



Hitotsubashi ICS-FS Working Paper Series

FS-2016-J-001

外国為替取引におけるクラスタ現象のモデル化

佐久間 吉行

一橋大学大学院国際企業戦略研究科

横内 大介

一橋大学大学院国際企業戦略研究科

First version: April 23, 2016

All the papers in this Discussion Paper Series are presented in the draft form. The papers are not intended to circulate to many and unspecified persons. For that reason any paper can not be reproduced or redistributed without the authors' written consent.

外国為替取引におけるクラスタ現象のモデル化

佐久間 吉行 *

横内大介 †

平成 28 年 4 月 23 日

本稿は為替 Tick データに現れる為替取引のクラスタ（塊）現象を説明する統計モデルを提案する論文である。先行研究である Shibata[2006] は、各クラスタ内の為替取引発生に対して定常ポアソン過程を仮定することで、クラスタの検出法の確立および為替の対数価格差がラプラス分布することを示した。そして、平常時のデータを用いて提案した統計モデルの当てはまり具合を確認した。本稿はその Shibata[2006] の仮定を一部変更し、取引発生に対してはより柔軟な複合ポアソン過程を、対数価格差に対しては事前に t 分布を仮定することで、新たなクラスタの検出方法を確立した。そして、先行研究では用いられなかった異常時 Tick データ、具体的には 2010 年のフラッシュクラッシュ時のデータを用いて、2 つのクラスタ検出法と各分布の適合度の比較を行った。その結果、取引が活発な時間帯では本稿提案のモデルのほうが近似精度が高いことを確認した。

1 はじめに

外国為替取引のマーケット¹において銀行などのレート配信業者は、自らが過度に為替のポジションリスクを負うことなく、安定したスプレッド収益を獲得することを目指している。しかしながら、外国為替取引にはさまざまな要因が複雑に絡み合っているため、単純な数理モデルや統計モデルだけでは為替取引のメカニズムを説明することは困難であり、実務の助けになるようなフレームワークはまだ十分に確立されていない。

この現象の複雑さを引き起こす要因の 1 つとして、為替取引のクラスタ現象が上げられる。この為替取引のクラスタ現象とは、取引の活況具合に応じて、約定発生の間隔 Δt が短くなる時間帯と長くなる時間帯に分かれる現象のことである。このクラスタ現象を捉えるには、約定発生の間隔 Δt を説明するモデルが必要になる。しかしながら、通常ファイナンスの研究の設定と異なり、クラスタ現象を捉えるモデルでは発生間隔 Δt が等間隔であるという仮定をおけない。そこで本稿では

*一橋大学大学院国際企業戦略研究科後期博士課程

†一橋大学大学院国際企業戦略研究科

¹外国為替取引の参加者は、為替の流動性を需要する輸出入などの実需に基づく参加者、為替に流動性を供給する銀行などのレート配信業者、投機的な参加者、中央銀行、投資ファンドなど多岐にわたる。取引の形態は、互いのクレジットを事前に承認したうえでの相対取引が原則となる。クレジットを承認し合った銀行間では、この相対取引を EBS などの電子プラットフォームを通じて行っている。EBS はこういった仕組みを束ねるところであり質が高いデータを提供している。また、実需に基づく大口参加者と銀行などのレート配信業者の間の取引も、レート配信業者が開設する電子ポータルサイトを通じて行うのが一般的となっている。こうした電子取引のスピードはより高速化の傾向にあり、1/1000 秒単位の為替レートの配信も行われるようになっている。

Δt に対して約定発生の順序に対応する添え字 i をつけて、不均一な約定発生間隔を Δt_i で表現し、モデルとして表現することを目指した。また本稿のモデルには価格変動のしやすさを表す仕掛けとして、モデルに対して Δt_i で基準化した単位時間あたりのボラティリティ σ^2 をとりこみ、各クラスタにおけるボラティリティ σ^2 を一定とする平易な仮定を与えた。そして、単位時間あたりのボラティリティ σ^2 と約定発生間隔 Δt_i の積をトータルボラティリティ Σ と定義した。さらに、為替レートを急激に変化させる何らかの情報（雇用統計の発表、中央銀行発表、テロ事件の勃発など）が発生した状態を異常時、それ以外を平常時と定義した。この仮定や定義をおくことで Δt_i の間隔の長さやボラティリティ σ^2 の大きさが為替レートに与える影響についての仮説を組み立てることが可能になり、ポジションリスクの計量化の助けとなる。たとえば、平常時かつ約定の取引回数が少ない時間帯は、ボラティリティ σ^2 一定の仮定の下では取引間隔が長くトータルボラティリティ Σ が増加する時間帯となる。レート配信業者がこの時間帯にポジションを取っていた場合、反対売買を行ってポジションを解消（カバー取引と呼ぶ）するには時間がかかる。これは、ポジションリスクの高い時間帯といえるので、スプレッド幅を広げ、顧客に消極的なレート提示を行うことでリスクを低減する戦略が立案できる。一方で、取引回数が多い時間帯は、ポジションを取っても直ぐにカバーに入ることができるので、ポジションリスクが低い時間帯となり、積極的なレート提示を行う戦略が立案できる。一方、異常時のレート配信の現場は、平常時に積み上げた収益を失ってしまうリスクを負う。これを避けるため、実務で使えるモデルは、異常時を平常時から区別することも必要である。今回のわれわれのモデルの設定では、異常時にはボラティリティ σ^2 の増加や Δt_i の分布の裾の拡大などを通じて、トータルボラティリティ Σ が大きくなることが予想できる。このような為替取引の変動特性を現すモデルが構築できれば、異常発生のシグナルを得ることができ、これを用いた戦略が立案できる。

このような本稿のテーマである為替取引のクラスタ現象のモデル化にあたり、我々はまず先行研究で使われているモデルをサーベイした。為替や株の取引発生メカニズムを説明する手法の違いに焦点をおいて既存の研究を区分すると、既存研究は、「計量経済・ファイナンスの理論を用いて為替や株の取引を説明する研究」、「統計力学の理論を用いて為替や株の取引を説明する研究」、「統計モデルを用いて為替や株の取引を説明する研究」の3つに区分できる²。

「計量経済・ファイナンスの理論を用いて為替や株の取引を説明する研究」は、マーケットマイクロストラクチャーとして知られている。高頻度データに、ゲームの理論や情報の非対称性などの理論を当てはめることで、為替や株の取引を説明する研究となる。Evance[2002] は、NCK（私的ニュース）の代理変数としてディーラーの買い注文と売り注文の差を用いることで、状態空間モデルを構築し、CN（公的ニュース）と NCK（私的ニュース）が為替レートの挙動に与える影響を調査した。Easley.et.al[1996] は、混合ポアソン分布の尤度関数を用いて、マーケット参加者のうち一部の人が知りえない私的情報の増加を市場データから検出する方法である PIN を開発した。その他、マーケットマイクロストラクチャー分野を中心とした研究の紹介は、杉原 [2011] などに詳しい。

「統計力学の理論を用いて為替や株の取引を説明する研究」は、Tick データや高頻度データに、カオス・フラクタルなど経済物理の理論を当てはめることで、為替や株の取引を説明する研究となる。高安 [2007] は、統計力学の手法を基に、PUCK モデルを示し、為替レートの予測の可能性について言及した。Sazuka[2008] は、ソニー銀行の USD/JPY レートの取引発生間隔の分布が指数分布

²本稿では、データの発生時間とその時のレートを共に記録したデータを Tick データとよび、データの発生間隔を 1 秒、1 分などの短い時間の単位で等間隔に加工したデータを高頻度データとよぶ

ではなく、ワイブル分布に従うことを示した後、統計力学の手法を用いて、長時間漸近領域での中でワイブル法則とべき法則の相転移現象を確認した。

「統計モデルを用いて為替や株の取引を説明する研究」は、Tick データへの点過程モデルの当てはめや高頻度データへの時系列モデルの当てはめを通じて為替や株の取引を説明する研究となる。Hawkes[1971], [1974] は、点過程モデルの強度に周期性やトレンドを組み込んだ自己励起モデルを提案した。Ogata[1999] は、地震の発生時刻の点過程によるモデル化を提案し、条件付き強度を用いて点過程のモデルが、地震活動の様々な統計的調査に役立つことを示した。加藤・丸茂 [1999] は、為替レートなどの取引発生回数が、定常ポアソンに従うかを検証した。そして、観測した市場で定常ポアソン過程からの乖離を検出した結果を受け、原因として取引間の相互作用の存在を予想した。Engel[2000] は、取引発生間隔の系列に ACD(自己回帰条件付デュレーション) と呼ばれる時系列構造を仮定し、取引タイミングのモデル化を行うとともに価格のボラティリティへの影響を計測した。Shibata[2006] は、為替 Tick データの取引発生頻度に対し、強度一定の定常ポアソン過程を区分的に当てはめるモデル化を提案した。

本稿では、為替取引のクラスタ現象を捉えるモデル化を行うことがテーマなので、一般に単位時間間隔に補正した高頻度データを使用する「計量経済・ファイナンスの理論を用いて為替や株の取引を説明する研究」の方法は利用できない。為替 Tick データの取引発生間隔をそのまま眺める研究には、「統計力学の理論を用いて為替や株の取引を説明する研究」と「統計モデルを用いて為替や株の取引を説明する研究」があるが、物理学の知識を利用する研究手法は、現場で実務家が応用するにはハードルが高いと考え、本稿では「統計モデルを用いて現象を説明する研究」に着目した。

「統計モデルを用いて現象を説明する研究」の中で挙げた Shibata[2006] は、為替取引のクラスタ現象をモデル化した論文となる。為替取引の発生頻度に対し、強度一定の定常ポアソン過程を区分的に当てはめるモデル化を提案し、二分木アルゴリズムを用いて平常時のデータを当てはめ、そのあてはまりの良さを確認している。この論文の中で、為替レートの対数価格差は正規分布に、取引発生間隔は指数分布に従うものとしてモデル構築を行っている。Shibata[2006] の発想は我々のねらいと良く一致しており、本研究ではこの論文を先行研究とした。ただし、平常時の Tick データで当てはまりの良かった Shibata[2006] に対し、本稿では異常時における説明力についても焦点をあてる。そのため、分析に用いた Tick データは、EBS 社の 2010/5/4~2010/5/7 の USD/JPY の為替約定取引データ³であり、米国の S&P500 先物の流動性供給が瞬間的に壊れて大暴落したフラッシュクラッシュの時間帯を含んでいる。

