

Міністерство освіти і науки України
Київський національний університет імені Тараса Шевченка
Кафедра прикладної статистики факультету кібернетики

ОПТИМАЛЬНА СТРАТЕГІЯ КУПВЛІ Ф'ЮЧЕРСНИХ КОНТРАКТІВ

**Текстова частина до випускної кваліфікаційної роботи магістра
за спеціальністю „Системи і методи прийняття рішень” 8.04030302**

Керівник випускної
кваліфікаційної роботи
д.ф.-м.н., проф. Лебедев Є.О.

(підпис)
“ ____ ” червня 2016 р.

Виконав студент Какоиченко А.І.
“ ____ ” червня 2016 р.

Робота заслухана на засіданні кафедри прикладної статистики та рекомендована до захисту в ЕК, протокол № 9 від 17 травня 2016 р.

Завідувач кафедри прикладної статистики проф. Лебедев Є.О. _____

Київ 2016

ЗМІСТ

Стор.

Анотація	3
ВСТУП	4
РОЗДІЛ 1: Побудова моделі короткострокової динаміки ф'ючерсного ринку	
1.1. Загальна модель ринку ф'ючерсів.....	6
1.2. Вибірки даних для проведення дослідження.	11
1.3. Програмне забезпечення для проведення дослідження.	11
РОЗДІЛ 2: Статистичний аналіз потоків транзакцій	
2.1. Аналіз залежності інтенсивності потоку транзакцій від цінового рівня потоку транзакцій.	14
2.2. Дослідження взаємозв'язків між потоками транзакцій.	17
2.3. Визначення типу розподілу, що описує час між однотипними подіями.	22
РОЗДІЛ 3: Модель з дискретним часом	
3.1. Дослідження взаємозв'язку між інтенсивностями різних потоків.	30
3.2. Марковська модель.	34
РОЗДІЛ 4: Побудова оптимальної стратегії купівлі ф'ючерсних контрактів	
4.1. Стратегія розбиття заявки на заявки мінімального обсягу.	38
4.2. Стратегія на основі марковської моделі.	39
4.1. Порівняння стратегій.	40
Висновки по роботі та рекомендації для подальших досліджень . . .	42
Література	43
Додаток А. Програма для статистичного дослідження потоків транзакцій.	45
Додаток Б. Програмний код та файли налаштувань модулів програмного комплексу Watcher для експорту історичних біржових даних	62
Додаток В. Програма для числового розрахунку аналітичної формули з розділу 3.2	66
Додаток Г. Програма для проведення імітаційного моделювання з розділу 4.....	67

Анотація

У даній роботі проведено дослідження процесу ціноутворення ф'ючерсного ринку шляхом вивчення характеристик потоків транзакцій на постановку, зняття та зміну параметрів заявок. Розроблена модель дозволяє уникнути обмежень, які присутні у роботах [4],[5] і пов'язані з проблемами, які виникають при спробі опису проміжків часу між двома послідовними однотипними транзакціями, як показникових випадкових величин, чи випадкових величин, розподілених за гамма-розподілом.

В рамках проведених досліджень динаміку ринку промодельовано за допомогою марковського процесу та отримано аналітичну формулу для знаходження ймовірностей зростання та спадання ціни у короткотерміновій перспективі. Також розроблено стратегію купівлі ф'ючерсних контрактів на основі даного підходу. Ефективність стратегії досліджено за допомогою методів імітаційного моделювання.

Вступ

Більшість підходів моделювання та прогнозування динаміки фінансових ринків основана на дослідженні цінових рядів. Стан ринку у деякий момент часу у таких роботах описується цінами фінансових інструментів. Наприклад, у роботі [1] досліджується стратегія парної торгівлі. У роботах [2],[3], виконаних раніше автором даного дослідження, розглянуто стратегії статистичного арбітражу. Головний недолік таких робіт полягає у тому, що за вхідну інформацію узято лише цінові ряди двох активів з високим рівнем кореляції. Таким чином ігнорується інформація про поточний стан та динаміку книги заявок, що робить такі підходи недостатньо ефективними для короткострокового прогнозування та розробки стратегій високочастотної торгівлі.

Однак розвиток високочастотної торгівлі у середині 2000-х років та потреба у алгоритмах для автоматизованого підтримання ліквідності призвів до появи робіт, основаних на аналізі стану книги заявок та характеристиках потоків надходження та видалення нових заявок. Одними з перших робіт на дану тематику є [4], [5], у яких проведено дослідження динаміки книги заявок. Наприклад, у роботі [6] ставиться задача визначення оптимальної стратегії маркет-мейкінгу на американських акціях. Тобто, алгоритм має одночасно виставляти дві заявки - одну на купівлю, а іншу на продаж. Автори роботи роблять припущення про те, що потік зустрічних заявок, що задовольняють виставлену заявку, можна описати, як пуассонівський процес з інтенсивністю, яка функціонально залежить від різниці ціни виставленої заявки та ціни заявки з таким самим напрямом та першим місцем у черзі заявок. Результатом дослідження стали формули для обрахунку оптимальних цін для постановки заявок на купівлю та продаж. Стверджується, що стратегія, яка основана на отриманих формулах, має більше математичне сподіванням та меншу дисперсією, ніж очевидна стратегія маркет – мейкінгу, така, як виставлення

заявок за цінами перших заявок у чергах заявок на купівлю та продаж. У статті [7] дослідники звертають увагу на потоки транзакцій на постановку, видалення та задоволення лімітних заявок на строковому ринку Московської біржі. Автори припускають, що проміжки часу між транзакціями мають гамма розподіл з параметрами, залежними від часу. Проте, статистична перевірка цього твердження показала його суперечливість. Результатом дослідження став суб'єктивний висновок про залежність динаміки ринку від оцінок параметрів гамма-розподілів потоків транзакцій. Роботи [8], [9] показують, що стан книги заявок містить інформацію про короткострокову динаміку ціни, і може бути корисним з точки зору прогнозування стану ринку.

Метою даної роботи є ґрунтовне дослідження потоків транзакцій, побудова моделі динаміки ринку ф'ючерсів в залежності від стану книги заявок та розробка стратегії оптимальної купівлі ф'ючерсних контрактів.

1. Побудова моделі короткострокової динаміки ф'ючерсного ринку

Динаміка ф'ючерсного ринку у короткостроковій перспективі є динамікою ринку з неперервним двостороннім аукціоном. Через це його доречно моделювати на основі поведінки книги заявок. Розглянемо принцип функціонування книги заявок та можливості з його дослідження.

