

Department of Social Systems and Management

Discussion Paper Series

No. 1214

流動性を予測できるローインテリジェンスモデル

by

木村博道、秋山英三

August 2008

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

流動性を予測できるローインテリジェンスモデル*

木村 博道

秋山 英三

概要

伝統的な経済学では、投資家は完全合理的な存在であることを仮定してきた（一般均衡理論）。そのようなモデルでは均衡価格の存在などについて述べることはできても、実務家にとって重要な、アスクとビッドの差（スプレッド）や1単位の注文に対する平均的な価格変動（Kyleの λ ）について予測することはできなかった。一方、近年、市場環境をまったく考慮せず一様・ランダムに注文が到着するというモデル（ゼロインテリジェンスモデル：ZIM）によって、スプレッド、価格変動の激しさ（ボラティリティ）を予測できることが分かってきた。しかしZIMによってもKyleの λ の予測はできなかった。そこで本研究では、注文到着率が板上の指値注文量に依存するなどの注文到着パターンの性質を実際のデータを用いて解明しZIMに導入した。このモデルをローインテリジェンスモデルと呼ぶことにする。すると、スプレッド、ボラティリティに加え、一般均衡理論やZIMでは難しかったKyleの λ の予測が可能となることが分かった。このことは、ローインテリジェンスモデルで考慮した注文到着パターンを利用することにより流動性コストを抑える新たな方法を提案できる可能性を示唆する。

1 はじめに

1.1 市場流動性と本研究の目標

本研究の目標は、注文到着のパターンを分析することで、各銘柄の流動性を予測できるモデルである「ローインテリジェンスモデル」を構築することである。なお、対象銘柄は東京証券取引所（東証）1, 2部で取引されている約760銘柄とした。詳細は後述することとし、ここではまず本目標の重要性について述べることにする。

流動性を予測することは、特に実務家にとって重要である。というのは、即時往復売買のコストは流動性の指標の一つである「スプレッド（アスクとビッドの差）」によって見積もることができ、このコストは特に取引頻度の高い投資家にとっては収益に直接関わるからである。また大口の成行注文を行う投資家は自分の注文行動によりアスク（ビッド）がどのぐらい動くのかを常に気にしつつ成行注文を行っている。「マーケットインパクト」は n 単位の成行注文によって平均的にどのぐらい価格が動くかを意味する関数で

*岸本一男先生ならびに Shiwei Zuo 氏からは、日本の株式市場のデータと貴重な御助言を頂きました。この場をお借りしてお礼を申し上げます。

あり、流動性指標である「Kyleの λ 」を一般化したものである。ここでKyleの λ とは、仮に1単位の成行注文が到着したとするとそれによって平均的にどのくらい価格が動くかを意味する指標 (Kyle, 1985; 大村・宇野, 1998) である。大口の注文を行う投資家にとって、マーケットインパクトは自身の行動が市場に与える影響を表現しているという意味で重要な関数である。

このように流動性は重要な概念であり、その先行研究は多岐にわたる。例えば各銘柄を時価総額の大小によってグループに分けた後、株価と売買高の関係を回帰分析したり (大村・宇野, 1998)、主成分分析を用いて新しい流動性指標の妥当性を検証したり (木村・秋山, 2007)、流動性指標に対して主成分分析を行い何が流動性を決めているのかを調べたり (Flemming, 2003) といった研究が挙げられる。以上のように、多数の先行研究が変数間の線形性を仮定し、回帰分析やそれに準ずる手法を用いている。しかし、実際の市場では非線形な関係も当然あるものと思われる。例えば、投資家の行動と流動性の関係はその例である (詳細は3.2節で触れる)。本研究ではそのような非線形な関係を考慮したモデルを提案する。

1.2 流動性の予測とゼロインテリジェンスモデル

流動性は指値注文により供給され、成行注文により消費される。したがって流動性がどのように決まるかを調べるには、投資家がどのような状況で指値・成行注文を行うのかを調べモデルを作るという方法が直接的である。つまり投資家の平均的な注文行動を調べモデルに組み込むのである。では投資家行動と市場の関係についての先行研究にはどのようなものがあるだろうか？

例えば、伝統的な経済学 (一般均衡理論) では、投資家が合理的であることを仮定し効用と予算制約を用いて投資家行動を説明してきた。この理論は、消費者の効用関数や生産者の生産関数の具体的な形状を仮定することなく、均衡の存在性・一意性を議論することができる。その反面、流動性を予測するのは難しい。その理由の一つとしては、一般均衡理論ではただ1回の取引が均衡価格で行われて終了してしまい、価格ゆらぎが無い世界を考えているからということが挙げられる。

ところが、近年、スプレッドを定量的に予測できるモデルが登場した (Smith et al., 2003)。このモデルは、投資家が合理的であることを仮定する一般均衡理論とは正反対に、投資家はあたかも何も考えていないかのように振る舞い、注文はランダム・独立に到着するということを仮定する。このモデルを「ゼロインテリジェンスモデル」と呼ぶ。特に、Smithらのゼロインテリジェンスモデルでは投資家の合理性を仮定しないにもかかわらず、スプレッドなどの量を予測できることが分かっている (Farmer et al., 2005a)。ここで「予測」とは未来の予測ではなく同時刻の予測を指す。つまりスプレッドの平均値をモデルを用いて計算し、現実の市場のスプレッドの平均値と比較するということを意味する。本研究では Farmer et al. (2005a) にならい、この意味で予測という言葉を用いる。

ゼロインテリジェンスモデルに関する研究は多岐に渡る。その最初期の研究として挙げられるのは Becker (1962) である。通常の経済学では需要供給曲線の存在を導く際、投資家は自分の予算内で効用を最大化するように行動すると仮定してきた。ところが、Beckerは、予算制約は考慮するが効用最大化は目指さない投資家 (すなわち、ゼロインテリジェンストレーダー) から成る市場でも需要供給曲線を導けることを解析的に示した。

また、Gode and Sunder (1993) は、被験者のみからなる市場、予算制約のあるゼロインテリジェンストレーダーのみからなる市場、予算制約のないゼロインテリジェンストレーダーのみからなる市場の三者を比較し、各々の市場で均衡に到達するかどうか、余剰は最大化されているかどうかを検証した。ゼロインテリジェンストレーダーはコンピュータエージェントとして実装された。その結果、前二者で価格は市場均衡に到達し、余剰はほぼ最大になった。このことは、市場均衡を考える際には投資家の効用最大化を仮定する必要は無いということを示唆する。人々が自分の欲望に従って行動する（効用最大化）と社会全体がよりよい状態になるという考え方が「見えざる手」である。言い換えると、「見えざる手」が均衡へ向かわせる力は、経済学者が考えていたよりもかなり強く、予算制約だけを考慮すれば十分説明できるのかもしれないということが Gode and Sunder (1993) の研究で示されたのである。

これら 2 つの研究は経済学の基礎に関わる。両研究は、本研究と直接は関係が無いものの、効用最大化のような仮定が無くてもある程度市場で観察される現象を説明することができるのではないかという期待を抱かせてくれる。つまりゼロインテリジェンスモデルの可能性を示していると言える。

一方、現実的な市場制度とゼロインテリジェンスを仮定して、価格変動の時系列的特徴—例えばファットテイル (Mantegna and Stanley, 1999; Sornette, 2004)—を再現できるかどうかという立場の研究もある (Maslov, 2000; Slanina, 2001)。Maslov (2000) のモデルは注文駆動方式¹により価格決定が行われるようなゼロインテリジェンスモデルである。そのモデルでは売り指値、買い指値、売り成行、買い成行のいずれが到着するかは 1/4 の確率で決まるものとしている。指値注文の指値は、売り注文の場合、市場価格 + Δ ($\Delta = 1, 2, 3, 4$) の値をとる可能性がある。 Δ の値は 1/4 の確率で一様に決まるとする。一注文の注文数量は、売り買い、指値成行に関わらず、常に単位数量とする。また、一定時間待って執行されなかった注文はキャンセルされるものとする。以上の設定の下で、Maslov は、価格変化率はベキ分布に従うということを示シミュレーションによって示した。

また、ゼロインテリジェンスモデルは簡単なモデルであるため、解析的な扱いが可能なお場合もある。例えば、Slanina (2001) は、指値注文が必ず現在の市場価格から一定の距離の位置に置かれる、キャンセルは生じないなどの条件の下で、価格変化率の密度関数はベキ分布に従い、その指数は 2 となることを示した。

このようにゼロインテリジェンス、つまり投資家の意志決定を考慮しないモデルであっても、市場現象をかなりの程度説明できることが分かっている。上述のモデル以外にも多種多様なゼロインテリジェンスモデルが提案されている (e.g. Mendelson, 1982; Tang and Tian, 1999)。しかし、定量的な予測を出せて、しかも現実の市場から得られるデータによく一致するモデルはまだ少ない。そのような数少ないモデルの一つが、Smith et al. (2003) のゼロインテリジェンスモデルである。

Smith らは「価格決定には注文駆動方式を用いる。」「売り（買い）指値注文はビッド（アスク）より高い（安い）半無限区間全体に一様に到着する。単位時間あたり単位価格あたり指値注文到着率を α とする。」「成行注文はランダム・独立に到着する。単位時間あたり成行注文到着率を μ とする。」「指値注文のキャンセルはランダム・独立に発生する。指値注文がキャンセルされるまでの平均寿命の逆数を δ_s ² とする。」「一注文あたり

¹なお注文駆動方式によってどのように価格が決まるのかについては付録 A にまとめた。

²Smith らは平均寿命の逆数を δ としたが、本研究では δ_s とする。本研究では Smith らのモデルをモンテ

株数は σ 株で一定とする。」などを仮定した。その結果、スプレッドとボラティリティを上述の 4 変数の関数として具体的に表現することができた。

さらに、Farmer et al. (2005a) は、ロンドン証券取引所 (LSE) の 11 銘柄を対象にスプレッドとボラティリティを計測し、Smith らのゼロインテリジェンスモデルの予測値と実測値を比較した。その結果、スプレッドに対する、モデルの予測値と実測値を回帰したときの決定係数は 0.96 と非常に高く、よい精度でスプレッドの予測を行えることが明らかになった。またボラティリティに対する決定係数は 0.76 となり、これもまたよい精度で予測できることが分かった。

スプレッドは流動性の、簡単で代表的な指標の一つである。ボラティリティは値動きの重さを意味し、市場の安定性の一つの指標となる。またオプションプライシングにも利用され重要度は高い。ボラティリティは流動性の指標とはされていないが、本研究でも触れる。Smith らのゼロインテリジェンスモデルによって、板の状態を示す 4 つのパラメータを用いてスプレッド、ボラティリティという市場の流動性・安定性を示す量を予測できるようになった。Smith らのゼロインテリジェンスモデルは本研究の元となるモデルである。以下では、Smith らによって提案されたゼロインテリジェンスモデルを Smith らのモデルと呼ぶことにする。

Smith らのモデルで予測できない量もある。それは「マーケットインパクト」である。マーケットインパクトとは仮に $n (\geq 1)$ 単位の成行注文が到着したときに平均的にどのくらい価格が動くかを意味する。投資家が成行注文を行うことでアスク・ビッドは動いてしまうが、その値幅、すなわちマーケットインパクトを最小にすることを実務家は求めている (日経金融新聞, 2006; Kim, 2007)。したがって、マーケットインパクトの予測は急務である。本研究では $n = 1$ のときのマーケットインパクト、すなわち Kyle の λ の予測を目指す³。

では、マーケットインパクトを予測できるようなモデルはどのようなものだろうか。

1.3 ローインテリジェンスモデル

現実の投資家は完全に合理的ではないが全く何も考えていないというわけでもないから、マーケットインパクトを予測できるようなモデルを構築するにはある程度のインテリジェンスを導入するという方向が考えられる。例えば 3.2.1, 3.2.2 節で触れるようにデータから確認できることであるが、投資家は「厚み」の大きさによって行動を変えている。ここで厚みとはアスク・ビッドに存在する指値注文量のことであり、特定の呼値にかかる指値注文量については例えば「セカンドアスクの厚み」のように呼ぶことにする。単に「厚み」と言った場合にはアスク・ビッドの厚みを指す。また投資家は売り (買い) 成行注文を連続して行いやすいことが分かっている。Bouchaud et al. (2004) はパリ証券取引所のデータを用いて買い (売り) 成行注文が到着したら $+1$ ($+1$) のように符号をつけたときに、この符号は長期相関があることを示している。また Lillo and Farmer (2004) は一般に成行注文に限らず、買い板または売り板側のどちらで指値・成

カルロシミュレーションにより検証するが、その際必要なパラメータは「指値注文がキャンセルされるまでの平均寿命の逆数 δ_s 」ではなく「ある指値注文の単位時間あたりキャンセル率 δ_c 」だからである。両者の違いについては付録 C.3 を参照。

³なお $n > 1$ のときのマーケットインパクトはローインテリジェンスモデルではあまりうまく予測できないことも分かった。

行注文，キャンセルが生じやすいかについて長期相関があることを報告している．本研究においては，成行注文だけではあるがこのことを確認した（3.2.1 節）．さらに木村・秋山（2007）によると，スプレッドが3 単位以上開いているときに投資家が売り（買い）指値注文をアスクとビッドの間に行う場合，アスク（ビッド）に近い呼値に注文する傾向がある（3.2.2 節）．

そこで本研究では以下のインテリジェンスを導入した．

- 売り（買い）成行注文とアスク（ビッド）の厚みの関係．
- 連続して到着する売り（買い）成行注文．
- アスク・ビッドからの相対位置に依存する指値注文到着率．
- アスク（ビッド）に到着する指値注文と，アスク（ビッド）の厚みとセカンドアスク（セカンドビッド）の厚みとの差の関係．
- 少しずつ縮小するスプレッド．

このモデルを「ローインテリジェンスモデル」と呼ぶことにする．その結果，スプレッド，ボラティリティに加え，Kyle の λ の予測が可能となることが分かった．この予測は一般均衡理論でもゼロインテリジェンスモデルでも不可能である．このことから，市場流動性を予測するには，一般均衡理論が想定するような完全合理的な投資家を考える必要は無く，その一方で，逆方向のアプローチであるゼロインテリジェンスモデルが想定するような完全にランダムに注文を行うような投資家では不十分であり，本研究のように注文到着パターンを詳しく調べる必要があることが示唆される．