本論文の 2 節では、まず先行研究である Shibata[2006] のモデルの詳細を紹介する。3 節では、Shibata[2006] のモデルの発生強度に対してガンマ分布を事前分布として仮定し、再構築した新たなモデルを提案する。4 節ではデータを双方のモデルに当てはめた結果を比較する。5 節では、Shibata[2006] のモデルと我々のモデルの 2 つを組み合わせたモデルの提案とあてはめ結果を紹介する。6 節では、本論文のまとめを行う。

³最小取引発生間隔 1 秒のタイムスタンプと約定レートからなる USD/JPY の約定取引 Tick データ

2 先行研究の概要

本稿の先行研究である Shibata[2006] は、各クラスタでの取引発生件数が強度一定の定常ポアソン過程に従うという仮定の下で、Tick データを意味のあるクラスタに区分する方法を提案している。そして、対数価格差を取引発生間隔の平方根で除して基準化した系列が正規分布に従うという仮定の下で、対数価格差がラプラス分布に従うこともあわせて示している。本節ではこの Shibata [2006] における為替 Tick データの統計モデリングの方法について概説する。

2.1 約定取引発生が定常ポアソン過程に従う場合のクラスタ区分法

あるクラスタにおける総取引数を n とし、為替取引の取引発生時刻 $t_1 \leq t_2 \leq \dots \leq t_n$ における取引発生間隔を $\Delta t_i = t_i - t_{i-1}$ とおく。取引発生が強度 λ の定常ポアソン過程に従うと仮定すると Δt_i は指数分布に従う。このときの約定取引の間隔 $\Delta t_2, \dots, \Delta t_n$ の対数尤度関数は

$$L(\Delta t_2, \dots, \Delta t_n; \lambda) = \log \left(\prod_{i=2}^n \lambda \exp(-\lambda \Delta t_i) \right) = (n-1) \log \lambda - \lambda(t_n - t_1)$$

となり、 L を最大にするパラメータ λ の最尤推定量 $\hat{\lambda}$ は、

$$\hat{\lambda} = \frac{n-1}{t_n - t_1}$$

となる。そして、このときの最大対数尤度 L_0 は

$$L_0 = L(\Delta t_2, \dots, \Delta t_n; \hat{\lambda}) = (n-1) \log \frac{n-1}{t_n - t_1} - (n-1)$$

となる。

一方、時刻 t_1, \dots, t_k に発生する約定取引が強度 λ_1 の定常ポアソン過程に従い、 t_k, \dots, t_n に発生する約定取引が強度 λ_2 の定常ポアソン過程に従うと仮定する。この仮定は言い換えれば、 t_1, \dots, t_n は 1 つのクラスタではなく、2 つのクラスタから構成されているということを意味する。このとき、前者の約定取引の間隔 $\Delta t_2, \dots, \Delta t_k$ の対数尤度関数は

$$L_{1,k}(\Delta t_2, \dots, \Delta t_k; \lambda_1)$$

となり、後者の対数尤度関数は

$$L_{2,k}(\Delta t_{k+1}, \dots, \Delta t_n; \lambda_2)$$

となる。そして、それぞれの最尤推定量は $\hat{\lambda}_1 = \frac{k-1}{t_k - t_1}$ と $\hat{\lambda}_2 = \frac{n-k}{t_n - t_k}$ となるので、このときの最大対数尤度は

$$\begin{aligned} L_{1,2,k} &= L_1(\Delta t_2, \dots, \Delta t_k; \hat{\lambda}_1) + L_2(\Delta t_{k+1}, \dots, \Delta t_n; \hat{\lambda}_2) \\ &= (k-1) \log \frac{k-1}{t_k - t_1} + (n-k) \log \frac{n-k}{t_n - t_k} - (n-1) \end{aligned}$$

とあらわせる。この k を動かしたとき、最も大きな $L_{1,2,k}$ を達成する k' が、尤度原理の観点から分割点の最有力候補になる。

もし尤度比 $L_{1,2,k'}/L_0$ の理論分布が求められるのであれば、尤度比検定を通じて分割を行うか否かの統計的な判定が実施できる。しかしながら、この検定統計量の理論分布を解析的には求めることは容易ではないことから、Shibata[2006] ではモンテカルロシミュレーションを用いて検定統計量の分布を取得している。

実際の分析では全取引データを 1 つのクラスタとみなして分割を開始する。そして尤度比検定の帰無仮説（具体的には「分割前と後では最大対数尤度に差はない」）が棄却できなくなるまで各クラスタに対して再帰的に尤度比検定を実施することで統計的に意味のある分割点を求めている。

2.2 対数価格差が従う分布

前節のアルゴリズムで得られた任意の 1 つのクラスタの要素数を n とする。そして取引発生時刻 t_i の為替レート $R_{t_i}, i = 1, 2, \dots, n$ に対する対数価格差 x_i を

$$x_i = \log \left(\frac{R_{t_i}}{R_{t_{i-1}}} \right), i = 2, 3, \dots, n$$

と表すこととする。そのとき、標準ブラウン運動 B_t と定数ボラティリティ σ を用いて対数価格差が $\log R_t = \sigma B_t$ と記述できるならば、その対数価格差 x_i は、

$$x_i = \log \left(\frac{R_{t_i}}{R_{t_{i-1}}} \right) = \sigma (B_{t_i} - B_{t_{i-1}}) = \sigma \sqrt{\Delta t_i} z, \quad z \sim N(0, 1)$$

となる。前節の仮定より Δt_i は指数分布に従うので、対数価格差 x_i は $N(0, \sigma^2 \Delta t_i)$ という複合分布に従う。

この指数分布と正規分布の複合分布は、 $\alpha = \frac{\sqrt{2\lambda}}{\sigma}$ となるパラメータをもつラプラス分布に従う。以下はその密度関数である。

$$f(x; \alpha) = \frac{\alpha}{2} \exp(-\alpha|x|) \quad i = 2, \dots, n$$

この結果については Andrews[1974] の中で Lemma として与えられている。詳しい説明は本稿の付録を参照されたい。

各クラスタにおける対数差収益率 x_i はラプラス分布に従うので、そのパラメータ α の最尤推定量は

$$\hat{\alpha} = \frac{n}{\sum \left| \log \frac{R_{t_i}}{R_{t_{i-1}}} \right|}$$

で与えられる。そして、 λ の最尤推定量は

$$\hat{\lambda} = \frac{n-1}{t_n - t_1}$$

であるから、各クラスタでのボラティリティ $\hat{\sigma}$ は

$$\hat{\sigma} = \frac{\sqrt{2\hat{\lambda}}}{\hat{\alpha}}$$

と推定することができる。

3 複合ポアソン過程を用いた為替レートのモデル化

Shibata[2006] は、平常時のデータを用いてモデルの当てはまり具合を検証していたが、後に説明するように価格が急激に変化する局面の当てはまりはあまりよくない。本稿ではこの問題点を解決すべく、Shibata[2006] における定常ポアソン過程の仮定を変更し、強度パラメータにリーズナブルな確率分布を仮定した複合ポアソン過程を利用したモデリングを提案する。

3.1 本稿モデルの仮定と取引発生間隔の分布

Shibata[2006] では、為替取引の発生間隔 Δt_i に対して指数分布を仮定し、さらに対数取引価格に対してブラウン運動を仮定することで、対数取引価格差 x_i がラプラス分布に従うことを導いていたが、本稿では x_i および $\frac{x_i}{\sqrt{\Delta t_i}}$ に対して、それぞれ t 分布と正規分布を仮定して、取引発生間隔 Δt_i の分布および発生強度に従う分布を導く方法をとる。ここで、 $\frac{x_i}{\sqrt{\Delta t_i}}$ が正規分布に従うという仮定は、Shibata[2006] において $\frac{x_i}{\sqrt{\Delta t_i}}$ が $N(0, \sigma^2)$ に従うという設定に準じている。これ以後本稿では、取引発生メカニズムに関係なく、取引の発生間隔の逆数 $\frac{1}{\Delta t_i}$ を発生強度 λ と定義するので注意されたい。

先に述べたように本稿モデルの仮定は $\frac{x_i}{\sqrt{\Delta t_i}}$ が正規分布 $N(0, \sigma_1^2)$ に従い、 x_i が t 分布 $st(0, \sigma_2^2, d_2)$ に従うことである。ここで $st(0, \sigma_2^2, d_2)$ は、 $N(0, \sigma_2^2)$ に従う確率変数を自由度 d_2 でスケール調整したカイ二乗分布 $\frac{\chi^2(d_2)}{d_2}$ に従う確率変数の平方根で除した商に従う確率分布であり、その密度関数は

$$f(x_i) = \frac{\Gamma(\frac{d_2+1}{2})}{\Gamma(\frac{d_2}{2})\sqrt{\pi d_2 \sigma_2^2}} \left(1 + \frac{x_i^2}{d_2 \sigma_2^2}\right)^{-\frac{d_2+1}{2}}, \quad i = 2, 3, \dots, n$$

である。この自由度でスケール調整したカイ二乗分布 $\frac{\chi^2(d_2)}{d_2}$ に従う確率変数を ϕ とすれば

$$x_i \sim N\left(0, \frac{\sigma_2^2}{\phi}\right), \quad i = 2, 3, \dots, n$$

が成立する。そして $w = \phi^{-1}$ は密度関数

$$f(w) = \frac{d_2^{\frac{d_2}{2}} w^{-\frac{d_2}{2}-1}}{\Gamma(\frac{d_2}{2})} \exp\left(-\frac{d_2}{2w}\right)$$

