

市場流動性を説明できるローインテリジェンスモデル

木村 博道 秋山 英三
筑波大学

(受理 2008年9月11日; 再受理 2009年4月14日)

和文概要 古典的な経済学では、通常、投資家は完全合理的な行動を行うと仮定する。そのような仮定をおいた理論では、市場均衡の存在を証明することが可能であるが、現実の投資家にとって重要な、スプレッド（アスクとビッドの差）やマーケットインパクト（注文に対して価格がどのくらい変動するか）を説明することは難しい。その一方で、近年、投資家の合理性をまったく仮定せず、注文はポアソン到着に従って到着すると仮定する「ゼロインテリジェンスモデル」が提案されている（Smith et al. 2003）。このモデルでは、スプレッドや価格の拡散係数を定量的に説明することが可能である（Farmer et al. 2005）。しかしながら、ゼロインテリジェンスモデルでは、マーケットインパクトを表現する量の一種である、Kyle の λ （単位株数の注文に対して平均してどのくらい価格が変動するか）を説明することはできなかった。本研究では、まず、東京証券取引所の実データを用いて、指値注文分布や直前の注文に依存して注文到着率がどのように変動するかなどの性質を調べた。次に、ゼロインテリジェンスモデルにそれらの性質を導入し、シミュレーションを行った。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。その結果、ローインテリジェンスモデルでは、スプレッドや拡散係数に加えて、Kyle の λ をも説明できることが分かった。このことは、金融市場の流動性を理解するには、注文と注文分布の相互作用、あるいは注文同士の相互作用が重要である可能性を示唆する。

キーワード: シミュレーション, ローインテリジェンスモデル, ゼロインテリジェンスモデル, Kyle の λ , スプレッド, 価格の拡散係数

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、5本気配値データ（データの構造については付録B.3を参照。）と呼ばれている、指値注文がどのように分布しているかについてのデータを元にして、注文到着率などの市場の状態を表現するパラメータの推定方法を提案し、それらのパラメータを用いた人工市場モデルによって各種流動性指標を説明することである。

流動性は、特に実務家にとって重要な量である。彼らは自分の注文行動によりアスク（ビッド）がどのく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くらいアスク（ビッド）が動くかを意味する関数であり、流動性指標である「Kyle の λ 」を一般化したものである。ここで本研究における 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 [8, 13] である。マーケットインパクトは投資家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 [5]。

一方、Smith らは、流動性指標の一つであるスプレッド（アスクとビッドの差）などを定量的に説明できるモデルを提案している [15]。Farmer らは、ロンドン証券取引所の 11 銘柄を対象にスプレッドと価格の拡散係数（以下、拡散係数）を計測し、Smith らのモデルの説

明¹ 能力がかなり高いことを示した [3]。Smithらのモデルでは、注文はポアソン到着すると仮定する。このように、投資家は何も考えていないかのように振る舞い、結果として注文がポアソン到着すると仮定したモデルを「ゼロインテリジェンスモデル」と呼ぶ。注文がポアソン到着することを仮定したモデルはすでに多数存在する [2, 11, 16] が、これらはスプレッドなどを定量的に説明することを特に目指した研究ではない。また、Smithらのモデルから影響を受けた研究としては Mike & Farmer [12] の研究が挙げられる。この研究は注文同士の相互作用やキャンセルが生じる確率をモデル化することにより、スプレッドの大きさの分布関数などを説明しているが、マーケットインパクトについては言及していない。Slanina [14] は注文駆動方式を用いた仮想市場における価格時系列の研究についてレビューを行っている。このように、ポアソン到着を仮定し市場の様々な性質を理解しようとする研究は活発に行われている。

Smithらのモデルではスプレッドについては説明できるものの、マーケットインパクトはあまり説明できない。そこで、本研究では、スプレッド、拡散係数に加え、小口注文に対するマーケットインパクト、すなわち Kyle の λ の説明を目指して、Smithらのモデルを拡張する。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なので、以下で概略を述べる。

1.2. Smithらのゼロインテリジェンスモデルの概略

Smithらは、注文駆動方式の市場（用語については付録 A，図 8 にまとめた。）を用いて、注文がポアソン到着するという仮定から、スプレッド、拡散係数などを解析的に求めることができることを示した。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、現実の取引価格の自然対数に相当し、任意の実数値をとることが可能である。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²。
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様に到着する。単位価格・単位時間あたりに到着する指値注文量は一定の値 α をとる。
3. 成行注文は一定の到着率 μ で到着する³。また、成行注文が売り注文になるか買い注文になるかは、 $1/2$ の確率で決まる。
4. 一注文は、指値・成行、売り・買いに関わらず一定の株数 σ からなる。
5. キャンセルは一定の割合でランダム・独立に行われる（ポアソン到着）。指値注文のキャンセル率は δ とする。

以上の仮定のうち、4. と 5. はローインテリジェンスモデルでも踏襲する。

1.3. ローインテリジェンスモデル

伝統的な経済学では、投資家は予算の範囲で効用を最大化するように行動するとされてきた。その一方で、ゼロインテリジェンスな投資家を仮定すると、スプレッドや拡散係数を説

¹ここで、「説明」とは回帰分析の「説明」変数と同様の意味である。すなわち、あるモデル（回帰式やシミュレーションモデル）を用いてスプレッドや拡散係数を計算し、実際の値と比較したという意味である。本研究でもこの意味で説明という語を用いる。なお、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。

²Smithらは呼値の刻み幅（価格変動の最小単位）が有限である場合も考えているが結果にはほとんど影響しない。

³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 [3, 15] 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。

明であることが分かった。しかし、マーケットインパクトを説明することはできなかった。

そこで、マーケットインパクト（本研究では、特に Kyle の λ ）を説明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 3.1.1, 3.1.2 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可能性がある。ここで厚みとはある呼値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。また、単に「厚み」と言った場合にはアスクまたはビッドの厚みを指す。

本研究では以下のインテリジェンスを導入した。

- 成行注文については（3.1.1 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（3.1.2 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「アスク +1（ビッド -1）」の厚みが厚いと、アスク（ビッド）に売り（買い）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。

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。ただし、ゼロインテリジェンスモデルでは価格は現実の取引価格の対数（実数値）をとるとしていたが、ローインテリジェンスモデルでは、以上のようにモデルを拡張するため、価格は整数値（実際の市場価格を呼値の刻みで割った値）をとることとする。そうすることによって、上述の指値注文分布と注文の相互作用を自然に組み込むことができる。

ローインテリジェンスモデルが Kyle の λ （以下、 λ ）、スプレッド、拡散係数をどのくらい説明できるかを調べるため、モンテカルロシミュレーションを行った。図 1 はその手法を図示したものである。我々は、まず、東証のデータから、 λ 、スプレッドと拡散係数の実測値、モデルに必要なパラメータを求めた。パラメータは、全て、5 本気配値という市場の状態の部分的な情報から推定することができる。その一方、各指標のモデル値を求め、実測値と比較した。ただし、ローインテリジェンスモデルは複雑なモデルであるため、シミュレーションによってモデル値を出力せざるを得ない。そこで、ゼロインテリジェンスモデルのモデル値もシミュレーションにより求めることにより、ローインテリジェンスモデルと比較した。

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 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、拡散係数については高い説明力を持っているが、マーケットインパクトについてはあまり説明できないという結果を得た [3]。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。