1.4 本論文の構成

本研究では，スプレッドとボラティリティ，Kyle の λ の予測を目指すことにする．

買い（売り）成行注文株数を n ，その注文によってアスク（ビッド）が平均的にどのくらい動くかを Δp で表すと，マーケットインパクトは $\Delta p = \lambda n^\nu$ の形で書くのが一般的である（e.g. Gabaix et al. , 2003; Lillo and Farmer , 2004）．素朴に n に比例して価格が動くと考えれば $\nu = 1$ である． $n = 1$ のときのマーケットインパクト，すなわち λ を Kyle の λ という．以下では単に λ と呼ぶことにする．

本論文は以下のように構成される．まず，2 節では Smith らのモデルのパラメータの説明を行い，そのモデルを東証で検証する．検証はモンテカルロシミュレーションによりスプレッドなどを求め，各銘柄ごとに実測値と比較した．この，Smith らのモデルを検証するためのモデルを「モンテカルロゼロインテリジェンスモデル」と呼ぶことにする．その結果，スプレッド・ボラティリティはよく説明できた（決定係数はそれぞれ 0.94，0.88）が， λ はあまり説明できなかった（同 0.55），つまり Farmer らの結論はわずか 11 銘柄に対する検証であったが，それと同様の結果を東証の約 760 銘柄でも確認できたことになる．

次に 3 節ではローインテリジェンスモデルの説明を行い，そのモデルを東証で検証する．その結果，スプレッドに加え，マーケットインパクトもある程度の精度で予測できるようになった．実際，スプレッド，ボラティリティの決定係数はそれぞれ，0.93，0.87 となりモンテカルロゼロインテリジェンスモデルとほぼ同程度の予測力を持っていること，また λ の決定係数は 0.81 となり同モデルの予測力を上回っていることが分かった．

2 Smithらのゼロインテリジェンスモデルの東証への適用

Farmer et al. (2005a) らは Smith らのモデルをロンドン証券取引所で取引されている銘柄に適用し、スプレッド、ボラティリティについては高い予測力を持っているが入についてはあまり予測できないという結果を得た。本節では Smith らのモデルを東証の各銘柄に適用し、同様の結論が東証においても成り立つことを確認する。

2.1 手法

2.1.1 Smith らのゼロインテリジェンスモデルの詳細

Smith らは注文駆動方式による価格決定を用いて注文がランダムに到着するという仮定から、スプレッド、ボラティリティなどを求めることができることを示した。これから具体的にどのようなモデルであるかを解説する。

まず Smith らのモデルでは次のような仮定が置かれている。

1. モデル上の価格は実数値をとり注文価格は任意の実数値をとることが可能である。これは現実の市場価格の自然対数に相当する。したがって全く同じ呼値に注文が複数ある可能性は実質的になく、時間優先原則は考えなくてよい⁴。
2. 売り（買い）指値注文はビッド（アスク）より高い（安い）半無限区間全体に一律に到着する（図 1）。つまり、単位価格幅あたり単位時間あたり到着する指値注文量は一定の値をとり、それを α とおく。
3. 成行注文は一定の到着率で到着する。つまり、単位時間あたり到着する成行注文量は一定の値をとり、それを μ とおく⁵。
4. 注文が売り注文になるか買い注文になるかは、指値成行に関わらず $1/2$ の確率で決まる。つまり、単位時間あたり到着する売り（買い）成行注文量は $\mu/2$ である。
5. 一注文は、指値・成行に関わらず一定の株数 σ からなる。
6. キャンセルは一定の割合でランダム・独立に行われる。指値注文がキャンセルされるまでの平均寿命の逆数を δ_s とする。

時間は、物理的な時間ではなく、成行注文が到着するたびに $+1$ することにより計る。これを約定タイム⁶と呼ぶことにする。

以上の設定の下で、Smith らはどのように市場を扱ったかについて詳しく述べることにする。一般にある量の間を知りたい時、その単位に注目するとその関係を支配

⁴Smith らは呼値の刻み幅が有限である場合も考えているがほとんど結果には影響しない。

⁵アスク（ビッド）よりも高く買いたい（安く売りたい）という指値注文を実効的成行注文といい、厳密には成行注文とは異なる。しかし、データから両者を区別することは難しく、先行研究 Smith et al. (2003); Farmer et al. (2005a) でも区別していない。そこで、本研究でも区別せずにまとめて成行注文と呼ぶことにする。

⁶Farmer et al. (2005a) では指値注文が到着したりキャンセルが生じたりしたときも入れて時間を計っている（イベントタイム）。本研究でイベントタイムを用いないのは、モンテカルロシミュレーションを行うには約定タイムのほうが簡単だからである。

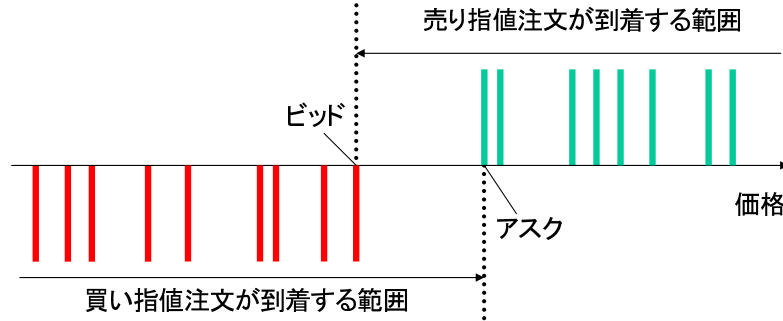


図 1: Smith らのモデルを前提としたときの、板状況の例．横軸が価格を意味する．価格は実数値をとり注文株数は一定値であるため、板はこの図のようになる．

表 1: Smith らのモデルのパラメーター一覧

変数	意味	次元
α	単位時間・単位価格幅あたりの指値注文到着量	株数/価格/時間
δ_s	指値注文がキャンセルされるまでの平均寿命の逆数	1/時間
μ	単位時間あたりの成行注文到着量	株数/時間
σ	一注文あたり注文株式数	株数

する方程式を扱いやすくなることがある．この手法を次元解析という．次元とは単位のことである．次元解析を以上の仮定の下で市場に適用すると次のようになる．

パラメータ $\alpha, \delta_s, \mu, \sigma$ を構成する次元は価格，株数，時間の 3 種類ある．各パラメータの次元を表 1 にまとめた．

この 4 つのパラメータを用いると、価格の次元を持つ量は μ/α ，株数の次元を持つ量は μ/δ_s と σ ，時間の次元を持つ量は $1/\delta_s$ と σ/μ しかない．また最も簡単な無次元量は $\delta_s\sigma/\mu$ である．以下， μ/α を p_c ， $\delta_s\sigma/\mu$ を ϵ と書く．

市場の平均スプレッド S とボラティリティ D は上記のパラメータ $\alpha, \delta_s, \mu, \sigma$ の関数として表現することができる．スプレッド S は価格の次元を持つので、無次元の関数 f を用いて、 $p_c f(\epsilon)$ の形をしているとしてよい．また、ボラティリティ D は価格²/時間の次元を持つので、無次元の関数 h を用いて、 $p_c^2 \delta_s h(\epsilon)$ の形をしているとしてよい． f, h の具体的な形は数値計算により求めることができる．その結果，Smith らは $\epsilon \gtrsim 0.05$ の範囲で

$$S = \frac{\mu}{\alpha} (0.28 + 1.86\epsilon^{0.75})$$

$$D_0 = \frac{\mu^2 \delta_s}{\alpha^2} (2\epsilon)^{-0.5}$$

$$D_{\infty} = \frac{\mu^2 \delta_s}{\alpha^2} (2\epsilon)^{0.5}$$

が成り立つことを示した．ここで Smith らのモデルではボラティリティは見ているタイムスケールに依存するが， D_0 の値はタイムスケールが非常に短い極限をとったときの， D_{∞} の値は非常に長い極限をとったときの値である．

なお，本研究の対象銘柄，対象期間の設定は付録 B に， α ， δ_s ， μ ， σ ，スプレッド，ボラティリティ， λ の推定法は付録 C にまとめた．次節では，これら 4 パラメータを用いてどのようにモンテカルロシミュレーションを行ったかについて述べる．

2.1.2 モンテカルロゼロインテリジェンスモデル

Smith らのモデルは解析的にスプレッドやボラティリティを求めることができるが，本研究ではモンテカルロシミュレーションによりスプレッド， λ ，ボラティリティを求めることにした．つまり，乱数を用いて注文を生成して取引を行わせ，データを実際の市場と比較した．これは，3 節で導入するローインテリジェンスモデルを解析的に扱うのは難しいため，両モデルを比較するにはモンテカルロシミュレーションを行ったほうが統一的に扱えるからである．

ここで注意すべき点は，モンテカルロゼロインテリジェンスモデルでは図 2 に示すように指値注文の到着範囲を有限の幅にせざるをえないという点である．Smith らのモデルでは，指値注文が到着する範囲は半無限区間，かつその到着率は単位時間・単位価格あたり一定を仮定している．モンテカルロシミュレーションを行うにあたってはこの 2 つの仮定は両立しない．なぜならこの 2 つを共に仮定してしまうと，単位時間に指値注文は無数個到着してしまうことになるからである．したがって指値注文が到着する範囲は有限であるとせざるをえない．

また，到着範囲を広く取りすぎると板上に指値注文が大量に存在することになり，シミュレーションが現実的な時間で終了しなくなってしまう．そこで本研究では売り（買い）指値注文はビッド（アスク）から w 離れた価格までの間に到着すると仮定する． w は付録 C.2 で述べる「幅」—これは，指値注文がどこに到着したかが，本研究で用いたデータから判定できる最大の範囲である—の 10 倍とする．

このような設定にしてしまうとアスク（ビッド）が存在しないことがありうる．つまり，特に指値注文があまり到着しないような銘柄の場合には，指値注文到着幅 w は有限なので単位時間に指値注文は有限個しか到着せず売り（買い）指値注文が無くなってしまい，アスク（ビッド）が定義できないことがありうる．そうした場合には，図 2 のようには指値注文の到着位置を決められない．そこでアスク・ビッドが無くなってしまったら，無くなる直前のアスク・ビッドを起点とし，そこから w だけ離れた位置までの間に指値注文が到着するものとする．

シミュレーションを行うにあたって，まず，初期の板，つまり最初の指値注文を用意する．本研究では，板が定常状態に速く到達するようにするために次のようにした．Smith らのモデルでは常に σ 株ずつ取引を行うということになっている．そしてアスク・ビッドから十分離れた領域には単位価格幅あたり α/δ_s 株の指値注文が存在することが分かっている．そこで初期売り板生成の第一段階として，単位価格幅あたり $\alpha/\sigma\delta_s$ 個の売り指値注文を置くことにする．つまり $N = \lceil \alpha w / \sigma \delta_s + 1 \rceil$ 個の売り指値注文を価格軸上の 0 から w まで一様・ランダムに並べる．+1 は，板上に最低一つの売り指値注文を置

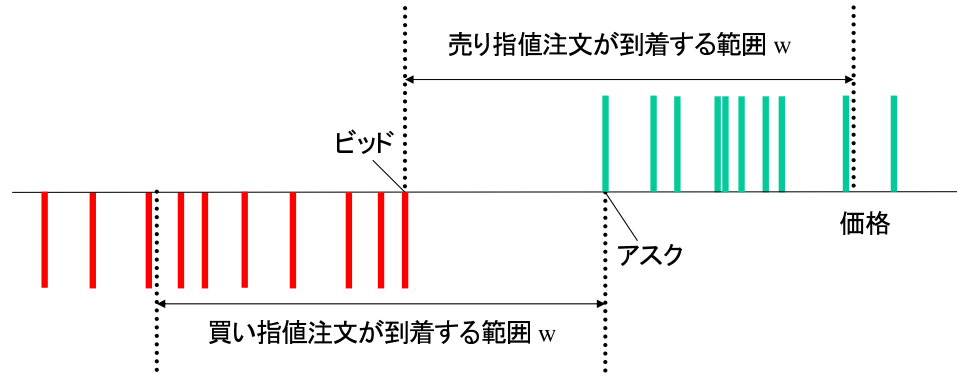


図 2: モンテカルロゼロインテリジェンスモデルでは指値注文はアスク・ビッドから一定の範囲 w 内で一様に到着すると仮定する．図 1 と比較せよ．

くためである．ここで $[\cdot]$ はガウス記号である．こうすると，単位価格幅あたり平均して $\alpha/\sigma\delta_s$ 個程度の売り指値注文を置くことができる．買い板も同様である．こうして生成された指値注文の分布は，シミュレーションの結果生成されるはずの分布よりも一様に近すぎてしまう．なぜならアスク・ビッド付近では成行注文の到着によってスプレッドが広がるからである．

そこで，初期板生成の第二段階として，成行注文が $[10/\epsilon + 1]$ 回到着してからスプレッド等の計算を始めることとする．この回数は Smith らのモデルを参考にして決めた．Smith らのモデルでは，成行注文が無かったとすると単位価格幅あたり平均して $\alpha/\sigma\delta_s$ 個程度の指値注文が存在することが分かっている．一方，成行注文が到着すると，アスク・ビッド付近の指値注文は成行注文によって消費されるため，アスク・ビッド付近の指値注文量は $\alpha/\sigma\delta_s$ 個よりも少なくなる．どのくらい少なくなるかは Smith らのモデルから計算できるため，その程度の回数の成行注文を到着させれば初期板が完成する．その回数は $1/\epsilon$ 回程度であるが，ここでは余裕をもって成行注文が $[10/\epsilon + 1]$ 回到着したら板がほぼ定常状態に到達したと見なすことにする．こうして初期板を用意することができた．

次に買い指値注文，売り指値注文，買い成行注文，売り成行注文，キャンセルの各イベントがどのようなタイミングで生じるのかを決める必要がある．つまり各イベントの生起率を求めることにする．イベント生起率が分かれば次のようにして各イベントが生じる確率も分かる．簡単に 2 種のイベント X_1, X_2 を考え，その単位時間あたり生起率はそれぞれ π_1, π_2 だったとする．イベント $X_i (i = 1, 2)$ が生じる確率 p_i は生起率 π_i に比例する．したがって比例定数を k とすると $p_i = k\pi_i$ である．そして必ずどちらかのイベントが生じるとすると $\pi_1 + \pi_2 = 1$ である．以上よりイベント $X_i (i = 1, 2)$ が生じる確率 p_i は $\pi_i/(\pi_1 + \pi_2)$ により求めることができる． $i > 2$ でも同様である．では各イベントの生起率を具体的に与えることにする．