1.1. Загальна модель ринку ф'ючерсів

Розглянемо принцип роботи ринку з неперервним двостороннім аукціоном. Кожен учасник такого ринку може виставляти у будь-який момент роботи ринку виставляти заявки двох типів – лімітні та ринкові. Лімітної заявка означає намір придбати або продати певний обсяг активу за певною ціною. Заявка додається до книги заявок (принцип формування книги заявок буде пояснено пізніше). Якщо після постановки заявки виникає ситуація, що ціна першої заявки у черзі заявок на купівлю більша, чи рівна ціні першої заявки у черзі заявок на продаж, то послідовно відбуваються угоди між новою заявкою та першою заявкою у черзі заявок з протилежним напрямом, доки ціна першої заявки у черзі заявок на купівлю не стане меншою за ціну першої заявки у черзі заявок на продаж. Під час знаходження у черзі лімітна заявка може бути задоволена повністю чи частково. При частковому задоволенні вона залишається стояти, проте її обсяг зменшується на розмір угоди. Ринкова заявка означає намір миттєво придбати або продати певний обсяг фінансового інструменту за ринковою ціною. При постановці ринкової заявки електронна система біржі послідовно проводить угоди між ринковою заявкою та першою заявкою у черзі з протилежним напрямом, доки не вичерпається обсяг ринкової заявки, чи усі лімітні заявки з протилежним напрямом не будуть вдоволені. Якщо обсяг усіх заявок з протилежним напрямом вичерпується раніше, аніж обсяг ринкової заявки, то

заявка задовольняється частково - лише на різницю між початковим та залишковим обсягом заявки, і така заявка більше у торгівельному процесі участі не бере. Учасник, що виставив лімітну заявку, може відправити транзакцію на її відміну. Якщо у момент надходження такої транзакції до торгівельної системи біржі відповідна заявка наявна у черзі (з повним чи частковим обсягом), то вона видаляється.

Біржова ціна фінансового інструменту є результуючою характеристикою торгівельної системи, поточний стан якої можна розглядати, як, так звану, книгу заявок. Книга заявок є чергою з пріоритетом. У даному випадку пріоритетом є ціна (у черзі заявок на купівлю більша ціна має більший пріоритет, а у черзі заявок на продаж менша ціна має більший пріоритет). Першою у книзі заявок є заявка, яка першою увійшла до книги серед заявок, що мають найбільший пріоритет. Будемо у кожний момент часу $t \in T$, де T -множина моментів часу у торговій сесії, розглядати множину заявок на купівлю $B(t)$ та на продаж $S(t)$.

$S(t) = \{s_1, s_2, \dots, s_{n_S(t)}\}$, $B(t) = \{b_1, b_2, \dots, b_{n_B(t)}\}$, де $n_S(t)$ —кількість лімітних заявок на продаж у момент часу t , $n_B(t)$ —кількість лімітних заявок на купівлю у момент часу t .

Лімітна заявка на продаж $s_i = \langle p_i^s, a_i^s, t_i^s \rangle$, де p_i – ціна заявки, a_i - залишковий обсяг заявки, t_i - момент часу постановки заявки у чергу.

Лімітна заявка на купівлю $b_i = \langle p_i^b, a_i^b, t_i^b \rangle$, де p_i – ціна заявки, a_i - залишковий обсяг заявки, t_i - момент часу постановки заявки у чергу.

Для $S(t)$ маємо : $i < j \leftrightarrow (p_i^s < p_j^s) \vee ((p_i^s = p_j^s) \wedge (t_i^s < t_j^s))$, для $B(t)$ маємо : $i < j \leftrightarrow (p_i^b > p_j^b) \vee ((p_i^b = p_j^b) \wedge (t_i^b < t_j^b))$.

Порядок у черзі серед заявок з однаковим пріоритетом визначається, як правило, за унікальним номером заявки. Він визначається торгівельною системою

автоматично для кожної нової заявки з тією умовою, що для довільної пари заявок менший номер має та, що надійшла до системи раніше, а двох заявок з однаковим номером не існує.

На динаміку книги заявок впливають такі події:

- Постановка лімітної заявки.
- Ринкова заявка. На деяких ринках (наприклад, на усіх ринках Московської та Української біржі) ринкові заявки не підтримуються електронною системою біржі. У даній роботі будемо вважати ринковими лімітні заявки такі, що після постановки цих заявок виникає ситуація, що ціна першої заявки у черзі заявок на купівлю більша, чи дорівнює ціні першої заявки у черзі заявок на продаж.
- Відміна лімітної заявки означає скасування та вилучення з книги заявок певної заявки.
- Зміна параметрів лімітної заявки означає відміну заявки та моментальну постановку нової з новими параметрами. Ми будемо розглядати таку транзакцію, як дві послідовні транзакції: зняття лімітної заявки та постановки нової.

Ці чотири потоки однозначно описують стан торгівельної системи у довільний момент часу.

Кращу ціну купівлі визначатимемо, як

$$b(t) = \begin{cases} \sup\{p_i^b : \langle p_i^b, a_i^b, t_i^b \rangle \in B(t)\}, & \text{якщо } n_B(t) > 0 \\ 0, & \text{якщо } n_B(t) = 0 \end{cases}.$$

Кращу ціну продажу визначатимемо, як

$$s(t) = \begin{cases} \inf\{p_i^s : \langle p_i^s, a_i^s, t_i^s \rangle \in S(t)\}, & \text{якщо } n_S(t) > 0 \\ p_{max}, & \text{якщо } n_B(t) = 0 \end{cases},$$

де p_{max} - деяке заздалегідь визначене велике число, таке, що ринкова ціна не може його перевищувати. Наприклад, при програмуванні моделі мовою C# та використанні типу `decimal` для опису цін, можна використовувати максимальне значення для цього типу – 79228162514264337593543950335.

