

# リミット・オーダー・ブックに対する実証分析と金融市場に見られる冪乗則について<sup>1</sup>

統計数理研究所統計思考院 荻原 哲平

Teppei Ogihara

School of Statistical Thinking, The Institute of Statistical Mathematics

大阪大学大学院基礎工学研究科 加藤 恭

Takashi Kato

Division of Mathematical Science for Social Systems, Graduate School of Engineering Science,  
Osaka University

東邦大学理学部情報科学科 高田 英行

Hideyuki Takada

Department of Information Science, Toho University

## 1 はじめに

近年, IT 技術の発展と共に金融市場において取引の高頻度化や自動化が急速に進展しており, それと共に, 大口取引の効率化・最適化に関わる理論研究も活発に行われている. 特に, 大量証券の売却(流動化)あるいは購入に関する最適執行スケジュール策定(最適執行問題)に関して様々な観点から数多くの研究がなされているが, そのような問題の考察においては市場流動性, 特に「取引主体(トレーダー)自身の取引行動が証券価格自体に影響を与えてしまう」というマーケットインパクト(Market Impact: 以下 MI と略記する, 価格インパクトとも呼ばれる)を考慮する事が重要となる. また [14, 15] では MI 関数の形状に焦点を当て, 理論・実証・実務のいずれの観点からも MI 関数の非線形性が本質的な意味を持つ事が指摘されており, 特に S 字型の形状を持つ MI 関数の重要性が示唆されている.

本稿では実際の金融市場における MI 関数の形状を調べるために高頻度複数気配データを用いた板及び MI の推定手法を提案し, 特に日経 225 先物に関する秒次複数気配データを用いた推定結果について報告する.

## 2 モデルの概要

本章では推定に使用した数理モデルについて解説する.

後で述べる様に, 本研究では大阪取引所にて上場・取引がなされている日経 225 先物(ラージ)を対象とした実証分析を行う. 日経 225 先物の取引にはオークション方式(個別競争取引)が採用されており, 東京証券取引所(以下東証)における株式市場(オーダードリブン方式)と同様, 板(Limit Order Book; LOB)<sup>2</sup>の上で主に指値注文(limit order), 成行注文(market order)といった発注によって取引が展開されている<sup>3</sup>. そこで, 本研究では買い(bid)・売り(ask)それぞれ

<sup>1</sup>本稿は将来公刊予定の論文の簡約版である.

<sup>2</sup>LOB, 指値・成行注文等の用語については [8, 14, 15] 等を参照せよ.

<sup>3</sup>制度上は更に最良指値注文とストップ注文があり, 更にストラテジー取引(カレンダーズプレッド)も採用されているが, これらは指値注文と成行注文の応用と考えられるため以下では指値・成行注文のみ扱う事とする.

れについて、各気配値に対する指値注文及び成行注文の到着枚数が複合 Poisson 過程に従うと仮定する。

具体的には、 $M_t$  を「初期時点 0 から時点  $t$  までの注文到着総枚数」を表す確率変数として、 $(M_t)_t$  が次のような確率過程に従うものとする。

$$M_t = \sum_{l=1}^{N_t} A_l = A_1 + \cdots + A_{N_t} \quad (2.1)$$

但し  $(N_t)_t$  は強度 (intensity) パラメーター  $\lambda$  を持つ Poisson 過程であり、 $(A_l)_l$  は独立同分布な正值確率変数列で  $(N_t)_t$  とは独立であると仮定する<sup>4</sup>。  $A_1$  の分布を  $\mu$  と記す。  $\mu$  の具体形については後述。この時、 $(M_t)_t$  は次の様にも表せる事に注意しておく。

$$M_t = \int_0^t \int_{(0,\infty)} z N(dtdz)$$

ここで  $N(dtdz)$  は平均測度  $\lambda dt \mu(dz)$  を持つ Poisson ランダム測度。

上の確率過程を、観測可能な  $n$  本の価格帯における指値注文、及び成行注文それぞれに対して適用する。ここでは、 $M_t^{0,c}$  を成行注文の  $t$  時点までの総到着枚数とし、 $M_t^{k,c}$  ( $k = 1, \dots, n$ ) で「最良気配から  $(k-1)$  tick 離れた指値注文」の総到着枚数を表す事とする。ここで  $c \in \{a, b\}$  であり、 $a$  は売り、 $b$  は買いを表す。  $N_t^{k,c}, A_t^{k,c}, \lambda^{k,c}, \mu^{k,c}$  ( $k = 0, \dots, n, c = a, b$ ) についても同様。

また指値注文は新規の注文到着だけでなく既存の注文に対するキャンセルも考慮する必要がある。ここでは、 $t$  時点までの指値注文キャンセル枚数が次の確率過程

$$\tilde{M}_t^{k,c} = \int_0^t \int_{(0,\infty)} z \tilde{N}_{V^{k,c}}^{k,c}(dtdz)$$

に従うと仮定する<sup>5</sup>。ここで、所与の関数  $w = (w_t)_{t \geq 0}$  に対して  $N_w^{k,c}(dtdz)$  は平均測度  $\tilde{\mu}_{w_t}^{k,c}(dz) \tilde{\lambda}^{k,c} dt$  を持つ Poisson ランダム測度であり、 $V^{k,c} = (V_t^{k,c})_t$  は各  $k \in \{1, \dots, n\}, c \in \{a, b\}$  に対応する価格帯における板上の既存の指値注文の枚数の過程を表す。  $\tilde{\lambda}^{k,c}$  は新たなキャンセル注文の到着頻度を表すパラメーターであり、板の厚さが  $x$  枚である時の一回あたりのキャンセル注文到着枚数の分布が  $\tilde{\mu}_x^{k,c}$  で表される。ここで  $\tilde{\mu}_x^{k,c}$  の台は  $(0, x]$  に含まれるものとする。

以上により総計  $2(n+1+n) = 4n+2$  個の確率過程が定義された。

### 3 データ

本研究では、大阪取引所にて上場・取引がなされている日経 225 先物の買い (bid)、売り (ask) 各々の秒次の 10 本複数気配データ、及び tick by tick の約定データを使用した。具体的には 2012

<sup>4</sup>金融機関の自己資本比率規制であるバーゼル II/III において、オペレーショナルリスクに対する先進的計測手法として標準的である損失分布手法 (Loss Distribution Approach; LDA) の多期間版のモデルと言える。詳細は [12] 等を参照せよ。

<sup>5</sup>後で述べる通り、キャンセル枚数の分布はキャンセル到着時の板の厚さ (既存の指値注文の枚数) に依存するため、新規注文到着過程のように厳密な意味で複合 Poisson 過程として定義する事は出来ない。しかし、板の厚さに対応する確率変数 (ないし確率過程) の下での条件付き確率の下では複合 Poisson 過程と同様の取り扱いが出来るため、本稿ではしばしば新規注文・注文キャンセル到着枚数を表す確率過程をまとめて単に複合 Poisson 過程と呼ぶ事にする (但し板の厚さも注文キャンセル枚数過程の影響を受ける事を考慮すると、これはあくまで擬似的な表現である事に注意が必要である)。

年4月2日から2012年12月28日までの日中取引のデータを用いるが、オープニング・(プレ) クロージングに近い時間帯は市場取引の構造がザラ場とは少し異なる可能性を考慮し、9時03分から15時07分までの各秒におけるデータを使用した<sup>6</sup>。

2章で触れたように、本研究では各時点におけるLOBのデータそのものではなく、指値・成行注文（及び指値キャンセル）到着枚数をモデル化するものであるため、上記のデータを加工して「新規注文・キャンセルの各時点における到着枚数」の形に整理する必要がある。

### 3.1 データ加工方法

全観測期間を  $\{0, d, 2d, \dots, Id\}$  と離散化し ( $d$  は時間間隔の最小単位, ここでは元となるLOBデータが秒次である事に鑑みて  $d$  を1秒と定める, また  $I$  は観測期間の全秒数), また呼値 (tick size) を  $\delta$  で表す (日経225先物の場合  $\delta$  は10円となる). そして  $y_i^{k,a}$  (及び  $y_i^{k,b}$ ) を,  $id$  時における, 最良気配値から  $(k-1)$  tick 離れた価格帯における指値売り (または買い) 注文の枚数<sup>7</sup>を表すものとし ( $k=1, \dots, n$ ), 対応する価格帯の気配値を  $z_i^{k,a}$  (及び  $z_i^{k,b}$ ) と書く<sup>8</sup>. また  $(s_j, p_j, v_j)_{j=1, \dots, J}$  を全観測期間における約定データの列とする. 具体的には,  $s_j, p_j, v_j$  は  $j$  回目の約定データに関する約定時刻, 約定価格, 累積出来高をそれぞれ表すものとする.

さて,  $x_i^{k,a}$  (及び  $x_i^{k,b}$ ) によって  $(i-1)d$  時から  $id$  時に到着した売り (及び買い) 注文の枚数を表す事とする ( $k=0$  の時は成行,  $k \geq 1$  の時は指値: 値が負の時はキャンセルを表すものとする). 上記の元データ  $(y_i^{k,c})_{i,k,c}$  及び  $(s_j, p_j, v_j)_j$  から,  $(x_i^{k,c})_{i,k,c}$  を抽出する事が本節の目的となる.

データの加工をするにあたり, 以下の二つの仮定を置く.

- (i) 成行注文, 及び各価格帯の指値注文・キャンセルの到着は単位時間  $d$  あたりそれぞれ高々1度しか生じない.
- (ii) 最良気配値の変更が生じた場合のデータは欠損値扱いとする, 即ち  $z_{i-1}^{1,a} \neq z_i^{1,a}$  または  $z_{i-1}^{1,b} \neq z_i^{1,b}$  の時は  $id$  時点のデータは無視する.

**Remark 1.** 仮定 (i) は, 高頻度取引 (High Frequency Trading; HFT) が蔓延している現代の金融市場においては現実的とは言えない仮定である<sup>9</sup>. ミリ秒・マイクロ秒単位の高速注文が頻繁に行われる近年の株式市場において, 仮定 (i) の設定は多数の注文の到着及びキャンセルを一つに集約してしまう事を意味しており, その際に大数の法則が働く事によっていくつかの不確実性要因を消滅させてしまっている可能性がある.

また仮定 (ii) においては, 大規模な成行注文による劇的な価格変化が生じた際のデータを捨象してしまう事になり, 分布の fat-tail 性の過小評価に繋がってしまうと考えられる.

上述の通り, これらの仮定は現実の市場の様相をモデルに反映させる意味で適切とは言えないが, 現状のデータの制約の下ではこれ以上の精緻なデータ整備は困難であり, より精度の高いデータを用いた実証分析は今後の課題としたい.

<sup>6</sup>厳密には15時00分が東京証券取引所における現物取引の終了時刻となるため, この時間についても市場の状況が他の時間帯と異なる可能性が考えられるが, 今回は特段の対処は行わなかった.

<sup>7</sup>1枚は先物価格の1,000倍に対応している.

<sup>8</sup>即ち  $z_i^{k,a} = z_{i-1}^{k,a} + (k-1)\delta$ ,  $z_i^{k,b} = z_{i-1}^{k,b} - (k-1)\delta$  が成立している.

<sup>9</sup>東証市場 (一部上場) におけるHFT取引の割合は2013年5月の段階で売買代金ベースで25.9%, 注文金額ベースで51.6%程度と言われている [10].