を持つ逆ガンマ分布 $Ga^{-1}\left(\frac{d_2}{2}, \frac{d_2}{2}\right)$ に従う。

一方で、 $\frac{x_i}{\sqrt{\Delta t_i}} \sim N(0, \sigma_1^2)$ を書き換えると

$$x_i \sim N(0, \sigma_1^2 \Delta t_i), \quad i = 2, 3, \dots, n$$

と表現できる。この2つの x_i が従う確率分布は等しくなければならないので、

$$\Delta t_i = \frac{\sigma_2^2}{\sigma_1^2 \phi}$$

という条件が導かれる。これは、 Δt_i が逆ガンマ分布 $Ga^{-1}\left(\frac{d_2}{2}, \frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2}\right)$ に従い、その逆数である発生強度 λ はガンマ分布 $Ga\left(\frac{d_2}{2}, \frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2}\right)$ に従うことを意味している。

3.2 取引約定回数の分布

クラスタ内で λ がガンマ分布 $Ga\left(\frac{d_2}{2}, \frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2}\right)$ に従うとき、その密度関数は、

$$h(\lambda) = \frac{\frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2} \frac{d_2}{2} \lambda^{\frac{d_2}{2}-1}}{\Gamma\left(\frac{d_2}{2}\right)} \exp\left(-\frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2} \lambda\right)$$

である。この仮定のもので、本節では時刻 t までに発生する約定回数 N_t の分布が負の 2 項分布となることを示す。

時刻 t までに発生する約定回数 N_t がポアソン分布に従うならばその確率関数は以下で与えられる。

$$P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t}, \quad k = 0, 1, \dots$$

本稿モデルでは λ はガンマ分布 $Ga\left(\frac{d_2}{2}, \frac{d_2}{2a}\right)$ に従うので、その確率関数は

$$\begin{aligned} P(N_t = k) &= \int_0^\infty \frac{(\lambda t)^k}{k!} e^{-\lambda t} h(\lambda) d\lambda \\ &= \int_0^\infty \frac{(\lambda t)^k}{k!} e^{-\lambda t} \frac{\left(\frac{d_2}{2a}\right)^{\frac{d_2}{2}}}{\Gamma\left(\frac{d_2}{2}\right)} \lambda^{\frac{d_2}{2}-1} \exp\left(-\frac{d_2}{2a} \lambda\right) d\lambda \\ &= \frac{t^k \frac{d_2}{2a} \frac{d_2}{2}}{\Gamma(k+1) \Gamma\left(\frac{d_2}{2}\right)} \int_0^\infty \lambda^{\frac{d_2}{2}+k-1} \exp\left(-\lambda\left(\frac{d_2}{2a} + t\right)\right) d\lambda \\ &= \frac{t^k \frac{d_2}{2a} \frac{d_2}{2}}{\Gamma(k+1) \Gamma\left(\frac{d_2}{2}\right)} \int_0^\infty \left(\frac{s}{\frac{d_2}{2a} + t}\right)^{\frac{d_2}{2}+k-1} \exp(-s) \frac{ds}{\frac{d_2}{2a} + t} \\ &= \frac{t^k \frac{d_2}{2a} \frac{d_2}{2}}{\Gamma(k+1) \Gamma\left(\frac{d_2}{2}\right)} \left(\frac{d_2}{2a} + t\right)^{-\frac{d_2}{2}-k} \int_0^\infty s^{\frac{d_2}{2}+k-1} \exp(-s) ds \\ &= \frac{t^k \frac{d_2}{2a} \frac{d_2}{2}}{\Gamma(k+1) \Gamma\left(\frac{d_2}{2}\right)} \left(\frac{d_2}{2a} + t\right)^{-\frac{d_2}{2}-k} \Gamma\left(\frac{d_2}{2} + k\right) \\ &= \frac{\Gamma\left(\frac{d_2}{2} + k\right)}{\Gamma(k+1) \Gamma\left(\frac{d_2}{2}\right)} \left(\frac{t}{\frac{d_2}{2a} + t}\right)^k \left(\frac{\frac{d_2}{2a}}{\frac{d_2}{2a} + t}\right)^{\frac{d_2}{2}} \\ &= \frac{\Gamma\left(\frac{d_2}{2} + k\right)}{\Gamma(k+1) \Gamma\left(\frac{d_2}{2}\right)} q_t^{\frac{d_2}{2}} (1 - q_t)^k \end{aligned}$$

となり、これは負の 2 項分布の確率関数である。ここで、 $s = \lambda\left(\frac{d_2}{2a} + t\right)$, $q_t = \frac{\frac{d_2}{2a}}{\frac{d_2}{2a} + t}$ と変形しているので注意されたい。

3.3 取引発生間隔 Δt_i が逆ガンマ分布に従う場合のクラスタリング

3.3.1 パラメータ推定

取引発生強度 λ がガンマ分布 $Ga\left(\frac{d_2}{2}, \frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2}\right)$ に従うとき、取引発生間隔 Δt_i は逆ガンマ分布 $Ga^{-1}\left(\frac{d_2}{2}, \frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2}\right)$ に従う。この時の約定取引の間隔 $\Delta t_i, i = 2, 3, \dots, n$ の密度関数は

$$g(\Delta t_i) = \frac{\frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2} \frac{d_2}{2} \Delta t_i^{-\frac{d_2}{2}-1}}{\Gamma\left(\frac{d_2}{2}\right)} \exp\left(-\frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2} \frac{1}{\Delta t_i}\right)$$

である。以後、表記を簡潔にするため $a = \frac{\sigma_1^2}{\sigma_2^2}$ とおく。

Δt_i に関する対数尤度関数は

$$\begin{aligned} L(\Delta t_2, \dots, \Delta t_n; d_2, a) &\equiv \frac{(n-1)d_2}{2} \log\left(\frac{d_2}{2a}\right) - \frac{d_2}{2} \sum_{i=2}^n \log(\Delta t_i) - \sum_{i=2}^n \log(\Delta t_i) \\ &\quad - (n-1) \log\left(\Gamma\left(\frac{d_2}{2}\right)\right) - \frac{d_2}{2a} \sum_{i=2}^n \frac{1}{\Delta t_i} \end{aligned}$$

と表せる。各々の 1 次偏導関数を求めると

$$\frac{\partial L}{\partial d_2} = \frac{n-1}{2} \log\left(\frac{d_2}{2a}\right) + \frac{n-1}{2} - \frac{1}{2} \sum_{i=2}^n \log(\Delta t_i) - \frac{n-1}{2} \frac{\Gamma'\left(\frac{d_2}{2}\right)}{\Gamma\left(\frac{d_2}{2}\right)} - \frac{1}{2a} \sum_{i=2}^n \frac{1}{\Delta t_i} \quad (3.1)$$

$$\frac{\partial L}{\partial a} = -a^{-1} \frac{(n-1)d_2}{2} + a^{-2} \frac{d_2}{2} \sum_{i=2}^n \frac{1}{\Delta t_i} \quad (3.2)$$

となり、偏導関数をそれぞれ 0 としてパラメータについて解けば最尤推定量が求まる。

a の最尤推定量は (3.2) 式より

$$\hat{a} = \frac{1}{n-1} \sum_{i=2}^n \frac{1}{\Delta t_i}$$

と求められる。

本稿のモデルでの未知パラメータは、 $\sigma_1^2, \sigma_2^2, d_2$ の 3 種類である。本稿モデルでは、 Δt_i が逆ガンマ分布 $Ga^{-1}\left(\frac{d_2}{2}, \frac{d_2}{2} \frac{\sigma_2^2}{\sigma_1^2}\right)$ に従い、 $\frac{x_i}{\sqrt{\Delta t_i}}$ が正規分布 $N(0, \sigma_1^2)$ に従い、 x_i が t 分布 $st(0, \sigma_2^2, d_2)$ に従うモデル設定なので、(3.2) 式で求めた \hat{a} を用いて、3 パラメータの推定量を求める方法は以下の、2 通り考えられる。

1 つ目の方法は、まず \hat{a} を (3.1) 式に代入することで

$$\frac{\Gamma'\left(\frac{d_2}{2}\right)}{\Gamma\left(\frac{d_2}{2}\right)} - \log\left(\frac{d_2}{2}\right) + \log\left(\frac{1}{n-1} \sum_{i=2}^n \frac{1}{\Delta t_i}\right) + \frac{1}{n-1} \sum_{i=2}^n \log(\Delta t_i) = 0$$

を導出し、ニュートンラフソン法などを用いて \hat{d}_2 を数値計算する。さらに x_i が $st(0, \sigma_2^2, d_2)$ に従うことを利用して $\hat{\sigma}_2^2$ を最尤推定する。そして $\sigma_1^2 = a\sigma_2^2$ という関係式を用い、求めた $\hat{\sigma}_2^2$ と \hat{a} を利用して $\hat{\sigma}_1^2$ を求める。

2つ目の方法は、 $\frac{x_i}{\sqrt{\Delta t_i}}$ が正規分布 $N(0, \sigma_1^2)$ に従うことから $\hat{\sigma}_1^2$ を求め、 \hat{a} と関係式 $\frac{\hat{\sigma}_1^2}{\hat{a}} = \sigma_2^2$ から $\hat{\sigma}_2^2$ を求める。そして、 x_i が $st(0, \sigma_2^2, d_2)$ に従うことを利用し \hat{d}_2 を最尤法で求める。

なお、本稿の分析では1つ目の推定方法を用いている。これは、次節のクラスタ分割は \hat{d}_2 と \hat{a} の値がわかれば実行できるので分割実行時の計算量が減ること、および1つ目の方法のアルゴリズムの方が Shibata[2006] の分割アルゴリズムとの比較に適していることがその理由である。

3.3.2 クラスタの分割

クラスタ内の全データから求めた最尤推定量 \hat{d}_2, \hat{a} とすると最大対数尤度は

$$\begin{aligned} L_0 &= L(\Delta t_2, \dots, \Delta t_n; \hat{d}_2, \hat{a}) \\ &= \frac{(n-1)\hat{d}_2}{2} \log\left(\frac{\hat{d}_2}{2\hat{a}}\right) - \frac{\hat{d}_2}{2} \sum_{i=2}^n \log(\Delta t_i) - \sum_{i=2}^n \log(\Delta t_i) - (n-1) \log\left(\Gamma\left(\frac{\hat{d}_2}{2}\right)\right) - \frac{\hat{d}_2}{2\hat{a}} \sum_{i=2}^n \frac{1}{\Delta t_i} \end{aligned}$$