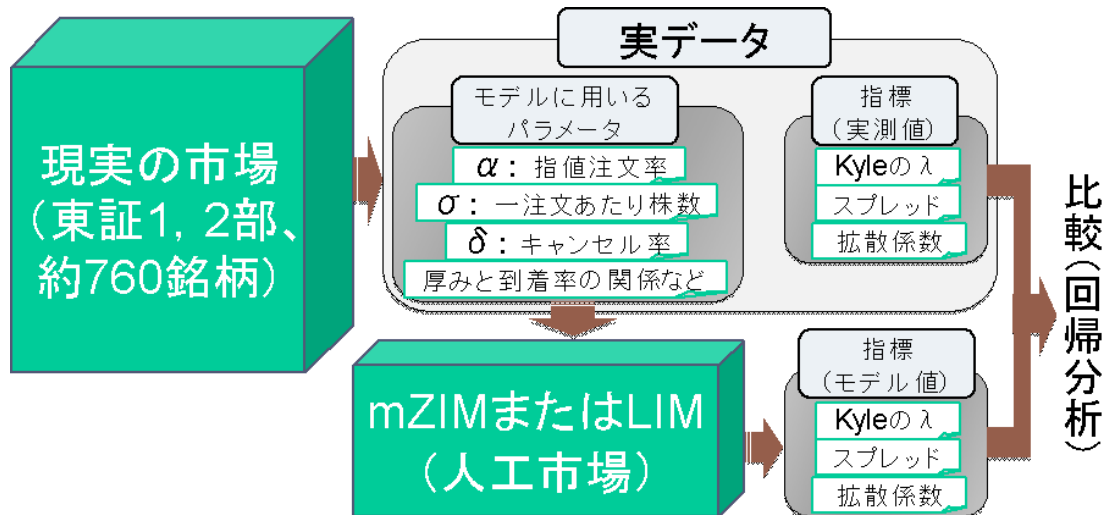


図 1: ローインテリジェンスモデルの検証法の概略

2.1. 手法

対象銘柄と対象期間

本研究で扱った銘柄は約 760 銘柄である。銘柄選定基準についての詳細は付録 B に記載した。概略を述べておくと、指値注文到着率などのパラメータ、 λ 、スプレッド、拡散係数を求められる程度の取引があり、かつ、株式分割・会社合併のようなモデルの考慮外の事象が起きていない銘柄を対象とした。本研究の対象期間は 2003 年 10 月 30 日から 2006 年 2 月 28 日までである。

Smith らのモデルとモンテカルロゼロインテリジェンスモデルの相違点

Smith らのモデルは解析的にスプレッドや拡散係数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 λ 、スプレッド、拡散係数を求めることにした。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。このモデルを特に、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。

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価格・単位時間あたり一定であることを仮定している。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限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市場に指値注文が多数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.3 で述べる「スコープ」—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。

w を有限と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、単位時間に指値注文は有限個しか到着しないので、市場に売り（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。

モンテカルロゼロインテリジェンスモデルのアルゴリズム

シミュレーションにおける時間は注文が約定するたびに +1 することにより測られる。これを約定タイムと呼ぶことにする。

シミュレーションを行うにあたり、最初に、 $N = \lceil \alpha w / \sigma \delta + 1 \rceil$ 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。 N は、Smith らのモデルでは、投資家は常に σ 株ずつ取引を行うという仮定と、アスク、ビッドから十分離れた領域には単位価格あたり α / δ 株の指値注文が存在することから導くことができる。ただし、最低一つの売り指値注文を置くために、1 を足している。また、 $[x]$ は x を超えない最大の整数値を表す。こうすると、単位価格あたり平均して $\alpha / \sigma \delta$ 個程度の売り指値注文を置くことができる。買い指値注文についても同様である。

次に、実際に注文を発生させ、取引を行わせる。そのために、ここで、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の生起率を求めておく。イベント生起率が分かれば、以下の要領で各イベントが生じる確率も求めることができる。 n 種のイベント X_1, X_2, \dots, X_n を考え、その単位時間あたり生起率はそれぞれ $\pi_1, \pi_2, \dots, \pi_n$ であるとする。イベント $X_i (i = 1, \dots, n)$ 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、いずれかのイベントが必ず生じるものとする $\sum_{j=1}^n \pi_j = 1$ である。以上より、イベント $X_i (i = 1, \dots, n)$ が生じる確率 p_i は $\pi_i / \sum \pi_j$ により求めることができる。

では、指値注文の生起率について考察する。 α の定義から、 σ 株の売り指値注文が bid から bid + w までの領域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。買い指値注文についても同様に、ask から ask - w までの領域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、成行注文の生起率について、 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 (C.1 節を参照)，買い(売り)成行注文生起率は $1/2$ となる。

最後に、キャンセルの生起率について述べる。モンテカルロゼロインテリジェンスモデルでは、 δ はある指値注文の単位時間あたりキャンセル率である。現在の全指値注文総数を 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 $n\delta$ に比例する。

以上の考察をまとめると、注文の時間発展を決めるアルゴリズムは次のようになる。全イベントの生起率 $2\alpha w / \sigma + 1 + n\delta$ を P 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 の領域に一様・ランダムにポアソン到着する。
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 の領域に一様・ランダムにポアソン到着する。
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 $n\delta / P$ 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。

λ , スプレッド, 拡散係数の計算は、成行注文が $[10/\delta + 1]$ 回到着してから、行い始めるものとする。この回数は Smith らのモデルを参考にして決めた。

シミュレーションは、成行注文がさらに 10000 回到着した後に終了するものとする。その後、 λ 、スプレッド、拡散係数を計算する。その計算法は付録 C.5 に記載した。

結果の出力

全ての銘柄についてシミュレーションと λ 、スプレッド、拡散係数の計算が終わったら、それらの実測値とモデル値が一致したかどうかを見るために回帰分析を行う。[3]によると、回帰係数の誤差は通常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こと。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、本研究のデータ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傾向がある。

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、モデル値 (predicted values) が実際の市場から得られた値 (actual values) とどの程度一致しているのかを検証する。そのために λ 、スプレッド、拡散係数の各指標に対し、 $\log(\text{actual values}) = A \times \log(\text{predicted values}) + B$ 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。モデル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。

図 11 (a) は各銘柄に対し、 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モデル値がどの程度一致しているかを示した図である。横軸がモデル値、縦軸が実測値を表し、各 + が一つの銘柄を示している。また、点線は実測値とモデル値が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 $A = 1.147 \pm 0.038$ 、 $B = 0.001 \pm 0.247$ 、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モデル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。つまり Smith らのモデルは λ をあまり説明できないことが分かる。

図 11 (c) は、スプレッドについて同様に示したものである。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説明できていることが分かる。

図 11 (e) は拡散係数について同様に示したものである。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。モデル値は実測値よりも小さすぎるとはいえ、モデル値と実測値は傾向としては一致していることが分かる。

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り、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。したがって、高流動性銘柄に対して λ の説明力が低いことはモデルに改良の余地があることを示唆すると考えられる。

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 らによると、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 $A = 0.99 \pm 0.10$ 、 $B = 0.06 \pm 0.29$ 、 $R^2 = 0.96$ 、拡散係数で $A = 1.33 \pm 0.25$ 、 $B = 2.43 \pm 1.75$ 、 $R^2 = 0.76$ であった [3]。したがって、LSE においてはある程度の精度でスプレッドと拡散係数を説明できることが分かっている。

一方，東証で取引されている 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説明できていることが分かった．本研究で扱った東証のデータと 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，拡散係数に関しては（モンテカルロ）ゼロインテリジェンスモデルのモデル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 についてはあまり高い精度では説明できない．特に，図 11 (a) 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 のモデル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，説明力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， λ ，スプレッド，拡散係数に対して高い説明力を持っていることを示す．

3.1. ローインテリジェンスモデルの詳細

ローインテリジェンスモデルでは，価格は対数をとらず整数値で考える．市場は注文駆動方式より取引が行われ，価格が決定する（モンテカルロ）ゼロインテリジェンスモデルでは価格は実数値をとり呼値の刻み幅は 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．

3.1.1. 成行注文到着パターン

以下では投資家の振る舞いを具体的に見ていく．まず，東証の実データを用いて，成行注文の到着パターンを調べた結果について述べる．

厚みと成行注文到着率

約定タイムを用いると成行注文到着率 μ は σ に等しい（付録 C.1 節参照）．しかし，実際の市場では μ はゆらいでいる．そこでどのような状況で μ が大きくなるのかを考える．