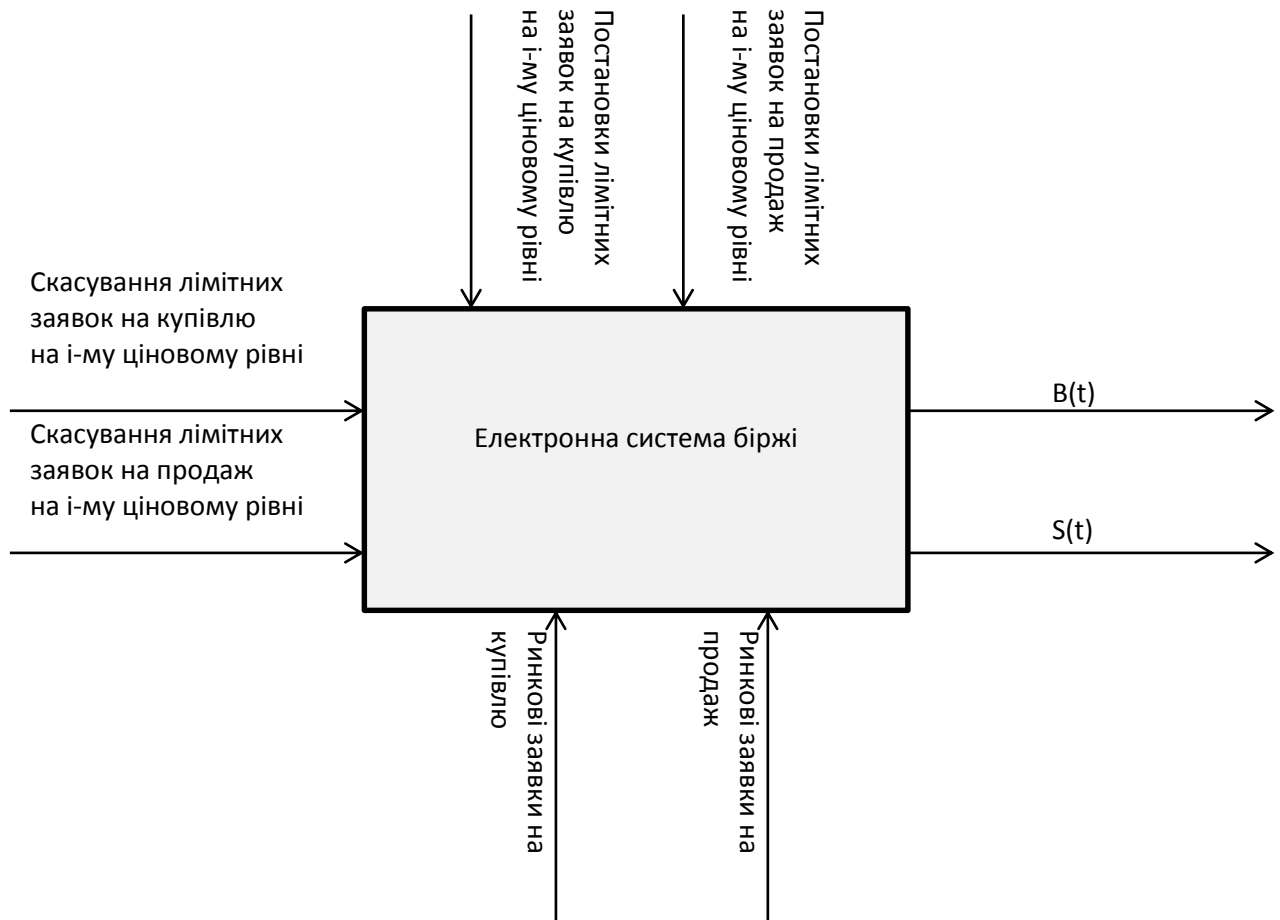
Процес ціни можна описати, як

$$P(t) = \frac{b(t) + s(t)}{2}.$$

Ціни біржових інструментів можуть бути як цілочисельні, так і дробові. Не обмежуючи загальності, ми розглядаємо у нашій моделі лише фінансові інструменти з цілочисельними цінами та мінімальним кроком ціни, рівним 1. Якщо інструмент має мінімальний крок ціни не рівний одиниці, то розділивши ціни всіх заявок, угод, та показники, що мають розмірність, аналогічну ціні, ми отримуємо інструмент з цілочисельними цінами та мінімальним кроком ціни, рівним 1.

i -м ціновим рівнем для заявок на купівлю будемо називати сукупність лімітних заявок на купівлю з ціною p , для яких $b(t) - p + 1 = i$, а для заявок на продаж i -м ціновим рівнем називатимемо сукупність лімітних заявок на продаж з ціною p , для яких $p - s(t) + 1 = i$.

Отже, ми будемо розглядати ринок, як наступну систему:



Таке представлення є дещо спрощеним, бо біржова система подає на вихід не лише інформацію про стан книги заявок, а й список угод, інформацію щодо кількості відкритих позицій, довідкову інформацію щодо торгового інструменту та торгової сесії та конфіденційну інформацію щодо стану заявок кожного учасника. Однак у нашій моделі ми використовуємо лише інформацію про стан книги заявок та потоки, які подаються на вхід електронної системи біржової системи.

Кожен з вхідних потоків можна описати як деякий випадковий процес. В першу чергу будемо аналізувати характеристики випадкових величин, які дорівнюють проміжкам часу між двома подіями одного типу.

1.2. Вибірки даних для проведення дослідження

Дані для досліджень будемо брати з повного журнал заявок строкового ринку Московської біржі за 20.01.2014. Ми розглядаємо заявки та угоди з найліквіднішого ф'ючерсного контракту Московської біржі - ф'ючерсного контракту на індекс РТС.

1.1. Програмне забезпечення для проведення дослідження

В межах цієї роботи ми будемо використовувати декілька модулів програмного комплексу Watcher та пакет для статистичного аналізу R.

Програмний комплекс Watcher призначений для накопичення та систематизації біржових історичних даних, аналізу, розробки, тестування, оптимізації та застосування торгових стратегій для строкового ринку. Розроблений автором даної роботи в рамках дослідження ціноутворення ринку ф'ючерсних контрактів. Використані мови програмування C# та C++. Підтримує біржові протоколи прямого доступу до серверів Української та Московської бірж, та пройшов сертифікацію на обох біржах.