とあらわせる。

一方、時刻 t_1, \dots, t_k に発生する約定取引が先に定義したパラメータ d_{2k_1}, a_{k_1} を持つガンマ分布に従い、 t_k, \dots, t_n に発生する約定取引がパラメータ d_{2k_2}, a_{k_2} を持つガンマ分布に従うと仮定する。このとき、前者の約定取引の間隔 $\Delta t_2, \dots, \Delta t_k$ の対数尤度関数は

$$L_{1,k}(\Delta t_2, \dots, \Delta t_k; d_{2k_1}, a_{k_1})$$

となり、後者の対数尤度関数は

$$L_{2,k}(\Delta t_{k+1}, \dots, \Delta t_n; d_{2k_2}, a_{k_2})$$

となる。

$L_{1,k}, L_{2,k}$ 各々で最尤推定量 $\hat{d}_{2k_1}, \hat{a}_{k_1}, \hat{d}_{2k_2}, \hat{a}_{k_2}$ を取得すると、それぞれの最大対数尤度は

$$\begin{aligned} L_{k_1} &= L_{1,k}(\Delta t_2, \dots, \Delta t_k; \hat{d}_{2k_1}, \hat{a}_{k_1}) \\ &= \frac{(k-1)\hat{d}_{2k_1}}{2} \log\left(\frac{\hat{d}_{2k_1}}{2\hat{a}_{k_1}}\right) - \frac{\hat{d}_{2k_1}}{2} \sum_{i=2}^k \log(\Delta t_i) - \sum_{i=2}^k \log(\Delta t_i) \\ &\quad - k \log\left(\Gamma\left(\frac{\hat{d}_{2k_1}}{2}\right)\right) - \frac{\hat{d}_{2k_1}}{2\hat{a}_{k_1}} \sum_{i=2}^k \frac{1}{\Delta t_i} \\ L_{k_2} &= L_{2,k}(\Delta t_{k+1}, \dots, \Delta t_n; \hat{d}_{2k_2}, \hat{a}_{k_2}) \\ &= \frac{(n-k)\hat{d}_{2k_2}}{2} \log\left(\frac{\hat{d}_{2k_2}}{2\hat{a}_{k_2}}\right) - \frac{\hat{d}_{2k_2}}{2} \sum_{i=k+1}^n \log(\Delta t_i) - \sum_{i=k+1}^n \log(\Delta t_i) \\ &\quad - (n-k) \log\left(\Gamma\left(\frac{\hat{d}_{2k_2}}{2}\right)\right) - \frac{\hat{d}_{2k_2}}{2\hat{a}_{k_2}} \sum_{i=k+1}^n \frac{1}{\Delta t_i} \end{aligned}$$

となるので、このときの最大対数尤度は

$$L_{1,2,k} = L_{1,k}(\Delta t_2, \dots, \Delta t_k; \hat{d}_{2k_1}, \hat{a}_{k_1}) + L_{2,k}(\Delta t_{k+1}, \dots, \Delta t_n; \hat{d}_{2k_2}, \hat{a}_{k_2})$$

と表せる。この k を動かしたとき、最も大きな $L_{1,2,k}$ を達成する k' が、尤度原理の観点から分割点の最有力候補になる。

もし尤度比 $L_{1,2,k'}/L_0$ の理論分布が求められるのであれば、尤度比検定を通じて分割を行うか否かの統計的な判定が実施できるが、本稿のモデルにおいても、この検定統計量の理論分布を解析的には求めることは容易ではない。このため、本稿のモデルでも、Shibata[2006] と同様にモンテカルロシミュレーションを用いて検定統計量の分布を取得した。

具体的には、 L_0 で取得した最尤推定量 \hat{d}_2, \hat{a} をもとに逆ガンマ分布のサンプルデータの系列 $\{\Delta t_i\}_{i=2,\dots,N}$ を

$$\Delta t_i \sim Ga^{-1}\left(\frac{d_2}{2}, \frac{d_2}{2a}\right)$$

より、発生させて経験分布関数を作ることで、尤度比の検定統計量が従う分布を取得した。そして得られた分布から、棄却域を設定し、実現値 $L_{1,2,k'}/L_0$ が、棄却域に入らない場合にはデータを分割しない一方で、実現値が、棄却域に入る場合には k' でデータを分割することとした。

4 先行研究と提案モデルとの比較

3節で見た通り、本稿が提案するモデルは λ が確率変数となっているので、Shibata[2006] のモデルよりも柔軟にトータルボラティリティ Σ を表現できる。この節では、Shibata[2006] のモデルと本稿のモデルの2つのモデルを用い、2010年5月に発生したフラッシュクラッシュ前後のデータに対する両者の当てはまり具合を比較した結果を述べる。

具体的には、まず各クラスタへの確率分布の適合度をコルモゴロフ・スミルノフ検定（以下、KS検定）及び、QQプロットで確認する。また、トータルボラティリティ Σ とボラティリティ σ^2 、約定期間 Δt_i について、平常時、異常時、取引の閑散時、活発時をグラフを通じて比較する。

なお、Shibata[2006] のモデルと本稿のモデルは、個々の取引発生に対して独立性を仮定している。そのため Ljung-Box 検定により、各クラスタ内での自己相関の有無も確認している。

4.1 使用するデータ

分析に用いた Tick データは、EBS 社による 2010/5/4 21:00:03~2010/5/7 05:25:49(GMT) の期間のドル円 (USD/JPY) の約定取引データであり、全部 36759 レコードからなる⁴。この期間は、米国の *S&P500* 先物の流動性供給が瞬間的に壊れて大暴落したフラッシュクラッシュの時間帯⁵を含んでいる。図1を見てわかるように、暴落の影響は為替にも波及し円高に振れている。

本稿の分析では、USD/JPY 為替約定取引データの約定の発生した時間、約定価格（売り買いは区別せず合算）をもとに、取引発生間隔と対応する対数価格差を取得し、Shibata[2006] モデルと本稿モデルに当てはめた⁶。

⁴日本時間 (JST) では 2010/5/5 06:00:03~2010/5/8 14:25:49、ニューヨーク夏時間 (EDT) では 2010/5/4 17:00:03~2010/5/8 01:25:49. (EDT) に対応。GMT 基準では、EDT+4h=GMT となる

⁵フラッシュクラッシュは、一般に 2010/5/6 14:30~14:45 (EDT) までの時間帯を指す。尚、本稿では、モデルの当てはめのみ焦点をあてており、フラッシュクラッシュの内容そのものには立ち入らない

⁶少数のデータを使用したことで過度に当てはまりが良く見えることを防ぐためと、パラメータ推計でエラーを出さずに推計するため、クラスタへの区分に際して、70 ticks を最小のクラスタの要素数とした。

4.2 コルモゴロフ-スミルノフ検定を用いたクラスタへの区分

これまでの議論からわかるようにクラスタの分割は尤度比検定の有意水準 γ に依存する。今回の分析では各クラスタの取引間隔のデータに対して、仮定する確率分布への適合度がもっとも高くなるようにパラメータ γ を決定し、Shibata[2006] のモデルと本稿のモデルの性能を比較した。

具体的にはパラメータ γ を 1% から 30% の範囲で 1% ずつ変化させて全クラスタに KS (コルモゴロフ-スミルノフ) 検定を実施した。そして、帰無仮説が棄却できない割合を計算し、もっとも高い割合のときの γ を採用した。なお、今回の分析では Shibata[2006] モデルに対して上記の手順で得られたパラメータ γ の値を先に求め、それと同じ値を本稿のモデルのクラスタ分割でも用いることとした。実際、この γ の値は Shibata[2006] モデルにおける Δt の分布の適合の割合を最も高くする最適値であり、もし同じ γ の値で Shibata[2006] モデルよりも本稿モデルのほうが適合の割合が高ければ、本稿のモデルの方が当てはまっていることを意味する。

表 1 は、Shibata[2006] のモデルで KS 検定の適合の割合を探索した結果である。Shibata[2006] のモデルでは、クラスタ内では、 Δt_i はパラメータ λ の指数分布、 $\frac{x_i}{\sigma\sqrt{\Delta t_i}}$ は標準正規分布、 x_i はパラメータ α のラプラス分布に従う。今回の探索で、 Δt_i が指数分布に従う割合が最も高かったパラメータ γ の値が 24% であることが読み取れる。この可変パラメータ γ に対し、 Δt_i が指数分布に従う適合割合は 39.3%、 $\frac{x_i}{\sqrt{\Delta t_i}}$ が正規分布に従う割合は 27.5%、 x_i がラプラス分布に従う割合は 20.5% となっている。

そこで、パラメータ γ を 24% とした場合の本稿のモデルでの KS 検定の適合の割合を調査した。本稿のモデルでは Δt_i が逆ガンマ分布 $Ga^{-1}(\frac{d_2}{2}, \frac{d_2}{2a})$ 、 $\frac{x_i}{\sqrt{\Delta t_i}}$ が正規分布 $N(0, \sigma_1^2)$ 、 x_i が t 分布 $st(0, \sigma_2^2, d_2)$ に従っており、それぞれ 40.9%、29.6%、23.3% という結果となった。

以上より、Shibata[2006] のモデルと本稿のモデルの比較では、全てにおいて、本稿のモデルの方が Shibata[2006] のモデルよりも KS 検定の適合の割合が高くなることがわかる。なお、Shibata[2006] のモデルでは 356 個のクラスタ、本稿のモデルでは 399 個のクラスタを得ている。

図 2 は Shibata[2006] での区分の結果、 λ 、および為替レートの推移を示し、図 3 は本稿のモデルでの区分の結果、 d_2 、および為替レートの推移を示している。どちらの図でも、 λ および d_2 が小さくなる取引発生間隔が疎なところでは、価格の変動が穏やかになる一方で、 λ および d_2 は大きくなる取引発生間隔が密なところでは、価格の変動は激しくなることが見て取れる。

4.3 クラスタ内での独立性の確認

Shibata[2006] モデルと本稿のモデルの Δt_i は独立性を仮定していることから、得られたクラスタについて、Ljung-Box 検定による自己相関の有無を調べた。

