場における株価変動の統計性」統計数理研究所共同研究リポート 360 巻「経済物理学とその周辺(12)」,pp.64-65, 2016 年 3 月（統計数理研究所）
[3] 田中美栄子,藤井猛,「アローヘッド市場における株価変動の統計性」,日本物理学会第 71 回年会（東北学院大学,2016 年 3 月 19-22 日)講演予稿集 71, 20BU-6.
[4] 柳川一貴,「短い時間での価格変動を起こす要因」日本大学大学院理工学研究科平成 9 年度修士論文；日本物理学会秋の分科会（神戸大学,1997 年 9 月 16 日）講演予稿集 52(2-3),8a-YD-9;総合研究大学院大学共同研究"新分野の開拓"1997-2002,小グループ「経済学」第 1 回研究会(1998 年 7 月 14 日東京駅ホテル)における講演．
[5] 田中美栄子「アローヘッド市場における価格変動の統計分布－Le´vy 分布再考－」電子情報通信学会非線形問題/複雑コミュニケーションサイエンス研究会,機械振興会館,東京,2016/6/13 (IEICE technical report: 信学技報 116(86),27-30,2016-06-13) http://ci.nii.ac.jp/ naid/ 40020881215

ⓒ 2016 Information Processing Society of Japan