**Remark 2.** 今回使用した金融高頻度データは所謂ビッグデータと呼ばれる類のデータであり、そこには特有の問題もいくつか存在する。例えば、蓄積されている約定データは、実際に市場で約定してからインターネット回線を介してデータベンダーに情報が届くまでに僅かながら時間の遅れがあり、その影響で蓄積データのタイムスタンプが必ずしも正しいとは限らず、場合によってはデータに矛盾が生じている（例えば累積出来高が時刻について単調増大となっていない）事もある。また、HFT による注文は僅かな時間のうちに発注とキャンセルが繰り返されるため、データ取得漏れの可能性も実際には無視出来ない。また [9, 22] の実証分析において、高頻度株価データの観測ノイズの存在が指摘されており、これを考慮した数理統計学的手法も研究されてきた。特に観測ノイズ下の株価ボラティリティの推定に関しては、[5, 21, 25] 等が知られている。しかし、今回の研究ではビッグデータ特有の問題に起因するデータの矛盾等は全て欠損値として処理した。

上記の仮定 (i)(ii) の下で、まず  $k \geq 2, c \in \{a, b\}$  に対して、即ち最良気配以外の価格帯における指値注文に関して  $(x_i^{k,c})_i$  を導出する事を考える。この場合、LOB の変化は指値新規注文またはキャンセルの到着によってしか生じない事になるので  $y_i^{k,c} = y_{i-1}^{k,c} + x_i^{k,c}$  が成り立つ。これを整理して

$$x_i^{k,c} = y_i^{k,c} - y_{i-1}^{k,c}, \quad i = 1, \dots, I \quad (3.1)$$

を得る。

最良気配の指値注文、及び成行注文についてはもう少し状況が複雑である。まず、注文の約定は基本的に成行注文によってのみ起こる事に注意する<sup>10</sup>。約定は単位時間につき 1 回しか起こらないという前提に基づくと、約定したのが成行買い注文によるものか、あるいは売り注文によるものかは、約定価格と最良気配値の比較によって判断する事が出来る。即ち、約定価格が最良買い気配値  $z_i^{1,b}$  と一致しているならばそれは成行売り注文が到着したと考えられ、逆に  $z_i^{1,a}$  と一致している場合は成行買い注文の到着と解釈される（どちらでもない場合はデータ不整合により欠損値と捉える）。

ここで約定 tick データ  $(s_j, p_j, v_j)_j$  を約定価格と約定枚数に関する秒次データに変換しておこう。各  $i = 0, \dots, I$  に対して  $j(i) = \sup\{j = 0, \dots, J; \lceil s_j/d \rceil \leq i\}$  と定め、 $q_i = p_{j(i)}$ ,  $w_i = v_{j(i)} - v_{j(i-1)}$  (但し  $v_0 = 0$ ) として  $(q_i, w_i)_i$  を定義する。ここで、 $\sup$  の中身が空集合となる場合は欠損値として処理する。

上を用いると、 $k = 0, 1$  の時の  $x_i^{k,a}$  及び  $x_i^{k,b}$  は次を満たす事となる。

- $z_i^{1,a} = q_i$  の場合、成行買い注文の到着による  $w_i$  枚の約定があったと考えられるため

$$x_i^{0,a} = 0, \quad x_i^{0,b} = w_i \quad (3.2)$$

であり、また  $y_{i-1}^{1,a}$  から  $y_i^{1,a}$  への変化は成行買い注文と最良気配指値売り注文の到着またはキャンセルによって生じる事から  $y_i^{1,a} = y_{i-1}^{1,a} + x_i^{1,a} - x_i^{0,b}$ , 即ち

$$x_i^{1,a} = y_i^{1,a} - y_{i-1}^{1,a} + w_i \quad (3.3)$$

最良気配買い注文に関しては成行注文が存在しないので

$$x_i^{1,b} = y_i^{1,b} - y_{i-1}^{1,b} \quad (3.4)$$

<sup>10</sup>例えば安値に指値売り注文等を入れる事によっても約定は成立するが、これらは本質的に成行売り注文と認識する。

- $z_i^{1,b} = q_i$  の場合, 上の逆の状況を考えれば良く

$$x_i^{0,a} = w_i, \quad x_i^{0,b} = 0, \quad (3.5)$$

$$x_i^{1,a} = y_i^{1,a} - y_{i-1}^{1,a}, \quad (3.6)$$

$$x_i^{1,b} = y_i^{1,b} - y_{i-1}^{1,b} + w_i \quad (3.7)$$

以上の (3.1)–(3.7) により, 各時点における注文到着枚数データ  $\mathcal{X} = \{x_i^{k,c}; k = 0, \dots, n, c = a, b, i = 1, \dots, I\}$  を仮定 (i)(ii) の下で抽出する事が出来る.

### 3.2 注文到着データの特徴：人間の恣意性と分布の歪みについて

本節では, 上節にて導出した注文到着枚数データの経験分布の特徴を調べる. まず各  $k = 0, \dots, n$  及び  $c = a, b$  に対して  $(x_i^{k,c})_{i=1, \dots, I}$  の定める経験分布の分布関数

$$F^{k,c}(z) = \frac{1}{I} \sum_{i=1}^I 1_{\{x_i^{k,c} \leq z\}}$$

を考え, 対応する裾確率  $\bar{F}^{k,c}(z) = 1 - F^{k,c}(z)$  の両対数プロット  $(\log z, \log \bar{F}^{k,c}(z))_z$  を表したものが 図 1-22 である. いずれについてもある程度強い fat-tail 性が観測出来ているが<sup>11</sup>, もう一つの特徴として, 例えば  $z = 10, 100$  等といった「人間が見てキリの良い数字」となっている時に裾確率関数の形状が歪んでいる傾向が見取れる. 例えば 図 13, 14 の  $z = 10$  や 図 21, 22 の  $z = 100$ , 図 2 の  $z = 1000$  等は顕著である.

そこで, 次に  $\mathcal{X}$  に含まれるキャンセルを除く全てのデータを用いたヒストグラムについて調べてみる. 図 23, 24 はこのヒストグラムを, 一回あたりの注文到着枚数 (横軸) が 100 枚以下のものと 101 枚以上のものに分けて描いたものである. これを見ると, やはりキリの良い数字のところでヒストグラムが単調性を失っている (キリの良い数字の枚数の注文が到着する頻度が相対的に高い) 事が見取れる. これより, 本研究においては「人間の恣意性に基づくキリの良い注文枚数」の集合として

$$S = \{s_1, \dots, s_m\} = \{10, 15, 20, 25, 30, 40, 50, 100, 150, 200, 250, 300, 400, 500, 1000\} \quad (3.8)$$

と定める事とし<sup>12</sup>, 注文到着枚数分布  $\mu$  の推定において他の数字と区別して扱う事とする.

## 4 分布の推定手法

本章では 2 章で与えた分布の推定手法について述べる. 推定は各  $k, c$  について独立に行うため, ここでは  $k, c$  の記号は省略する.

<sup>11</sup>分布の裾が両対数プロットにて直線状になっている時, 裾確率は冪乗に減衰する (regularly varying).

<sup>12</sup>今回は 図 23, 24 を見て「決め打ち」で与えたが, 例えばヒストグラムの単調性が大きく崩れる枚数によって  $S$  を定義する等によりシステマティックに定める事も出来る.

#### 4.1 Poisson 強度 $\lambda, \tilde{\lambda}$ の推定手法

2章で与えた複合 Poisson 過程 (2.1) に対して, 観測期間  $[0, T]$  における実際の注文到着枚数データ  $(x_{t_i})_{i=1, \dots, N}$  を用いてパラメーター推定を行う事を考える. ここで,  $x_{t_i}$  は  $t_i$  時点における注文到着枚数を表し, 値が負である時はキャンセル枚数を表すものとする<sup>13</sup>. また  $0 \leq t_1 < \dots < t_N \leq T$  である (または前章の記号を用いて,  $N = I$  として  $t_i = id$  と読み替えても良い). ここでは Poisson 強度パラメーター  $\lambda$  の推定手法として最尤法を採用する. この場合,  $\lambda$  の推定量は

$$\hat{\lambda} = \frac{1}{NT} \sum_{i=1}^N 1_{\{x_{t_i} > 0\}} \quad (4.1)$$

で与えられる.

キャンセルについても同様であり, 対応する最尤推定量は

$$\hat{\tilde{\lambda}} = \frac{1}{NT} \sum_{i=1}^N 1_{\{x_{t_i} < 0\}} \quad (4.2)$$

である.

#### 4.2 $\mu$ の推定手法

3.2 節で得られた知見に従い, 一回あたり注文到着枚数分布  $\mu$  については通常 (body) の分布の他に「 $S$  にのみ値を取る分布」を考慮した混合分布として定義する事とする. 具体的には, 2章で与えた  $A_k$  が

$$\begin{aligned} A_k &= A_k^1 1_{\{U_k^1 < r\}} + A_k^2 1_{\{U_k^1 \geq r\}}, \\ A_k^1 &\sim \text{Exp}(\lambda_1), \quad A_k^2 = \sum_{i=1}^m s_i 1_{\{p_{i-1} \leq U_k^2 < p_i\}}, \\ p_i &= \frac{\sum_{j=1}^i \exp(-\lambda_2 s_j)}{\sum_{j=1}^m \exp(-\lambda_2 s_j)} \end{aligned}$$

と与えられると仮定する. ここで  $U_k^1, U_k^2$  は一様分布に従う確率変数であり,  $\text{Exp}(\lambda^1)$  はパラメーター  $\lambda^1$  の指数分布を表す ( $A_k^1, U_k^1, U_k^2$  は全て独立とする). この時  $A_k$  は指数分布と多項分布の混合分布に従う事となる.

上で推定すべきパラメーターは  $r \in (0, 1), \lambda_1, \lambda_2 > 0$  の3個である. ここではこれらを擬似最尤法によって推定する事を考える. 即ち, 次で定義される擬似対数尤度

$$H_1(\lambda_1, \lambda_2, r) = \sum_{i=1}^N \log \left( r \lambda_1 e^{-\lambda_1 x_{t_i}} + (1-r) \sum_{j=1}^m p_j 1_{\{x_{t_i} = s_j\}} \right)$$

に対する次の最適化問題

$$\underset{(\lambda_1, \lambda_2, r) \in K}{\text{argmax}} H_1(\lambda_1, \lambda_2, r) \quad (4.3)$$

<sup>13</sup>成行注文に関しては  $x_{t_i}$  は負値となる事は無い.