まず指値注文について， α の定義から， σ 株の売り指値注文が bid から $\text{bid} + w$ までの範囲のどこかに到着するとしたときの，売り指値注文生起率は $\alpha w/\sigma$ である． σ 株の買い指値注文についても同様に， ask から $\text{ask} - w$ までの範囲のどこかに到着するとし

たときの、買い指値注文生起率は $\alpha w/\sigma$ である。

次に成行注文について、 μ の定義から、買い（売り）成行注文が起きる確率は、売り買い等確率としているので $\mu/2\sigma$ に比例する。本研究では約定タイムを採用したので $\mu = \sigma$ となり（付録 C.1 を参照。）、買い（売り）成行注文生起率は $1/2$ である。

最後にキャンセルについて、Smithらのモデルでキャンセルに関係するパラメータは δ_s である。これは、約定しなかった指値注文がキャンセルされるまでの平均寿命の逆数である。キャンセルをシミュレーションに実装するためには、ランダム・独立に、かつ平均寿命が δ_s^{-1} となるようにキャンセルが起きる必要がある。しかし、指値注文は成行注文の到着によっても消滅するため、単純にある一つの注文がキャンセルされるというイベントの生起率は δ_s^{-1} ではない。つまり δ_s というパラメータを用いてはシミュレーションを行いにくい。そこで単位時間あたりのある注文のキャンセル率 δ_c というパラメータを導入する。 δ_s と δ_c がどのくらい異なるかは付録 C.3 で論じる。板上に存在する全株数を n とすると、次に起こるイベントがキャンセルとなる確率は $n\delta_c$ に比例する。

以上の考察をまとめると、板の時間発展を決めるアルゴリズムは次のようになる。各イベントが生じる確率の分母を $P = 2\alpha w/\sigma + 1 + n\delta_c$ とおくと

1. 確率 $\alpha w/\sigma P$ で売り指値注文到着イベントが発生し、bid から bid + w の範囲に一樣・ランダムに到着する。
2. 確率 $\alpha w/\sigma P$ で買い指値注文到着イベントが発生し、ask から ask - w の範囲に一樣・ランダムに到着する。
3. 確率 $1/2P$ で売り成行注文到着イベントが発生する。
4. 確率 $1/2P$ で買い成行注文到着イベントが発生する。
5. 確率 $n\delta_c/P$ でキャンセルイベントが発生する。実際にどの注文がキャンセルされるかは、ランダムに決まる。

シミュレーションは成行注文が 10000 回到着したら終了するものとする。その後、スプレッド、ボラティリティ⁷、 λ を計算する。

図 3 はシミュレーションの流れを図示したものである。

シミュレーションが終わったら、実測値と予測値が一致したかどうかを見るために回帰分析の結果を出力する。Farmer et al. (2005b) によると、回帰係数の誤差は通常の最小二乗法ではなくブートストラップ法に基づいて評価すべきだという。なぜならパラメータが自己相関していたり、銘柄間で相関があったりするからである。しかしモンテカルロシミュレーションを用いて彼らの評価法を適用するのは難しい。その評価法を実行するには、対象期間を分割してその各期間内でパラメータを求める必要がある。すると特に流動性があまり高くない銘柄に対してはデータが足りなくなってしまうからである。そこで本研究では参考として通常の最小二乗法によって求めた標準偏差を示す。これはブートストラップ法により求めた誤差より小さくなる。

⁷本研究では D_0 に相当するボラティリティのみを対象とし、 D_∞ に相当するボラティリティは対象としない。理由は二つある。一つ目は連続したデータの大きさは最大でも 2 時間半（後場での取引）なのであまり大きくはできず、データ量も限られてしまうからである。二つ目はタイムスケールを大きくすると、市場外からの情報の効果が現れ始め投資家はゼロインテリジェンスから外れた行動を示すようになり誤差も大きくなるだろうと思われるからである。

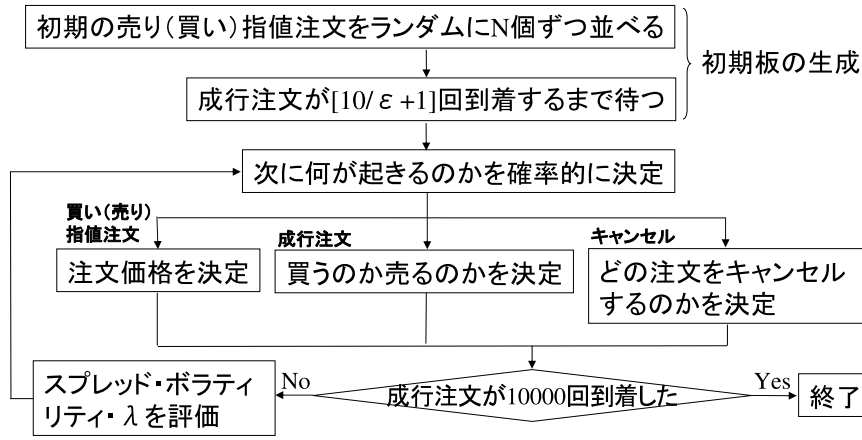


図 3: Smith らのモデルをモンテカルロシミュレーションにより検証する際の手順をフローチャート化した。

2.2 結果

本節では, Smith らのモデルのモンテカルロシミュレーションを行うことによって, モデルの予測が実際の市場から得られた値とどの程度一致しているのかを検証する. そのためにスプレッド, λ , ボラティリティについては $\log(\text{actual values}) = A \times \log(\text{predicted values}) + B$ という形の回帰分析を行う. なお各指標の値が 0 となっている銘柄があればそれを除いて回帰を行う. 予測値と実測値が完全に一致したときには $A = 1, B = 0, R^2 = 1$ である.

2.2.1 スプレッドに対する検証

図 11 (a) は, 各銘柄に対し, スプレッドの実測値と, モンテカルロゼロインテリジェンスモデルによる予測値がどの程度一致しているかを示した図である. スプレッドとはアスクとビッドの値の差であり, 即時往復売買のコストを示す. 横軸が予測値, 縦軸が実測値を表し, 各 + が一つの銘柄を示している. 点線はモンテカルロゼロインテリジェンスモデルの予測と完全に一致したときに乗る線, 実線は回帰直線である. 回帰係数は $A = 0.968 \pm 0.009, B = 0.046 \pm 0.053$, 決定係数は 0.94 となった. このことから, かなりの精度でスプレッドを予測できていることが分かる.

2.2.2 ボラティリティに対する検証

図 11 (c) はボラティリティについて同様に示したものである. ボラティリティとは, 価格変動がランダム・ウォークに従うと仮定すると τ 時間後の価格と現時点の価格の 2 乗差は τ に比例するが, その比例定数のことである. 回帰係数は $A = 0.822 \pm 0.011,$

$B = -1.443 \pm 0.158$ ，決定係数は 0.88 となった．予測値は実測値よりも小さすぎるとはいえ，予測値と実測値は傾向としては一致していることが分かる．

2.2.3 λ に対する検証

図 11 (e) は λ について同様に示したものである． λ とは，1 単位の成行注文が到着したとするとアスク・ビッドが平均的にどのくらい動くかを示す指標である．回帰係数は $A = 1.147 \pm 0.038$ ， $B = 0.001 \pm 0.247$ ，決定係数は 0.55 となった．回帰係数はかなり $A = 1$ ， $B = 0$ に近いものの，決定係数が小さいことが分かった．特に λ が小さい，すなわち流動性が高い銘柄については予測からのずれが非常に大きく，中には 2 桁程度異なるような銘柄もある．そして点線は予測値と実測値が完全に一致したときに乗る線 ($y = x$) なので，モデルの予測は全体的に大きすぎる．つまり Smith らのモデルは λ をあまり予測できないことが分かる．

流動性が高い銘柄は取引量も多いためデータ数は十分あるはずであり，したがってパラメータ推定の誤差もそれほど大きくないと考えられる．またモデルの応用を考えるにあたってあまりに流動性の低い銘柄は投資対象とはなりづらいため，流動性の高い銘柄で精度が低いのは歓迎すべきことではない．

2.3 Smith らのゼロインテリジェンスモデルの検証：まとめ

Farmer et al. (2005a) によると，Smith らのモデルをロンドン証券取引所 (LSE) の 11 銘柄に対して検証した結果，スプレッドで $A = 0.99 \pm 0.10$ ， $B = 0.06 \pm 0.29$ ， $R^2 = 0.96$ ，ボラティリティで $A = 1.33 \pm 0.25$ ， $B = 2.43 \pm 1.75$ ， $R^2 = 0.76$ であった．したがって，LSE においてはスプレッドとボラティリティの予測はある程度の精度で可能であることが分かっている．

一方，東証で取引されている約 760 銘柄に対してモンテカルロゼロインテリジェンスモデルを検証した結果，全体としてやはりよく予測できていることが分かった．本モデルで扱う東証のデータと LSE のデータは構造が異なるので，直接比較はできないが，それでも東証においても，スプレッド・ボラティリティに関しては (モンテカルロ) ゼロインテリジェンスモデルの予測値と実測値はある程度一致したと言ってよいだろう．

しかしながら， λ の予測についてはあまり高い精度で予測することはできない．特に，図 11 (e) を詳しく見ると分かるように，流動性が高い銘柄ではモンテカルロゼロインテリジェンスモデルを用いると λ の予測値は大きすぎる傾向があることが分かる．

そこで，我々は，モンテカルロゼロインテリジェンスモデルを拡張し，予測の精度を高めることができるかどうかを考えた．

3 ローインテリジェンスモデルの東証への適用

本節では，まずローインテリジェンスモデルの詳細を説明し，次に東証の各銘柄に適用する．そしてスプレッド，ボラティリティ， λ に対して高い予測力を持っていることを示す．

3.1 はじめに

前節まででモンテカルロゼロインテリジェンスモデルは、スプレッド、ボラティリティに対しては有効なモデルであること、しかしながら λ に対しては、特に流動性が高い銘柄ではあまりうまく予測できないことが明らかになった。なぜ流動性が高い銘柄の λ をうまく予測できないのだろうか。

その一つの理由として、モンテカルロゼロインテリジェンスモデルではアスク・ビッドに存在する注文株数（厚み）を考慮できないということが挙げられる。現実の市場で流動性が十分高い銘柄ではアスク・ビッドに多数の注文が存在することになるが、ゼロインテリジェンスモデルではどんなに流動性が高い銘柄でもアスク・ビッドに存在する注文株数は確率1で σ 株しかない。これは「指値注文は常に σ 株ずつ到着する。」「指値注文の呼値は任意の実数値をとりうる。」と仮定したためである。したがって、もし投資家が厚みを考慮して成行注文を行うなら、モンテカルロゼロインテリジェンスモデルで自然にこの行動を組み込むことは難しい。そして実際、注文到着率は厚みに依存することを後で示す。

本研究で考慮している（モンテカルロ）ゼロインテリジェンスモデルで無視され、しかし現実の市場で観察されている板と注文、または注文同士の相互作用は以下の通りである。

- 成行注文については（3.2.1 節）
 1. アスク（ビッド）の厚みが厚いと、売り（買い）成行注文を出しやすい。
 2. 直前の成行注文が売り（買い）であれば、売り（買い）成行注文を出しやすい。
- 指値注文については（3.2.2 節）
 1. アスク・ビッドから遠く離れた呼値よりもアスク・ビッド付近の呼値に指値注文を出しやすい。
 2. アスク（ビッド）の厚みよりセカンドアスク（セカンドビッド）の厚みが厚いと、アスク（ビッド）に指値注文を出しやすい。
 3. スプレッドが3呼値単位以上開いているときにスプレッド内部に売り（買い）指値注文を出すとき、アスク（ビッド）から近い呼値に出しやすい。つまりスプレッドは少しずつ縮小する傾向がある。

以上の仮定に用いるパラメータは全てデータから求めることができる。このようなモデルをローインテリジェンスモデルと呼ぶことにする。

（モンテカルロ）ゼロインテリジェンスモデルでは価格は現実の市場価格の対数（実数値）をとるとしていたが、「厚みを考慮する」などの拡張を考慮するため、ローインテリジェンスモデルでは価格は整数値をとることとする。そうすることによって1.1 節で述べたような、板と注文の相互作用を自然に組み込むことができる。

3.2 ローインテリジェンスモデルの詳細

前節で述べたように、価格は対数をとらず整数値で考える。市場は注文駆動型の取引により売買が進む（モンテカルロ）ゼロインテリジェンスモデルでは価格は実数値をとり呼値の幅は0なので、注文の優先順序は成行優先と価格優先のみを考えればよく、注文

駆動型市場のもう一つの原則である時間優先原則を考える必要は無かった．一方，ローインテリジェンスモデルでは価格は整数値をとるので時間優先原則を考慮しなければならない．そして価格が整数値になったことに伴い，各パラメータ，スプレッド，ボラティリティ， λ を求めなおす必要がある．しかしその方法はモンテカルロゼロインテリジェンスモデルで用いられたパラメータを求める方法とほとんど変わらない．単に $\log(\text{価格})$ となっているところを対数をとらずに価格そのままの値を用いて計算すればよいだけである．

3.2.1 成行注文到着パターン

以下では投資家の振る舞いをさらに具体的に見ていく．まず成行注文がどのような到着パターンを持っているかについて述べる．

厚みと成行注文到着率 2.1.1 節で述べたように，約定タイムを用いると成行注文到着率 μ は σ に等しい．しかし実際の市場では μ はゆらいている．そこでどのような状況で μ が大きくなるのかを考える．

投資家は完全に合理的ではないにせよ，インテリジェンスを全く持っていないと仮定することも極端である．例えば図 4 (a) のようにビッドに指値注文が大量に存在する状況下で，ある株を急いで購入したい投資家がいたとする．この投資家は A：成行注文によって直ちに購入するか，B：ビッドに指値を行い売り成行注文が到着するまで待つかの意志決定を迫られている．ここで B の選択を行ったとすると，約定するまでに時間がかかったり悪くすると約定しないことがありうる．したがって図 4 (a) のような状況では投資家は A の選択，すなわち成行注文を行いやすくなるだろう．逆に図 4 (b) のようにビッドに存在する指値注文がそれほど多くない場合には，B の選択を行えば約定する確率は高く，しかも A の選択を行った場合に比べ 1 呼値単位分安く株式を購入することができる．つまりビッドに買い指値注文が多くあればあるほど，買い成行注文が増えるという傾向があることが推測できる．