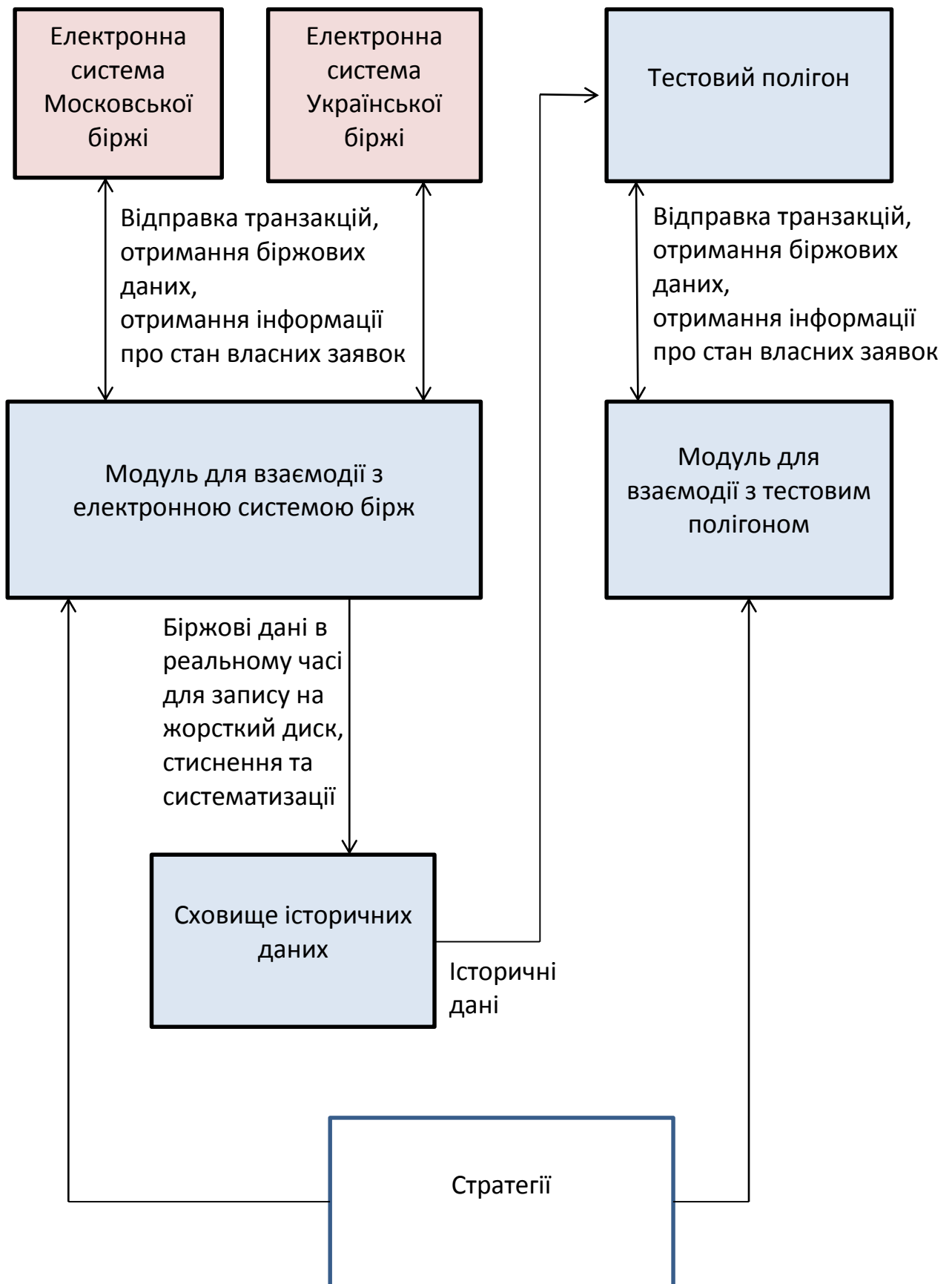
Програмний комплекс складається з чотирьох основних модулів:

- модуль для взаємодії з електронною системою бірж (застосування стратегій та накопичення даних);
- сховище історичних даних;
- тестовий полігон для аналізу історичних даних, розробки, тестування, та оптимізації стратегій;
- модуль для взаємодії з тестовим полігоном.

Модулі, що входять до програмного комплексу та їх програмні коди не призначені для розповсюдження, за винятком програмних кодів модулів, створених в межах даного дослідження.

У даній роботі ми будемо створювати декілька стратегій та застосовувати їх до історичних даних для отримання вибірок у зручному вигляді та для аналізу алгоритмів, запропонованих у цій роботі.

Схематично програмний комплекс можна зобразити таким чином:



2. Статистичний аналіз потоків транзакцій

Проаналізуємо статистичні характеристики різних потоків транзакцій, та проаналізуємо розподіли емпіричних проміжків часу між двома послідовними подіями одного типу на близькість між відповідними розподілами, отриманими з різних потоків та відповідність цих величин певним розподілам.

2.1. Аналіз залежності інтенсивності потоку транзакцій від цінового рівня потоку транзакцій

Розглянемо розподіл кількості надходжень нових заявок в залежності від цінового рівня.

Аналогічно для H_1

$$\chi^2 = 2889.507 > \chi_{0.95}^2(149) = 179.5806$$

і для H_2

$$\chi^2 = 3449.871 > \chi_{0.95}^2(149) = 179.5806$$

Всі три гіпотези відхиляються, що означає, що серед цих трьох вибірок даних немає хоча б однієї пари однаково розподілених.

Неважко показати, що і для заявок на продаж серед вибірок X^s , Y^s та Z^s не знайдеться жодної пари однаково розподілених.

2.3. Визначення типу розподілу, що описує час між однотипними подіями

Отже, ми маємо визначити розподіли для вибірок X^b , Y^b , Z^b , X^s , Y^s та Z^s .

В якості допустимих розподілів будемо використовувати експоненційний, гамма-розподіл та розподіл Вейбулла.

Експоненційний розподіл має такі характеристики:

$$X \sim \text{Exp}(\lambda), \lambda > 0, \text{ якщо } f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

$$E[X] = \lambda^{-1}, \text{Var}(X) = \lambda^{-2}$$

Для гамма-розподілу

$$X \sim \text{Ga}(k, \theta), k > 0, \theta > 0, \text{ якщо } f_X(x) = \begin{cases} x^{k-1} \frac{e^{-x/\theta}}{\theta^k \Gamma(k)}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

де $\Gamma(k) = \int_0^{\infty} x^{k-1} e^{-x} dx$ – Гамма-функція Ейлера.

$$E[X] = k\theta, \text{Var}(X) = k\theta^2$$

Для розподілу Вейбулла

$$X \sim W(k, \lambda), k > 0, \lambda > 0, \text{якщо } f_X(x) = \left(\frac{k}{\lambda}\right) \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k}$$

Використаємо метод максимальної вірогідності

Суть метода полягає у наступному. Нехай є деяка вибірка x_1, x_2, \dots, x_n з розподілу P_{θ} , де $\theta \in \Theta$ -невідомі параметри. Нехай $L(x|\theta): \Theta \rightarrow \mathbb{R}$ -функція вірогідності, де $x \in \mathbb{R}^n$. Точкова оцінка $\hat{\theta}_{\text{МП}} = \hat{\theta}_{\text{МП}}(x_1, x_2, \dots, x_n) = \arg \max_{\theta \in \Theta} L(x_1, x_2, \dots, x_n|\theta)$ називається оцінкою максимальної вірогідності параметра θ .