を数値的に解く事で推定を行う<sup>14</sup>。ここで  $K = [10^{-5}, 1-10^{-5}] \times [10^{-5}, 1-10^{-5}] \times [10^{-5}, 1-10^{-5}]$  は数値最適化における打ち切りを考慮したパラメーターの空間を表す。

### 4.3 $\tilde{\mu}$ の推定手法

次にキャンセル枚数の推定方法について述べる。今、 $k$  回目の注文キャンセル到着枚数  $\tilde{A}_k$  が届いた時の同価格帯における板の厚さを  $V_k$  とすると、常に  $\tilde{A}_k \leq V_k$  の関係が成立している。即ち、既存の指値注文の枚数を超えるキャンセルが到着する事は無く、 $\tilde{\mu}$  は有界な台を持つ分布として定義する必要がある。そこで、本研究では以下を仮定する。

$$\begin{aligned}\tilde{A}_k &= \tilde{A}_k^1 1_{\{\tilde{U}_k^1 < \tilde{\tau}\}} + \tilde{A}_k^2 1_{\{\tilde{U}_k^1 \geq \tilde{\tau}\}}, \\ \tilde{A}_k^1 &\sim B(V_k, \tilde{p}), \quad \tilde{A}_k^2 = V_k \tilde{U}_k^2\end{aligned}$$

ここで  $\tilde{U}_k^1, \tilde{U}_k^2$  は一様分布に従う確率変数であり、 $B(v, p)$  は試行回数  $v$ 、成功確率  $p$  の二項分布を表す。また上で現れる ( $\tilde{A}_k$  以外の) 確率変数は全て独立とする。

キャンセル枚数が基本的に二項分布で表される事を前提とした定式化であるが、分析に用いたデータを見る限り、二項分布では発生確率が非常に低いと考えられる大量キャンセルが生じる事が度々あり、そのような低頻度事象を捕捉し分布にある程度の fat-tail 性を持たせるために  $\tilde{A}_k^2$  を追加した混合分布の形で  $\tilde{\mu}$  を表現している。

推定すべきパラメーターは  $\tilde{\tau}$  と  $\tilde{p}$  の二つであり、観測期間  $[0, T]$  における実際のキャンセル枚数到着データ  $(\tilde{x}_{\tilde{t}_i})_{i=1}^{\tilde{N}}$  とその直前の板の厚さデータ  $(v_{\tilde{t}_i})_{i=1}^{\tilde{N}}$  を用いて擬似対数尤度関数は次で与えられる。

$$H_2(\tilde{p}, \tilde{\tau}) = \sum_{i=1}^{\tilde{N}} \log \left( \tilde{\tau} \cdot \frac{v_{\tilde{t}_i}!}{\tilde{x}_{\tilde{t}_i}!(v_{\tilde{t}_i} - \tilde{x}_{\tilde{t}_i})!} \cdot \frac{\tilde{p}^{\tilde{x}_{\tilde{t}_i}}(1-\tilde{p})^{v_{\tilde{t}_i}-\tilde{x}_{\tilde{t}_i}}}{1-(1-\tilde{p})^{v_{\tilde{t}_i}}} + \frac{1-\tilde{\tau}}{v_{\tilde{t}_i}} \right)$$

よって、パラメーター推定のためには次の最大化問題を数値的に解けば良い。

$$\operatorname{argmax}_{(\tilde{p}, \tilde{\tau}) \in \tilde{K}} H_2(\tilde{p}, \tilde{\tau}) \quad (4.4)$$

但し  $\tilde{K} = [10^{-5}, 1-10^{-5}] \times [10^{-5}, 1-10^{-5}]$ 。

## 5 推定結果

### 5.1 Poisson 強度 $\lambda, \tilde{\lambda}$

前章で紹介した (4.1) 及び (4.2) の手法を用いて 1 日毎に推計した  $\lambda, \tilde{\lambda}$  の時系列推移を調べると、これらの推定値はいずれも、分次対数収益率を用いて計算した実現ボラティリティ (Realized Volatility; RV)<sup>15</sup> と似通った変動をしている事が分かる<sup>16</sup>。表 1, 2 は、上の各パラメーターの前

<sup>14</sup>(4.3) は多変数の非線形最適化問題であり、数値解法として滑降シンプレックス法 (downhill simplex method, Nelder-Mead 法とも呼ばれる)、焼きなまし法 (Simulated Annealing; SA)、逐次二次計画法 (Sequential Quadratic Programming; SQP) 等様々なアルゴリズムが知られている。本研究では滑降シンプレックス法を用いた。

<sup>15</sup>[3] で指摘されるように秒次の収益率から計算し RV はマーケットマイクロストラクチャーノイズの影響を強く受けるため、ここでは分次に丸めたものを使用した。

<sup>16</sup>なお、本章における各パラメーター推計値は全て 1 時間当たりのものとなるようスケージングを施している (Poisson 強度は  $T$  倍法、ボラティリティは  $\sqrt{T}$  倍法)。

日との差分値に関して RV 差分値との相関係数を計算したものである。いずれのパラメータに関しても、RV の変動の高さ（即ち市場の不安定性）と取引活況度との間に正の相関関係がある事が分かる。

Poisson 強度パラメーターに関して、もう一つ極めて興味深い推計結果が得られる。各  $c = a, b$  に対して、1 日毎ではなく全期間  $i = 1, \dots, K$  のデータを用いて  $\lambda^{k,c}, \tilde{\lambda}^{k,c}$  を  $k$  毎に推計し、 $(k, \lambda^{k,c})_{k=1, \dots, 10}$  及び  $(k, \tilde{\lambda}^{k,c})_{k=1, \dots, 10}$  のそれぞれについて両対数プロットを描いてみると、図 25, 26 のように綺麗な直線関係を観測する事が出来る。これは一般に冪乗則 (power law) と呼ばれる現象であり、自然現象や社会現象において度々観測される経験則である [19, 20, 24]。

注文到着及びキャンセル到着それぞれにおいて、 $k = 1, \dots, 10$  の 10 個のデータを用いて  $(\log k, \log \lambda^{k,c})_k$  及び  $(\log k, \log \tilde{\lambda}^{k,c})_k$  のそれぞれについて線形回帰を行うと

$$\begin{aligned} \log \lambda^{k,a} &= -1.72 \log k + 6.06 \quad (R^2 = 0.994), \\ \log \lambda^{k,b} &= -1.72 \log k + 6.06 \quad (R^2 = 0.993), \\ \log \tilde{\lambda}^{k,a} &= -1.63 \log k + 5.78 \quad (R^2 = 0.990), \\ \log \tilde{\lambda}^{k,b} &= -1.65 \log k + 5.80 \quad (R^2 = 0.988) \end{aligned}$$

となっており、いずれについても極めて高い決定係数となっている。ここで、例えば指値買い注文到着強度  $\lambda^{k,b}$  に関して、上の関係式は定数  $C (= e^{6.06})$  を用いて  $\lambda^{k,b}$  が、最良気配値（あるいは証券時価）近辺からの「距離」を表すパラメーター  $k$  について

$$\lambda^{k,b} \approx C k^{-1.72}$$

という冪関数として近似的に表される事を意味している。

上の結果は、推計を行う期間を一時間毎にした場合（即ち使用する時点を  $\{0d, \dots, Kd\} \cap \{\text{対応する時刻が } h \text{ 時台}\} (9 \leq h \leq 14)$  に制限した場合）や四半期毎にした場合（ $\{0d, \dots, Kd\} \cap \{\text{対応する日付が } m \sim m+2 \text{ 月}\} (m = 4, 7, 10)$  とした場合）についても同様の傾向が観測されている（表 3-6 を参照せよ）。

**Remark 3.** 上の様な冪乗則が観測された要因については現時点ではまだ明らかにされていないが、一つの可能性として、HFT の蔓延による金融市場の多様性の低下が挙げられるかもしれない。即ち、人間が市場の動向を目で見て自身の意思で発注を行うのに対して、HFT によるアルゴリズムトレーディングにおいては、アルゴリズム自体は統一的では無いとしてもやはり取引の画一化は避けられず、それに伴い何らかの意味での大数の法則が働き、図 25, 26 のような極めて人工的と言える現象が生じているのかもしれない。

一方で、仮定 (i) による制約が背後の一因となっている可能性もある [1]。即ち、本来であれば 1 秒間に数十回あるいはそれ以上の回数、新規注文到着及びキャンセルが繰り返されている（かもしれない）のに対し、本研究で使用したデータはそれらをまとめて一つの注文と整理しているため、そこでやはりノイズの消滅が起こっているかもしれない。

上記のいずれかを仮定した上でも尚、両対数プロットにおいて「株価からの距離」と「注文到着頻度（強度）」が直線関係にある理由に対する明確な解答を得る事は難しいが、[20] で紹介されている他の様々な自然現象・社会現象に見られる冪乗則と同様、大変興味深い現象と言える。



## 5.2 注文到着分布 $\mu$

$\mu^{k,c}$  を特徴付ける各パラメーター  $r^{k,c}, \lambda_1^{k,c}, \lambda_2^{k,c}$  ( $k = 0, \dots, 10, c = a, b$ ) に対して, (4.3) を数値的に解く事で 1 日毎の推計値を算出し, その中央値を表 7 にまとめた.

まず,  $r^{k,c}$  の中央値が上限値である 1 にかなり近い値となっている事が分かる. これは, 混合分布のうち指数分布が表れる確率が多項分布よりも相当に高い事を意味している. しかし,  $r^{k,c}$  の推定値の時系列推移 (図 27-48) を見るとこれは必ずしも正しいとは言えない. 特に, 図 39, 40 では 2012 年 6 月頃に  $r^{6,a}, r^{6,b}$  の値が大きく低下している様子が見て取れる. これは図 13, 14 で見られた, 注文枚数が 10 枚のところでは生じる分布の歪みに対応している. 実際, 元データを見るとこの時期に丁度 10 の注文が頻繁に (しかし不定期に) 出されている事が分かる. このように, 一回あたり注文枚数分布として多項分布を用いた混合分布を採用した事によって, 実際の株式市場で散見される「キリの良い数字の枚数の注文」という特徴的な性質を正しく捉える事が出来ていると言える.

しかし, 多項分布における各  $s_1, \dots, s_m$  の現れやすさを表す重み付けパラメーター  $\lambda_2^{k,c}$  に関しては数値的最適化の計算の際の下限値である  $10^{-5}$  に張り付いてしまっており, 最適化計算の手法の改善は今後の課題の一つである.

## 5.3 指値注文キャンセル分布 $\tilde{\mu}$

前節と同様,  $\tilde{\mu}^{k,c}$  を特徴付けるパラメーター  $\tilde{r}^{k,c}, \tilde{p}^{k,c}$  について, (4.4) を解いて 1 日毎の推計値を算出したものの中央値が表 8 にまとめられている. これを見ると, 平均的には  $\tilde{r}^{k,c}$  は 1 に非常に近い値となっており, fat-tail 性を取り込むために採用した一様分布はほとんど選ばれていない事が分かる. また, 前節の  $r^{k,c}$  と異なり, いずれの期間においてもほぼ安定的に 1 に近く, 混合分布を採用した事による効果はさほど見られない. 更に,  $\tilde{p}^{k,c}$  もまたかなり大きな値となっている様子が見られる.  $\tilde{p}^{k,c}$  は概ね, キャンセルが到着した時に LOB 上の指値注文のうちどれだけの割合がキャンセルされるかを表す値と言える. 最良気配値から離れた価格帯におけるキャンセル率が 9 割程度というのは, 実際の市場における現象と比較して高過ぎると考えられ,  $\tilde{\mu}^{k,c}$  の推計手法については今後改善の余地が残っている.

**Remark 4.** 実は, 頻度という意味では HFT による指値注文のキャンセル率は非常に高く, およそ 8 ~ 9 割の指値注文はすぐにキャンセルされるのが実態であるという [1]. これは, 超短期アルファを収益の源泉とする HFT の多くのアルゴリズムでは僅かでも発注が遅れた際は機会損失を避けるためにすぐにキャンセルを行う事が多いためである.

しかし, 本研究においては仮定 (i) の制約により, 上記のような高頻度の発注・キャンセルの多くは相殺されていると考えられ, 本節における推計結果が上述のような HFT の実態を捉えたものであるとは言い難い.

## 6 期待 LOB 及び対応する期待 MI 関数に対するシミュレーション分析

前章で得られたパラメーター推定値を用いて, [15] の 6 章と同様の手法に基づき期待 LOB に対するモンテカルロシミュレーションを行い, 対応する (平均的) ワンショット MI 関数を再現してその形状を調べる. 但し, [15] では注文到着を表す確率過程として拡張 Delattre–Robert–Rosenbaum モデルが採用されているのに対し, 本研究では 2 章で提案した複合 Poisson 過程を用いる. また

パラメーターは前章で得られた日次の推定値 (Poisson 強度は一時間当たりの値にスケールリングする) の中央値を用いて与える. また先物価格  $(P_t)_t$  は幾何 Brown 運動に従うとし, ドリフトは価格過程がマルチンゲールになるように設定する. 即ち

$$P_t = P_0 \exp\left(-\frac{\sigma^2}{2}t + \sigma B_t\right), t \in [0, T]$$

であり,  $(B_t)_t$  は標準 Brown 運動. ここで  $\sigma$  は分次収益率を用いて計算した RV である (Poisson 強度と同じく一時間当たりのものにスケールリングし, またやはり日次で推定したものの中央値を採用した).