この結果、有意水準 1% では、Shibata[2006] のモデルの 93.2% のクラスタが、そして本稿のモデルの 94% が自己相関を有していないという帰無仮説を棄却できなかった。なお、全データに Ljung-Box 検定を実施すると、有意水準 1% で帰無仮説は棄却され自己相関は存在するという結果になる。つまり、適当なクラスタに分割することで取引発生間隔の自己相関はほぼ解消されており、どちらのモデルもよく当てはまっていることがわかった。

4.4 QQプロットへの当てはめ

本節では、約定取引が低迷していた時間帯のクラスタと約定取引が比較的活発であった時間帯のクラスタを任意に抽出し、仮定している確率分布への当てはまりをQQ-プロットを通じて視覚的に比較する。

本分析では約定取引が低迷していた時間帯のクラスタとして、Shibata[2006]のモデルの当てはめ結果からは、強度 $\lambda = 0.15$ を持つ 16920 番目から 16993 番目のクラスタを選び、本稿のモデルの当てはめ結果からは、これとほぼ対応した 16919 番目から 16989 番目のクラスタを抽出した。また、約定取引が比較的活発であった時間帯のクラスタでも同様に、Shibata[2006]のモデルでの強度 $\lambda = 0.54$ を持つ 26649 番目から 26745 番目のクラスタと、これとほぼ対応した本稿のモデルによる 26648 番目から 26745 番目のクラスタを抽出した。そして、これらのクラスタに対してQQプロットに当てはめた。

図4が低迷時のShibata[2006]のモデルによる当てはめ結果に対するQQプロット、図5が低迷時の本稿のモデルによる当てはめ結果に対するQQプロットである。また図6が活発時のShibata[2006]のモデルのそれに、図7が本稿のモデルのそれに対応している。これらの図には上から順に Δt_i , $\frac{x_i}{\sqrt{\Delta t_i}}$, x_i の3種類のデータに対するQQプロットが描かれている。

図4および図6の上段のグラフは λ の指数分布、中段のグラフは標準正規分布、下段のグラフはパラメータ $\hat{\alpha}$ のラプラス分布に当てはめたQQプロットである。そして、図5および図7の上段のグラフは逆ガンマ分布 $Ga^{-1}(\frac{\hat{d}_2}{2}, \frac{\hat{d}_2}{2\hat{\alpha}})$ 、中段のグラフは $\frac{x_i}{\sqrt{\Delta t_i}}$ を正規分布 $N(0, \hat{\sigma}_1^2)$ に、下段のグラフは t 分布 $st(0, \hat{\sigma}_2^2, \hat{d}_2)$ に当てはめたQQプロットである。

約定取引が低迷していた時間帯のQQプロットでは、本稿モデルの図5の1段目の分布の裾の当てはまりが悪く、また3段目のグラフも全体的に当てはまりがよくない。一方でShibata[2006]のモデルの図4の当てはまりが良く見える。

約定取引が比較的活発であった時間帯のQQプロットについては、Shibata[2006]のモデルの図6の1段目の裾の当てはまりがあまり良くないが、本稿モデルの図7のグラフではそのようなはずれ具合は見られない。

以上のことから、取引の状態によって2つのモデルの当てはまりは逆転することが明らかになった。

4.5 相場の状態とトータルボラティリティの傾向

本節では、平常時と異常時、取引が低迷する時間帯と取引が活発な時間帯からなる状態に分類したときのトータルボラティリティ Σ 、ボラティリティ σ^2 、および約定発生間隔 Δt_i の特性をグラフによる視覚化を通じて確認する。

図8はボラティリティ σ^2 、約定発生間隔 Δt_i の推移を示したグラフであり、図9はトータルボラティリティ Σ とUSD/JPYの為替レートの推移を示したグラフである。

平常時、トータルボラティリティ Σ におけるボラティリティ σ^2 の影響は小さく、約定発生間隔 Δt_i の大きさに影響を受けている。フラッシュクラッシュ発生の異常時では、約定発生間隔 Δt_i は小さくなる一方で、ボラティリティ σ^2 そのものが大きくなる時間帯があった。そしてフラッシュクラッシュの直後には、 Δt が大きく、また、ボラティリティ σ^2 が大きくなる時間帯が生じていた。

また全体的にトータルボラティリティ Σ の大きな時間帯は、本稿のモデルの当てはまりが良い場合が多く、 Σ 小さな時間帯は Shibata[2006] のモデルの当てはまりが良い場合が多かった。

5 独立異分布としてのモデル化

前節の分析では、本稿のモデルは KS 検定を通じたモデルの検証では Shibata[2006] よりもよい結果を示したが、取引の状態によってはモデルの当てはまりは、Shibata[2006] よりも悪くなることがわかった。この節ではこの問題を改良する方法を提案する。

提案する方法は単純で、 Δt_i が従う分布をクラスタごとに指数分布ないし逆ガンマ分布の2つ分布を尤度原理で切り替える、いわば独立異分布モデルの当てはめである。本節ではこの独立異分布モデルを本データに当てはめた結果を紹介する。

4.2 節と同様の手順で独立異分布の適合の度合いを調べた結果、 Δt_i の適合の割合は 45.0%、 $\frac{x_i}{\sqrt{\Delta t_i}}$ の適合の割合は 35.2%、 x_i の適合の割合は 27.6% となっており先の2つのモデルを単独で用いるよりも改善した。特に取引発生頻度の低い時間帯には指数分布が、取引発生頻度の高い時間帯には逆ガンマ分布が選ばれていた。

図 10 は、指数分布と逆ガンマ分布のどちらが選ばれたかを時系列で図示している。指数分布に対して当てはまりが良かったのは、為替レートがあまり動いていなかった時間帯であり、逆ガンマ分布の当てはまりが良かったのは為替レートの変化が激しかった時間帯となっている。つまり、本節で提案するモデル化の方法は、前節で示唆された2つのモデルの当てはまりの特性を利用しており、2つの異なる確率分布の切り替えが正しく行われていることを示唆している。

6 まとめ

本論文では、同一クラスタにおける為替取引の発生が強度一定のポアソン過程に従うという Shibata[2006] の仮定を緩め、同一クラスタ内で強度がガンマ分布に従う複合ポアソン過程の仮定をおき、取引発生時間間隔と為替レートの関係性の再構築を試みた。

Shibata[2006] のモデルと本稿のモデルでの当てはめ結果を比較したところ、 Δt_i 、 $\frac{x_i}{\sqrt{\Delta t_i}}$ 、 x_i のすべての系列で、本稿のモデルの方が Shibata[2006] のモデルよりも適合の割合が高かった。

QQ プロットによる適合度の検証では、取引発生回数度の低いクラスタにおいては Shibata[2006] のモデルの当てはめよりも本論文のモデルの当てはまりが悪く見えるクラスタが多く存在した。この問題を解決するために、2つのモデルの仮定を組み合わせ、取引発生間隔が指数分布と逆ガンマ分布の2つの分布に対して独立異分布で従う仮定をおいたところ、Shibata[2006] のモデル及び本稿のモデルよりも高い当てはまりを示した。

今後の課題としては、クラスタ間関係に対する取り扱いがあげられる。本稿では、クラスタ間関係を扱わない Shibata[2006] の仮定を受け、取引発生間隔と為替レートとの関係性のみ焦点を充ててモデル化を行ってきたため、クラスタ間関係についての分析は行っていない。これについては今後の課題としたい。

A 正規分布に指数分布の平方根を乗じた分布がラプラス分布となることの証明

以下では, Shibata[2006] のモデルで, 正規分布に指数分布の平方根を乗じた分布がラプラス分布となることの証明に使用した以下の等式を導出する.

$$\int_0^{\infty} \exp\left(-\frac{1}{2}(a^2u^2 + b^2u^{-2})\right) du = \left(\frac{\pi}{2a^2}\right)^{\frac{1}{2}} \exp(-|ab|) \quad a, b > 0$$

A.1 ラプラス変換

以下にラプラス変換の定義と幾つかの公式を掲載する.

定義 1 : ラプラス変換

$$L\{F(t)\}(s) \equiv \int_0^{\infty} F(t) \exp(-st) dt$$

定義 2 : 指数関数的位数

関数 $F(t)$ に対して, $t \geq T$ である全ての t に対して

$$|\exp(-\sigma t)F(t)| < M$$

または,

$$|F(t)| < \exp(\sigma t)M$$

となるような, $\sigma, M(> 0), T(> 0)$ を見つけることができた時, $F(t)$ は, $t \rightarrow \infty$ で指数関数的位数 σ を持つという.

定義 3 : クラス A の関数 $F(t)$ が, $t \rightarrow \infty$ である指数関数的位数 σ を持ち, $t \geq 0$ のあらゆる有限区間で連続である場合, 関数 $F(t)$ はクラス A に属するといひ以下で書く.

$$F(t) \in A$$

公式

公式 1 : $\frac{df(t)}{dt} \in A$ の時, $L\{\frac{df(t)}{dt}\}(s) = sL\{f(t)\}(s) - f(0)$

公式 2 : $f(t) \in A$ の時, $L\{tf(t)\}(s) = -\frac{dL\{f(t)\}(s)}{ds}$

公式 3 : $f(t), g(t) \in A$ の時, $L\{f(t)\} = L\{g(t)\}$ ならば, $f(t), g(t)$ の不連続点を除いて, $f(t) = g(t)$

公式 4 : $L\{\int_0^t F(x)dx\}(s) = s^{-1}L\{F(t)\}(s)$

公式 5 : $L^{-1}\{\exp(-y^{\frac{1}{2}})\}(t) = \frac{\exp(-\frac{1}{4t})}{2\sqrt{\pi t^3}}$

A.2 命題とその証明

命題 :

$$L\left(\frac{1}{\sqrt{\pi t}} \exp\left(-\frac{a^2}{4t}\right)\right) = \frac{1}{\sqrt{s}} \exp(-a\sqrt{s}) \quad a \geq 0$$

証明： $y(s) = \frac{1}{\sqrt{s}} \exp(-a\sqrt{s})$ $a \geq 0$ と置く。この時、

$$y(s) = \frac{1}{\sqrt{s}} \exp(-a\sqrt{s}) = s^{-\frac{1}{2}} \exp(-as^{\frac{1}{2}})$$