そこでアスク(ビッド)の厚み d_m/σ と， d_m が与えられたときの条件付 μ_{d_m}/σ —すなわち単位時間あたり売り(買い)成行注文到着株数を σ で割ったもの— の関係を考える． μ_{d_m} は (厚みが d_m であったときの条件付成行注文株数) / (厚みが d_m であったときの条件付約定タイム) により求められる．ゼロインテリジェンスを仮定すると $\mu_{d_m}/\sigma = 1$ ，すなわち μ_{d_m}/σ は厚みによらず定数値となる．

図 5 (a) はキッコーマン (2801) ,(b) は NTT ドコモ (9437) の各銘柄に対して，横軸に厚み d_m/σ ，縦軸に μ_{d_m}/σ をプロットしたものである．キッコーマンは流動性が低く板上にあまり指値注文がない銘柄の典型例である．また NTT ドコモは流動性が非常に高く板上に指値注文が大量に存在する銘柄の典型例である．

図 5 を見ると 2 つのことが分かる．第一に厚みが小さいところでは右上がりになっている．これは先ほどの推測どおりである．第二に μ_{d_m}/σ は厚みに対して単調に増加するのではなく，厚みが大きいところでは分散が大きくなり，図 5 (b) では減少する傾向さえ観察される．なぜこのようなグラフになるのかはよく分からないが，このグラフは投資家の非線形な行動を示しているといえる．

この曲線を表現する関数として本研究では次のものを仮定する． μ_{d_m}/σ を y ， d_m/σ

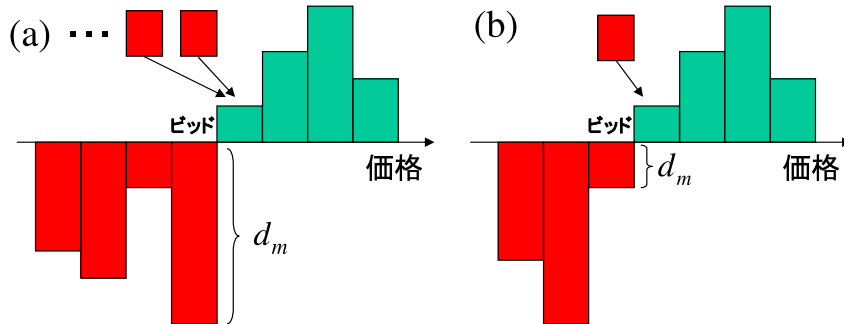


図 4: ビッドに買い指値注文が (a) 大量にあるときと (b) そうでないときの，成行注文到着率の違い．

を x とおき， $A, \gamma_1, \gamma_2, \gamma_3$ を回帰分析によって求める係数とすると

$$y = A \exp(\gamma_1 x^{0.5} + (\gamma_2 + \text{err.})x + \gamma_3 x^{1.5}). \quad (1)$$

ここで err. は誤差項である．これらの係数を求めるために両辺の対数を取り x で割ると，

$$\frac{\log(y)}{x} = \frac{\log(A)}{x} + \gamma_1 x^{-0.5} + \gamma_2 + \text{err.} + \gamma_3 x^{0.5}.$$

したがって銘柄ごとに $\log(y)/x$ を $1/x, x^{-0.5}, x^{0.5}$ により回帰すればよい．この回帰分析により自由度修正済み決定係数が 0.8 を超える銘柄は全体の 60% を超えるため，1 式は多数の銘柄に対して十分な精度で厚みと成行注文到着率の関係を表現できていると考えよう．

連続して到着する成行注文 売り成行注文が到着したら -1 を，買い成行注文が到着したら $+1$ を割り当てることにより符号列を作ると，その列は長時間相関を持つことが知られている (Bouchaud et al., 2004; Lillo and Farmer, 2004)．つまり売り (買い) 成行注文の後には売り (買い) 成行注文が続きやすい．このことをモデルに組み込むことを考える．

ローインテリジェンスモデルでは，モンテカルロゼロインテリジェンスモデルと同様，注文は必ず σ 単位で行われると仮定しているが，実際の市場では σ 株よりも大口の成行注文が到着することもしばしばある．そのような注文は板の形状を大きく変えてしまうため，流動性に大きな影響を与えていることは十分想像できる．

そこで先行研究 Bouchaud et al. (2004); Lillo and Farmer (2004) では成行注文の符号しか見ていなかったが，本研究では注文の大きさも見ることにする．つまり N 株の売り (買い) 成行注文が到着したら， N を σ で割り四捨五入した値の回数だけ連続して売り (買い) 成行注文が到着したと解釈することにする．

図 6 は，上述の意味で n 回連続して到着した売り (買い) 成行注文の次に到着した成行注文が売り (買い) 注文となる確率 P を示している．横軸が n ，縦軸が P を示す．

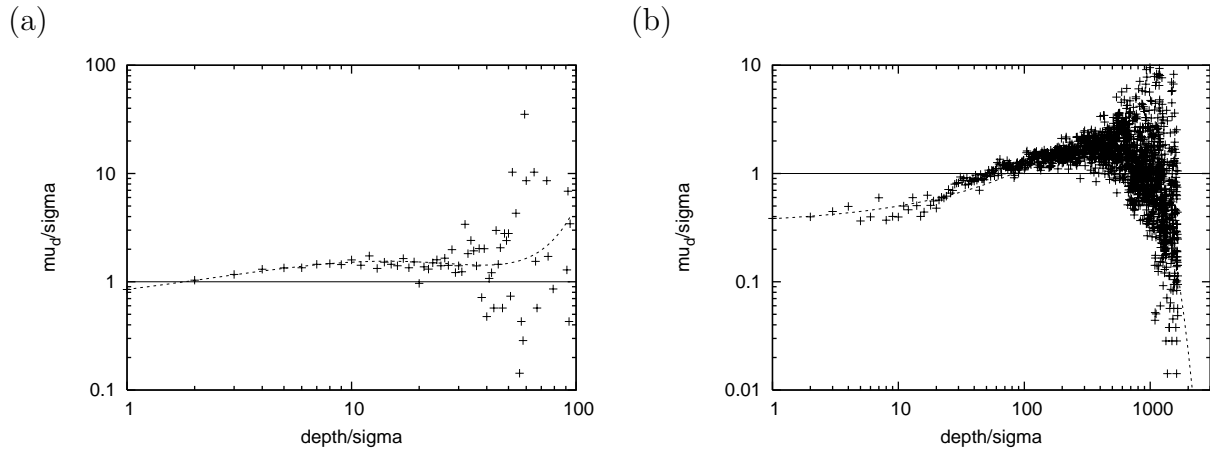


図 5: ある厚み d_m の下での、条件付成行注文到着率 μ_{d_m} . 横軸は d_m/σ , 縦軸は μ_{d_m}/σ である . 実線はゼロインテリジェンスを仮定したときに得られる線 , 点線は (1) 式を用いて回帰したときに得られる線である . (a) はキッコーマン (2801) , (b) は NTT ドコモ (9437) のデータを用いた .

やはり連続して売り (買い) の成行注文が続くと次の成行は売り (買い) 注文になりやすい . これは先行研究を支持する結果である .

この関係を表現するために次のような回帰式を用いた .

$$P(n) = a + b \log(n) + \frac{c}{n} . \quad (2)$$

ここで a, b, c は回帰係数であり銘柄ごとに求める . n があまり大きくなるとデータ数が少なくなり回帰の精度が悪くなるので , 本研究では $n < 40$ の範囲で回帰を行った . 図 6 中の点線は回帰結果を図示したものである . よく近似できていることがわかる . 実際 , 全体の 89% 以上の銘柄で自由度修正済み決定係数が 0.8 を超えるので , 2 式は「売り (買い) 成行注文が n 回連続して到着した後 , 次も売り (買い) 成行注文である確率」を多数の銘柄に対して十分な精度で表現できていると考えてよいだろう .

以上の考察から成行注文到着率を次のように決める . まず時刻 t における「売り (買い) 成行注文が n 回連続して到着した後 , 次も売り (買い) 成行注文である確率」 P_t を

$$P_t = \begin{cases} P(n) & (0 < P(n) < 1) \\ P_{t-1} & (\text{otherwise}) \end{cases}$$

のように決める . つまり P_t は確率なので 0 より小さくなったり 1 より大きくなったりしないよう , そのような場合には直前の P_{t-1} を用いるようにするということである . また時刻 0 での確率 P_0 は $1/2$ と仮定する .

以上述べてきたことをまとめると成行注文到着パターンは次のように書ける . μ_{d_m} と P_t は独立に変動すると仮定すると , アスクの厚みが d_m で売り成行注文が直前に n 回連

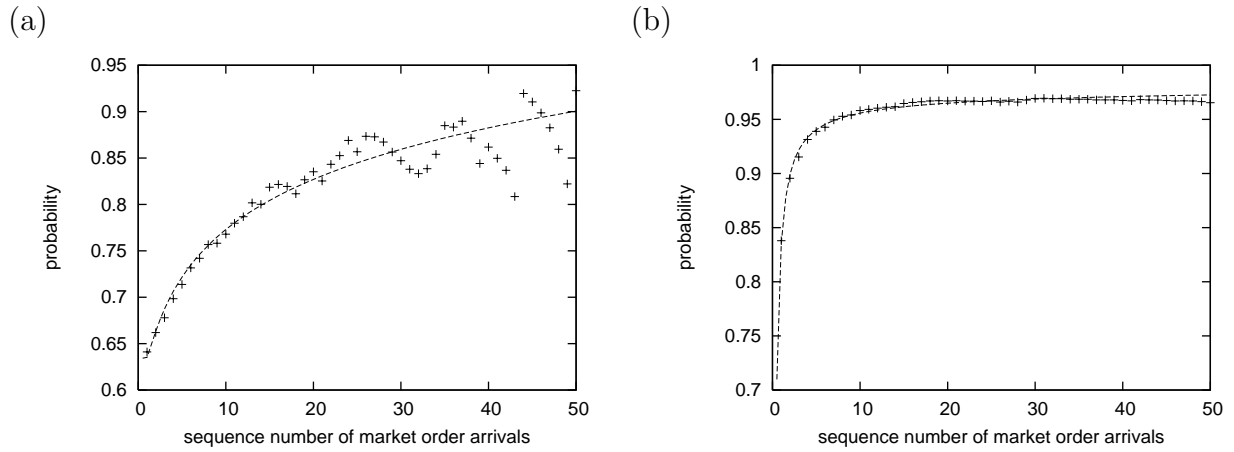


図 6: 売り（買い）成行注文が n 回連続して到着した後の成行注文が売り（買い）注文となる確率 P . 横軸が n , 縦軸が P を示している . 点線は (2) 式を用いて回帰したときに得られる線である (a) はキッコーマン (2801) (b) は NTT ドコモ (9437) のデータを用いた .

続していたとき , 次の時刻における売り成行注文到着率 μ_{sell} は

$$\mu_{\text{sell}} = A \exp \left(\gamma_1 \left(\frac{d_m}{\sigma} \right)^{0.5} + \gamma_2 \left(\frac{d_m}{\sigma} \right) + \gamma_3 \left(\frac{d_m}{\sigma} \right)^{1.5} \right) P_t$$

となる . 同様の条件で買い成行注文が直前に n 回連続していたときには次の時刻における μ_{sell} は

$$\mu_{\text{sell}} = A \exp \left(\gamma_1 \left(\frac{d_m}{\sigma} \right)^{0.5} + \gamma_2 \left(\frac{d_m}{\sigma} \right) + \gamma_3 \left(\frac{d_m}{\sigma} \right)^{1.5} \right) (1 - P_t)$$

で与える . 買い成行注文到着率 μ_{buy} についても同様である .

ただし板上に買い（売り）指値注文が一つも無かったときは売り（買い）成行注文が到着しても約定できないため , 売り（買い）成行注文到着確率は 0 とする .

3.2.2 指値注文到着パターン

アスク・ビッドからの相対位置に依存する指値注文到着率 次に指値注文の到着パターンについて考える . 先に「アスク・ビッドから遠く離れた呼値よりもアスク・ビッド付近に指値注文を出しやすい」と述べた . これをモデルに取り入れるため本研究では売り指値注文を次の 4 種類に分ける .

1. アスクとビッドの間に到着する注文 . 到着率は α_{in} で一定値とする .
2. アスクの厚みがセカンドアスクの厚みよりも小さいときにアスクに到着する注文 . 到着率は一定値ではなく , $d_l (\equiv \text{セカンドアスクの厚み} - \text{アスクの厚み})$ に依存する . 平均到着率は α_{low} とする .

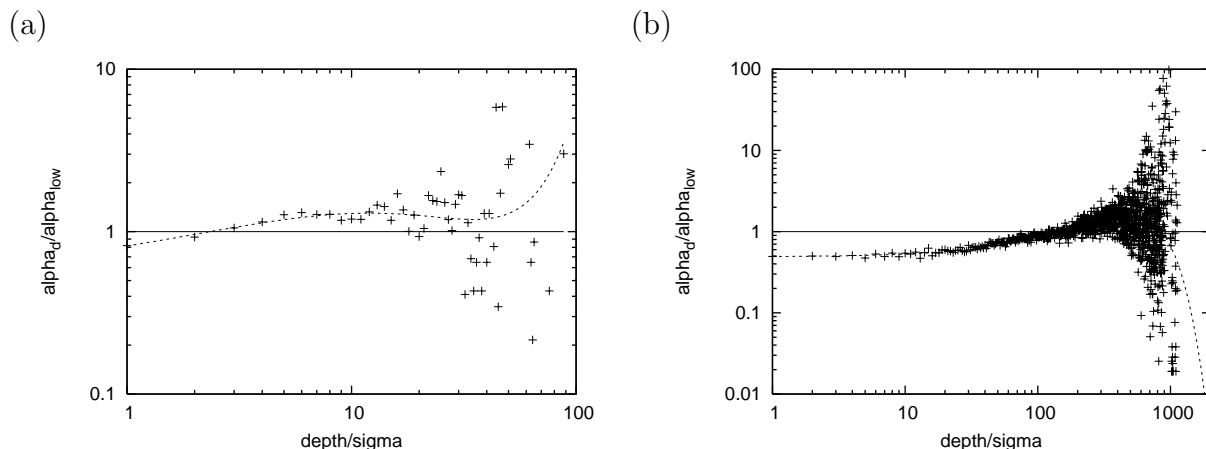


図 7: アスク (ビッド) に到着する指値注文の到着率と , アスク (ビッド) の厚みとセカンドアスク (セカンドビッド) の厚みの差 $d_l (> 0)$ の関係 . 横軸は d_l , 縦軸は $\alpha_{d_l} / \alpha_{low}$ である . (a) キッコマン (b) NTT ドコモに対する結果を示す . 点線は (3) 式を用いた回帰により得られた曲線である . また実線は α_{d_l} が d_l によらず一定の値 α_{low} であると仮定したときに得られる線である .