例えば，図 2 (a) のようにビッドに指値注文が多数存在する状況下で，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 2 (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傾向があるかもしれない．逆に，図 2 (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いほど，買い成行注文は増加す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である．

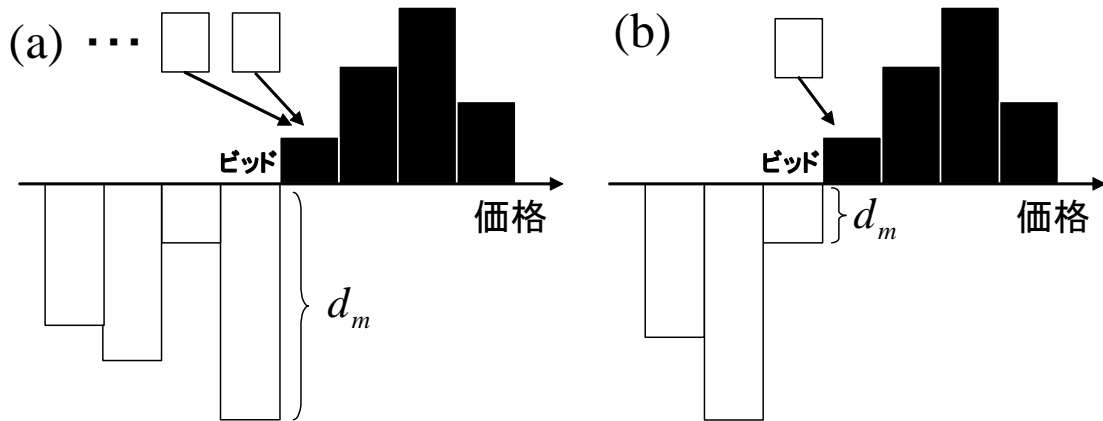


図 2: ビッドに買い指値注文が (a) 多数あるときと (b) そうでないときの、成行注文到着率の違い

図 3 (a) はキッコーマン (銘柄コードは 2801) (b) は NTT ドコモ (同 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。前者は流動性が低い銘柄の、また後者は流動性が非常に高い銘柄の典型例である。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 (3.2) 式を用いて回帰したときに得られる線である。

図 3 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、図 3 (b) では減少する傾向さえ観察される。

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⁴。 μ_{d_m}/σ を y 、 d_m/σ を x とおき、 A_m 、 $\gamma_{1,m}$ 、 $\gamma_{2,m}$ 、 $\gamma_{3,m}$ を回帰分析によって求める係数とすると

$$y = A_m \exp \left[\gamma_{1,m} x^{0.5} + (\gamma_{2,m} + \text{err.}) x + \gamma_{3,m} x^{1.5} \right]. \quad (3.2)$$

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、

$$\frac{\log(y)}{x} = \frac{\log(A_m)}{x} + \gamma_{1,m} x^{-0.5} + \gamma_{2,m} + \text{err.} + \gamma_{3,m} x^{0.5}.$$

そこで、銘柄ごとに $\log(y)/x$ を $1/x$ 、 $x^{-0.5}$ 、 $x^{0.5}$ により回帰することにする。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 (3.2) 式は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てよいだろう。

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、その列は長期記憶を持つことが知られている [1, 9]。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。この性質をモデルに組み込むことを考える。

⁴次の式

$$\frac{\log(y)}{x} = \frac{\log(A)}{x} + \gamma_1 x^{-0.5} + (\gamma_2 + \text{err.}) + \gamma_3 x^{0.5} + \gamma_4 \frac{\log(x)^{0.5}}{x} + \gamma_5 \frac{\log(x)}{x} + \gamma_6 \frac{\log(x)^{1.5}}{x} \quad (3.1)$$

を仮定して赤池情報量規準 (AIC) を各銘柄に適用すると、最良のモデル (AIC が最小となるような回帰式) 中の項のうち、最も多くの銘柄に含まれていた項 (定数項を除く) は第 1 項である。また、2 番目、3 番目に多くの銘柄に含まれていた項は順に第 2 項、第 4 項であった。そこで、本研究では (3.2) 式を用いた。

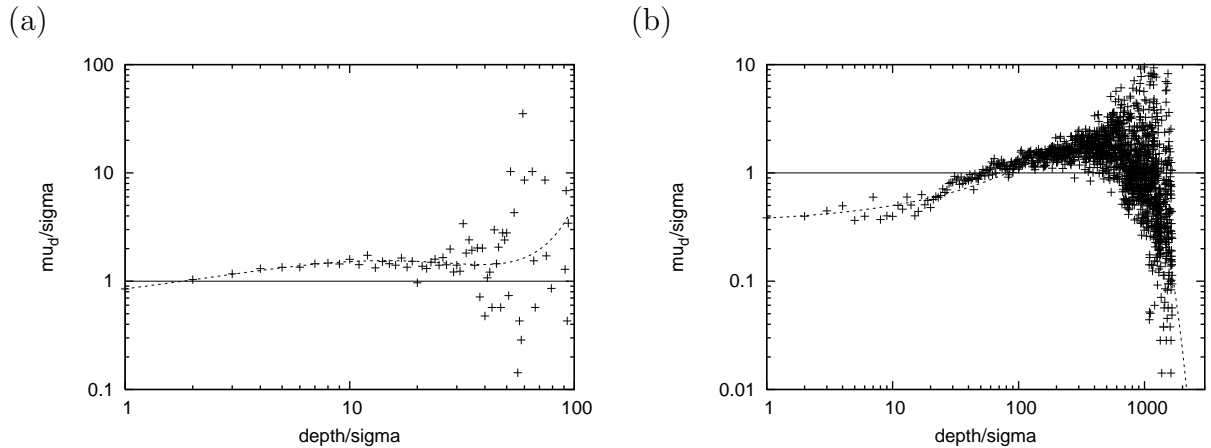


図 3: ある厚み d_m の下での、条件付成行注文到着率 μ_{d_m}

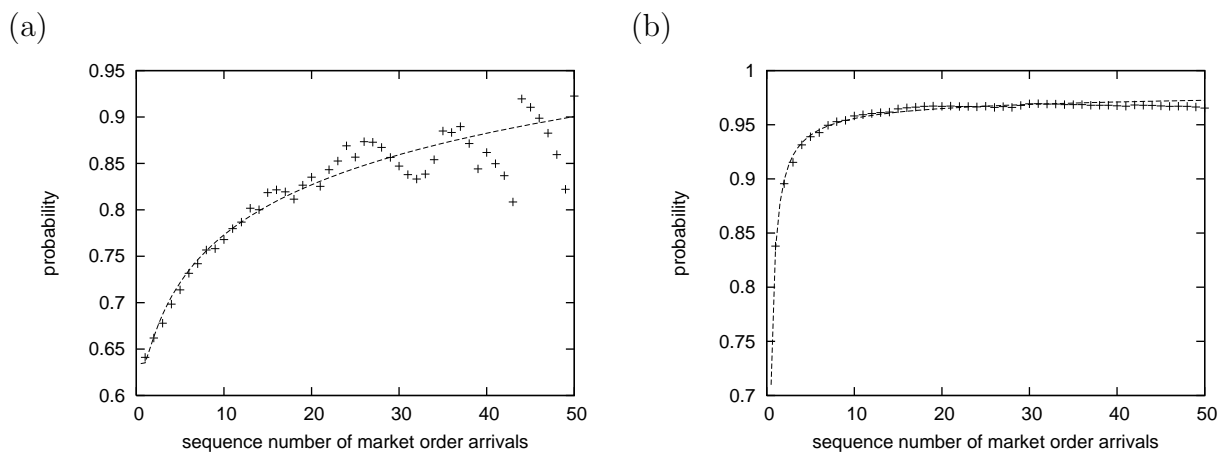


図 4: 売り（買い）成行注文が n 回連続して到着した後の成行注文が売り（買い）注文となる確率 P

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。そのような注文は指値注文分布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 [1, 9] では成行注文の符号のみを考慮していたのに対し、本研究では注文の大きさも見ることにする。つまり、データ上、 N 株の売り（買い）成行注文が到着していたとすると、モデル上では、 N を σ で割り四捨五入した値 n 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 4 は (a) はキックマン (2801) (b) は NTT ドコモ (9437) に対して、上述の意味で売り（買い）成行注文が n 回連続して到着したとき、そ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。点線は (3.3) 式を用いて回帰したときに得られる線である。これらの図から、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とが分かる。これは先行研究を支持する結果である。