Для експоненційного розподілу функція вірогідності для параметра λ має такий вигляд:

$$L(\lambda) = \prod_{i=1}^n \lambda \exp(-\lambda x_i) = \lambda^n \exp\left(-\lambda \sum_{i=1}^n x_i\right) = \lambda^n \exp(-\lambda n\bar{x}),$$

де $x = (x_1, \dots, x_n)$ – реалізація вибірки незалежних однаково розподілених випадкових величин, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ – вибіркове середнє.

Візьмемо похідну від логарифму функції вірогідності:

$$\frac{d}{d\lambda} \ln(L(\lambda)) = \frac{d}{d\lambda} (n \ln(\lambda) - \lambda n\bar{x}) = \frac{n}{\lambda} - n\bar{x} \begin{cases} > 0, & 0 < \lambda < \frac{1}{\bar{x}}, \\ = 0, & \lambda = \frac{1}{\bar{x}}, \\ < 0, & \lambda > \frac{1}{\bar{x}}. \end{cases}$$

Таким чином, оцінкою максимальної вірогідності для параметра λ експоненційного розподілу є

$$\bar{\lambda} = \frac{1}{\bar{x}}.$$

Для Гамма-розподілу функція вірогідності для N незалежних однаково розподілених реалізацій (x_1, x_2, \dots, x_N) виглядає наступним чином:

$$L(k, \theta) = (k - 1) \prod_{i=1}^N f(x_i; k, \theta),$$

де $f(x_i; k, \theta)$ - значення щільності гамма-розподілу з параметрами k та θ .

Ми можемо розрахувати логарифмічну функцію правдоподібності:

$$\ell(k, \theta) = (k - 1) \sum_{i=1}^N \ln(x_i) - \sum_{i=1}^N \frac{x_i}{\theta} - Nk \ln(\theta) - N \ln(\Gamma(k))$$

Узявши похідну по θ та прирівнявши її нулеві, отримуємо оцінку максимальної вірогідності для параметра θ :

$$\hat{\theta} = \frac{1}{kN} \sum_{i=1}^N x_i.$$

Підставляємо цей результат до логарифмічної функції вірогідності і отримаємо

$$\ell(k, \hat{\theta}) = (k - 1) \sum_{i=1}^N \ln(x_i) - Nk - Nk \ln\left(\frac{\sum_{i=1}^N x_i}{kN}\right) - N \ln(\Gamma(k)).$$

Візьмемо похідну по k та прирівняємо її до нуля. Отримаємо:

$$\ln k - \psi(k) = \ln\left(\frac{1}{N} \sum_{i=1}^N x_i\right) - \frac{1}{N} \sum_{i=1}^N \ln x_i,$$

де $\psi(k) = \frac{d}{dx} \ln \Gamma(x)$ – дигамма-функція. На жаль, немає аналітичного розв’язку для k . Чисельне значення k може бути знайдене методом Ньютона. Початкове значення k можна знайти з такого наближення:

$$\ln k - \psi(k) \approx \frac{1}{2k} \left(1 + \frac{1}{6k + 1} \right).$$

Якщо покласти $s = \ln \left(\frac{1}{N} \sum_{i=1}^N x_i \right) - \frac{1}{N} \sum_{i=1}^N \ln x_i$, тоді, згідно з [10],

$$k \approx \frac{3 - s + \sqrt{(s - 3)^2 + 24s}}{12s},$$

з похибкою, що не перевищує 1.5% від істинного значення.

Для параметрів розподілу Вейбулла оцінки методу максимальної вірогідності виглядають наступним чином: оцінка для параметра λ в степені k :

$$\hat{\lambda}^k = \frac{1}{N} \sum_{i=1}^N (x_i^k - x_N^k),$$

де $x_1 > x_2 > \dots > x_N$ N найбільші N спостережень.

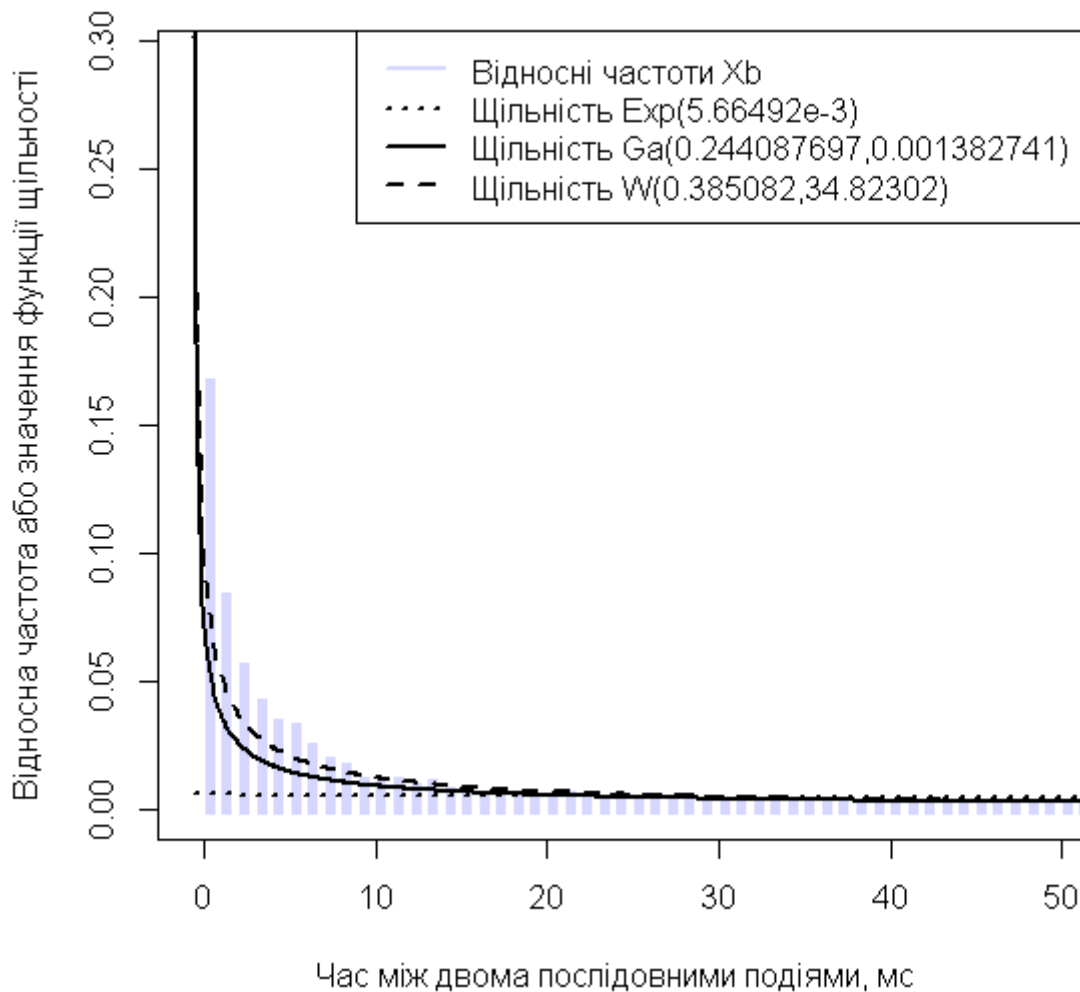
Оцінка максимальної вірогідності для параметра k :