3. アスクの厚みがセカンドアスクの厚みよりも大きいときにアスクに到着する注文 . 到着率は α_{high} で一定値とする .
4. アスクよりも高い呼値に到着する注文 . 到着率は α_{out} で一定値とする .

到着率 α_{in} , α_{low} , α_{high} , α_{out} の推定法は C.2 節で説明した α の推定法とほぼ同じで , 単に対象となる指値注文について算出すればよいだけである . 買い指値注文についても同様なので以下では売り指値注文についてのみ述べる .

まず「1. アスクとビッドの間に到着する注文」については一定の到着率 α_{in} で到着するものとする . しかしモンテカルロゼロインテリジェンスモデルと異なり板に一樣に到着するのではなく , 後述するようにアスク側から到着すると仮定する . 「3. アスクの厚みがセカンドアスクの厚みよりも大きいときにアスクに到着する注文」 , 「4. アスクよりも高い呼値に到着する注文」についてはそれぞれ一定の到着率 α_{high} , α_{out} で , モンテカルロゼロインテリジェンスモデルと同様に一樣・ランダムに到着すると仮定する . ただし 2.1.2 節で論じたようにあまりにアスク・ビッドから遠い呼値への注文は意味が無いので , $bid + w$ よりも遠い呼値には指値注文は到着しないものとする . w は「幅」(C.2 節で説明した「幅」と定義は同じで , 価格を整数化したもの .) の 10 倍である .

最後に「2. アスクの厚みがセカンドアスクの厚みよりも小さいときにアスクに到着する注文」について , この注文が最も重要である . なぜなら α_{in} , α_{low} , α_{high} , α_{out} の中で α_{low} が最大となる銘柄が全体の 86% を超えるからである . そこで「2. アスクの厚みがセカンドアスクの厚みよりも小さいときにアスクに到着する注文」を詳しくモデル化することにする . このときの到着率は一定値ではなく , d_l (\equiv セカンドアスクの厚み - アスクの厚み) に依存すると仮定する . 具体的にどのように扱うのかについて次に述べる .

以上の状況をまとめたものが図 8 である .

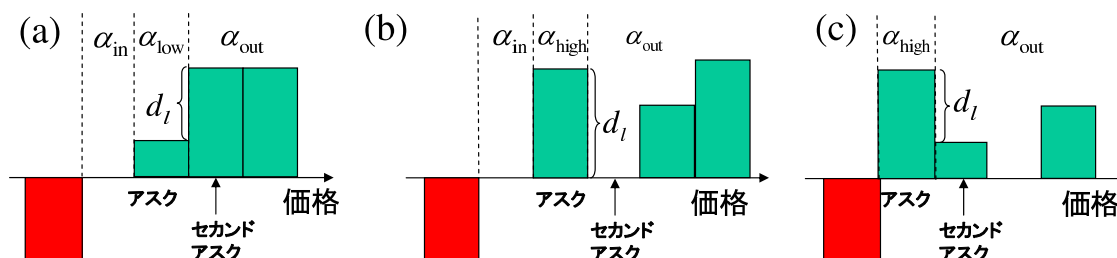


図 8: d_l , α_{in} , α_{low} , α_{high} , α_{out} の定義 (a)(b) ではスプレッドが 2 以上空いているため、アスクとビッドの間に指値注文が到着しうる。その到着率は α_{in} である。一方 (c) ではスプレッドは 1 であるため、 α_{in} の到着率で指値注文が到着する領域はない。また (a) $d_l (\equiv$ セカンドアスクの厚み - アスクの厚み) > 0 のときはアスクへの指値注文到着率 α_{d_l} は d_l に依存し、その平均値は α_{low} である。しかし (b)(c) のように $d_l < 0$ のときはアスクへの指値注文到着率は定数 α_{high} である (d_l に依存しない) と仮定する。

アスクに到着する指値注文の到着率 売り注文を出そうとしている投資家がいるとして、どのような状況なら指値注文をアスクに出しやすいかを考えよう。特にその投資家はアスクに指値として出すか成行によりすぐ売るかの決断をせまられているとする。アスクに存在する指値注文量 (つまりアスクの厚み) が十分少なければアスクに売り指値を出すのもよいだろう。

では投資家はアスクの厚みの大きさをどのように感知しているのだろうか。ここではセカンドアスクの厚みとの相対的な大きさの差を感知していると考えよう。そうだとすると $d_l (\equiv$ セカンドアスクの厚み - アスクの厚み) が大きいときアスクに指値注文が到着しやすいことになる。図 7 は $d_l > 0$ のときの (a) キッコーマン (b) NTT ドコモに対する d_l と指値注文到着量 α_{d_l} の関係を図示したものである。縦軸は α_{low} で割ってある。これらの図によると先の予想通り d_l が大きいときアスクに指値注文が到着しやすい傾向があることが分かる。

この現象をモデルに取り入れることにする。まず $d_l > 0$ のときを考える。成行注文の場合と同様、次のような回帰式を仮定する。 $\alpha_{d_l}/\alpha_{low}$ を y , d_l/σ を x とおき、 A , γ_1 , γ_2 , γ_3 を回帰分析によって求める係数とすると

$$y = f(x) \equiv A \exp(\gamma_1 x^{0.5} + (\gamma_2 + \text{err.})x + \gamma_3 x^{1.5}). \quad (3)$$

回帰は (1) 式と同様に、両辺の対数を取り x で割ってから行う。この回帰式を用いると全体の 7 割以上の銘柄に対して自由度修正済み決定係数が 0.8 を超える。したがって (3) 式は「アスクの厚みとセカンドアスクの厚みの差と、指値注文到着率」を多数の銘柄に対して十分な精度で表現できていると考えてよいだろう。

$d_l \leq 0$ のときは指値注文は d_l に依存せず一定の到着率 α_{high} で到着すると仮定する。

少しずつ縮小するスプレッド 木村・秋山 (2007) によると、スプレッドが 3 単位以上開いているときにアスクとビッドの間に到着する売り (買い) 指値注文はアスク (ビッ

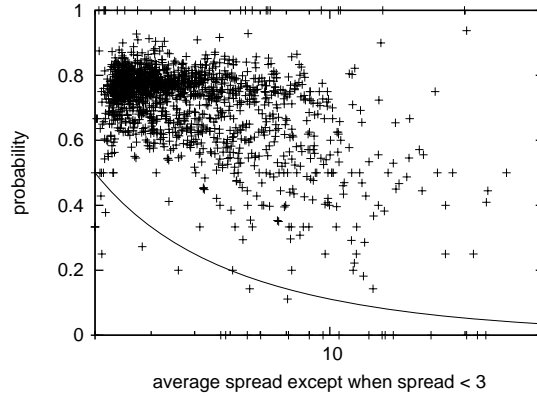


図 9: スプレッドが 3 単位以上開いていたときに，アスクとビッドの間に到着した売り（買い）指値注文が $\text{ask} - 1$ ($\text{bid} + 1$) に到着していた確率とそのときの平均スプレッドを銘柄ごとにプロットした．横軸はスプレッドが 3 以上開いているという条件の下でのスプレッドの平均値，縦軸はその到着確率を示す．実線はアスクとビッドの間どの呼値にも等確率で到着すると仮定したときに得られる線である．ほとんどの銘柄で実線を上回っているので，アスクとビッドの間に到着する指値注文はスプレッドを少しずつ縮めるように到着することが分かる．

ド）から近い呼値に到着しやすい．つまりスプレッドは少しずつ縮小する傾向がある．図 9 は，スプレッドが 3 単位以上開いていたときに，売り（買い）指値注文が $\text{ask} - 1$ ($\text{bid} + 1$) に到着する確率とそのときの平均スプレッドを銘柄ごとにプロットしたものである．実線はどの呼値に到着するかが完全にランダムであると仮定したときに得られる線である．98% 以上の銘柄がこの線を上回っているのでアスクとビッドの間に到着する指値注文はスプレッドを少しずつ縮めるように到着することが分かる．図 10 (a) はそのような状況を描いたものである．

そこで本研究では図 10 (b) に示すように，アスクとビッドの間に到着する売り（買い）指値注文は必ず $\text{ask} - 1$ ($\text{bid} + 1$) に到着すると仮定する．

以上で述べたことをまとめると指値注文到着率は次のように決まる．売り指値注文の到着率 $\alpha_{\text{sell}}(p)$ は呼値 p に依存して次のように決まる．

$$\alpha_{\text{sell}}(p) = \begin{cases} 0 & (p < \text{ask} - 1), \\ (\text{ask} - \text{bid} - 1)\alpha_{\text{in}} & (p = \text{ask} - 1), \\ \alpha_{\text{low}}f(d_l/\sigma) & (p = \text{ask} \ \& \ d_l > 0), \\ \alpha_{\text{high}} & (p = \text{ask} \ \& \ d_l \leq 0), \\ \alpha_{\text{out}} & (\text{ask} < p < \text{bid} + w), \\ 0 & (p \geq \text{bid} + w). \end{cases}$$

ここで $p = \text{ask} - 1$ のとき α_{in} に $\text{ask} - \text{bid} - 1$ がかかっているのは $\text{bid} < p < \text{ask}$ に到着するはずだった売り指値注文を全て $\text{ask} - 1$ に到着するようにしているからである．

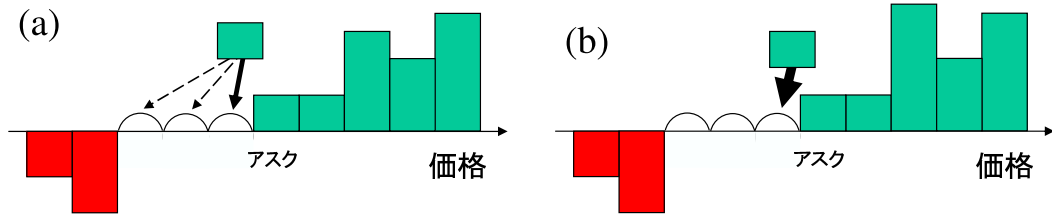


図 10: (a) 現実の市場における，アスクとビッドの間に到着する指値注文の模式図．売り指値注文の場合には $\text{ask} - 1$ に到着する確率が高い (b) ローインテリジェンスモデルにおける，アスクとビッドの間に到着する指値注文の模式図．売り指値注文の場合には必ず $\text{ask} - 1$ に到着する．

$f(\cdot)$ は (3) 式によって与えられる関数である．買い指値注文到着率 $\alpha_{\text{buy}}(p)$ についても同様に

$$\alpha_{\text{buy}}(p) = \begin{cases} 0 & (p > \text{bid} + 1), \\ (\text{ask} - \text{bid} - 1)\alpha_{\text{in}} & (p = \text{bid} + 1), \\ \alpha_{\text{low}}f(d_l/\sigma) & (p = \text{bid} \ \& \ d_l > 0), \\ \alpha_{\text{high}} & (p = \text{bid} \ \& \ d_l \leq 0), \\ \alpha_{\text{out}} & (\text{ask} - w < p < \text{bid}), \\ 0 & (p \leq \text{ask} - w) \end{cases}$$

によって与える．

3.2.3 注文到着パターン：まとめ

こうして成行・指値注文到着率を与えることができたので実際にどのように成行・指値・キャンセルの各イベントが生じるのかについてまとめる．なおキャンセルについてはモンテカルロゼロインテリジェンスモデルと同様，キャンセル率 δ_c でランダム・独立に生じるとする．

各イベントは以下のように生成する．板上に存在する全指値注文数を n ，各イベントが生じる確率の分母を $P = \sum_p \alpha_{\text{sell}}(p) + \sum_p \alpha_{\text{buy}}(p) + \mu_{\text{sell}} + \mu_{\text{buy}} + n\delta_c$ とおくと

1. 確率 $\alpha_{\text{sell}}(p)/\sigma P$ で呼値が p の売り指値注文到着イベントが発生する．
2. 確率 $\alpha_{\text{buy}}(p)/\sigma P$ で呼値が p の買い指値注文到着イベントが発生する．
3. 確率 μ_{sell}/P で売り成行注文到着イベントが発生する．
4. 確率 μ_{buy}/P で買い成行注文到着イベントが発生する．
5. 確率 $n\delta_c/P$ でキャンセルイベントが発生する．実際にどの注文がキャンセルされるかは，ランダムに決まる．

他の設定はモンテカルロゼロインテリジェンスモデルと同様である．

3.3 結果

本節ではスプレッド，ボラティリティ， λ に対するローインテリジェンスモデルの予測と実測値を比較する．まずはスプレッドからである．

3.3.1 スプレッドに対する検証

図 11 (b) は，スプレッドから 1 を引いた値の，実測値と本モデルによる予測値がどの程度一致しているかを示した図である．ここでスプレッドとはアスクとビッドの差である．価格は整数であるとしたため，スプレッドの最小値は 1 である．そこで図 11 (b) ではスプレッドから 1 を引いた値を図示してある．各 + が一つの銘柄を示している．点線は本モデルの予測と完全に一致したときに乗る線，実線は回帰直線である．回帰係数は $A = 1.033 \pm 0.010$ ， $B = 0.285 \pm 0.018$ ，決定係数は 0.93 となり，ゼロインテリジェンスモデルと同程度の予測能力を持っていることが分かる．

3.3.2 ボラティリティに対する検証

図 11 (d) はボラティリティについて同様に示したものである．ここでボラティリティとは，価格がランダムウォークに従うとしたとき τ 時間後の価格分散は τ に比例するが，その比例係数のことである．回帰係数は $A = 0.786 \pm 0.011$ ， $B = 0.127 \pm 0.024$ ，決定係数は 0.87 となり，スプレッドと同様ボラティリティについても，ゼロインテリジェンスモデルと同程度の予測能力を持っていることが分かる．

3.3.3 λ に対する検証

図 11 (f) は λ について同様に示したものである．ここで λ とは，仮に σ 株の成行注文が到着したら平均して価格はどれくらい動くかを示す指標である．回帰係数は $A = 0.929 \pm 0.016$ ， $B = -0.133 \pm 0.027$ ，決定係数は 0.81 となった．ゼロインテリジェンスモデルの場合には決定係数は 0.55 であったので，ローインテリジェンスモデルはうまく λ を予測できていることが分かる．

3.3.4 モンテカルロゼロインテリジェンスモデルとローインテリジェンスモデルの比較

2 節では東証のデータを用いて Smith らのモデルをモンテカルロシミュレーションにより検証した．その結果，スプレッドに関しては決定係数が 0.94 となり，Smith らのモデルはかなり有効なモデルであることが分かった．しかし，仮に 1 単位 (σ 株) の成行注文が到着したときに平均してどのくらい価格が動くかを示す λ については決定係数はそれぞれ 0.55 とそれほど高くならなかった．つまり Smith らのモデル及びモンテカルロゼロインテリジェンスモデルでは λ を予測することは難しい．

一方，ローインテリジェンスモデルではスプレッド，ボラティリティの決定係数はほ

表 2: モンテカルロゼロインテリジェンスモデル (mZIM) とローインテリジェンスモデル (LIM) の比較. $\log(\text{実測値}) = A \times \log(\text{予測値}) + B$ の形に回帰したときの B , A , 決定係数 R^2 を示す () 内は標準偏差.