[15] の 6.2 節の手順 1. ~ 5. によって, 将来時刻  $T$  における LOB  $(F_T(l))_{l \in \mathbb{N}}$  を得る事が出来る<sup>17</sup>. ここで  $F_T(l)$  は  $T$  時点における価格帯  $l$  における板の厚さであり, 正の時は指値買いを, 負の時は売りをそれぞれ表すものとする. 但し日経 225 先物の呼値は 10 円である事から, ここでは「価格帯  $l$ 」とは  $l \times 10$  円の価格帯を表す事とする.

上の手順を  $m$  回繰り返す, その平均値を取る事によって  $T$  時における期待 LOB  $(\bar{F}_T(l))_{l \in \mathbb{N}}$  が計算出来る. ここでは  $m = 50,000$  とし, また時間の離散化に関しては  $[0, T]$  を 10,000 分割した, 即ち時刻の刻み幅  $d$  として  $1/10000$  時間 = 0.36 秒と設定した. 最終時刻は  $T = 6$  時間としたが, これは一日の日中取引が行われるおおよその時間を表している. また初期時点における先物価格は 10,000 円とし ( $P_0 = 1000$ ), 各価格帯における初期 LOB は  $F_0(l) = 1(l < 1000), 0(l = 1000), -1(l > 1000)$  と設定した.

**Remark 5.**  $t$  時点における LOB  $F_t = (F_t(l))_l$  を  $\mathbb{R}^{\mathbb{N}}$  値確率変数とみなした時に, 上の手順で得られる期待 LOB  $(\bar{F}_T(l))_{l \in \mathbb{N}}$  は,  $T$  が十分大きいならば  $F_t$  の不変分布の期待値  $(E[F_\infty(l)])_l$  の近似値となっていると考えられる. 厳密にこれを主張するためには, 確率過程  $(F_t)_t$  のエルゴード性を示す必要があるが, 今後の課題としたい. なお, 同様の考え方により LOB の長期挙動を示した研究として [6] がある.

図 49 は上記のシミュレーション結果を表している. 横軸が板の厚さを表しているが, 最も厚い価格帯においても指値注文枚数はせいぜい実際の市場における最小注文単位である 1 枚程度であり, 実態から明らかにかけ離れた結果となっている.

一方, 指値注文のキャンセルが起こらない, 即ち  $\lambda^{k,c} = 0, k = 1, \dots, n, c = a, b$  とした場合, シミュレーション結果は図 50 のようになり, 実際の板の厚さと概ね整合的となる. これより, 現状では指値注文キャンセル枚数分布  $\mu$  の推定手法にまだ改善の余地がある事が確認出来る.

図 50 で示される期待 LOB に対して [15] の (2.2) 式を用いて計算される (ワンショット) MI 関数  $g(\psi)$  を描くと図 51, 52 のようになり, 一回の成行注文に対応する MI 関数が S 字型となっている様子が見て取れる. これは [14, 15] で示唆された事と整合的である.

最後に, 上で推計された MI 関数  $g(\psi)$  の下でのリスク中立的なトレーダーに対する期待売却執行コスト最小化問題を解き, 最適分割執行戦略によって生じる Implementation Shortfall (IS) コスト ([4, 15] を参照) を計算しておこう. まず図 52 の関数に対して  $\psi \in [0, 6000]$  の上で切片固定 (3 次) 多項式回帰を行うと次の結果を得る<sup>18</sup>.

$$g(\psi) = 1.44 \cdot 10^{-13} \psi^3 - 9.98 \cdot 10^{-10} \psi^2 + 4.55 \cdot 10^{-6} \psi, R^2 = 0.999 \quad (6.1)$$

<sup>17</sup>厳密には, 上の手順に更に指値注文キャンセル到着のシミュレーションを追加する必要があるが, 新規注文到着と同様であるため詳細は省略する.

<sup>18</sup> $\psi$  の上限を 6,000 で打ち切ったのは, 期待 LOB 算出のためのシミュレーションにおける初期の LOB の影響を排除するためである. 即ち, 今回のシミュレーションでは LOB 初期値として板の厚さを 1 枚と薄く設定しており, も

なお、これは [13] における離散時間モデルの MI 関数  $g_n(\psi)$  ( $n$  は時間分割回数であり、本論文における  $1/d$  に対応している) に対応するものである。(6.1) を離散時間 MI 関数と捉えた場合の対応する連続時間 MI 関数は

$$\hat{g}(\zeta) = d^{-1}g(d\psi)$$

となる。ここで  $\zeta$  は無限小時間における瞬間的執行枚数、即ち執行速度を表している (詳細は [13] を参照せよ)。この時、 $\psi$  枚の売却分割執行を行う際の最適期待売却代金は次の値関数で与えられる。

$$\begin{aligned} V(\psi, P_0) = & \sup_{(\zeta_r)_{r \in \mathcal{A}_T(\psi)}} \mathbb{E}[W_T] \\ \text{s.t. } & dW_r = \zeta_r P_r dr, \quad W_0 = 0, \\ & dP_t = P_t(\sigma dB_t - \hat{g}(\zeta_t) dt) \end{aligned} \quad (6.2)$$

ここで  $\mathcal{A}_T(\psi)$  は許容執行戦略全体を表す集合。対応する IS コストは

$$\text{IS}(\psi) = -\log \frac{V(\psi, P_0)}{\psi P_0}$$

と表される。

図 53 はワンショットの MI 関数  $g(\psi)$  と分割執行に対する MI 関数 (IS コスト関数)  $\text{IS}(\psi)$  を表している。大量執行において、分割執行によって執行コストを大きく削減出来る事、 $g(\psi)$  は S 字型であるのに対し  $\text{IS}(\psi)$  は緩やかな凹型である事等が見て取れる。

**Remark 6.** 上の最適執行問題 (6.3) は 2 章で与えたモデルと厳密には対応していない。2 章のモデルは各時点における指値・成行注文の到着一つ一つをモデル化した micro scale のモデルとなっている。そのような micro scale における指値・成行注文執行を考慮した最適分割執行問題は [7, 11] 等で扱われており、そこでは最早、約定価格という意味での証券価格を古典的な伊藤過程として記述する事は出来ない。

## 7 まとめと今後の課題

本稿では高頻度複数気配データ・約定データを用いた LOB の推定手法を提案し、日経 225 先物の秒次データを用いて実証分析を行った。そして、注文 (及びそのキャンセル) 到着の頻度を表す Poisson 強度パラメーターと最良気配値 (または証券の時価) からの距離の間に強い冪乗則が現れる事を発見した。また、一回あたり注文到着枚数分布に混合分布を適用する事で、人間の恣意性に基づく「キリが良い数字」における分布の歪みを抽出する事が出来た。更に、推計した各パラメーターの下での期待 LOB に対応する (ワンショット) MI 関数が S 字型となっている事をシミュレーション分析によって示した。

指値注文キャンセル分布の推定手法については改善の余地があると言える。キャンセル枚数の上限を設定するために今回は二項分布を用いたモデルを扱ったが、他の手法として、キャンセルし Remark 5 で指摘している通りのエルゴード性が保証されているならば  $T \rightarrow \infty$  の極限移行によって LOB 初期値の影響は排除されるはずであるが、実際には  $T$  は有限値であるために  $P_0 = 1000$  から離れた価格帯における指値注文の枚数は初期値である  $\pm 1$  程度となり、 $\psi$  が十分大きい場合の MI 関数は急激に増大してしまう。このような初期値の影響を避けるために  $\psi$  を 6,000 枚程度で打ち切るのが妥当と判断した。

枚数自体ではなくキャンセル率の分布を推定する事が考えられる。一回あたりキャンセル率に対する適切な分布の設定及び推定手法の構築は今後の主要な課題の一つである。

また今回、新規注文・注文キャンセル到着枚数を表す確率過程は全て独立であるという仮定を置いていた。しかし実際の市場においてはこれらの間には依存関係が観測されると考えられる。これらの確率過程の間の相関を考慮した推定手法の構築も今後の課題としたい。

今回は日経 225 先物に対する LOB 推定を行ったが、同様の分析を個別銘柄に対して行う事も興味深いテーマである。特に、銘柄規模や流動性の違いによる推計結果の差異や、業種毎の特性等を調べる事は実務への応用上も重要であると言える。また、流動性の低い小型銘柄であれば HFT が占める割合も低く、本研究における仮定 (i) の制約が現実的となるような分析も可能となるかもしれない。

理論的な課題として、下記の二つの意味での極限移行に関するテーマが挙げられる。

一つ目は Remark 5 で記したようなエルゴード性（長期漸近挙動）に関するものである。エルゴード性が保証されるならば、十分大きな  $T$  における期待 LOB を実際に長期的な平均 LOB として特徴付ける事が出来、対応する平均的 MI 関数を、[13] で扱っているような時間的一様な非線形関数として特徴付ける事が出来る。

二つ目は Remark 6 で取り上げた、「古典的な」最適執行問題 (6.3) との整合性に関するものである。2 章で構築したモデルが micro scale であるのに対し、(6.3) で与えられているのはもう少し長い時間間隔（mezzo scale と呼ぶ事とする）における執行問題であり、所謂 optimal scheduling (slicing) に対応している。何らかの意味での time scaling により、micro scale と mezzo scale それぞれにおける最適執行問題を関係付ける事は理論的に重要な課題である。その際、呼値の幅に関しても scaling を考える事も重要となる可能性がある。近年、東証において呼値の単位の段階的な適正化が進んでおり、大型銘柄については呼値の縮小が進んでいる。また海外においても呼値の縮小に関する議論は活発であり、特に米国においては 2010 年 5 月のフラッシュ・クラッシュ ([16, 17])、及び 2014 年 3 月の “Flash Boys [18]” の出版等を受けて呼値の適正化に関する議論が活発となっている<sup>19</sup>。取引の高頻度化と共に呼値の縮小は市場の健全性・公平性に繋がる可能性があると考えられており ([2])、理論的にも、time scaling と tick size scaling の両方を考慮する事によって mezzo scale model への橋渡しを行う事が出来るかもしれない。本質的に大きな意味のある今後の研究テーマと考えられる。

## 謝辞

本論文の執筆にあたり数々の有益な御助言をいただきました立命館大学の足立 高德教授と University of California, Berkeley の山本 寛道氏に感謝致します。

## References

- [1] 足立 高德, 研究集会「ファイナンスの数理解析とその応用 (FMA2014)」におけるディスカッション (2014).
- [2] 足立 高德, 高頻度取引の実際, 研究集会「金融工学・数理計量ファイナンスの諸問題 2014」における講演 (2014).