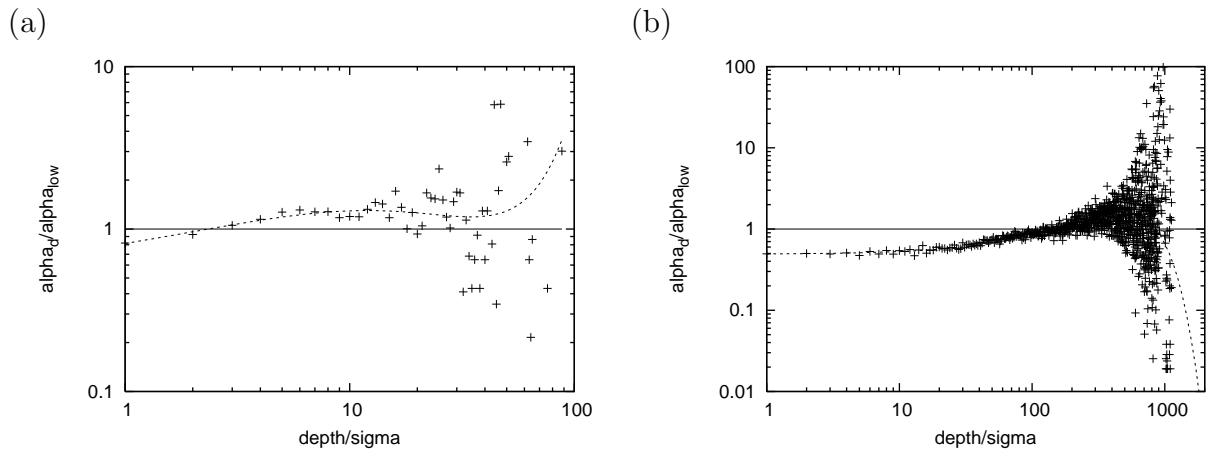


図 5: アスク (ビッド) に到着する指値注文の到着率と , アスク (ビッド) の厚みとセカンドアスク (セカンドビッド) の厚みの差 $d_l (> 0)$ の関係

この関係を表現するために次のような回帰式を用いた⁵ .

$$P(n) = a + b \log(n) + \frac{c}{n} . \quad (3.3)$$

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 4 中の点線は回帰結果を示したものである . 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 (3.3) 式は $P(n)$ を十分な精度で表現できていると考えてよいだろう .

ただし市場に買い (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) 成行注文到着率は 0 とする .

以上で導入されたパラメータ $A_m, \gamma_{1,m}, \gamma_{2,m}, \gamma_{3,m}, a, b, c$ は買い側・売り側の注文をまとめて計算し , 銘柄ごとに求める .

3.1.2. 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 . 1.3 節では , 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 . この性質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 .

⁵次の式

$$P(n) = a + b \log(n) + \frac{c}{n} + \frac{d}{n^2} + e \log(n)^2 + f \frac{\log(n)}{n}$$

を仮定して赤池情報量規準 (AIC) を各銘柄に適用すると , 最良のモデル (AIC が最小となるような回帰式) 中の項のうち , 最も多くの銘柄に含まれていた項は第 2 項である (定数項は除く) . また , 2 番目に多くの銘柄に含まれていた項は第 3 項であった . そこで , 本研究では (3.3) 式を用いた .

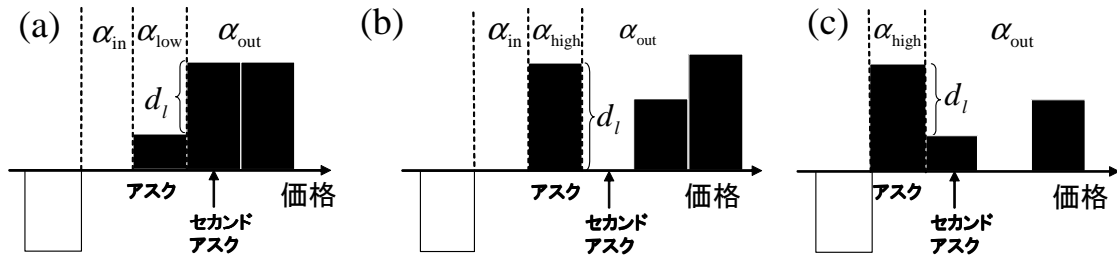


図 6: d_l , α_{in} , α_{low} , α_{high} , α_{out} の定義

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.3 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どの呼値にも一様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．

次に、「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様にポアソン到着すると仮定する．ただし 2.1 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「スコープ」(C.3 節で説明した「スコープ」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 6 である．(a)(b) ではスプレッドが 2 開いているため，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) $d_l > 0$ 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 $d_l \leq 0$ 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アスクに指値注文を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に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引量 (つまりアスクの厚み) が十分少なければ約定しやすくなるので，投資家はアスクに指値注文を行いやすくなる

かもしれない。

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、そうだとすると d_l が正に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5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 と指値注文到着率 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。点線は (3.4) 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。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。

この傾向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⁶。 $\alpha_{d_l}/\alpha_{low}$ を y 、 d_l/σ を x とおき、 $A_l, \gamma_{1,l}, \gamma_{2,l}, \gamma_{3,l}$ を回帰分析によって求める係数とすると

$$y = f(x) \equiv A_l \exp \left[\gamma_{1,l} x^{0.5} + (\gamma_{2,l} + \text{err.}) x + \gamma_{3,l} x^{1.5} \right]. \quad (3.4)$$

回帰は (3.2) 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 (3.4) 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。また、ここで新たに定義されたパラメータ $A_l, \gamma_{1,l}, \gamma_{2,l}, \gamma_{3,l}, \alpha_{in}, \alpha_{low}, \alpha_{high}, \alpha_{out}$ は買い側・売り側の注文をまとめて計算し、また銘柄ごとに求める。

少しずつ縮小するスプレッド

木村・秋山 [6] 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。そこで、本研究では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「アスク -1（ビッド +1）」に到着すると仮定する。

3.1.3. ローインテリジェンスモデルの定義

実データから分かる上述の傾向をモンテカルロゼロインテリジェンスモデルに導入した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。なお、このモデルのシミュレーションの方法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が、以下の3点が異なる。

1. 価格は整数値をとる。
2. 成行注文到着率を (3.2) 式 (3.3) 式に基づいて決定する。
3. 指値注文到着率を (3.4) 式などに基づいて決定する。

1点目は、ローインテリジェンスモデルでは、価格是对数をとらず整数値であるとし、呼値の刻みを単位とする点である。Smithらのモデルおよび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刻み幅は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い。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

⁶成行注文の場合と同様に (3.1) 式を仮定して赤池情報量規準 (AIC) を各銘柄に適用すると、最良のモデル (AIC が最小となるような回帰式) 中の項のうち、最も多くの銘柄に含まれていた項 (定数項を除く) は第1項である。また、2番目、3番目に多くの銘柄に含まれていた項は順に第2項、第4項であった。したがって、本研究では (3.4) 式を仮定する。

則を考慮しなければならない。これは、同呼値に到着している指値注文が2つ以上あった場合、先に到着した指値注文から約定するということを意味する。

そして、価格が整数値になったことに伴い、各パラメータ、 λ 、スプレッド、拡散係数を求めなおす必要がある。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各パラメータ、各指標を推定する方法とほぼ同じで、 $\log(\text{価格})$ となっているところを対数をとらずに、価格を呼値の刻みで割った値を用いて計算する。