モデルと検証した市場	指標	B (SD)	A (SD)	R^2
mZIM を東証で検証	スプレッド	0.046(0.053)	0.968(0.009)	0.94
	ボラティリティ	-1.443(0.158)	0.822(0.011)	0.88
	λ	0.001(0.247)	1.147(0.038)	0.55
LIM を東証で検証	スプレッド	0.285(0.018)	0.1.033(0.010)	0.93
	ボラティリティ	0.127(0.024)	0.786(0.011)	0.87
	λ	-0.133(0.027)	0.929(0.016)	0.81
SmithらのモデルをLSEで 検証 (Farmer et al., 2005a)	スプレッド	0.06(0.29)	0.99(0.10)	0.96
	ボラティリティ	2.43(1.75)	1.33(0.25)	0.76

ば変わらず, λ の決定係数は上昇した. 表 11 は, 以上の結果をまとめたものである.

3.4 議論

ゼロインテリジェンスモデルは注文がランダム・独立に到着し注文のキャンセルも確率的に生じるというモデルである. そのようなモデルの中で 2.1.1 節で紹介したような, 指値注文が価格軸上無限遠まで一様に到着するというモデルを Smith らのモデルと呼ぶことにする.

2 節では, 東証のデータを用いて, Smith らのモデルをモンテカルロシミュレーションにより検証した. その結果, スプレッド, ボラティリティについては高い決定係数を得ることができた. しかし σ 株の成行注文が到着したときにどのぐらい価格が動くかを示す指標である λ についてはあまり予測することができなかった.

そこで本研究では「板と注文の相互作用」や「注文間相互作用」を記述する方法を考察しそれをモデルに組み込んだ. このモデルをローインテリジェンスモデルと呼ぶことにする. 3 節ではローインテリジェンスモデルを検証するため, モンテカルロシミュレーションによりスプレッド, ボラティリティ, λ を予測できるかどうかを調べた. その結果, 三者の予測が可能であることが明らかになった.

Farmer et al. (2005a) らの研究では, ロンドン証券取引所で取引されている 11 銘柄に対して全ての注文を入力データとして扱い, 高い決定係数を得た. 彼らが用いているデータは, ある注文が最終的に約定したのかキャンセルされたのかまで追うことができる, 高頻度データである. しかし, 本研究では板情報, すなわちある呼値に何株の注文があるかというデータしか扱えない. さらにその板情報も 5 本値しか存在しない. このことはキャンセル率の推定に影響を及ぼす.

また Farmer et al. (2005a) らが用いていた方法では, 単位時間・単位価格あたりの

指値注文到着率 α を求めるときに、アスク・ビッドから極端に遠い呼値をとる指値注文をカウントしていない。また逆にアスク・ビッドに極端に近い呼値に到着した指値注文もカウントしていない。しかし、本研究で扱ったデータは5本値しかないので、そのような処理を行うことはできない。

本研究で用いたデータには以上のような様々な制約があるにもかかわらず、Smithらのモデルはスプレッド・ボラティリティ共によく予測できることが分かった。さらにローインテリジェンスモデルでは λ をも予測できるようになった。しかも本研究は約760銘柄と、Farmer et al. (2005a)らが対象とした11銘柄よりもずっと多数の銘柄を対象とした。この中には様々な業種・規模の企業が含まれている。このことはSmithらのモデル及びローインテリジェンスモデルの頑健性を示すものと思われる。こうして、スプレッド、ボラティリティ、 λ については「板と注文の相互作用」と「注文同士の相互作用」だけを考慮したモデルであるローインテリジェンスモデルで十分予測できることが分かった。

では、なぜSmithらのモデルはマーケットインパクト（あるいは λ ）を予測できないのだろうか。大きな理由の一つは、Smithらのモデルでは指値注文が一様に到着すると仮定したことである。3.2.2節で触れたようにアスク・ビッド付近に到着する指値注文はそれ以外の呼値に到着する指値注文と比べてかなり多い。 λ はアスク・ビッドに存在する指値注分量の大きな影響を受けることはその定義から分かるので、指値注文が一様に到着するという仮定が非現実的すぎたのである。

ところがそれを考慮しただけではまだ λ を予測できない。アスク・ビッドに到着する指値注分量をデータに基づいて増やすと、特に流動性の高い銘柄で増やしすぎになってしまう。アスク・ビッドの厚みが現実の市場よりも厚くなってしまふのである。そこでローインテリジェンスモデルでは売り（買い）成行注分量はアスク（ビッド）の厚みに依存するとした（3.2.1節）。このようにすると、アスク（ビッド）の厚みが厚すぎる場合には売り（買い）成行注文が生じる確率が増加し、それ以上のアスク（ビッド）への指値注文の増加が抑えられるのである。

その結果 λ をうまく予測できるようになることはかなり意外である。投資家はマーケットインパクトが小さくなるように取引を行っているはずである（日経金融新聞、2006; Kim, 2007）。もしもそうなら、買い（売り）成行注文を行う投資家はアスク（ビッド）の厚みを見ているということになる。つまり売り（買い）成行注分量はビッド（アスク）の厚みに依存するとしたほうがよさそうである。ところがモデルの上では売り（買い）成行注分量はアスク（ビッド）の厚みに依存するとして十分うまく予測できる。この理由はよく分からないが、とにかく板と注文の相互作用や注文と注文の間の相互作用が流動性に影響を与えていることは確かである。

最後に、本研究で解明した、これらの相互作用の応用可能性について述べておく。これらの相互作用を利用することにより、以下のようにして成行注文を行う際のマーケットインパクトを小さくすることが可能かもしれない。ここでは買い成行注文を行うとする。まずアスクの厚みがセカンドアスクの厚みより厚くなるのを待つ。次にアスクの厚みがなくなる直前までできるだけゆっくり買い成行注文を出す。アスクの厚みがなくなる直前まで来たら買い成行注文を出すのをやめ、アスクの厚みがセカンドアスクの厚みと同程度になるまで待つ。つまりアスクの厚みがセカンドアスクの厚みより薄いときには、アスクの厚みの回復力は強いので、アスクを動かさないように成行注文を出し続けるという戦略である。

また以下のようにして指値注文のマーケットインパクトも低減できるかもしれない。ここでは売り指値注文をアスクに行うとする。最適なタイミングはビッドに多数の注文

があるときである．このときは原則として多数の買い成行注文が入りやすい．したがって売り指値注文は約定しやすいのである．

このように本研究で解明した投資家行動の特徴は実際的な応用を持つ可能性がある．またその特徴を考慮すると，投資家の効用のような測定できない量を用いずに，流動性のある程度説明することができた．このことはゼロインテリジェンスモデル及びローインテリジェンスモデルの可能性を示している．

参考文献

- Becker, G. (1962), "Irrational behavior and economic theory," *Journal of Political Economy*, **70**, 1, pp. 1–13.
- Blume, M., MacKinlay, A. and Terker, B. (1989), "Order Imbalance and Stock Price Movements on October 19 and 20 1987," *Journal of Finance*, **44**, pp. 827–848.
- Bouchaud, J.-P., Y. Gefen, M. Potters and M. Wyart (2004), "Fluctuations and response in financial markets: the subtle nature of 'random' price changes," *Quantitative Finance*, **4**, pp. 176–190.
- Cont R and Bouchaud J-P (2000), "Herd behavior and aggregate fluctuations in financial markets," *Macroeconomic Dynamics*, **4**, pp. 170–196.
- D. Gode and S. Sunder (1993), "Allocative efficiency of markets with zero intelligence traders: Markets as a partial substitute for individual rationality," *Journal of Political Economy*, **101**, pp. 119–137.
- Didier Sornette (2004), *Why Stock Markets Crash: Critical Events in Complex Financial Systems*, Princeton University Press.
- Lillo F and Farmer J D (2004), "The Long Memory of the Efficient Market," *Studies in Nonlinear Dynamics & Econometrics*, **8**, 3.
- Farmer J D, P Patelli and I I Zovko (2005), "The Predictive Power of Zero Intelligence in Financial Markets," *Proceedings of the National Academy of Sciences*, **102**, 6, pp. 2254–2259.
- Farmer J D, P Patelli and I I Zovko (2005), "The Predictive Power of Zero Intelligence in Financial Markets Supplementaty Information," *Proceedings of the National Academy of Sciences*, **102**, 6, pp. 2254–2259.
- Slanina F (2001), "Mean-field approximation for a limit order driven market model," *Physical Review E*, **64**, 056136.
- I. Zovko and J. D. Farmer (2002), "The power of patience: a behavioral regularity in limit order placement," *Quantitative Finance*, **2**, 387.

- Kendall Kim (2007), *Electronic and Algorithmic Trading Technology: The Complete Guide (Complete Technology Guides for Financial Services)*, Academic Press.
- 木村 博道, 秋山 英三 (2007), 「流動性指標に見るトレーダの行動」『情報処理学会論文誌：数理モデル化と応用』48, SIG 19(TOM 19) .
- Kyle, A. S. (1985), "Continuous Auctions and Insider Trading," *Econometrica*, **53**, 6.
- Mantegna R N and Stanley H E (1999), *Introduction to Econophysics: Correlations and Complexity in Finance*, Cambridge, MA: Cambridge University Press.
- Mendelson H (1982), "Market behavior in a clearing house," *Econometrica*, **50**, 6, pp. 1505–1524.
- Michael J. Fleming (2003), "Measuring treasury market liquidity," *FRBNY Economic Policy Review*, 83-99.
- 日経金融新聞 (2006) , 「アルゴリズム取引上陸 (上) 効率売買で顧客争奪 クレディ・スイス, ゴールドマン」日経金融新聞, 2006/2/15 .
- 大村 敬一, 宇野 淳 (1998), 「株価と売買高」『早稲田商学』第 376 号 .
- S. Maslov (2000), "Simple model of a limit order-driven market," *Physica A*, **278**, pp. 571–578.
- Smith E, J D Farmer, L Gillemot and S Krishnamurthy (2003), "Statistical Theory of the Continuous Double Auction," *Quantitative Finance*, **3**, pp. 481–514.
- Tang L-H and Tian G-S (1999), "Reaction-diffusion-branching models of stock price fluctuations," *Physica A*, **264**, pp. 543–550.
- V. Plerou, P. Gopikrishnan, X. Gabaix, H.E. Stanley (2002), "Quantifying Stock Price Response to Demand Fluctuations," *PHYSICAL REVIEW E*, **66**, 027104
- Xavier Gabaix, Parameswaran Gopikrishnan, Vasiliki Plerou and H Eugene Stanley (2003), "A Theory of Power Law Distributions in Financial Market Fluctuations," *Nature*, **423**, pp. 267–270.