<sup>19</sup>一方で、画一的な呼値の縮小は小型銘柄の流動性を低下させる可能性がある事も指摘されている [23]。また我が国においても、TOPIX100 採用銘柄以外に対する呼値縮小に関しては慎重な姿勢が取られている。

- [3] Ait-Sahalia, Y., Mykland, P. and Zhang, L., How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise, *The Review of Financial Studies*, **18**(2) (2005), 351–416.
- [4] Almgren, R., F. and Chriss, N. Optimal execution of portfolio transactions, *Journal of Risk*, **18** (2000), 57–62.
- [5] Barndorff-Nielsen, O. E., Hansen, P. R. Lunde, A. and Shephard, N., Designing realized kernels to measure the ex post variation of equity prices in the presence of noise, *Econometrica* **76**(6) (2008), 1481–1536.
- [6] Cont, R., Stoikov, S. and Talreja, R., A stochastic model for order book dynamics, *Operations Research*, **58**(3) (2010), 549–563.
- [7] Crisafi, M. A. and Macrina, A., Optimal execution in lit and dark pools, arXiv preprint.
- [8] Gould, M. D., Porter, M. A., Williams, S., McDonald, M., Fenn, D. and Howison, S., Limit order books, Preprint (2012).
- [9] Hasbrouck, J., Assessing the quality of a security market: A new approach to transaction-cost measurement. *The Review of Financial Studies*, **6**(1) (1993), 191–212.
- [10] 保坂 豪, 東京証券取引所における High-Frequency Trading の分析, *証券アナリストジャーナル*, **52**(6) (2014), 73–82.
- [11] 家田 雅志, 有限時間ホライゾンにおける複合確率制御の数値計算, *日本応用数理学会 2013 年度年会予稿集* (2013).
- [12] Kato, T., Quantitative operational risk management: properties of operational Value at Risk (Op-VaR), *RIMS Kokyuroku*, **1818**, Financial Modeling and Analysis (FMA2012) (2012), 98–112.
- [13] Kato, T., An optimal execution problem with market impact, *Finance and Stochastics*, **18**(3) (2014), 695–732.
- [14] 加藤 恭, On convexity/concavity of market impact functions: empirical and simulation-based studies of limit order books, *RIMS Kokyuroku*, **1886**, Financial Modeling and Analysis (FMA2013) (2014), 95–117.
- [15] 加藤 恭, マーケットインパクト関数の非線形性について: 凸性・凹性に関する実証・シミュレーション分析と最適執行モデルの導出, *日本応用数理学会論文誌*, **24**(3) (2014), 203–237.
- [16] 加藤 恭, 山中 卓, システミックリスクに纏わる数理モデルについて, *日本応用数理学会論文誌*, **24**(4) (2014), 397–443.
- [17] Lehalle, C-A., Market microstructure knowledge needed for controlling an intra-day trading process, *Handbook on Systemic Risk*, Eds Fouque, J. P. and Langsam, J. (2013), 549–578.

- [18] Michael, L., *Flash Boys (A Wall Street Revolt)*, W. W. Norton & Co. Inc. (2014).
- [19] 長藤 剛, 中田 貴之, 神崎 有吾, 邦銀 18 行のオペレーショナルリスク損失データの分布形状の共通性とそれを利用したオペレーショナルリスクの簡易な計算式について, 金融庁金融研究センター調査レポート (2011).
- [20] Newman, M. E. J., Power laws, Pareto distributions and Zipf's law, *Contemporary Physics*, **46**(5) (2005), 323–351.
- [21] Podolskij, M. and Vetter, M., Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps, *Bernoulli*, **15**(3) (2009), 634–658.
- [22] Roll, R., A simple implicit measure of the effective bid-ask spread in an efficient market, *Journal of Finance*, **39** (1984), 1127–1139.
- [23] 清水 葉子, アメリカのティックサイズ拡大議論, 日本証券経済研究所『証研レポート』, **1679** (2013), 38–47.
- [24] Yamamoto, K., Stochastic model of Zipf's law and the universality of the power-law exponent, *Phys. Rev. E.*, **89**(4) (2014).
- [25] Zhang, L., Mykland, P. A. and Ait-Sahalia, Y., A tale of two time scales: determining integrated volatility with noisy high-frequency data, *Journal of the American Statistical Association*, **100**(472) (2005), 1394–1411.

School of Statistical Thinking, The Institute of Statistical Mathematics  
10-3 Midori-cho, Tachikawa-shi, Tokyo 190-8562, Japan  
E-mail address: ogihara@ism.ac.jp

統計数理研究所・統計思考院 荻原 哲平

Division of Mathematical Science for Social Systems, Graduate School of Engineering Science  
Osaka University, 1-3 Machikaneyama, Toyonaka-shi, Osaka 560-8531, Japan  
E-mail address: kato@sigmath.es.osaka-u.ac.jp

大阪大学・大学院基礎工学研究科 加藤 恭

Department of Information Science  
Toho University, 2-2-1 Miyama, Funabashi-shi, Chiba 274-8510, Japan  
E-mail address: hideyuki.takada@is.sci.toho-u.ac.jp

東邦大学・理学部情報科学科 高田 英行

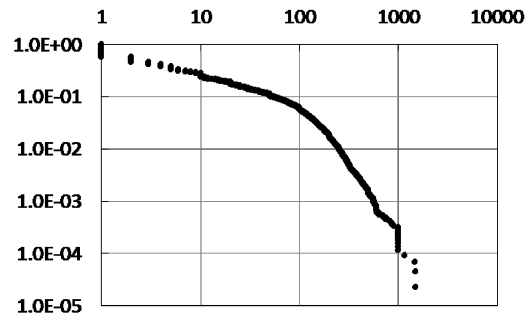
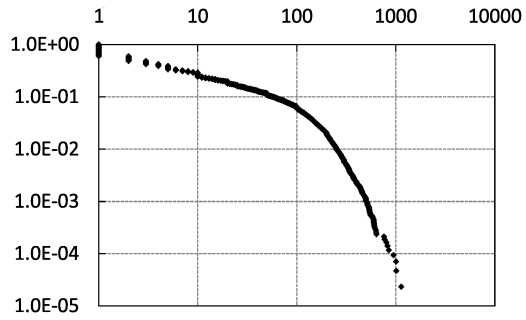


図 1: 裾確率の両対数プロット: 成行買い ( $k = 0$ ). 図 2: 裾確率の両対数プロット: 成行売り ( $k = 0$ ).

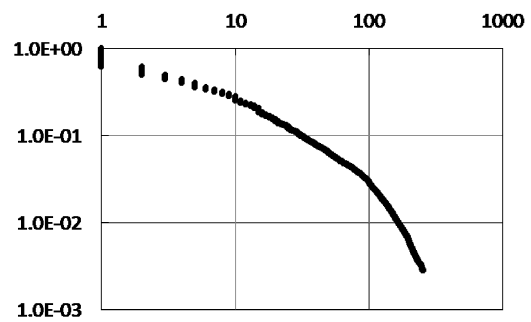
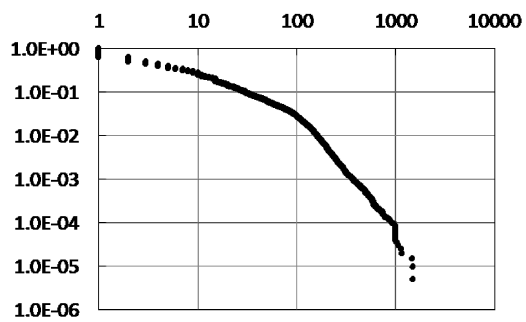


図 3: 裾確率の両対数プロット: 最良気配指値買い ( $k = 1$ ). 図 4: 裾確率の両対数プロット: 最良気配指値売り ( $k = 1$ ).

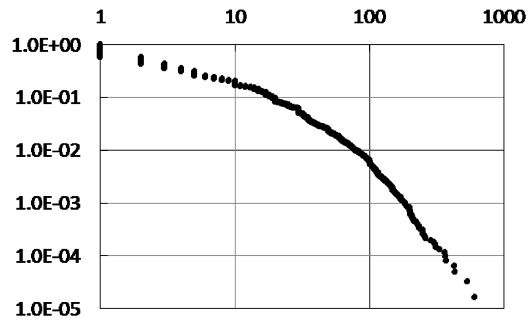
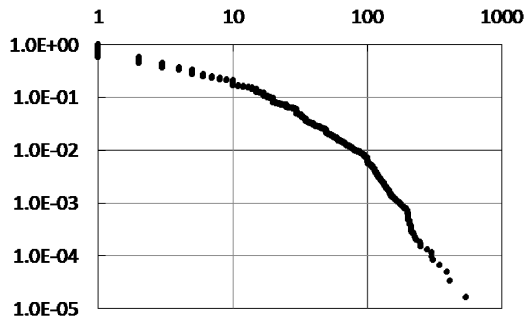


図 5: 裾確率の両対数プロット: 指値買い ( $k = 2$ ). 図 6: 裾確率の両対数プロット: 指値売り ( $k = 2$ ).

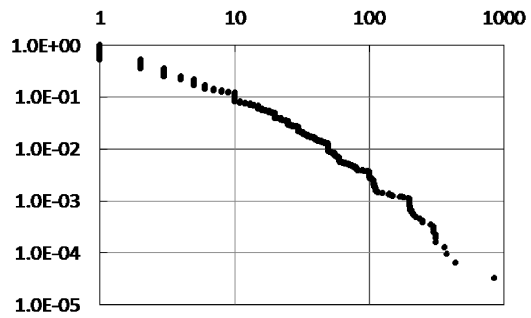
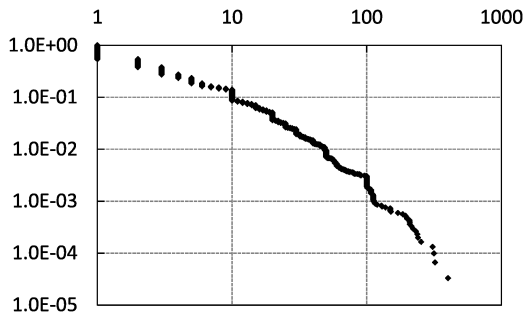


図 7: 裾確率の両対数プロット: 指値買い ( $k = 3$ ). 図 8: 裾確率の両対数プロット: 指値売り ( $k = 3$ ).

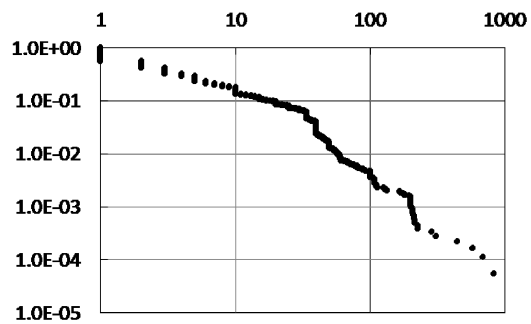
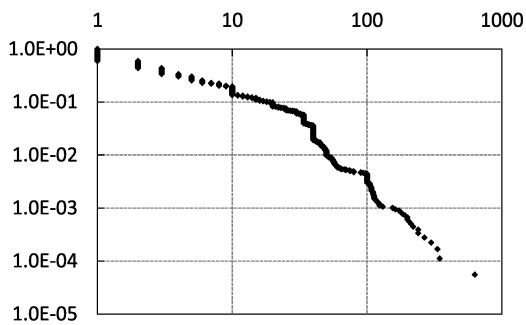


図 9: 裾確率の両対数プロット: 指値買い ( $k = 4$ ).

図 10: 裾確率の両対数プロット: 指値売り ( $k = 4$ ).



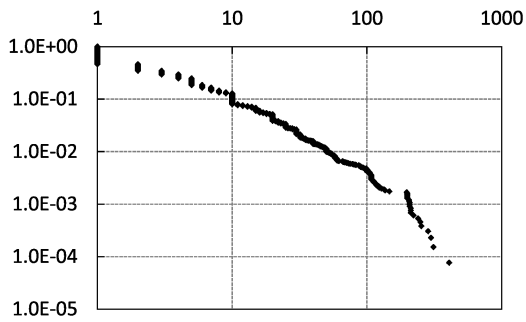


図 11: 裾確率の両対数プロット: 指値買い ( $k = 5$ ).

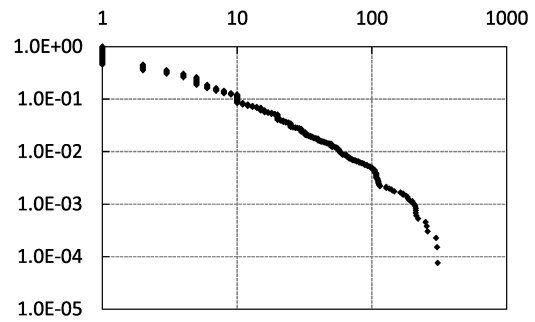


図 12: 裾確率の両対数プロット: 指値売り ( $k = 5$ ).