成行注文到着率の決定

2点目は、成行注文は3.1.1節で実証した性質に基づいて到着するとした点である。以下ではこの点について詳述する。ローインテリジェンスモデルでは、成行注文到着率を次のように決める。

まず時刻 t (t は約定タイム、すなわち注文が約定するたびに +1 することにより測る。) 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。 $P(n)$ は(3.3)式により定義される。 P_t は確率なので、 $P(n) \leq 0$ または $P(n) \geq 1$ となった場合には、直前の P_{t-1} を用いるようにする。また、時刻0での確率 P_0 は $1/2$ と仮定する。

そして、 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

$$\mu_{\text{sell}} = A_m \exp \left[\gamma_{1,m} d_m^{0.5} + \gamma_{2,m} d_m + \gamma_{3,m} d_m^{1.5} \right] P_t$$

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

$$\mu_{\text{sell}} = A_m \exp \left[\gamma_{1,m} d_m^{0.5} + \gamma_{2,m} d_m + \gamma_{3,m} d_m^{1.5} \right] (1 - P_t)$$

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。

指値注文到着率の決定

3点目は、指値注文は3.1.2節で実証した性質に基づいて到着するとした点である。以下では、この点について詳述する。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。

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

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。 $f(\cdot)$ は(3.4)式によって与えられる関数である。買い指値注文到着率 $\alpha_{\text{buy}}(p)$ についても同様である。

ローインテリジェンスモデルのアルゴリズム

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述べる。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 でランダム・独立に生じる（ポアソン到着）とする。

各イベントは以下のように生成する。市場に存在する全指値注文数を n 、全イベント生成率 $\sum_p \alpha_{\text{sell}}(p) + \sum_p \alpha_{\text{buy}}(p) + \mu_{\text{sell}} + \mu_{\text{buy}} + n\delta$ を P 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 $n\delta/P$ 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。

3.2. 結果

本節では、 λ 、スプレッド、拡散係数に対するローインテリジェンスモデルのモデル値と実測値を比較する。

図 11 (b) は λ について、実測値と本モデルによるモデル値がどの程度一致しているかを示した図である。横軸がモデル値、縦軸が実測値を表す。各 + が一つの銘柄を示している。また、点線はモデル値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 $A = 0.929 \pm 0.016$, $B = -0.133 \pm 0.027$, 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 を説明できていることが分かる。

図 11 (d) は、スプレッドから 1 を引いた値について同様に示した図である。価格は整数値（市場価格を呼値の刻みで割った値）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 (d) ではスプレッドから 1 を引いた値を図示してある。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説明力を持っていることが分かる。

図 11 (f) は拡散係数について同様に示したものである。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、拡散係数についても、ゼロインテリジェンスモデルと同程度の説明力を持っていることが分かる。

3.2.1. 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、モンテカルロゼロインテリジェンスモデルでは λ を説明することは難しい。

一方、ローインテリジェンスモデルではスプレッド、拡散係数の決定係数はほとんど変わらず、 λ の決定係数は上昇した。表 1 は、 $\log(\text{実測値}) = A \times \log(\text{モデル値}) + B$ の形に回帰したときの B , A , 決定係数 R^2 を示したものである（）内は標準偏差である。

表 1: 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

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	λ	0.001(0.247)	1.147(0.038)	0.55
	スプレッド	0.046(0.053)	0.968(0.009)	0.94
	拡散係数	-1.443(0.158)	0.822(0.011)	0.88
LIM を東証で検証	λ	-0.133(0.027)	0.929(0.016)	0.81
	スプレッド	0.285(0.018)	1.033(0.010)	0.93
	拡散係数	0.127(0.024)	0.786(0.011)	0.87
Smithらのモデルを LSEで検証 [3]	スプレッド	0.06(0.29)	0.99(0.10)	0.96
	拡散係数	2.43(1.75)	1.33(0.25)	0.76

3.3. 議論

3.3.1. 市場制度と投資家行動

ゼロインテリジェンスモデルは注文がポアソン到着し注文のキャンセルも確率的に生じるといふモデルである。Smithらのモデルは、そのようなモデルの一つであり、スプレッド、拡散係数についてはうまく説明することができた。しかし σ 株の成行注文が到着したときにどのくらい価格が動くかを示す指標である λ についてはあまり説明できなかった。

そこで本研究では「指値注文分布と注文の相互作用」や「注文間相互作用」を記述する方法を考案しそれをモデルに組み込んだ。3節ではこのモデル—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 λ を説明できるかどうかを調べた。その結果、高い説明力を持つことが分かった。

Farmer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指値注文、成行注文およびキャンセルを入力データとして用いて、高い決定係数を得た[3]。彼らが用いているデータは、指値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本気配値しか存在しない。このことは指値注文到着率とキャンセル率の推定に影響を及ぼす。

本研究で用いたデータには以上のような制約があるにもかかわらず、ローインテリジェンスモデルは、スプレッド、拡散係数に加え、 λ をも説明できる。つまり、ローインテリジェンスモデルは、5本気配値データから求められたパラメータと流動性指標の関係を接続できるモデルとなっている。しかも本研究は約760銘柄と、[3]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、 λ 、スプレッド、拡散係数については「指値注文分布と注文の相互作用」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説明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)を説明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。アスク、ビッドに存在する指値注文量は λ に大きな影響を与える。指値注文が一様に到着するという仮定が非現実的であるために、Smithらのモデルは λ を説明できない

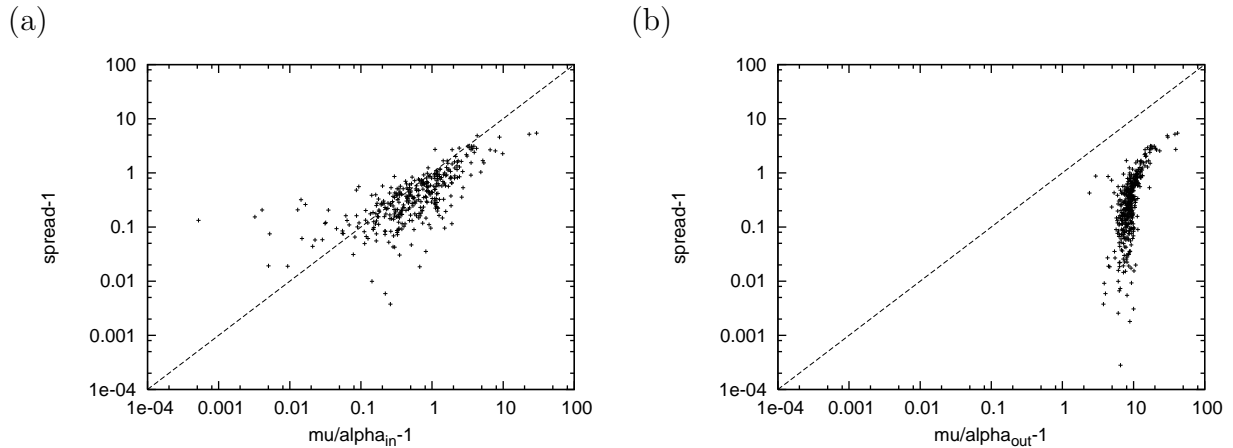


図 7: (a)「スプレッド-1」と「 $\sigma/\alpha_{in}-1$ 」(d)「スプレッド-1」と「 $\sigma/\alpha_{out}-1$ 」の関係

のである。

本研究で導入した各指値注文到着率はアスクやビッドからの距離に依存する。それらの到着率（特に α_{in} と α_{out} ）はスプレッドにどのくらい影響を与えるのかについて考える。図 7 は、Smith らのモデルによるスプレッドのモデル値の主要項（ $\approx \mu/\alpha = \sigma/\alpha$ ）から 1 を引いた値とスプレッドの実測値から 1 を引いた値を図示したものである。縦軸は「スプレッド-1」の実測値、横軸は (a) $\alpha = \alpha_{in}$ (b) $\alpha = \alpha_{out}$ に対して示した「スプレッドのモデル値の主要項-1」である。破線は $y = x$ の直線である。各グラフのスケールはそろえてある。(b) ではモデル値の値が大きく実測値から外れる一方 (a) では両者はそれほど大きくはずれないことがわかる。したがって、スプレッドの大きさにはアスク、ビッド付近に到着する指値注文が大きく影響することが分かる。また、 λ にも大きく影響することが推測できる。以上の詳しい議論は、[7] に記載した。