A 注文駆動方式による価格決定

東証で用いられている価格決定方式を簡単に述べておく。証券市場には大きく分けて二種類の注文方法がある。売買価格を提示するタイプの注文（指値注文）と現在出ている指値注文の中で最良の価格で売買するという注文（成行注文）である。指値注文は注文執行前であればキャンセルすることも可能である。価格別の指値注文量の情報を板、指値注文の指している価格を気配もしくは呼値、売り（買い）注文板上の最安値（最高値）をアスク（ビッド）、アスクとビッドの差をスプレッドという。そして注文は以下の原則に従って処理される。

1. 成行優先原則，すなわち成行注文は指値注文よりも優先的に取引される．
2. 価格優先原則，すなわち売り（買い）指値注文の中ではより安い（高い）注文が優先的に取引される．
3. 時間優先原則，すなわち同価格の指値注文の中ではより注文時刻の早い注文が優先的に取引される．

このような価格決定方式を注文駆動方式という（図 12 を参照）．

本研究で扱うデータは指値注文が存在する 5 番目の呼値まで存在する．これを 5 本値という．そして，通常，セカンドアスク（セカンドビッド）とは注文が存在する呼値だけを見たときの，2 番目の呼値を指すが，本研究では呼値に注文があるかないかに関わらず常に $ask + 1$ ($ask - 1$) の価格のことをさすことにする．

本研究では，注文駆動方式に基づくゼロインテリジェンスモデルを扱う．

B 推定の対象

B.1 推定の対象銘柄

本研究において対象とする銘柄は，東証 1, 2 部で取引されている銘柄である．ただし，ETF や東証が主市場となっていない銘柄は除いてある．そのような銘柄は株価指数や東証以外の市場の動きに左右され，ゼロインテリジェンスモデル及びローインテリジェンスモデルでは扱いにくくなるからである．また REIT も含んでいない．さらに，対象銘柄は以下で述べる条件を満たす必要がある．

データが存在する期間中，合併を経験した会社について，合併後に存続したほうの会社に関してはデータから除き，被合併側の会社のみを対象とする．つまり，合併前のデータしか扱わない．これは，モデル上，合併は異なる $\alpha, \mu, \delta, \sigma$ を持つ会社を混合することに相当し，今回の分析の目的に沿わないからである．なお，新規上場・上場廃止された会社はデータの存在する期間に対してそのまま分析対象とする．

呼値の刻みは価格水準ごとに決まっている．例えば価格が 2000 円以下では 1 円，2001 円以上 3000 円以下では 5 円単位で注文するようになっている．この変更はゼロインテリジェンスモデルの予測に影響するので，対象期間中，呼値の刻み変更の価格に一度でも触れた銘柄は分析対象外とする．

株式はあるまとまった単位で取引される．これを売買単位という．売買単位は，例えば 1000 株単位から 100 株単位のように，変更することが可能である．この変更はゼロインテリジェンスモデルの予測に影響するかどうか不明であるため，売買単位変更のあった銘柄は分析対象外とする．

企業は 1 株をいくつかに分割したり（株式分割），新株を発行したりすることで（増資），発行済株式数を増やすことがある．小規模な株式分割や増資はゼロインテリジェンスモデルの予測にあまり影響しないと考えられるため，そのまま分析対象とする⁸．

以上で述べたことを全て満たす銘柄は約 1500 銘柄存在する．そのうち λ を計算でき

⁸大幅な株式分割を行った銘柄はかなりの影響があるかもしれないが，このケースは分割前後で株価が急落することになり，「呼値の刻みが変更された銘柄は分析対象外」というルールによって除かれる．

表 3: 約定データの構造

日付	時刻	約定データか 気配データか	約定価格	約定株数	売り注文か 買い注文か
20031030	0903	0	398	10	0

る銘柄は約 760 銘柄である．この約 760 銘柄を対象に計算を行う⁹．データは日経メディアマーケティング株式会社から購入した，ティックデータ（個別株・5 本気配板）（ヒストリカル）を使用した．

B.2 推定の対象期間

次に対象期間について述べる．データが存在する期間は 2003 年 10 月 30 日から 2006 年 2 月 28 日までである．東証の場合，一日の最初の取引，昼休み直前・直後，一日の最後の取引は注文駆動方式ではなく需給の一致する価格によって取引される（板寄せ）のでこれは除く．

市場には急激な価格変動に対処するためにいろいろな制度が設けられている．例えば，大口の買い成行注文が入ってきたとし，そのまま売買が成立してしまうと価格が大きく動いてしまうとする．このような場合，取引所はその呼値の周知を図り，新規の売り注文が入ってくるのを待つ．このときの気配を特別気配という．また，一日の値動きの幅には制限がある．この制限幅まで買われたときをストップ高，売られた時をストップ安という．ゼロインテリジェンスモデルではそうした制度は考慮されていないため，このような状況はできるだけ排除し，データの精度をなるべくそろえる必要がある．そこでこうした制度が発動するたびにデータを区切り，普通の気配（一般気配）のときだけを対象にパラメータを計算する．つまり，時間的に連続していて，かつ一般気配のみからなるデータを作る．この区切られたデータをデータブロックと呼ぶことにする．

B.3 板情報の処理

データブロックは二種類のデータから成る．一つは注文成立時に記録されるデータで，約定価格，約定株数などが記されている．これを約定データという（表 3）．もう一つは板上の注文が増減するたびに記録されるデータで，気配データと呼ばれる（表 4）．本研究で用いる気配データは，アスクまたはビッドから 5 番目の気配までしか記録されていない．これを 5 本値という．

東証における株取引は原則，9:00 から 11:00（前場）の間と 12:30 から 15:00（後場）の間に行われる．前後場をまとめて「場」と呼ぶことにする．場は取引が連続的に行わ

⁹正確には，モンテカルロゼロインテリジェンスモデルの検証では 760 銘柄，ローインテリジェンスモデルの検証では 757 銘柄が対象となる．両方で対象銘柄が異なるのは，流動性が非常に低い銘柄に対しては，必要なパラメータが全て計算できるとは限らないからである．

表 4: 気配データの構造

日付	時刻	約定データか 気配データか	気配	気配にかかる 株数	売り板か 買い板か
20031030	0900	1	399	42	0
20031030	0900	1	400	62	0
20031030	0900	1	401	25	0
20031030	0900	1	402	82	0
20031030	0900	1	403	59	0
20031030	0900	1	398	81	1
20031030	0900	1	397	502	1
20031030	0900	1	396	201	1
20031030	0900	1	395	12	1
20031030	0900	1	394	10	1

れる最大単位となっており、データブロックの最大単位にもなる。

場中では時に、一般気配以外の気配が出現することがある。付録 B.2 で説明した特別気配、ストップ高安などである。また指値注文が少なすぎてアスク・ビッドが存在しないこともある。ゼロインテリジェンスモデルではこのような制度・状況は考慮していないため、本研究では取り除くことにした。取り除くたびにデータの連続性は失われることになる。こうして残ったデータがデータブロックである。

データブロックは必ず気配データから始まるようにする。それは次のような理由による。気配データは、指値注文、成行注文、キャンセルによって変化する。図 13 はこれらをどのように区別するかを示したものである。指値注文とキャンセルは気配データの差分により判別されるため、データブロックは必ず気配データから始まるようにする。例えば特別気配の直後に約定データが入っていたらそれは取り扱わずに捨てることにする。

約定データには、買い成行注文によって約定したのか、売り成行注文によって約定したのかは記録されていない(表 3 は説明のためにつけてある)。そこで、それを判別するために Blume et al. (1989) の方法を用いる。これはアスクでの取引は成行買い注文、ビッドでの取引は成行売り注文とし、それ以外の取引は無視するというものである。「それ以外の取引」はたまたま同時にアスクとビッドの間に注文が来る場合に発生するが、そのようなことはほとんどない。

C モンテカルロゼロインテリジェンスモデルに関する各パラメータ, スプレッド, ボラティリティ, λ の推定法

本節ではモンテカルロゼロインテリジェンスモデルに関する $\sigma, \mu, \alpha, \delta_s, \delta_c$, スプレッド, ボラティリティ, λ の推定法について述べる. なおローインテリジェンスモデルについても同様であり, 変更点は価格の対数をとらずに価格そのものの値を用いるという点だけである.

C.1 σ, μ の推定法

σ の定義は一注文あたりの株数である. 注文には指値と成行の 2 種類がある. 本研究では指値注文のみを考えて σ を求める. 成行注文を考慮しないのは見かけ上 σ が小さくなってしまふからである. というのは, アスク (ビッド) に存在する株数より大口の成行注文が入ってきた時には, その注文はアスク (ビッド) 内で処理しきれず次の気配にかかることになるが, それは見かけ上別の注文として扱われてしまうからである. したがって, σ の定義は (ある期間中の, 把握できた全指値注文株数) / (把握できた指値注文の回数) となる. この式を用いてデータブロックごとに σ を求め, その σ をデータブロック中の約定タイムで重みをつけて平均したものを, その銘柄の σ とする.

μ の定義は単位時間あたりの成行注文数である. ここで, 時間は約定タイム, すなわち成行注文が到着するたびに +1 するというやり方を用いて計っているため, μ は, 結局, 成行注文の 1 回あたり注文株数に相当する. ところが, この値はそのまま求めると上述したように過小評価になってしまう. そこで, ゼロインテリジェンスモデルの想定どおり成行注文の 1 回あたり注文株数は, 指値注文の 1 回あたり注文株数 σ に等しいと仮定する. つまり $\mu = \sigma$ である. したがってモンテカルロゼロインテリジェンスモデルの場合には μ の値を求める必要はない.

C.2 α の推定法

次に α の推定法について述べる. α の定義は単位時間・単位価格あたりの指値注文株数である. α の推定は μ, σ ほど簡単ではない. なぜなら, 本研究で用いたデータは 5 本値なので, 全ての指値注文を把握できるわけではないからである. つまり, 直前の気配データに無い気配が次のデータで出現したり (図 14 左), あったはずの気配が消えていたり (図 14 右) ということが起きる.

そこで次のように考える. 図 14 のような売り板があったとする. $p_1 =$ 直前の売り 5 本値気配の中で最高値 (対数をとる前の, 現実の市場価格), $p_2 =$ 直後の売り 5 本値気配の中で最高値 (同じく, 対数をとる前の, 現実の市場価格), $\min = \min(p_1, p_2)$ とおく. 売り気配板の場合, $(\log(\text{bid}) + \log(\text{bid} + 1))/2$ から $(\log(\min) + \log(\min + 1))/2$ までの価格幅を単に「幅」と呼ぶことにする. 買い気配板の場合も同様である.

このとき, 本来なら全価格軸上で α を推定したいが, データの都合上そのようなことはできないので, たまたまある幅のところだけがデータに残っていて, データに残らなかったところは残ったところと同じように指値注文が到着していたのだ, と考える. 図

15 は、幅の時間変化のイメージ図である。横軸が $\log(\text{価格})$ を表し、縦軸が約定タイムを表している。幅の外のイベントに関してはデータが無いので、 α の計算に寄与しない。つまり灰色に塗られた部分の面積 A の内部に到着した指値注文量をカウントし、それを A で割れば α を求めることができる。 $A = \sum (\text{幅} \times \text{指値注文が発生しなかった時間})$ なので、 α は (ある期間中の、把握できた全指値注文株数) / A となる。ここでゼロインテリジェンスモデルでは売りと買いは対称なので、売りと買いの指値注文を区別無くまとめて注文量や A を計算する。そしてある銘柄の α は、データブロックごとに求めた α の約定タイムによる重み付き平均によって求める。

C.3 δ_s, δ_c の推定法

本研究で用いたデータでは個々の注文の行方は追えないため、注文ごとにその寿命を求めることは不可能である。そこで、各キャンセルは独立に生じ板は定常状態にあると仮定して δ_c, δ_s を求めることにする。

本文中で δ_c は単位時間あたりにある注文がキャンセルされる確率であるとした。一方、Smith et al. (2003) では、 δ_s はある指値注文が最終的に約定しなかった時の、その注文の寿命の逆数と定義している。両者は異なる値をとる。

以下では、時刻 1 にデータブロックが始まり時刻 T にデータブロックが終わるとする。つまりデータブロック中 T 回成行注文が到着する。また、以下で扱う量の中に単位が株数のものがあれば、そのような量は全て、データブロック中のデータから求められた σ で規格化する。

まず、 δ_c を約定・気配データから求める方法を考える。時刻 t における幅内部の指値注文量を N_t 、幅内部に到着する指値注文の到着率を α' ¹⁰、成行注文到着率を μ' ¹¹ とすると、

$$N_t - N_{t-1} = \alpha' - \mu' - \delta_c N_{t-1} .$$

$t = 2$ から T まで辺々加え合わせると、

$$N_T - N_1 = (\alpha' - \mu')(T - 1) - \delta_c \sum_{t=1}^{T-1} N_t . \quad (4)$$

したがって、

$$\delta_c = \frac{(\alpha' - \mu')(T - 1) - (N_T - N_1)}{\sum_{t=1}^{T-1} N_t} = \frac{(\alpha' - \mu') - \frac{N_T - N_1}{T-1}}{\langle N \rangle} . \quad (5)$$

$\langle N \rangle$ は板上の指値注文量の時間平均である。こうしてデータブロック中における δ_c を求めることができた。

¹⁰ α' がついている理由は、 α は単位価格あたり単位時間あたりの指値注文到着率であり両者を区別するためである。

¹¹ 本研究では $\mu = \sigma$ であるが、それは有限期間に対しても成り立っているとは限らず、一般には μ の値はゆらいている。そこでデータブロック中における成行注文量を T で割ったものを μ' とおく。

次に δ_s を求める方法について考える．ある時刻 $t-1$ から t までの間の指値注文到着量は α' ，成行注文到着量は μ' ，注文キャンセル量は平均的に $\delta_c N_{t-1}$ に比例する．したがってある時刻 $t-1$ から t までの間に板上の注文が減る（成行注文が到着するかキャンセルが生じる．）回数と板上の注文が増減する（指値・成行注文が到着するかキャンセルが生じる．）回数の比は

$$\frac{\mu' + \delta_c N_{t-1}}{\alpha' + \mu' + \delta_c N_{t-1}}$$

で与えられる．ここで，注目している指値注文がランダムに選択され，約定またはキャンセルにより消滅すると仮定する¹²．その注文が時刻 $t-1$ から t までの間に板上から消滅する確率 P_t は

$$P_t = \frac{\mu' + \delta_c N_{t-1}}{\alpha' + \mu' + \delta_c N_{t-1}} \frac{1}{N_{t-1}} \approx \frac{\mu' + \delta_c \langle N \rangle}{\alpha' + \mu' + \delta_c \langle N \rangle} \frac{1}{\langle N \rangle}$$

程度である．ここで (4) 式を思い出すと

$$\delta_c \langle N \rangle = \alpha' - \mu' - \frac{N_T - N_1}{T-1}.$$

したがって，

$$P_t = \frac{\alpha' - \frac{N_T - N_1}{T-1}}{2\alpha' - \frac{N_T - N_1}{T-1}} \frac{1}{\langle N \rangle} (\equiv P)$$

となり， P_t は t によらない．すると，ある指値注文がちょうど時刻 t でキャンセルされる確率は，その指値注文は時刻 $t-1$ まで板上に存在し時刻 t でキャンセルされるということから，

$$(1-P)^{t-1} \delta_c$$

に比例する．したがって，ある指値注文が時刻 0 に板に到着したとすれば，その注文がキャンセルされるまでの平均寿命 δ_s^{-1} は

$$\delta_s^{-1} = \frac{\sum_{t=1}^{\infty} t(1-P)^{t-1} \delta_c}{\sum_{t=1}^{\infty} (1-P)^{t-1} \delta_c} = \frac{1}{P}$$

ゆえに $\delta_s = P$ である．

ある日ある場の δ ($= \delta_c, \delta_s$) はデータブロックごとに求めた δ を，約定タイムで重みを付けた平均である．そしてある銘柄の δ は，場ごとの重み付き平均のさらに重み付き平均とした．これは α の場合と同じである．ただし (5) 式によると δ_c が負になることもありうる．そのようなデータブロックは除いて計算を行った．

¹²この仮定は正しくない．なぜなら指値注文執行は優先順位が決まっており指値注文はランダムに選択されて約定するわけではないからである．しかし近似としては十分である．