より、

$$\begin{aligned} y'(s) &= -\frac{1}{2}s^{-\frac{3}{2}} \exp(-as^{\frac{1}{2}}) + s^{-\frac{1}{2}} \left(-\frac{a}{2}\right) s^{-\frac{1}{2}} \exp(-as^{\frac{1}{2}}) \\ &= -\frac{1}{2}s^{-\frac{3}{2}} \exp(-as^{\frac{1}{2}}) - \frac{a}{2}s^{-1} \exp(-as^{\frac{1}{2}}) \\ &= -\frac{1}{2}s^{-1} (s^{-\frac{1}{2}} \exp(-as^{\frac{1}{2}}) + as^{\frac{1}{2}} s^{-\frac{1}{2}} \exp(-as^{\frac{1}{2}})) \\ &= -\frac{1}{2}s^{-1} (y(s) + as^{\frac{1}{2}} y(s)) \\ &= -\frac{1}{2}s^{-1} y(s) (1 + as^{\frac{1}{2}}) \end{aligned}$$

以上より、以下が示せた。

$$y'(s) = -\frac{1}{2}s^{-1} y(s) (1 + as^{\frac{1}{2}})$$

次に、 $y'(s)$ を使って、以下を示す。

$$(sy'(s))' = -\frac{1}{2}y'(s) + \frac{a^2}{4}y(s)$$

まず、 $sy'(s) = -\frac{1}{2}y(s)(1 + as^{\frac{1}{2}})$ より、

$$\begin{aligned} (sy'(s))' &= -\frac{1}{2}(y(s)(1 + as^{\frac{1}{2}}))' = -\frac{1}{2}(y(s) + y(s)as^{\frac{1}{2}})' \\ &= -\frac{1}{2}(y'(s) + y'(s)as^{\frac{1}{2}} + y(s)a\frac{1}{2}s^{-\frac{1}{2}}) = -\frac{1}{2}y'(s) - \frac{1}{2}as^{\frac{1}{2}}(y'(s) + \frac{1}{2}y(s)s^{-1}) \\ &= -\frac{1}{2}y'(s) - \frac{1}{2}as^{\frac{1}{2}} \left(-\frac{1}{2}as^{-\frac{1}{2}}y(s)\right) \\ &= -\frac{1}{2}y'(s) + \frac{a^2}{4}y(s) \end{aligned}$$

以上より、以下を得る。

$$(sy'(s))' = -\frac{1}{2}y'(s) + \frac{a^2}{4}y(s)$$

次に, $L\{f(t)\}(s) \equiv y(s)$ と置くことで, 上の式をラプラス変換の形に書き換える.

$$\begin{aligned}
 L\{t(tf(t))'\}(s) &= -\frac{d}{ds}(L\{tf(t)\}'(s)) \quad (\text{公式 2}) \\
 &= -\frac{d}{ds}(sL\{tf(t)\}(s)) \quad (\text{公式 1}) \\
 &= \frac{d}{ds}\left(s\frac{dL\{f(t)\}(s)}{ds}\right) \quad (\text{公式 2}) \\
 &= \frac{d}{ds}\left(s\frac{dy(s)}{ds}\right) \\
 &= (sy'(s))'
 \end{aligned}$$

よって

$$L\{t(tf(t))'\}(s) = -\frac{1}{2}y'(s) + \frac{a^2}{4}y(s)$$

ここで, $y(s)$ は, 定義 $L\{f(t)\}(s) = y(s)$ と公式 2 より,
 $L\{tf(t)\}(s) = -L\{f(t)\}'(s) = -y'(s)$
 よって, 以下を得る.

$$L\{t(tf(t))'\}(s) = \frac{1}{2}L\{tf(t)\}(s) + \frac{a^2}{4}L\{f(t)\}(s)$$

よって, 以下の常微分方程式を得る.

$$t(tf(t))' = \frac{1}{2}tf(t) + \frac{a^2}{4}f(t)$$

式を変形して常微分方程式を解く.

$$\begin{aligned}
 f(t) + tf'(t) &= (tf(t))' \\
 &= \frac{1}{2}f(t) + \frac{a^2}{4}\frac{1}{t}f(t)
 \end{aligned}$$

よって,

$$\begin{aligned}
 tf'(t) &= -\frac{1}{2}f(t) + \frac{a^2}{4}\frac{1}{t}f(t) \\
 f'(t) &= -\frac{1}{2}\frac{1}{t}f(t) + \frac{a^2}{4}\frac{1}{t^2}f(t) \\
 \frac{f'(t)}{f(t)} &= -\frac{1}{2}\frac{1}{t} + \frac{a^2}{4}\frac{1}{t^2}
 \end{aligned}$$

ここまで変形すれば、変数分離型として以下のように解くことができる。

$$\begin{aligned}
 \frac{d \log f(t)}{dt} &= \frac{f'(t)}{f(t)} = -\frac{1}{2} \frac{1}{t} + \frac{a^2}{4} \frac{1}{t^2} \\
 \log f(t) &= \int \frac{d \log f(t)}{dt} dt = -\frac{1}{2} \int \frac{1}{t} dt + \frac{a^2}{4} \int \frac{1}{t^2} dt \\
 &= -\frac{1}{2} \log(t) - \frac{a^2}{4} t^{-1} + C \\
 f(t) &= \exp\left(-\frac{1}{2} \log(t) - \frac{a^2}{4} t^{-1} + C\right) \\
 &= \exp\left(\log(t^{-\frac{1}{2}})\right) \exp\left(-\frac{a^2}{4t}\right) \exp(C) \\
 &= c(t^{-\frac{1}{2}}) \exp\left(-\frac{a^2}{4t}\right) \\
 &= \frac{c}{\sqrt{t}} \exp\left(-\frac{a^2}{4t}\right)
 \end{aligned}$$

次に c を求める。 c を決めるために、 $s \rightarrow \infty, t \rightarrow \infty$ とすると。

$$\begin{aligned}
 y(s) &= \frac{1}{\sqrt{s}} \exp(-a\sqrt{s}) & s \rightarrow \infty \Rightarrow y(s) &= \frac{1}{\sqrt{s}} \\
 f(t) &= \frac{c}{\sqrt{t}} \exp\left(-\frac{a^2}{4t}\right) & t \rightarrow \infty \Rightarrow f(t) &= \frac{c}{\sqrt{t}}
 \end{aligned}$$

以上、 $v^2 \equiv st$ と置くと、

$$\begin{aligned}
 \frac{1}{\sqrt{s}} &= L\left(\frac{c}{\sqrt{t}}\right) \\
 &= \int_0^\infty \frac{c}{\sqrt{t}} \exp(-st) dt \\
 &= \int_0^\infty \frac{c\sqrt{s}}{vs} \exp(-v^2) 2v dv & (2v dv = s dt) \\
 &= \frac{2c}{\sqrt{s}} \int_0^\infty \exp(-v^2) dv & \left(\int_0^\infty \exp(-v^2) dv = \sqrt{\pi}/2\right) \\
 &= \frac{2c}{\sqrt{s}} \frac{\sqrt{\pi}}{2} \\
 &= \frac{c\sqrt{\pi}}{\sqrt{s}}
 \end{aligned}$$

以上より、 $c = \frac{1}{\sqrt{\pi}}$
よって、以下となる。

$$f(t) = \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{a^2}{4t}\right)$$

$L\{f(t)\}(s) \equiv y(s) \equiv \frac{1}{\sqrt{s}} \exp(-a\sqrt{s}) \quad a \geq 0$ と, 置いて計算してきたので, 上をこれに代入すると以下を得る.

$$L\left\{\frac{1}{\sqrt{\pi t}} \exp\left(-\frac{a^2}{4t}\right)\right\}(s) = \int_0^\infty \frac{1}{\sqrt{\pi t}} \exp\left(-\frac{a^2}{4t}\right) \exp(-st) dt = \frac{1}{\sqrt{s}} \exp(-a\sqrt{s}) \quad a \geq 0, s > 0$$

この式を用いると, 本文中で使用した以下の式変形を得る.

$$\int_0^\infty \exp\left(-\frac{1}{2}(a^2 u^2 + b^2 u^{-2})\right) du = \left(\frac{\pi}{2a^2}\right)^{\frac{1}{2}} \exp(-|ab|) \quad a, b > 0$$

この変形の詳細は以下の通り. ただし, $a = \sqrt{2s}, b = \frac{v}{\sqrt{2}}, u^2 = t$

$$\begin{aligned} \int_0^\infty \exp\left(-\frac{1}{2}(a^2 u^2 + b^2 u^{-2})\right) du &= \frac{1}{2u} \int_0^\infty \exp\left(-\frac{1}{2}\left(2st + \frac{v^2}{2} \frac{1}{t}\right)\right) dt \\ &= \frac{1}{2\sqrt{t}} \int_0^\infty \exp\left(-\frac{v^2}{4} \frac{1}{t}\right) \exp(-st) dt \\ &= \frac{1}{2} \sqrt{\frac{\pi}{s}} \exp(-v\sqrt{s}) \\ &= \frac{\sqrt{\pi}}{\sqrt{2a}} \exp(-ab) \\ &= \left(\frac{\pi}{2a^2}\right)^{\frac{1}{2}} \exp(-ab) \quad a, b > 0 \end{aligned}$$

参考文献

- [1] D. F. Andrews and C. L. Mallow. [1974] Scale Mixtures of Normal Distributions. J. R. Statist. , Soc. B36, 99-122, 1974.
- [2] Easley, D. , N. M. Kiefer, M. O'Hara, J. B. Paperman. [1996] Liquidity, Information, and Infrequently Traded Stocks. The Journal of Finance, Vol 51, Issue 4 (Sep. , 1996), pp1405-1436.
- [3] Robert F. Engle [2000], The econometrics of ultra-high-frequency data Econometrica, Vol.68, No.1, pp.122
- [4] Evans. M. D. , [2002] FX Trading and Exchange Rate Dynamics The Journal of Finance Vol.57, No.6 (Dec. , 2002), pp.2405-2447
- [5] A. G. Hawkes. [1971] Spectra of some self-exciting and mutually exciting point processes Biometrika, 58: 83?90, 1971