ところがそれを考慮しただけではまだ λ を説明できない。アスク、ビッドの指値注文到着率を実データに基づいた値にすると、特に流動性の高い銘柄で λ は小さすぎる値になってしまう。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っているのである。そこでローインテリジェンスモデルでは売り（買い）成行注文量はアスク（ビッド）の厚みに依存するとした (3.1.1 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。

こうして、 λ をうまく説明することができた。つまり、投資家の注文傾向は λ にかなり影響している。したがって、注文分布を考慮できる程度のインテリジェンス、すなわち、ローインテリジェンスな投資家を考察することが重要であることが分かった。

3.3.2. 本研究で導入した投資家行動の意味

以上のような投資家行動を考慮した結果、 λ をうまく説明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と考えられる [5]。もしもそうなら、売り（買い）成行注文量はビッド（アスク）の厚みに依存するとしたほうがよさそうである。ところが本研究から売り（買い）成行注文量はアスク（ビッド）の厚みに依存するとして λ を説明できることが分かった。

このことは次のように解釈できる。今、売り注文を確実に行いたいと考えている投資家がいるとする。そのためのもっともよい方法は売り成行注文を行うことである。成行注文を

行くと、市場に買い指値注文がある限り、確実に注文を実行できる。その代わりに、指値注文を行う場合と比較して安く売ることになってしまい、余分なコストがかかることになる。では、成行注文ではなく指値注文によって売ることにしたらどうだろうか。ここでは、投資家はアスクに指値注文を出すことを考えているものとする。また、スプレッドは1、すなわちアスクのすぐ左隣がビッドであるとする。すると、アスクに指値注文が多数あるときには、約定までの待ち時間が長くなったり、約定しないこともありうる。逆に、アスクに指値注文があまりないときには、そのような心配は少ない。したがって、投資家は、アスクに注文が多数あるときには、アスクに指値注文を行うよりも売り成行注文を出してしまうということが考えられる。つまり、ローインテリジェンスモデルが、売り（買い）成行注文量はアスク（ビッド）の厚みに依存するとして十分うまく説明できるのは、投資家の「早く確実に注文を実行したい」という行動を導入したからであると考えられる。

4. おわりに

本研究ではローインテリジェンスモデルを提案した。そして、指値注文分布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が明らかになった。

ローインテリジェンスモデルでは株式のファンダメンタルズや価格のトレンド、ニュース、あるいは投資家の効用などは考慮していない。にもかかわらず、 λ の説明力が高いということは、流動性を理解するにあたってファンダメンタルズなどの要素はあまり重要でなく、指値注文分布と注文、あるいは注文同士の相互作用を理解することが重要であることを示唆している。

謝辞

岸本一男先生（筑波大学大学院システム情報工学研究科教授）ならびに岸本研究室 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いた。両氏には深く感謝の意を表したい。なお、本研究は科研費（17103002,18700220）の助成を受けたものである。

参考文献

- [1] J.-P. Bouchaud, Y. Gefen, M. Potters and M. Wyart: Fluctuations and response in financial markets: the subtle nature of ‘random’ price changes. *Quantitative Finance*, 4 (2004), 176–190.
- [2] 遠藤操, 左士イ, 岸本一男: 2重待ち行列による日中株価変動のモデル化とその検証. 日本応用数学会論文誌, 16-3 (2006), 305–316.
- [3] J.D. Farmer, P. Patelli and I.I. Zovko: The predictive power of zero intelligence in financial markets. *Proceedings of the National Academy of Sciences*, 102-6 (2005), 2254–2259.
- [4] X. Gabaix, P. Gopikrishnan, V. Plerou and H.E. Stanley: A theory of power law distributions in financial market fluctuations. *Nature*, 423 (2003), 267–270.
- [5] K. Kim: *Electronic and Algorithmic Trading Technology: The Complete Guide (Complete Technology Guides for Financial Services)* (Academic Press, 2007).
- [6] 木村博道, 秋山英三: 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 (2007), SIG 19 (TOM 19).

- [7] 木村博道, 秋山英三: 市場流動性を説明できるローインテリジェンスモデル 補論. <http://abel.sk.tsukuba.ac.jp/~kimura/relations.pdf>
- [8] A.S. Kyle: Continuous auctions and insider trading. *Econometrica*, **53-6** (1985), 1315–1335.
- [9] F. Lillo and J.D. Farmer: The long memory of the efficient market. *Studies in Nonlinear Dynamics & Econometrics*, **8-3** (2004), article 1.
- [10] F. Lillo, J.D. Farmer and R.N. Mantegna: Master curve for price-impact function. *Nature*, **421** (2003), 129–130.
- [11] H. Mendelson: Market behavior in a clearing house. *Econometrica*, **50-6** (1982), 1505–1524.
- [12] S. Mike and J.D. Farmer: An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics and Control*, **32-1** (2008), 200–234.
- [13] 大村敬一, 宇野淳: 株価と売買高. 早稲田商学, **376** (1998), 31–60.
- [14] F. Slanina: Critical comparison of several order-book models for stock-market fluctuations. *European Physical Journal B*, **61** (2008), 225–240.
- [15] E. Smith, J.D. Farmer, L. Gillemot and S. Krishnamurthy: Statistical theory of the continuous double auction. *Quantitative Finance*, **3** (2003), 481–514.
- [16] L.-H. Tang and G.-S. Tian: Reaction-diffusion-branching models of stock price fluctuations. *Physica A*, **264** (1999), 543–550.

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．金融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（これを約定という）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情報，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8）．

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本気配値という．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 $\text{ask} + 1$ ($\text{bid} - 1$) の価格のことをさすことにする．

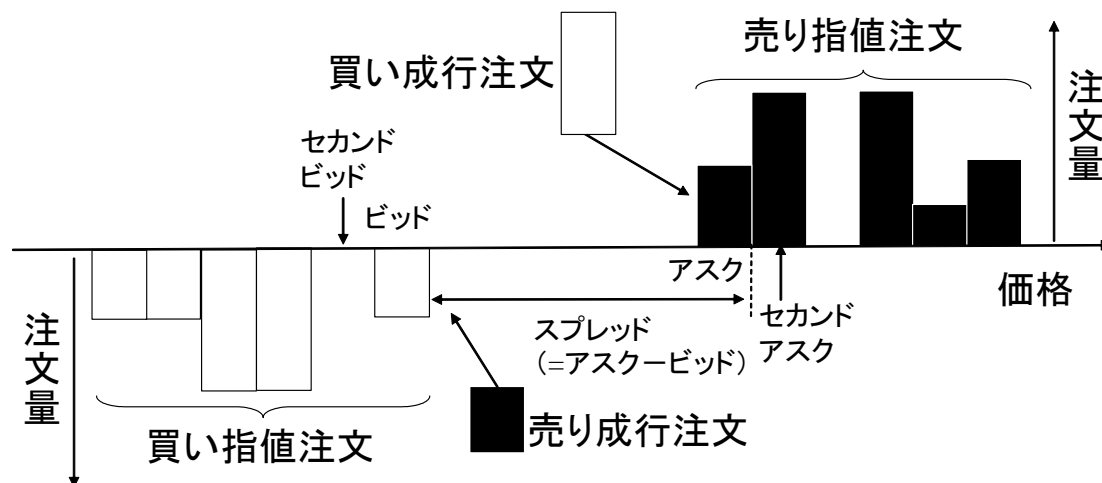


図 8: 板情報の例

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。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に影響するかどうか不明であるため、本研究では、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⁷。