$$\hat{k}^{-1} = \frac{\sum_{i=1}^N (x_i^k \ln x_i - x_N^k \ln x_N)}{\sum_{i=1}^N (x_i^k - x_N^k)} - \frac{1}{N} \sum_{i=1}^N \ln x_i$$

X	\bar{X}	S_N^2	Оцінка параметрів експ. розп.	Оцінка параметрів Гамма- розп.	Оцінка параметрів розп. Вейбулла
X^b	176.416	371795.6	$Exp(5.66492 \cdot 10^{-3})$	$Ga(0.244087697, 0.001382741)$	$W(0.385082, 34.82302)$
X^s	165.0767	329087.1	$Exp(6.05397 \cdot 10^{-3})$	$Ga(2.50064 \cdot 10^{-1}, 1.513882 \cdot 10^{-3})$	$W(0.392235, 34.83134)$
Y^b	342.6614	970257.3	$Exp(2.9172 \cdot 10^{-3})$	$Ga(0.227472, 0.000663581)$	$W(0.352236, 64.41688)$
Y^s	347.1032	1020286	$Exp(2.87984 \cdot 10^{-3})$	$Ga(0.231753, 0.000667415)$	$W(0.357146, 68.69595)$
Z^b	994.1599	8376713	$Exp(1.00585 \cdot 10^{-3})$	$Ga(0.340643, 0.000342638)$	$W(0.478171, 436.8373)$
Z^s	1041.363	8318343	$Exp(9.60256 \cdot 10^{-4})$	$Ga(0.329491, 0.000316396)$	$W(0.4645912, 450.8756)$

Порівняємо гістограми відносних частот з функціями щільностей відповідних розподілів

Розглянемо вибірку X^b та щільності відповідних розподілів



Неважко побачити, що, на жаль, жодний із запропонованих розподілів не описує адекватно випадкову величину, що спостерігається.

Проте, щоб переконатися в цьому припущенні, застосуємо критерій згоди Пірсона. Розглянемо гіпотезу H_0 яка полягає в тому, що вибірка $X = \{x_1, x_2, \dots, x_n\}$ є реалізаціями незалежних однаково розподілених випадкових величин з функцією розподілу $F(x)$. Тоді інтервал $[a, b]$, у який можуть потрапити випадкові величини з функцією розподілу $F(x)$, розділимо на k інтервалів $(a_i, b_i], i = 1 \dots k$, що не перетинаються. Нехай n_j – кількість спостережень в j -му інтервалі. $p_j = F(b_j) - F(a_j)$ – ймовірність потрапляння спостережень в j -й інтервал за умови виконання гіпотези H_0 . $E_j = np_j$ – очікувана кількість потраплянь в j -й інтервал. Розраховується статистика

$\chi^2 = \sum_{j=1}^k \frac{(n_j - E_j)^2}{E_j}$, яка асимптотично має розподіл χ^2 з $k-1$ степенями свободи.

Гіпотеза H_0 приймається, якщо $\chi^2 < \chi_{1-\alpha}^2(k-1)$, $\chi_{1-\alpha}^2(k-1)$ – квантиль розподілу χ^2 з $k-1$ степенями свободи рівня $1-\alpha$, де α – рівень значущості.

Для рівня значущості α , щорівнює 0.05, для експоненційного розподілу маємо:

$$\chi^2 = 98979.52 > \chi_{0.95}^2(149) = 179.5806.$$

Для гамма-розподілу:

$$\chi^2 = 8635.533 > \chi_{0.95}^2(149) = 179.5806.$$

Для розподілу Вейбулла:

$$\chi^2 = 4568.96 > \chi_{0.95}^2(149) = 179.5806.$$

Це вказує на те, що вибірка X^b не є реалізацією випадкових величин з відповідними розподілами.

Аналогічні тести проведемо над іншими п'ятьма вибірками. Значення χ^2 критеріїв подані у наступній таблиці:

Вибірка	Наближення експоненційним розподілом	Наближення Гамма-розподілом	Наближення розподілом Вейбулла
X^b	98979.52	8635.533	4568.96
X^s	86890.05	7974.429	4199.48
Y^b	213478.9	6997.937	4280.451
Y^s	201956	7072.643	4360.132
Z^b	39758.62	1666.954	998.5645
Z^s	57240.75	1700.663	1254.702

З цих тестів можна зробити 2 висновки:

- Жоден із запропонованих розподілів не описав величини, що досліджуються.
- Для кожної з випадкових величин найбільш близькою апроксимацією є випадкова величина з розподілом Вейбулла.

3.2. Марковська модель

Розглянемо стан черг заявок на купівлю та продаж за кращими цінами купівлі та продажу відповідно, у момент, коли різниця цін кращих заявок на купівлю та продаж рівна одному крокові ціни. Згідно з [11] основний потік транзакцій транзакції припадає на ціни кращої купівлі та продажу.