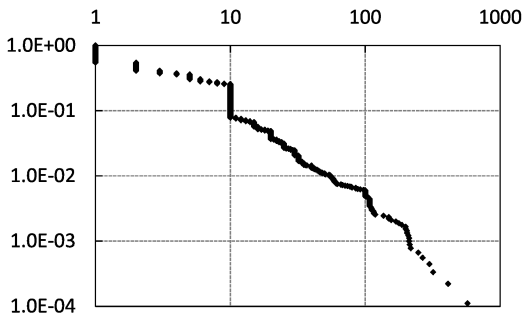


図 13: 裾確率の両対数プロット: 指値買い ( $k = 6$ ).

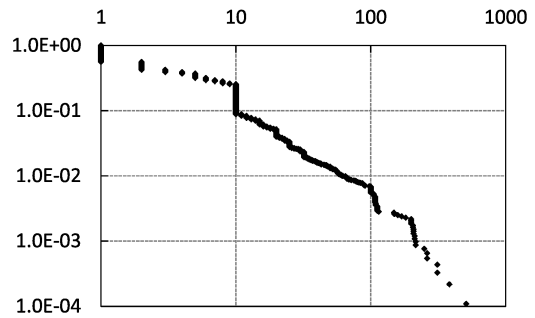


図 14: 裾確率の両対数プロット: 指値売り ( $k = 6$ ).

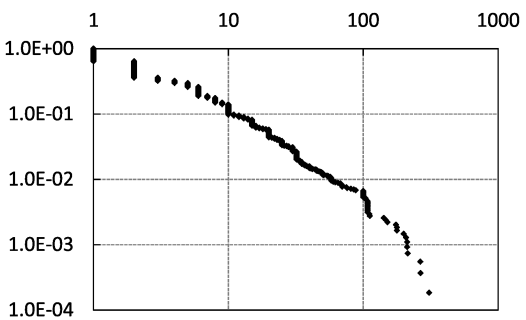


図 15: 裾確率の両対数プロット: 指値買い ( $k = 7$ ).

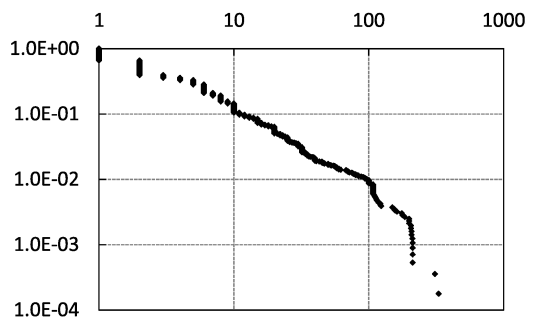


図 16: 裾確率の両対数プロット: 指値売り ( $k = 7$ ).

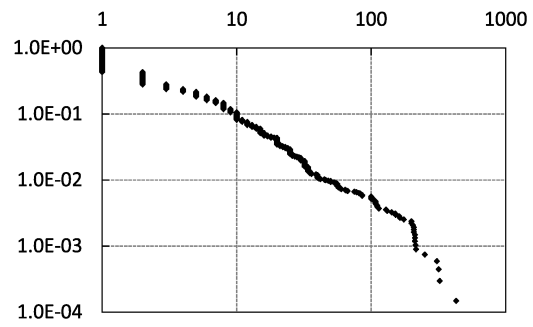
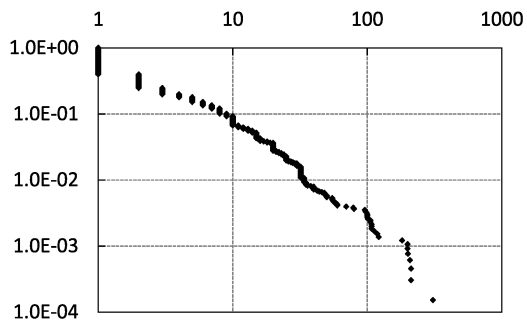


図 17: 裾確率の両対数プロット: 指値買い ( $k = 8$ ). 図 18: 裾確率の両対数プロット: 指値売り ( $k = 8$ ).

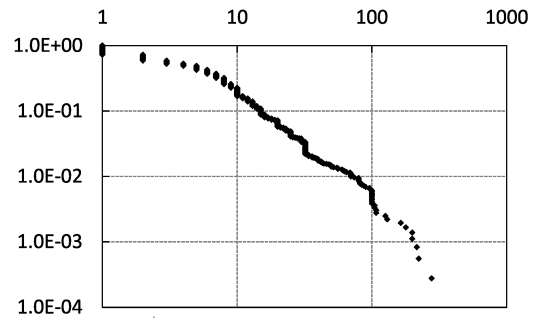
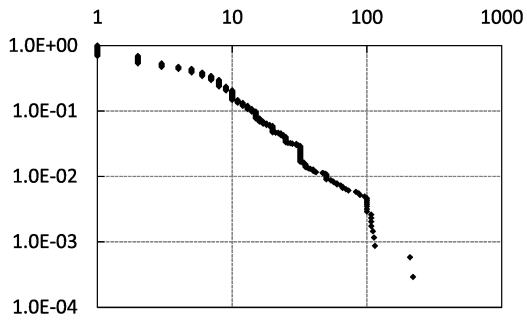


図 19: 裾確率の両対数プロット: 指値買い ( $k = 9$ ). 図 20: 裾確率の両対数プロット: 指値売り ( $k = 9$ ).

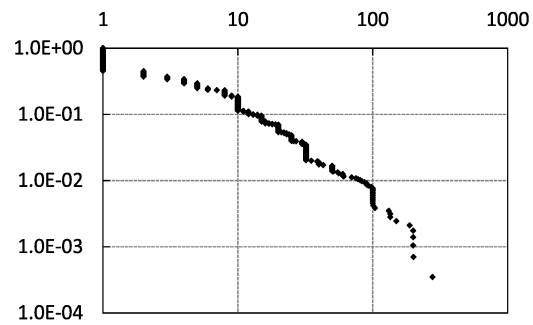
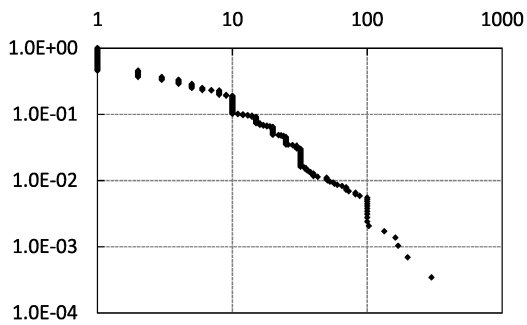


図 21: 裾確率の両対数プロット: 指値買い ( $k = 10$ ). 図 22: 裾確率の両対数プロット: 指値売り ( $k = 10$ ).

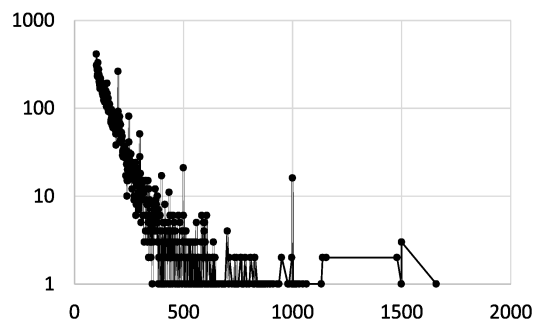
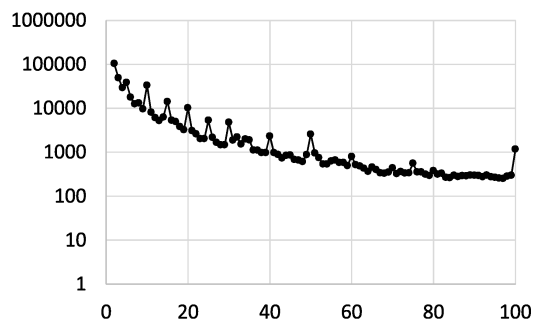


図 23: 注文到着 (キャンセル除く) 全データのヒストグラム (到着枚数 100 枚以下). 縦軸は対数目盛.  
 図 24: 注文到着 (キャンセル除く) 全データのヒストグラム (到着枚数 101 枚以上). 縦軸は対数目盛.

	market	limit1	limit2	limit3	limit4	limit5	limit6	limit7	limit8	limit9	limit10
bid	0.591	0.629	0.652	0.675	0.631	0.476	0.429	0.521	0.359	0.585	0.578
ask	0.516	0.664	0.701	0.714	0.629	0.500	0.447	0.565	0.482	0.608	0.579

表 1: 注文到着強度パラメーター  $\lambda$  (前日差分) と RV (前日差分) との相関係数.

	limit1	limit2	limit3	limit4	limit5	limit6	limit7	limit8	limit9	limit10
bid	0.543	0.615	0.624	0.564	0.450	0.317	0.441	0.281	0.408	0.434
ask	0.586	0.665	0.688	0.506	0.382	0.305	0.451	0.389	0.436	0.441

表 2: 注文到着強度パラメーター  $\lambda$  (前日差分) と RV (前日差分) との相関係数.

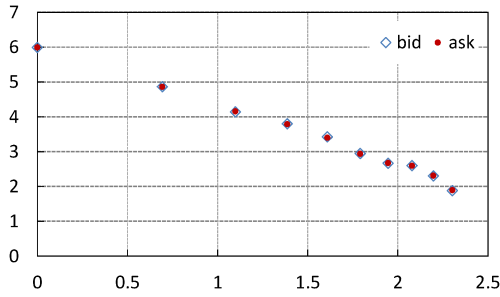


図 25: 指値注文到着強度  $\lambda^{k,c}$  ( $c = a, b$ ) の両対数プロット. 横軸は  $\log k$  を, 縦軸は  $\log \lambda^{k,c}$  をそれぞれ表している.

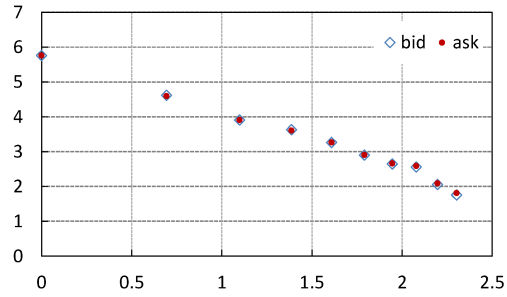


図 26: 指値注文キャンセル到着強度  $\tilde{\lambda}^{k,c}$  ( $c = a, b$ ) の両対数プロット. 横軸は  $\log k$  を, 縦軸は  $\log \tilde{\lambda}^{k,c}$  をそれぞれ表している.

	Slope(bid)	Slope(ask)	$R^2$ (bid)	$R^2$ (ask)
9 時台	-1.60	-1.59	0.994	0.994
10 時台	-1.75	-1.74	0.992	0.993
11 時台	-1.83	-1.80	0.988	0.988
12 時台	-1.60	-1.63	0.989	0.991
13 時台	-1.83	-1.84	0.993	0.995
14 時台	-1.79	-1.78	0.988	0.992

表 3: 指値注文到着強度  $\lambda^{k,c}$  ( $c = a, b$ ) の両対数プロット  $(\log k, \log \lambda^{k,c})_{k=1, \dots, 10}$  に対する単回帰係数と決定係数 (時間帯毎).

	Slope(bid)	Slope(ask)	$R^2$ (bid)	$R^2$ (ask)
9 時台	-1.67	-1.66	0.991	0.988
10 時台	-1.73	-1.70	0.983	0.986
11 時台	-1.61	-1.61	0.987	0.988
12 時台	-1.68	-1.61	0.986	0.989
13 時台	-1.77	-1.65	0.985	0.990
14 時台	-1.54	-1.50	0.984	0.988

表 4: 指値注文キャンセル到着強度  $\tilde{\lambda}^{k,c}$  ( $c = a, b$ ) の両対数プロット  $(\log k, \log \tilde{\lambda}^{k,c})_{k=1, \dots, 10}$  に対する単回帰係数と決定係数 (時間帯毎).