- [6] A. G. Hawkes and D. Oakes. [1974] A cluster process representation of a self-exciting point process. *J. Appl. Probab.* , 11: 493-503, 1974.
- [7] Ogata, Y. [1999] Seismicity analysis through point-process modeling A review, *Pure and Applied Geophysics*, Vol.155, 471-507.
- [8] Naoya Sazuka[2008] On the gap between an empirical distribution and an exponential distribution of waiting times for price changes in a financial market February 2, 2008
- [9] Ritei Shibata[2006] Modelling FX new bid prices as a clustered marked point process *COMP-STAT 2006 - Proceedings in Computational Statistics: 17th Symposium Held in Rome, Italy, Physica*, pp.1565-1572
- [10] 井上純一 [2008] 相互作用するミクロなトレーダ群のマクロな振る舞い *日本神経回路学会誌 The Brain & neural networks* 15(4), 272-288, 2008-12-05
- [11] 加藤俊康・丸茂幸平 [1999] 市場取引発生タイミングに関する実証分析 *IMES Discussion Paper*, No.99, J-27 日本銀行金融研究所
- [12] 杉原慶彦 [2011] 執行戦略と取引コストに関する研究の進展 *IMES Discussion Paper*, No.2011-J-12 日本銀行金融研究所
- [13] 松永健太 山田健太 高安秀樹 高安美佐子 [2012] スプレッドディーラーモデルの構築とその応用 *ファイナンスにおける人工知能応用 人工知能学会論文誌* 27(6), 365-375, 2012

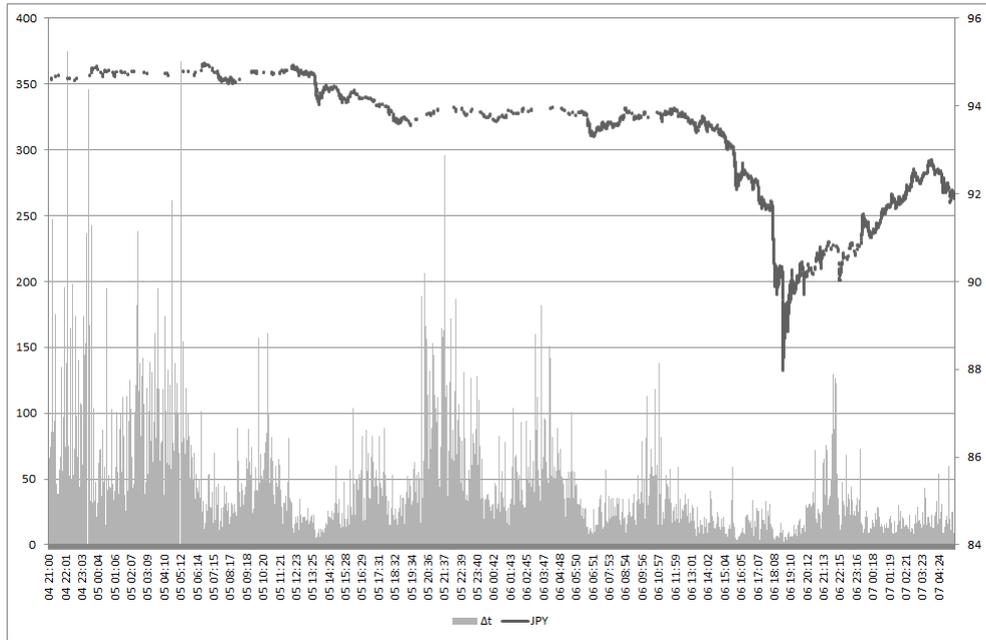


図 1: USD/JPY とその取引発生間隔 $\{\Delta t_i\}$ の時系列推移：図は実時間に沿った USD/JPY の推移（破線）(縦軸右:円) と共に, 取引約定間隔 $\{\Delta t_i\}$ (縦軸左:秒数) の推移を示したもの

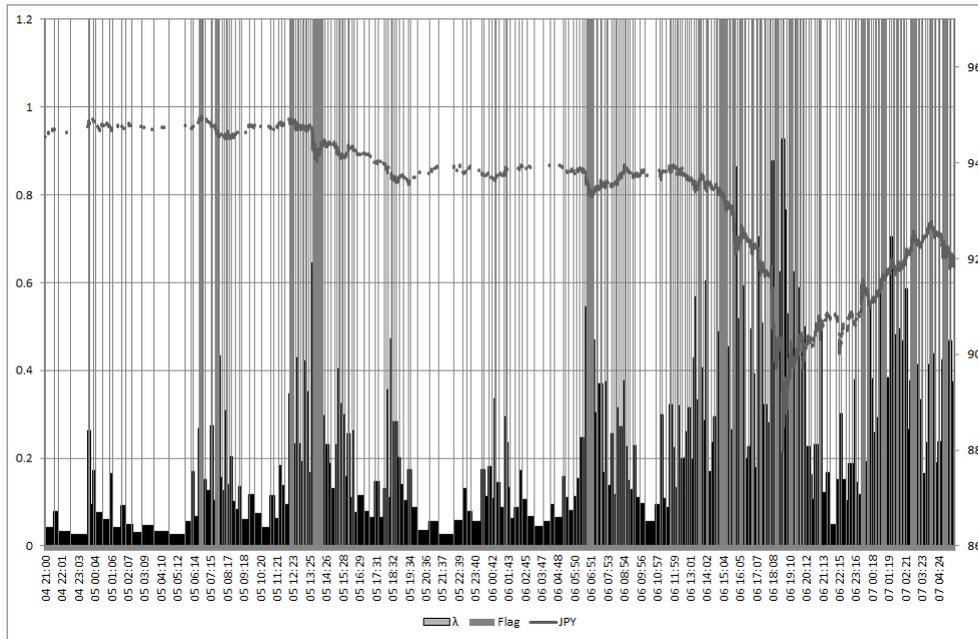


図 2: Shibata[2006] のモデルを用いたクラスタリング, USD/JPY とパラメータ λ の時系列推移: 図は実時間に沿った USD/JPY の推移 (破線) (縦軸右:円) と共に, Shibata[2006] のモデルを用いたクラスタリング (縦線) 及び当該クラスタ内での強度 λ (黒棒) (縦軸左: 1 秒当たりの約定発生件数) の推移を示したもの. 取引頻度が増す強度 λ が大きくなる時間帯に, 価格が比較的良く動くことが観察できる

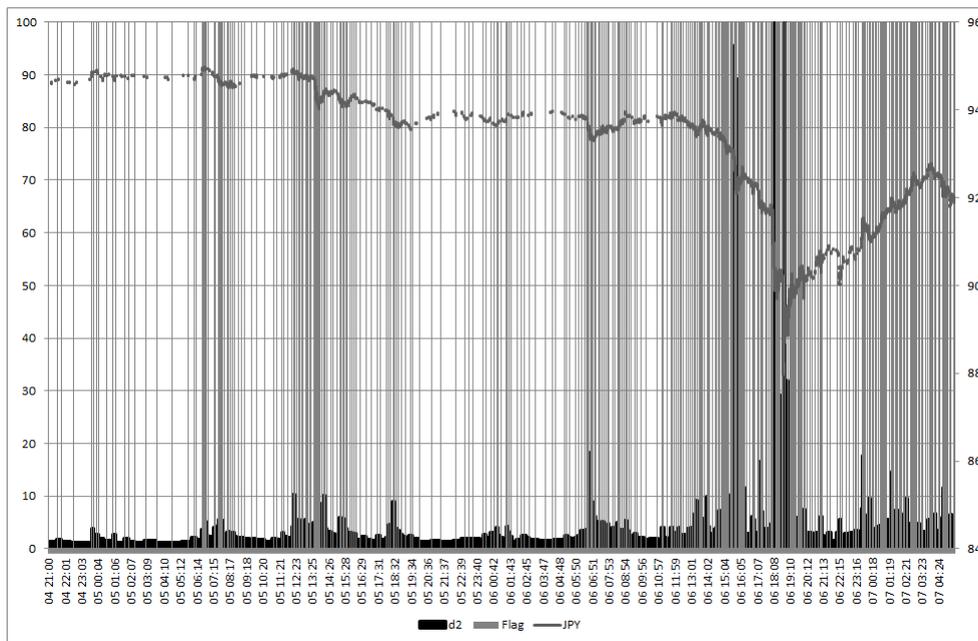


図 3: 本稿のモデルを用いたクラスタリング, USD/JPY とパラメータ d_2 の時系列推移: 図は実時間に沿った USD/JPY の推移 (破線) (縦軸右:円) と共に, 本稿のモデルを用いたクラスタリング (縦線) 及び当該クラス内でのパラメータ d_2 (黒棒) (縦軸左:t 分布の自由度) の推移を示したものの. 取引頻度が増すパラメータ d_2 が大きくなる時間帯に, 価格が比較的良く動くことが観察できる

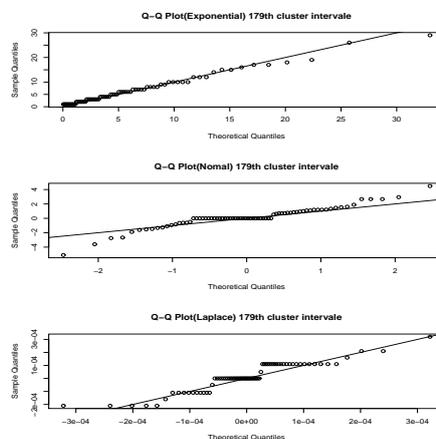


図 4: 約定頻度が低い時間帯のデータ ($\lambda = 0.15$): Shibata[2006] モデルの 16920 番目から 16993 番目の約定が対応するクラスで QQ プロットを調査したもの. 上から順に, $\{\Delta t_i\}$ を指数分布でモデル化した場合の QQ プロット, $\{\frac{x_i}{\sqrt{\Delta t_i}}\}$ を正規分布でモデル化した場合の QQ プロット, $\{x_i\}$ をラプラス分布でモデル化した場合の QQ プロット