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⁸。データは日経メディアマー

⁷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、「呼値の刻みが変わった銘柄は分析対象外」というルールによって除かれる。

⁸厳密にはモンテカルロゼロインテリジェンスモデルでは760銘柄、ローインテリジェンスモデルでは759銘柄

表 2: 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

表 3: 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	売り板情報か 買い板情報か
20031030	0900	1	399	42	0

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

本研究の対象期間は2003年10月30日から2006年2月28日である。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及びロー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。データブロックの統計量は[7]に記載した。

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2）。もう一つは存在する指値注文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3）。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本気配値という。したがって、実際の気配データは、一つの気配データにつき売り板情報・買い板情報合わせて高々10行からなる。

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

柄である。銘柄数が異なるのはローインテリジェンスモデルで用いられるパラメータを求めることができる銘柄は少ないからである。

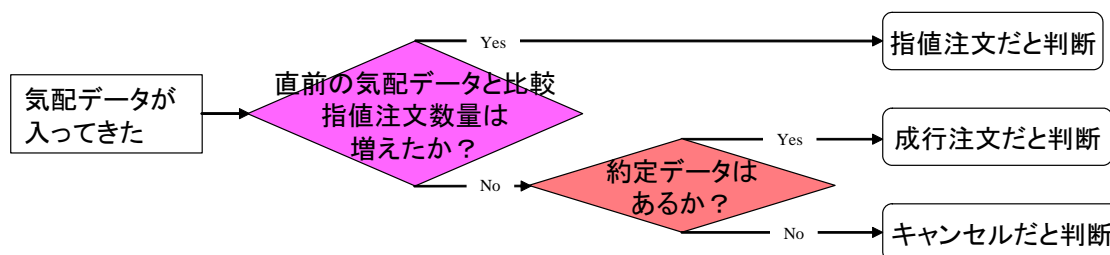


図 9: 気配データと約定データからどのような注文が入ったのかを判定するためのフローチャート

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 9 はこれらをどのように区別するかを示したものである。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。また, 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。

C. 本研究で用いる各パラメータ, λ , スプレッド, 拡散係数の推定法

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 σ , μ , α , δ , λ , スプレッド, 拡散係数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を呼値の刻みで割った値を用いるという点だけである。

C.1. σ の推定法

σ は一指値注文あたりの株数として定義される。注文には指値と成行の 2 種類がある。本研究では, Farmer ら [3] にならい, 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いうからである。というのは, アスク(ビッド)に存在する株数より大口の買い(売り)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, それはデータ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。

C.2. μ の推定法

μ の定義は単位時間あたりの成行注文株数である。ここで, 時間は約定タイム, すなわち注文が約定するたびに +1 するというやり方を用いて測っているため, μ は, 結局, 一注文あたり成行注文株数に相当する。ところが, この値は約定データからそのまま求めると C.1 節で述べ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一注文あたり成行注文株数は, 一注文あたり指値注文株数 σ に等しいと仮定する。した

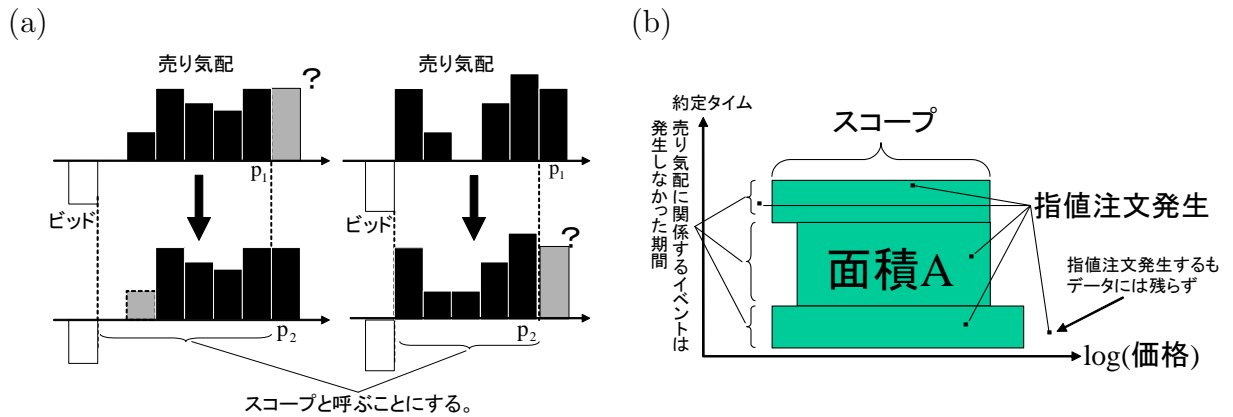


図 10: (a) 気配データの変化の例と「スコープ」の定義 (b) スコープの変化の例と面積 A の定義

がって、本研究では μ の値を求める必要はない。

C.3. α の推定法

次に α の推定法について述べる。 α の定義は単位価格・単位時間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 本気配値なので、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 10 (a) 左）、あったはずの気配が消えていたり（図 10 (a) 右）ということが起きる。これは Farmer らの研究 [3] にはない困難である。

そこで次のような方法を考案した。図 10 のような売り板情報があったとする。 $p_1 =$ 「直前の売り 5 本気配値の中で最高値（対数をとる前の、現実の取引価格）」、 $p_2 =$ 「直後の売り 5 本気配値の中で最高値（同じく、対数をとる前の、現実の取引価格）」、 $\min = \min(p_1, p_2)$ とおく。売り側の場合、 $(\log(\text{bid}) + \log(\text{bid} + 1))/2$ から $(\log(\min) + \log(\min + 1))/2$ までの価格幅を「スコープ」と呼ぶことにする。買い側の指値注文についても同様である。

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、たまたまスコープ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 10 は、スコープ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。スコープ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 の内部に到着した全指値注文株数をカウントし、それを A で割れば α を求めることができる。 $A = \sum(\text{スコープ} \times \text{指値注文が発生しなかった時間})$ 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。この操作をデータブロックごとに行う。その結果、得られた値の、約定タイムによる重み付き平均がある銘柄の α となる。

C.4. δ 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、注文ごとにキャンセルされるまでの平均時間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指値注文分布は定常状態にあると仮定して δ を求めることにする。

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。ま

た，以下で扱う量の中で株数の次元を持つものは σ を単位として測る．

まず， δ を約定データと気配データから求める方法を考える．スコープ内部に到着する指値注文の到着率を α' ⁹，成行注文到着率を μ' ¹⁰，時刻 t における幅内部の指値注文量を N_t ，時刻 $t-1$ と t の間にキャンセルされる指値注文量を $N_{c,t-1}$ とすると，平均的には

$$N_t - N_{t-1} \approx (\alpha' - \mu') \cdot 1 - N_{c,t-1} \approx \alpha' - \mu' - \delta N_{t-1} .$$

$t = 2$ から T まで辺々加え合わせると，

$$N_T - N_1 \approx (\alpha' - \mu')(T - 1) - \delta \sum_{t=1}^{T-1} N_t . \quad (\text{C.1})$$

したがって，

$$\delta \approx \frac{(\alpha' - \mu')(T - 1) - (N_T - N_1)}{\sum_{t=1}^{T-1} N_t} = \frac{(\alpha' - \mu') - \frac{N_T - N_1}{T-1}}{\bar{N}} . \quad (\text{C.2})$$

\bar{N} はスコープ内に存在する指値注文株数の平均である．こうしてあるデータブロック中における δ を求めることができた．

ある日ある前（後）場の δ 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とする．そして，ある銘柄の δ は，前（後）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 (C.2) 式によると δ 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．

C.5. λ ，スプレッド，拡散係数の推定法

λ の推定法

まず， 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．まず売り側の指値注文に注目して考える．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．マーケットインパクトは n のべき的な関数で記述できることが知られている [4, 10] ので， $\Delta p(n) = \lambda_1 n^\kappa$ の形を仮定する．そして，両辺の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・買いが対称なモデルなので，この操作は売り・買い区別無くまとめて行うことにする．