	Slope(bid)	Slope(ask)	$R^2$ (bid)	$R^2$ (ask)
4~6 月	-1.81	-1.80	0.990	0.990
7~9 月	-1.74	-1.74	0.989	0.990
10~12 月	-1.61	-1.61	0.985	0.987

表 5: 指値注文到着強度  $\lambda^{k,c}$  ( $c = a, b$ ) の両対数プロット  $(\log k, \log \lambda^{k,c})_{k=1, \dots, 10}$  に対する単回帰係数と決定係数 (四半期毎).

	Slope(bid)	Slope(ask)	$R^2$ (bid)	$R^2$ (ask)
4~6月	-1.79	-1.73	0.975	0.975
7~9月	-1.68	-1.63	0.978	0.983
10~12月	-1.48	-1.51	0.986	0.989

表 6: 指値注文キャンセル到着強度  $\tilde{\lambda}^{k,c}$  ( $c = a, b$ ) の両対数プロット  $(\log k, \log \tilde{\lambda}^{k,c})_{k=1, \dots, 10}$  に対する単回帰係数と決定係数 (四半期毎) .

$r^{k,c}$	market	limit1	limit2	limit3	limit4	limit5	limit6	limit7	limit8	limit9	limit10
Bid	1.000	1.000	0.990	0.966	0.968	0.977	0.972	0.982	0.985	1.000	0.955
Ask	1.000	1.000	0.993	0.968	0.978	0.974	0.977	0.987	1.000	0.978	0.945

$\lambda_1^{k,c}$	market	limit1	limit2	limit3	limit4	limit5	limit6	limit7	limit8	limit9	limit10
Bid	0.051	0.065	0.142	0.236	0.191	0.216	0.199	0.213	0.222	0.159	0.223
Ask	0.052	0.065	0.135	0.239	0.192	0.214	0.194	0.190	0.192	0.146	0.203

$\lambda_2^{k,c}$	market	limit1	limit2	limit3	limit4	limit5	limit6	limit7	limit8	limit9	limit10
Bid	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05
Ask	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05	1.0E-05

表 7: 成行・指値注文到着強度  $\mu^{k,c}$  ( $k = 0$  (成行),  $1, \dots, 10, c = a, b$ ) のパラメーター推定値 (1日毎に推定したものの中央値を掲載) .

$\tilde{r}^{k,c}$	limit1	limit2	limit3	limit4	limit5	limit6	limit7	limit8	limit9	limit10
Bid	0.995	0.998	0.998	0.998	0.997	0.997	0.997	0.997	0.995	0.996
Ask	0.995	0.998	0.998	0.998	0.997	0.997	0.997	0.997	0.995	0.996

$\tilde{p}^{k,c}$	limit1	limit2	limit3	limit4	limit5	limit6	limit7	limit8	limit9	limit10
Bid	0.593	0.799	0.877	0.853	0.894	0.882	0.899	0.874	0.882	0.875
Ask	0.584	0.790	0.882	0.859	0.900	0.893	0.895	0.881	0.881	0.887

表 8: 指値注文キャンセル到着強度  $\tilde{\mu}^{k,c}$  ( $k = 1, \dots, 10, c = a, b$ ) のパラメーター推定値 (1日毎に推定したものの中央値を掲載) .

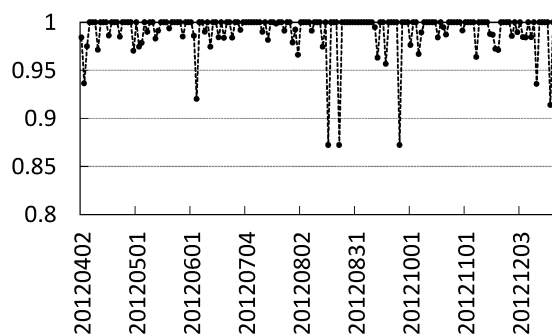
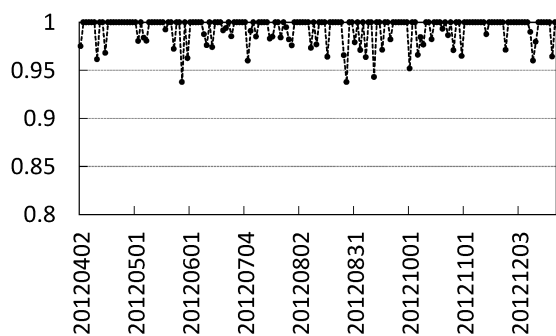


図 27: パラメーター  $r$  の時系列推移: 成行買い ( $k = 0$ ).

図 28: パラメーター  $r$  の時系列推移: 成行売り ( $k = 0$ ).

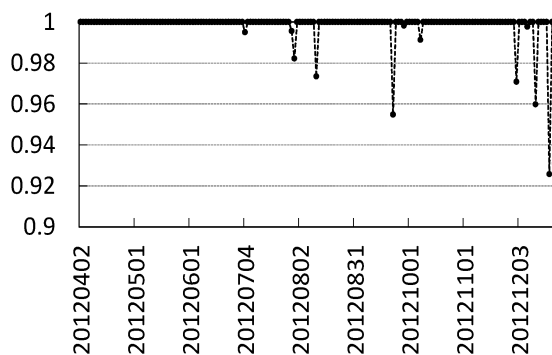
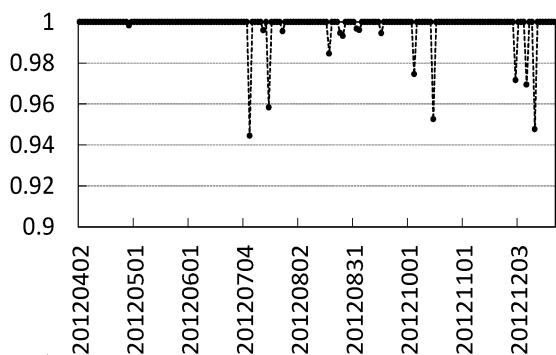


図 29: パラメーター  $r$  の時系列推移: 最良気配指値買い ( $k = 1$ ).

図 30: パラメーター  $r$  の時系列推移: 最良気配指値売り ( $k = 1$ ).

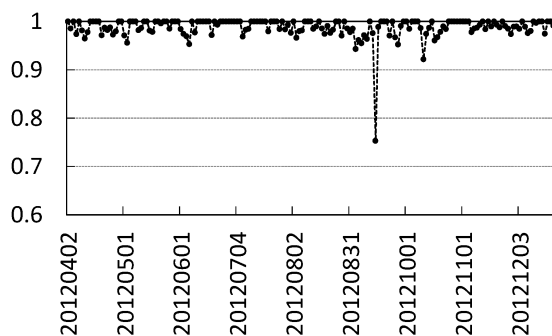
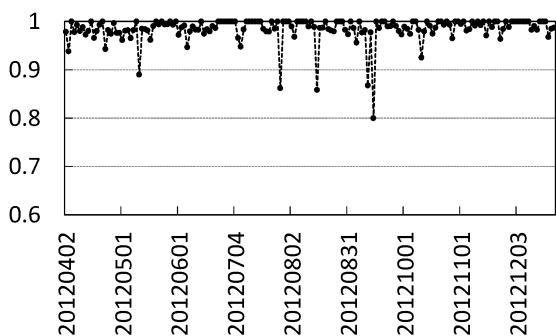


図 31: パラメーター  $r$  の時系列推移: 指値買い ( $k = 2$ ).

図 32: パラメーター  $r$  の時系列推移: 指値売り ( $k = 2$ ).

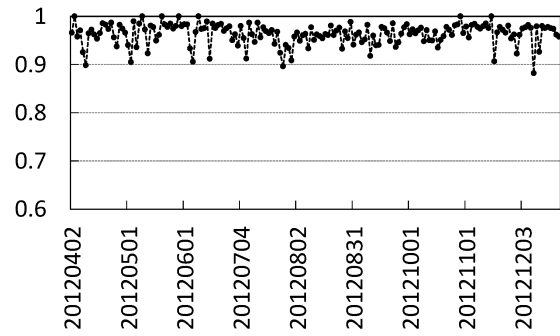
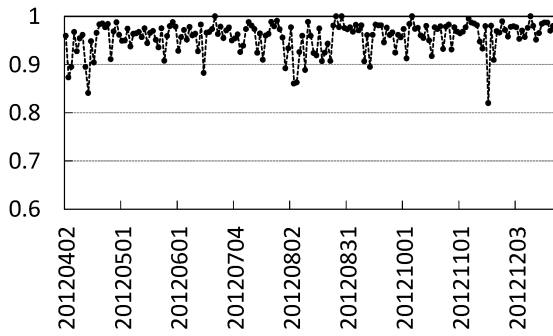


図 33: パラメーター  $r$  の時系列推移: 指値買い (  $k = 3$  ).

図 34: パラメーター  $r$  の時系列推移: 指値売り (  $k = 3$  ).

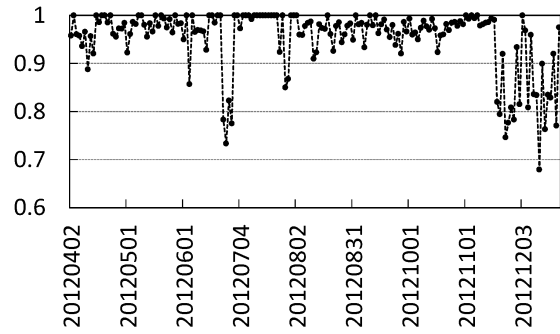
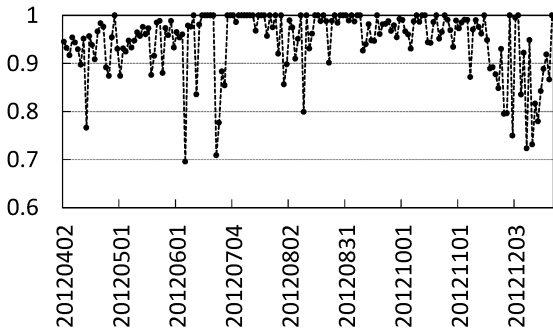


図 35: パラメーター  $r$  の時系列推移: 指値買い (  $k = 4$  ).

図 36: パラメーター  $r$  の時系列推移: 指値売り (  $k = 4$  ).

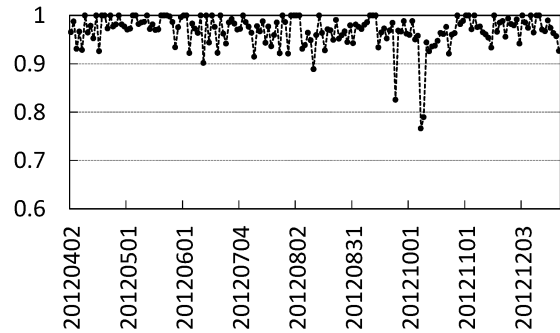
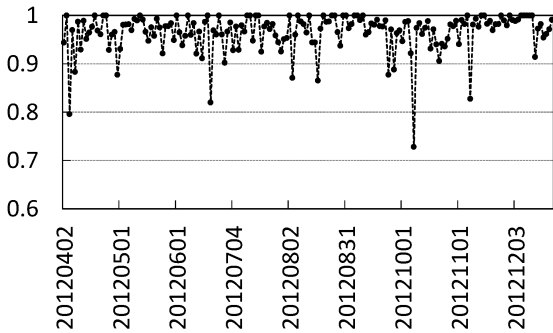


図 37: パラメーター  $r$  の時系列推移: 指値買い (  $k = 5$  ).

図 38: パラメーター  $r$  の時系列推移: 指値売り (  $k = 5$  ).