表 1: 指数分布に基づいたクラスタへの区分. 可変パラメータ γ 別の KS 検定の適合の割合の比較

パラメータ γ	Δ_{t_i}	$x_i/\sqrt{\Delta_{t_i}}$	x_i
0.01	0.378	0.144	0.105
0.02	0.377	0.160	0.130
0.03	0.377	0.169	0.136
0.04	0.385	0.187	0.144
0.05	0.378	0.195	0.141
0.06	0.370	0.210	0.156
0.07	0.372	0.207	0.151
0.08	0.369	0.210	0.149
0.09	0.369	0.209	0.150
0.10	0.369	0.212	0.154
0.11	0.366	0.210	0.155
0.12	0.375	0.215	0.154
0.13	0.377	0.218	0.152
0.14	0.377	0.224	0.153
0.15	0.379	0.236	0.173
0.16	0.383	0.243	0.178
0.17	0.383	0.254	0.184
0.18	0.390	0.253	0.183
0.19	0.390	0.257	0.188
0.20	0.388	0.259	0.190
0.21	0.393	0.268	0.194
0.22	0.392	0.270	0.196
0.23	0.392	0.270	0.196
0.24	0.393	0.275	0.205
0.25	0.392	0.275	0.202
0.26	0.391	0.279	0.207
0.27	0.390	0.276	0.206
0.28	0.388	0.280	0.205
0.29	0.389	0.293	0.214
0.30	0.384	0.293	0.211

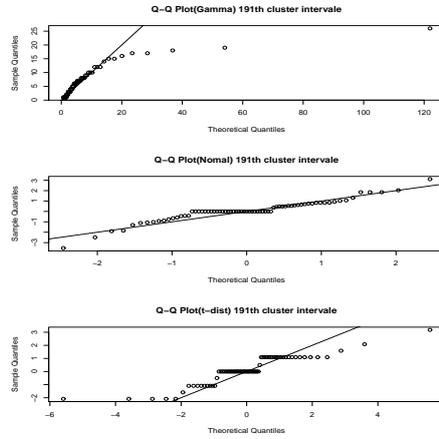


図 5: 約定頻度が低い時間帯のデータ: 本稿のモデルの 16919 番目から 16989 番目の約定が対応するクラスタでの QQ プロットを調査したもの. 上から順に, $\{\Delta t_i\}$ を逆ガンマ分布でモデル化した場合の QQ プロット, $\{\frac{x_i}{\sqrt{\Delta t_i}}\}$ を正規分布でモデル化した場合の QQ プロット, $\{x_i\}$ を t 分布でモデル化した場合の QQ プロット

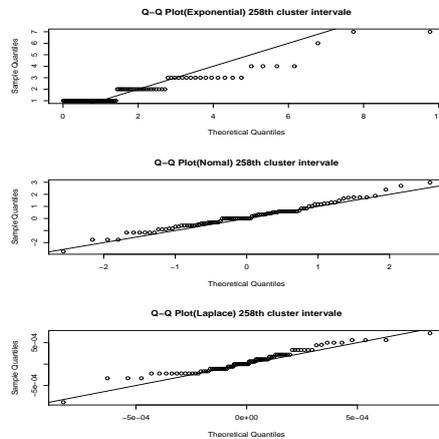


図 6: 約定頻度が高い時間帯のデータ ($\lambda = 0.54$): Shibata[2006] モデルの 26649 番目から 26745 番目の約定が対応するクラスタでの QQ プロットを調査したもの. 上から順に $\{\Delta t_i\}$ を指数分布でモデル化した場合の QQ プロット, $\{\frac{x_i}{\sqrt{\Delta t_i}}\}$ を正規分布でモデル化した場合の QQ プロット, $\{x_i\}$ をラプラス分布でモデル化した場合の QQ プロット

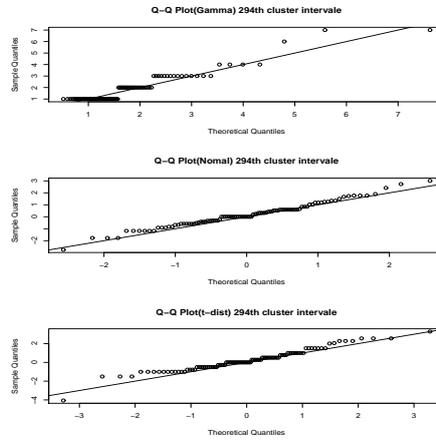


図 7: 約定頻度が高い時間帯のデータ: 本稿のモデルの 26648 番目から 26745 番目の約定が対応するクラスタでの QQ プロットを調査したもの. 上から順に $\{\Delta t_i\}$ を逆ガンマ分布でモデル化した場合の QQ プロット, $\{\frac{x_i}{\sqrt{\Delta t_i}}\}$ を正規分布でモデル化した場合の QQ プロット, $\{x_i\}$ を t 分布でモデル化した場合の QQ プロット

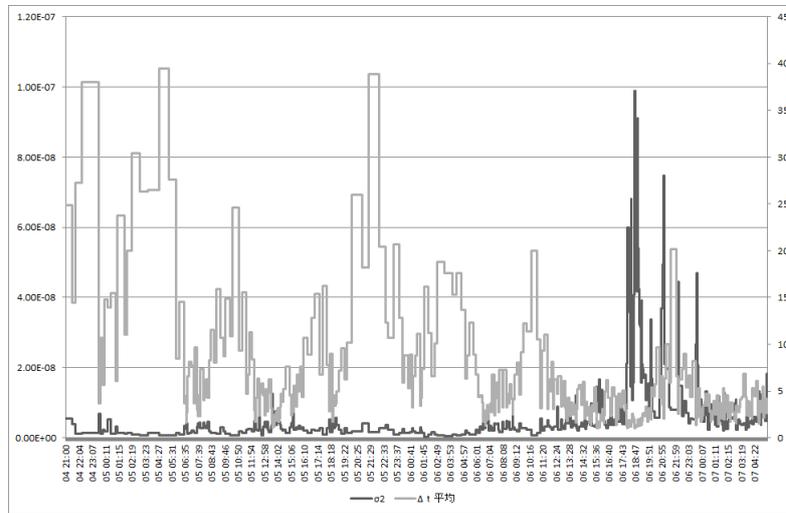


図 8: 本稿のモデルでのボラティリティとクラスタ内の平均約定間隔の時系列推移

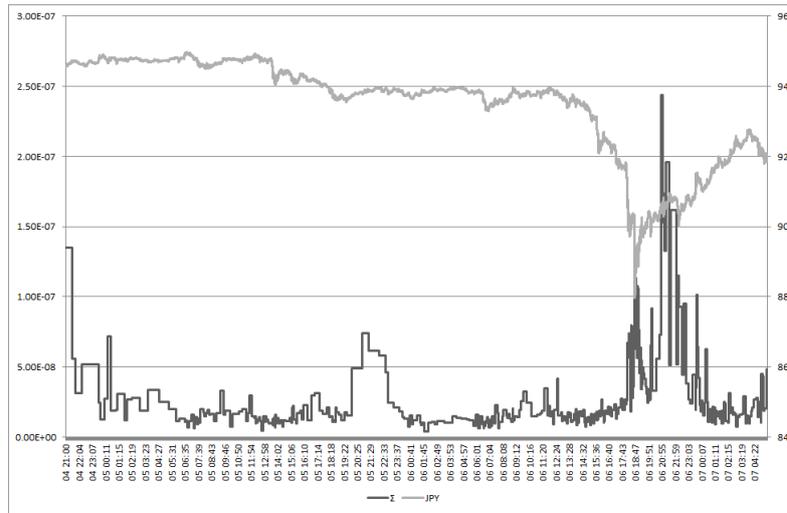


図 9: 本稿のモデルでのトータルボラティリティと為替レートの時系列推移

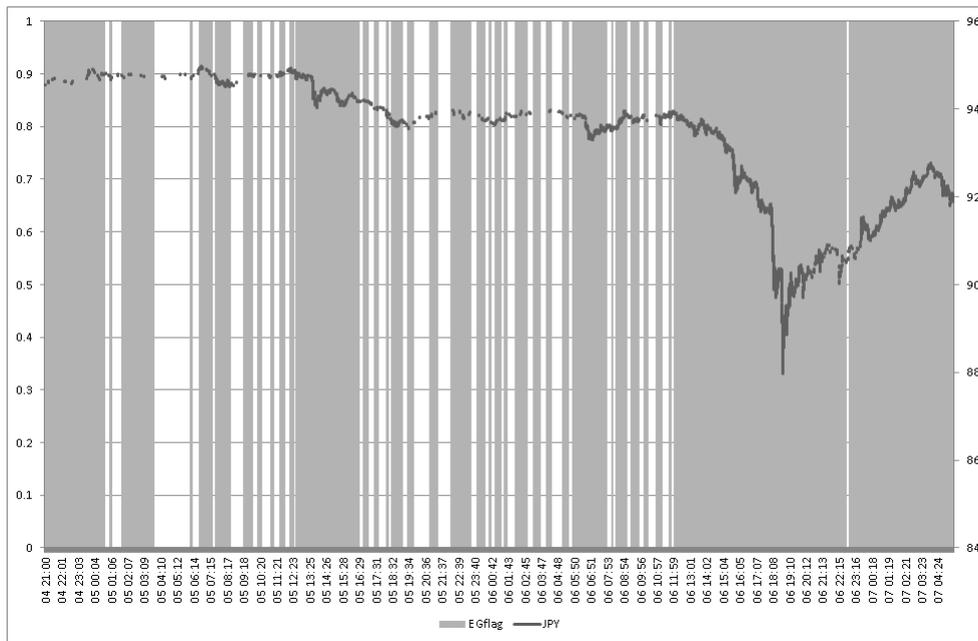


図 10: 指数分布と逆ガンマ分布を用いたクラスタリング (2005/10/4 GMT21:00:00-2005/10/7 GMT05:25:49)。レートと分布の推移。取引発生間隔の疎なところでは指数分布(無色), 取引発生間隔が密なところでは逆ガンマ分布(有色)が当てはまっていることがわかる。

表 2: Sibata[2006] モデルと本研究モデルの比較

	Sibata[2006]	本研究
クラスタ内での取引発生強度の分布	一定	ガンマ分布
取引発生間隔	指数分布	逆ガンマ分布
対数価格差を取引発生間隔で基準化した系列	正規分布	正規分布
対数価格差の系列	ラプラス分布	t 分布
クラスタ内での取引発生	定常ポアソン分布	負の2項分布