回帰の範囲 (n の最大値) は以下のように決める．本研究で扱うデータは 5 本気配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指値注文は一つも無くなってしまい，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，売り板情報・買い板情報を通して，存在する指値注文が最も少ないときを見つける．そのときの指値注文数を N_{\min} とおく．すると N_{\min} 株までの大口成行注文ならアスクを推定できる．そこで

⁹ α' がついている理由は， α は「単位価格・単位時間あたりの指値注文到着率」である一方， α' は「単位時間あたりの，スコープ内部への指値注文到着率」であり，両者を区別するためである．

¹⁰ 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 の値はゆらいでいる．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．

n の最大値を 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 λ_1, κ を求められない．そこで，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．モデル値の λ も同様に求めることができるが，注文は常に σ 単位ずつ到着するので， $\sigma/u = 1$ とする．つまりモデル値の場合は λ_1 が λ そのものである．

スプレッドの推定法

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は対数をとって考えるので，ある時刻 t のスプレッドは $\log(\text{ask}_t/\text{bid}_t)$ である．この値は時間とともに変動するのでその重み付き平均値をもってある銘柄のスプレッドとすることにする．スプレッドの実測値については，あるアスク，ビッドの組が存在した時間 w (約定タイムで測る．) で重みをつけて次の式で与える．

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

スプレッドのモデル値も同様に重みを付けて平均することにより求める．

拡散係数の推定法

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．その比例係数を拡散係数という．すなわち，ある時刻 t の価格 (仲値) を p_t とすると，価格時系列がランダムウォークに従うとき， $\langle (p_{\tau+t} - p_t)^2 \rangle \equiv \langle (\Delta p)^2 \rangle = D\tau$ が成り立つ． D が拡散係数である．最小二乗法を用いると D は以下のように書ける．

$$\text{actual diffusion rate} = D = \frac{\sum_{\tau=1}^{80} \langle (\Delta p)^2 \rangle \tau}{\sum_{\tau=1}^{80} \tau^2} .$$

ここで， τ の値の範囲は銘柄によらず 1 から 80 までとした．拡散係数のモデル値についてはこうして求めた D をそのまま用いた．一方，実測値については， D をデータブロックごとに求め，約定タイムで測ったデータブロックの大きさによる重み付き平均値を用いた．

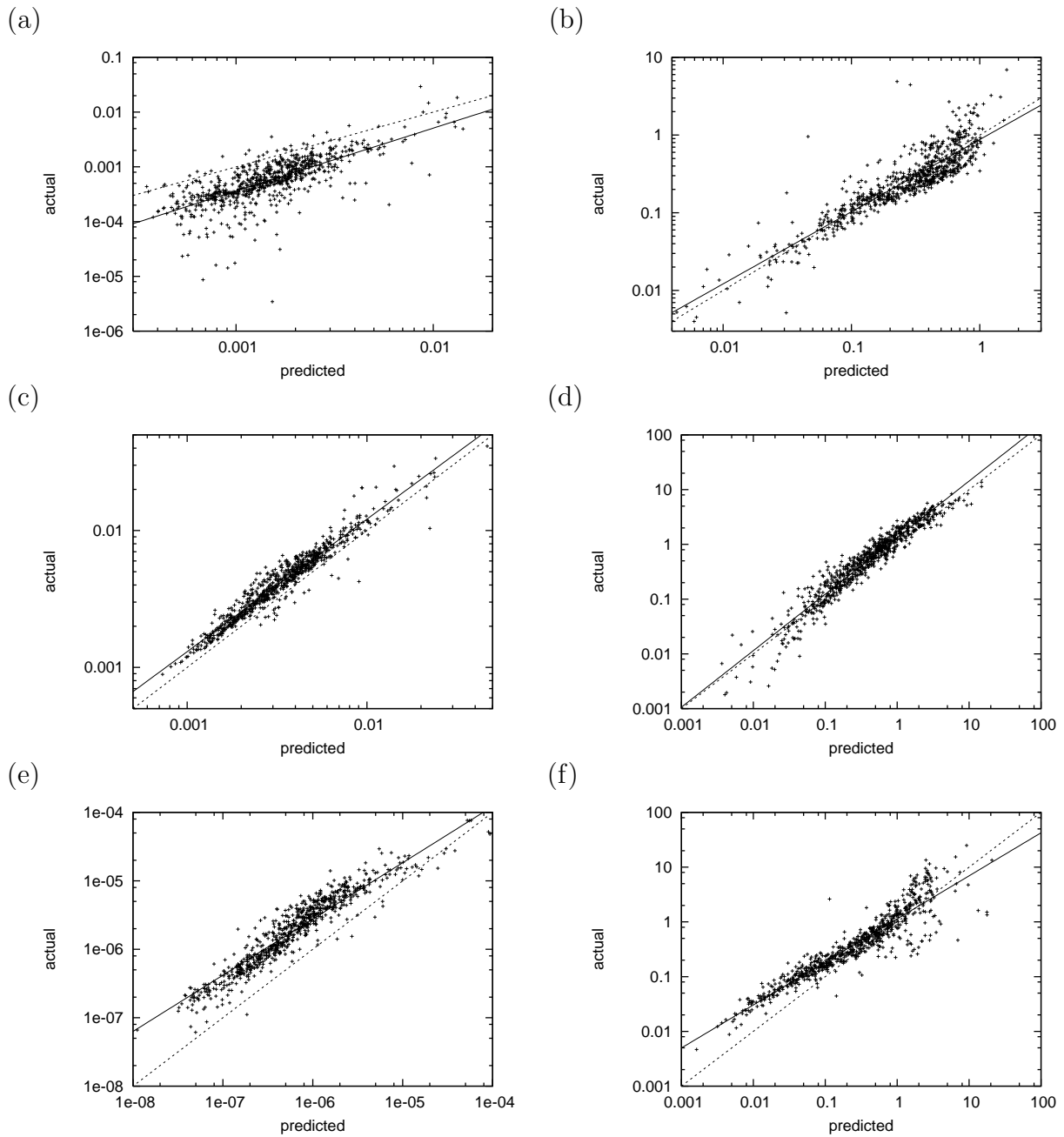


図 11: それぞれ (a) λ (c) スプレッド (e) 拡散係数の実測値とモンテカルロゼロインテリジェンスモデルによるモデル値の比較, および, それぞれ (b) λ (d) スプレッド (f) 拡散係数の実測値とローインテリジェンスモデルによるモデル値の比較

木村博道

筑波大学大学院システム情報工学研究科

〒 305-0006 茨城県つくば市天王台 1-1-1

E-mail: hkimura@sk.tsukuba.ac.jp

ABSTRACT

LOW INTELLIGENCE MODEL WHICH CAN EXPLAIN MARKET
LIQUIDITY

Hikomichi Kimura Eizo Akiyama
University of Tsukuba

Classical economists often assume perfect rationality of traders when they study financial markets. Though such theory can prove the existence of equilibrium prices, it cannot say anything about the quantities that are important for attendees in real stock markets, such as the spread (the size of gap between ask and bid) and market impact function (which describes the response of quoted prices to the arrival of new orders). On the other hand, recently, it has been shown that so-called “zero intelligence model (Smith et al. 2003)” in which traders have no intelligence and issue orders randomly (Poisson processes), can predict the spread and the diffusion rate of real stock markets (Farmer et al. 2005). However, zero intelligence model cannot predict market impact such as Kyle’s λ , which shows the average price change when a unit order arrives. In our study, first, we clarified the nature of order arrival pattern observed in the empirical data (for example, order arriving rate depends on the distribution of limit orders in the order book). Second, we introduced the nature of order arrival pattern into the zero intelligence model. We call such a model “low intelligence model”. We showed that low intelligence model can explain, in addition to spread and diffusion rate, Kyle’s λ , explanation of which is difficult for either general equilibrium theory or the zero intelligence model. This result implies that, in order to know liquidity in a stock market, it is important to understand the interaction between orders and the book, and also that between some orders and other orders.