C.4 スプレッド, ボラティリティ, λ の推定法

アスクとビッドの差をスプレッドという。ゼロインテリジェンスモデルでは価格是对数をとっているので, スプレッドは $\log(\text{ask}/\text{bid})$ である。そこでスプレッドの実測値はあるアスク・ビッドの組が存在した時間 w で重みをつけて次の式で与える。

$$\text{actual spread} = \frac{\sum w \log \frac{\text{ask}}{\text{bid}}}{\sum w}.$$

スプレッドの場合も α, δ と同様に約定タイムで重みを付けて平均をとった。スプレッドの予測値も同様に求める。

価格時系列がランダムウォークに従うとした時に τ 時間後の価格分散は τ に比例する。その比例係数をボラティリティという。本研究では次のようにボラティリティを求めることにする。価格 $p = (\log(\text{ask}) + \log(\text{bid}))/2$, $\tau = (\text{p が変化しなかった時間})$ とし, 最小二乗法を用いると, 次のようになる。

$$\text{actual diffusion rate} = \frac{\sum \langle (\Delta p)^2 \rangle \tau}{\sum \tau^2}.$$

ボラティリティの場合も約定タイムで重みを付けて平均をとった。ボラティリティの予測値も同様に求める。

最後に λ の推定法について述べる。1 単位の注文が平均してどのぐらい価格を動かすかを示す量を λ という。ただしゼロインテリジェンスモデル及びローインテリジェンスモデルでは注文は常に σ 株ずつ行われるとするので, λ の定義もそれに合わせて σ 株の注文が平均してどのぐらい価格を動かすかを示す量ということにする。まず売り板に注目して考える。今, n 株の成行注文が到着したと仮定する。このときアスクがどのぐらい移動するかは気配データからわかる。その新しく出現したアスクを p とする。次に各 n に対し, $\Delta p(n) (\equiv \log(p) - \log(\text{ask}))$ を求め平均をとる。さらに $\Delta p(n) = \lambda_1 n^\kappa$ の形を仮定し, 対数をとって回帰を行い, λ_1, κ を求める。ゼロインテリジェンスモデルは売り板・買い板が対称なモデルなので, この操作は売り板・買い板区別無くまとめて行うことにする。

回帰の範囲は以下のように決める。本研究で扱うデータは 5 本値なのであまりに大口の成行注文が到着するとしたときには板上の指値注文は無くなってしまい, アスクはどこに出現するか分からない。そこで, ある銘柄の λ を計算する際には, 売り板・買い板通して, 板上の指値注文が最も少ないときを見つける。そのときの板上の指値注文数を N_{\min} とおく。すると N_{\min} 株までの大口成行注文なら板上に指値注文が存在するのでアスクを算出できる。そこで回帰の範囲は N_{\min} とする。 $N_{\min} < 3$ となってしまうような銘柄は回帰分析が不可能なので本研究の対象外とする。こうして残った銘柄が約 760 銘柄である。

そして単位株数を u とすると λ は $\lambda_1 (\sigma/u)^\kappa$ により求めることができる。モデルによる予測値も同様に求めることができるが, 常に σ 単位で注文は到着するので, $\sigma/u = 1$ とする。つまり予測値の場合は λ_1 が λ そのものである。

こうして $\sigma, \mu, \alpha, \delta_s, \delta_c$ の各パラメータ, スプレッド, ボラティリティ, λ の実測値を計算することができるようになった。

(木村 博道: 筑波大学大学院システム情報工学研究科)

(秋山 英三: 筑波大学大学院システム情報工学研究科)

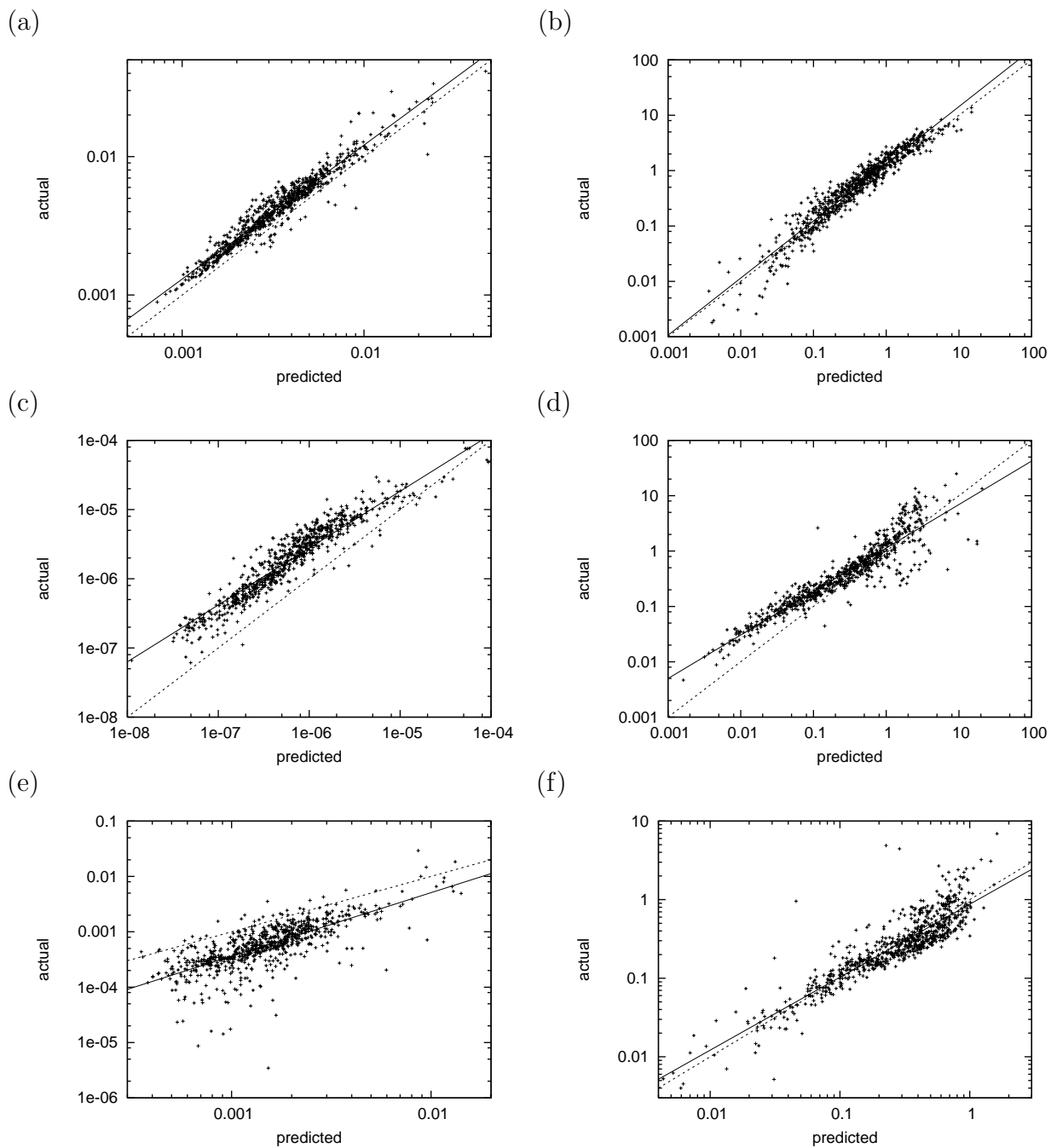


図 11: (a)(c)(e) はそれぞれスプレッド, ボラティリティ, λ の実測値とモンテカルロ
 ゼロインテリジェンスモデルによる予測値を銘柄ごとにプロットしたものである. また (b)
 (d)(f) はそれぞれ各指標の実測値とローインテリジェンスモデルによる予測値を銘柄ごと
 にプロットしたものである. 横軸が予測値, 縦軸が実測値を表す. 点線はモデルの予測と完全
 に一致したときに乗る線, 実線は回帰直線である.

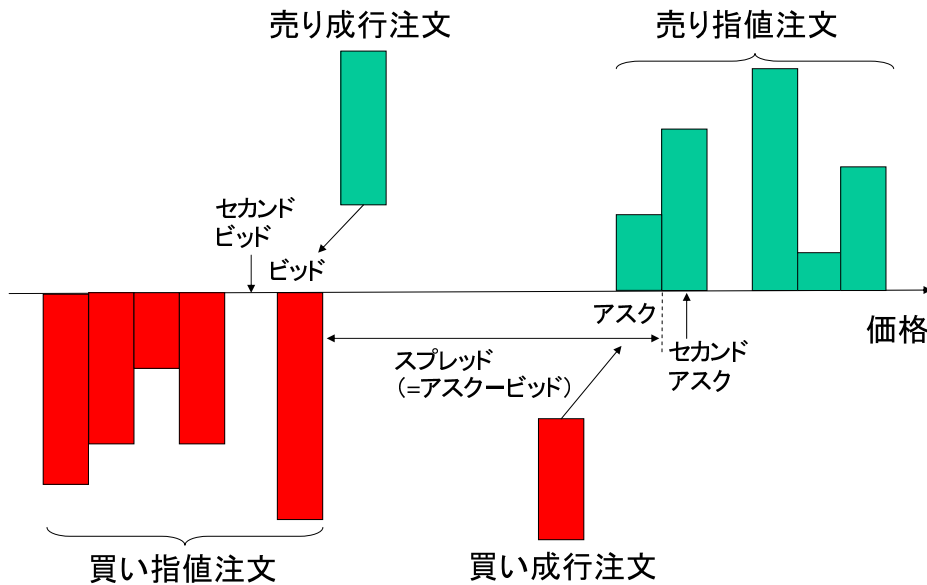


図 12: 板の例．横軸が価格，縦軸がその価格で注文された指値注文の量である．

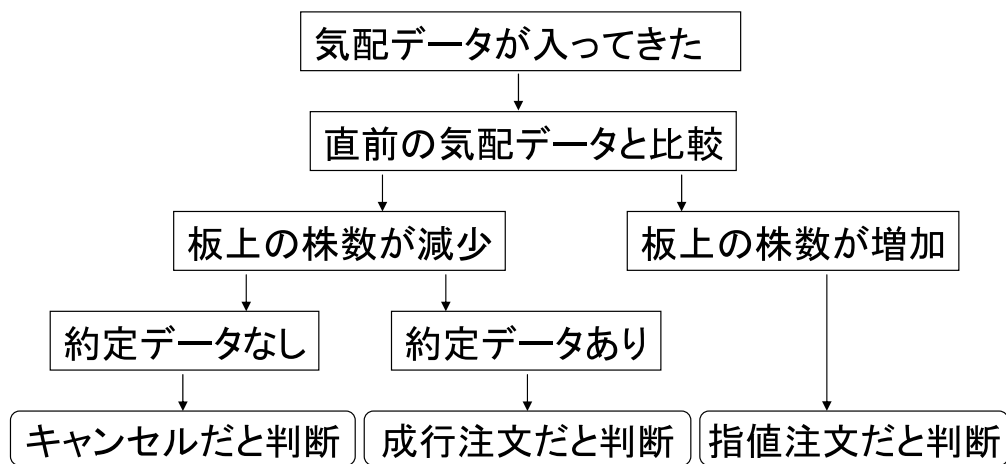


図 13: 気配データと約定データからどのような注文が入ったのかを判定するためのフローチャート．図中、「約定データあり/なし」は、現気配データと直前の気配データの間約定データがあるかないかを意味する．

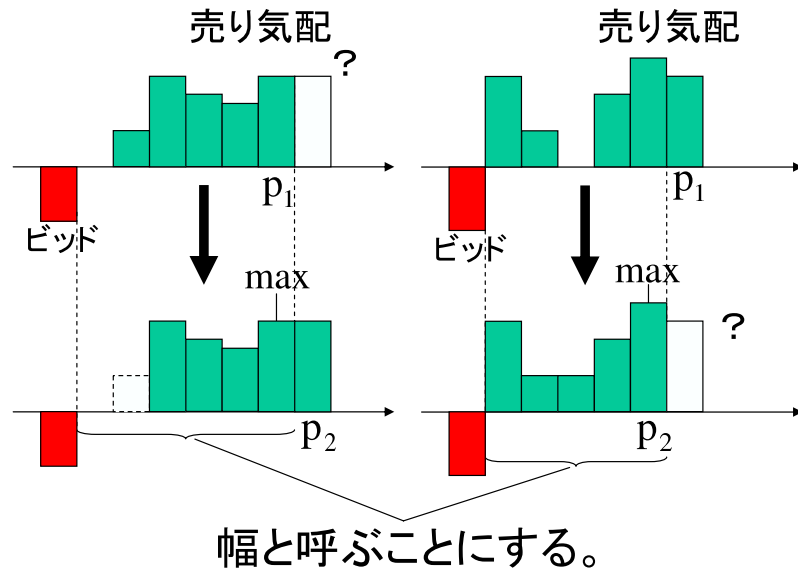


図 14: 気配データの変化の例と「幅」の定義。

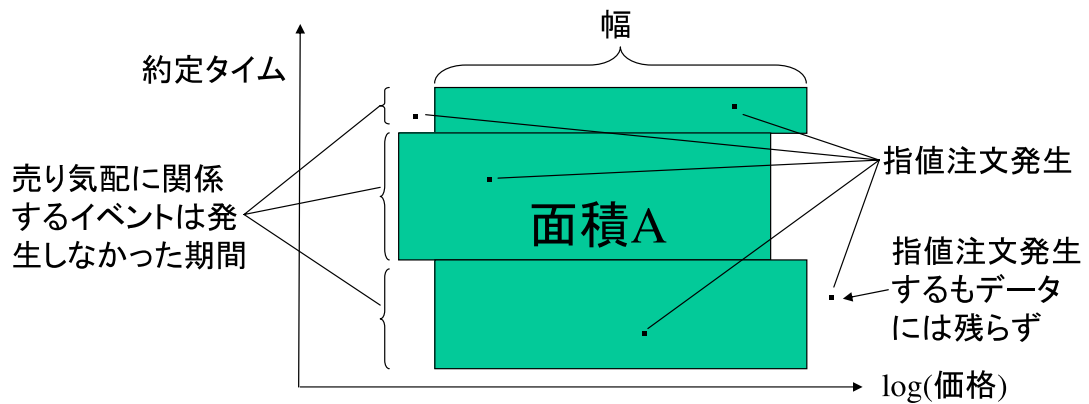


図 15: 幅の変化の例と面積 A の定義。