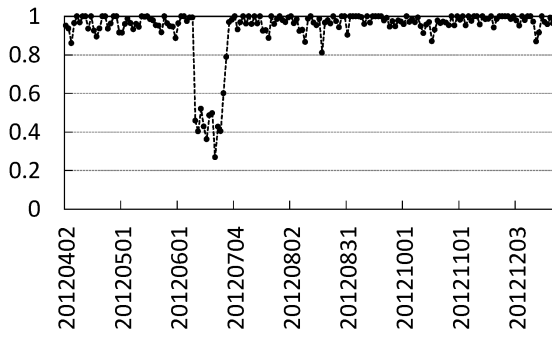


図 39: パラメーター  $r$  の時系列推移: 指値買い ( $k = 6$ ).

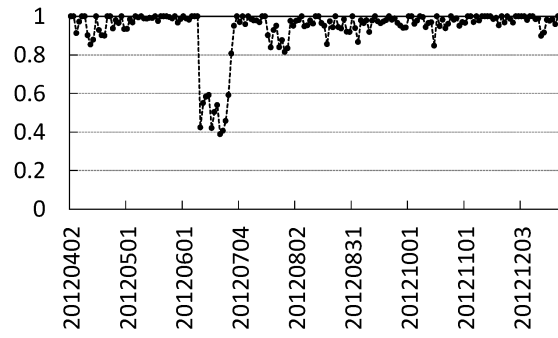


図 40: パラメーター  $r$  の時系列推移: 指値売り ( $k = 6$ ).

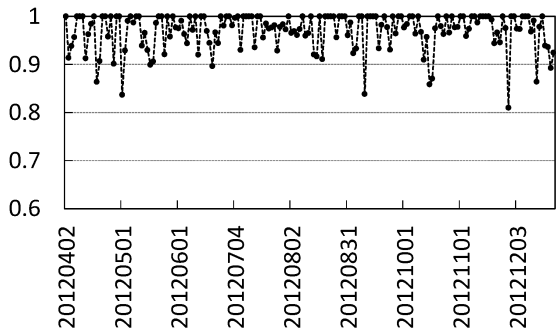


図 41: パラメーター  $r$  の時系列推移: 指値買い ( $k = 7$ ).

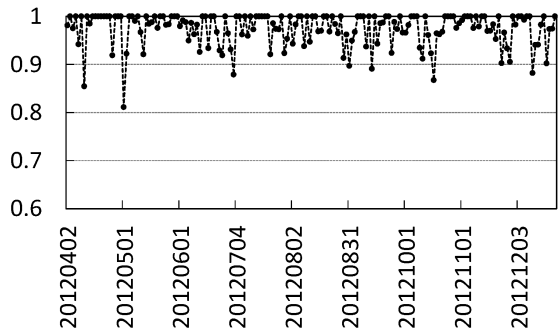


図 42: パラメーター  $r$  の時系列推移: 指値売り ( $k = 7$ ).

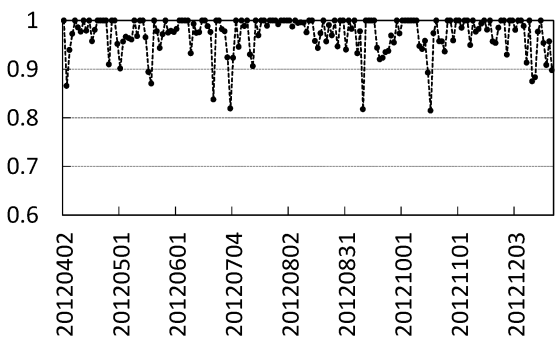


図 43: パラメーター  $r$  の時系列推移: 指値買い ( $k = 8$ ).

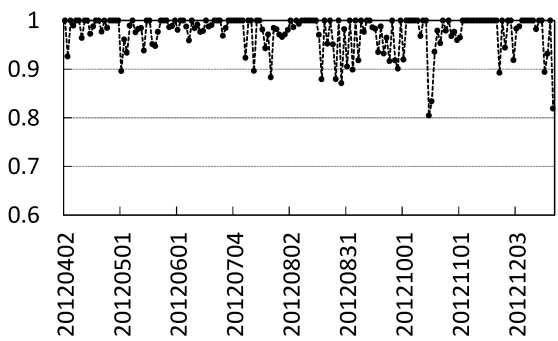


図 44: パラメーター  $r$  の時系列推移: 指値売り ( $k = 8$ ).



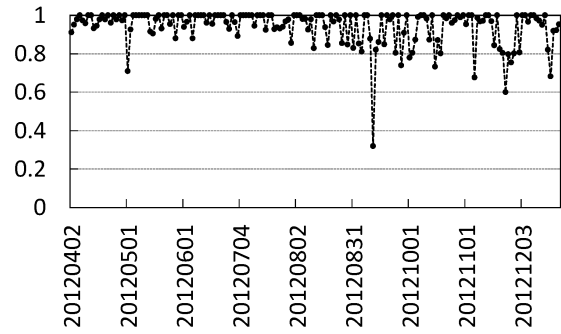
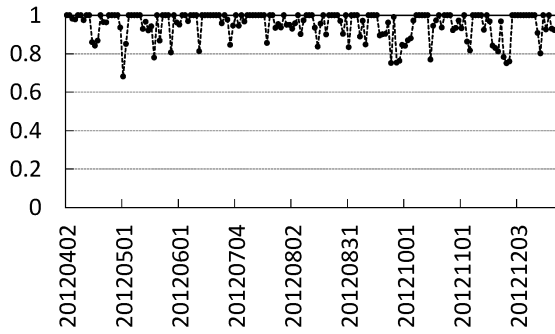


図 45: パラメーター  $r$  の時系列推移: 指値買い (k = 9). 図 46: パラメーター  $r$  の時系列推移: 指値売り (k = 9).

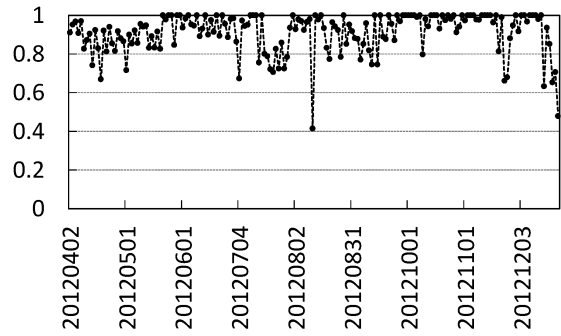
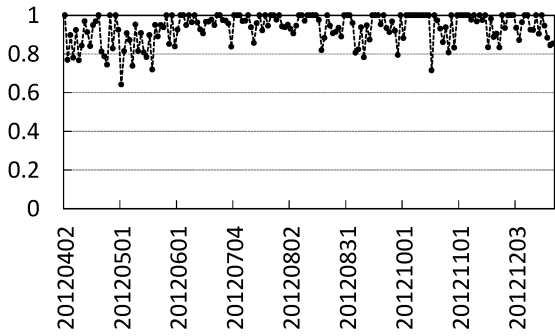


図 47: パラメーター  $r$  の時系列推移: 指値買い (k = 10). 図 48: パラメーター  $r$  の時系列推移: 指値売り (k = 10).

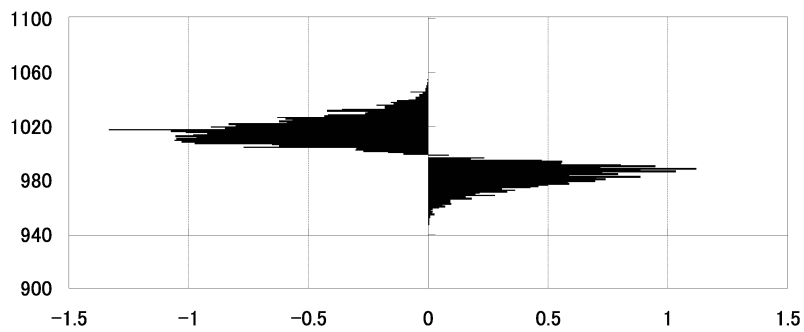


図 49: モンテカルロシミュレーションによる期待 LOB (指値注文キャンセルあり). 横軸は指値注文枚数 (正は bid, 負は ask). 縦軸は先物価格を 10 で割った値.

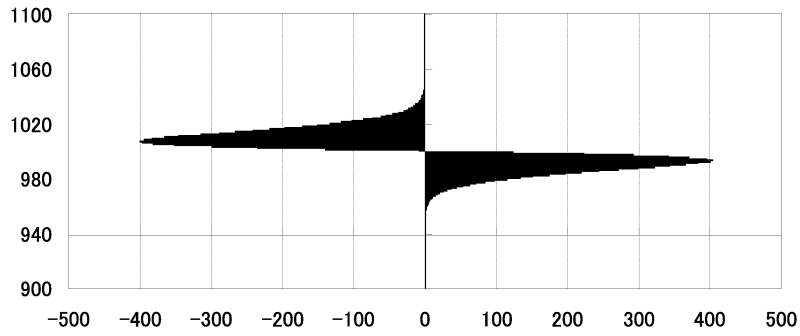


図 50: モンテカルロシミュレーションによる期待 LOB (指値注文キャンセル無し). 横軸は指値注文枚数 (正は bid, 負は ask). 縦軸は先物価格を 10 で割った値.

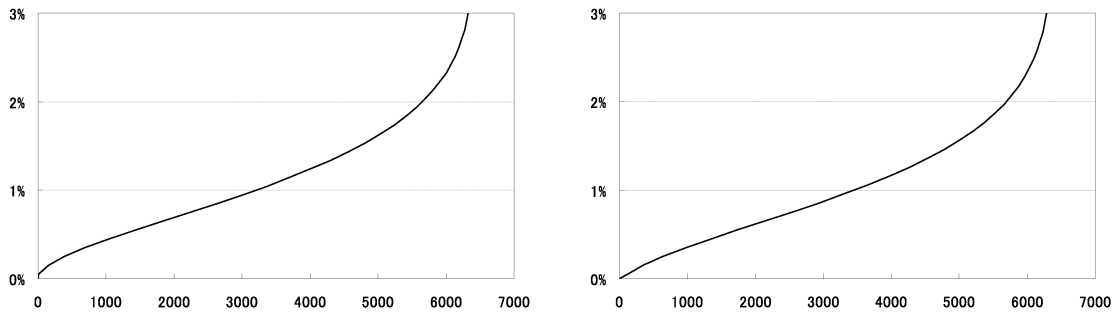


図 51: 期待 LOB から推計した成行買い執行に対する MI 関数. 横軸は執行枚数. 縦軸はワンショットインパクトコスト.  
 図 52: 期待 LOB から推計した成行売り執行に対する MI 関数. 横軸は執行枚数. 縦軸はワンショットインパクトコスト.

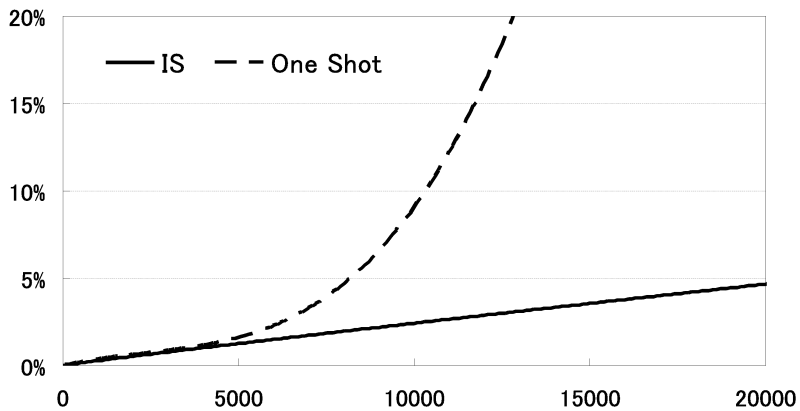


図 53: ワンショット MI 関数  $g(\psi)$  と IS コスト関数  $IS(\psi)$  の比較. 横軸は執行枚数  $\psi$ . 縦軸は執行コスト.