Розглянемо наступний ланцюг Маркова:

Множина станів $I \in \{X_{ij}\}_{i \in \mathbb{Z}_0, j \in \mathbb{Z}_0}$ та множина поглинаючих станів $\hat{I} \in \{X_{i0}\}_{i \in \mathbb{Z}_0} \cup \{X_{0j}\}_{j \in \mathbb{Z}_+}$. Те, що система знаходиться у стані X_{ij} означає, що за кращою ціною купівлі знаходиться i заявок на купівлю, а за кращою ціною продажу знаходиться j заявок на продаж.

Ймовірності переходів з непоглинаючих наступні:

$$p_{X_{ij}X_{i+1j}} = P\{\text{наступна транзакція – постановка заявки на купівлю}\} = P_{ab},$$

$$p_{X_{ij}X_{ij+1}} = P\{\text{наступна транзакція – постановка заявки на продаж}\} = P_{as},$$

$$\begin{aligned} p_{X_{ij}X_{i-1j}} &= P\{\text{наступна транзакція} \\ &\quad \text{– зняття заявки на купівлю, чи ринкова заявка на продаж}\} \\ &= P_{db} + P_{ms}, \end{aligned}$$

$$\begin{aligned} p_{X_{ij}X_{ij-1}} &= P\{\text{наступна транзакція} \\ &\quad \text{– зняття заявки на продаж, чи ринкова заявка на купівлю}\} \\ &= P_{ds} + P_{mb}, \end{aligned}$$

Припустимо, що потоки заявок збалансовані ($P_{ab} = P_{as} = P_{db} + P_{ms} = P_{ds} + P_{mb}$).

Зсувом ціни назвемо момент, коли кількість заявок на купівлю за кращою ціною купівлі чи кількість заявок на продаж за кращою ціною продажу знижується до нуля. Що призводить до зміни $p(t)$.

Теорема За умов Марківської моделі та наявності n заявок на купівлю за ціною $b(t)$ та p заявок на купівлю за ціною $s(t)$, та за умови, що $s(t) - b(t) = 1$, ймовірність, що наступним зсувом ціни буде зростання описується наступною формулою:

$$p^{up}(n, p) = \frac{1}{\pi} \int_0^\pi \left(2 - \cos(t) - \sqrt{((2 - \cos(t))^2 - 1)}\right)^p \frac{\sin(nt) \cos(\frac{t}{2})}{\sin(\frac{t}{2})} dt$$

Доведення

Розглянемо симетричне випадковим блуканням у позитивному ортанті \mathbb{Z}_+^2 із зупинкою на границі. Ймовірність того, що наступним зсувом ціни буде зростання за умови, що є n заявок на купівлю за ціною $b(t)$ та p заявок на купівлю за ціною $s(t)$ рівна ймовірності того, що випадкове блукання, що починається з (n, p) досягне вісі x раніше, ніж вісі y . Таким чином при всіх $n \geq 1$ та $p \geq 1$:

$$4p^{up}(n, p) = p^{up}(n + 1, p) + p^{up}(n - 1, p) + p^{up}(n, p + 1) + p^{up}(n, p - 1),$$

з граничними умовами: $p^{up}(0, p) = 0$ для всіх $p \geq 1$ та $p^{up}(n, 0) = 1$ для всіх $n \geq 1$. Ця задача також відома, як дискретна задача Діріхле [12]. Розв'язок, поданий у [13], який називається дискретною гармонічною функцією, показує, що для усіх $t \geq 0$ функції

$$f_t(x, y) = e^{xr(t)} \sin(yt),$$

$$\bar{f}_t(x, y) = e^{-xr(t)} \sin(yt),$$

$$\text{де } r(t) = \cosh^{-1}(2 - \cos t)$$

є розв'язками. Також у [13] показано, що ймовірність, що просте випадкове блукання $(M_k, k \geq 1)$, що починається у $(n, p) \in \mathbb{Z}^+ \times \mathbb{Z}^+$ досягне вісі в $(x, 0)$, дорівнює

$$\frac{2}{\pi} \int_0^\pi e^{-r(t)p} \sin(nt) \sin(tx) dt,$$

$$p^{up}(n, p) = \sum_{k=1}^{\infty} \frac{2}{\pi} \int_0^\pi e^{-r(t)p} \sin(kt) \sin(tx) dt,$$

Звідки

$$\sum_{k=1}^m \sin(kt) = \frac{\sin\left(\frac{mt}{2}\right) \sin\left(\frac{(m+1)t}{2}\right)}{\sin\left(\frac{t}{2}\right)} = \frac{\cos\left(\frac{t}{2}\right) - \cos\left(\left(m + \frac{1}{2}\right)t\right)}{2\sin\left(\frac{t}{2}\right)},$$

використовуючи інтегрування частинами, маємо наступний інтеграл:

$$\int_0^\pi \overbrace{\frac{e^{-r(t)p} \sin(nt)}{\sin\left(\frac{t}{2}\right)}}^{g(t)} \cos\left(\left(m + \frac{1}{2}\right)t\right) dt$$

$$= -\frac{1}{m + \frac{1}{2}} \int_0^\pi g'(t) \sin\left(\left(m + \frac{1}{2}\right)t\right) dt \xrightarrow{m \rightarrow \infty} 0.$$

Так, як g' обмежена. Отже:

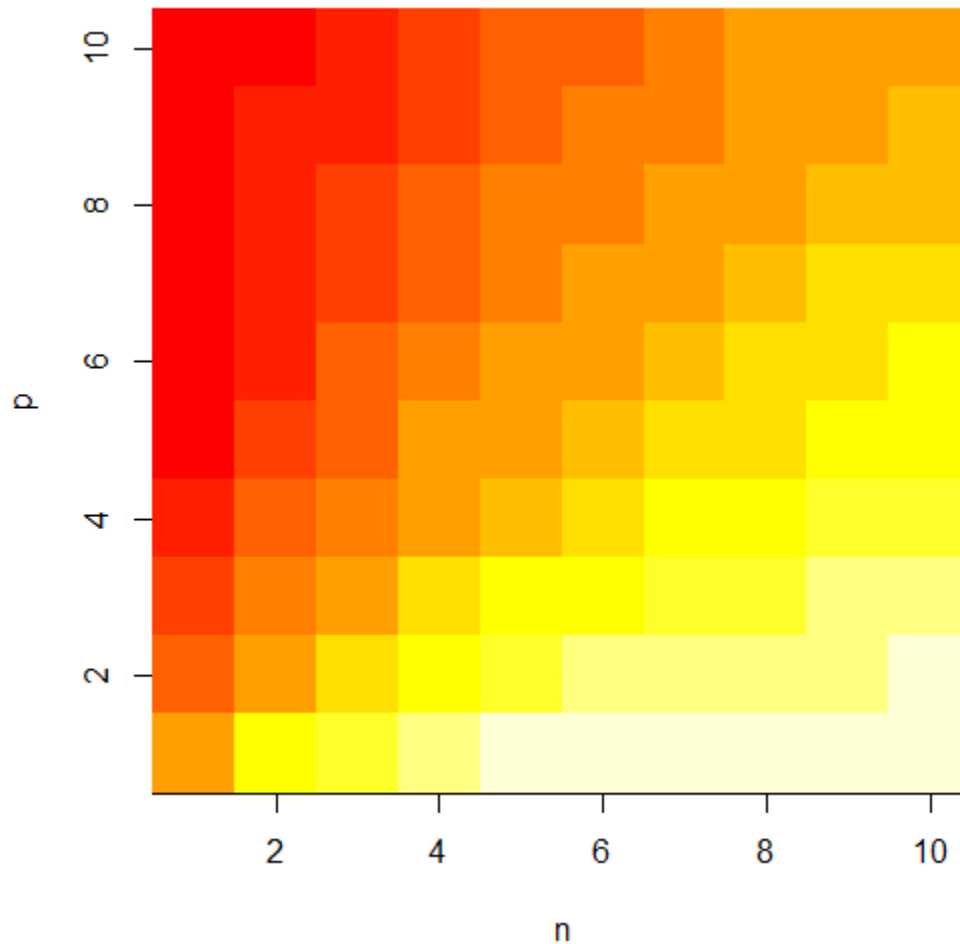
$$p^{up}(n, p) = \frac{1}{\pi} \int_0^\pi e^{-r(t)p} \frac{\sin(nt) \cos\left(\frac{t}{2}\right)}{\sin\left(\frac{t}{2}\right)} dt,$$

а підставивши

$$e^{-r(t)p} = (2 - \cos(t) - \sqrt{(2 - \cos(t))^2 - 1})$$

маємо:

$$p^{up}(n, p) = \frac{1}{\pi} \int_0^\pi \left(2 - \cos(t) - \sqrt{((2 - \cos(t))^2 - 1)} \right)^p \frac{\sin(nt) \cos(\frac{t}{2})}{\sin(\frac{t}{2})} dt.$$



Ймовірності, що наступний зсувом ціни буде зростання, в залежності від n та p (1- білий, 0-червоний) обраховані за отриманою формулою

4. Побудова оптимальної стратегії купівлі ф'ючерсних контрактів

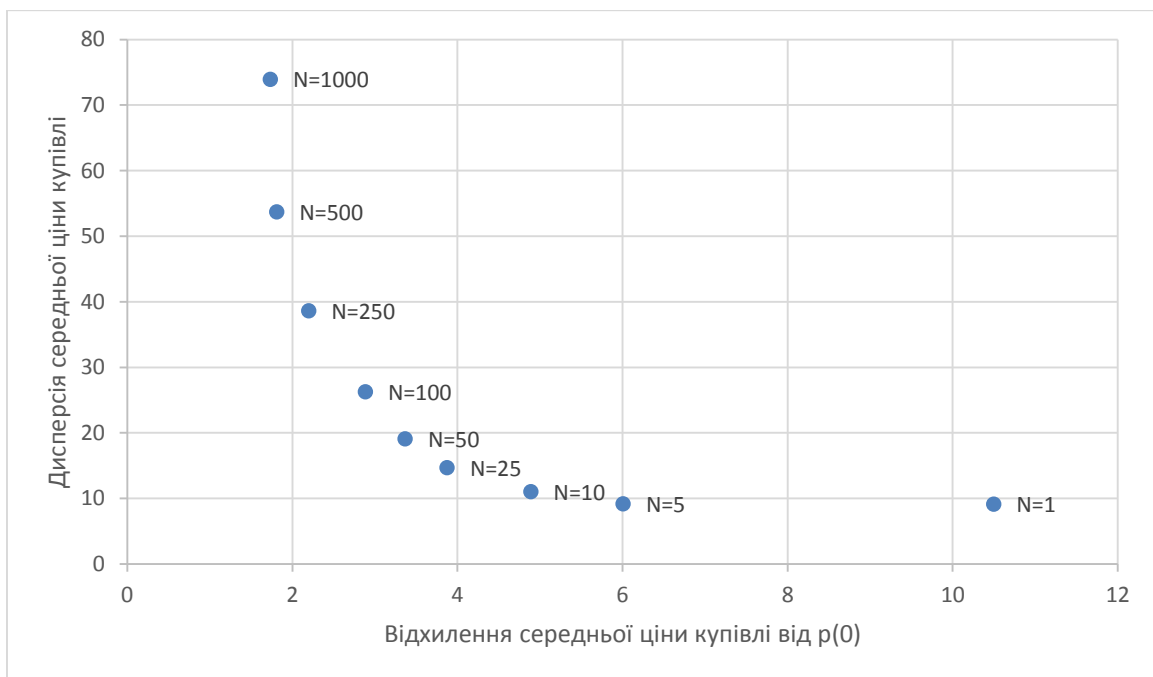
Розглянемо наступну задачу. В початковий момент часу на ринку, що описується марковською моделлю з попереднього розділу, відсутні заявки. Вважаємо, що початковою ціною на ринку є деяка стартова ціна $p(0)$. Якщо першою заявкою, що надійде на постановку, буде заявка на купівлю, то вона надійде за ціною на один крок ціни нижчою, аніж $p(0)$. Якщо це буде заявка на продаж, то, відповідно, вона надійде за ціною на один крок ціни вищою, аніж $p(0)$. Наступні заявки на купівлю надходять за ціною на один крок ціни меншою, ніж ціна кращої заявки на продаж, а наступні заявки на продаж надходять за ціною на один крок ціни вищою, ніж ціна кращої заявки на купівлю.

Нехай за деякої стратегії S , середня ціна, за якою буде придбано K ф'ючерсних контрактів, описується неперервною випадковою величиною $\xi(S, K)$.

Спробуємо, використовуючи аналітичну формулу, отриману в попередньому розділі, розробити стратегію, що буде мінімізувати матсподівання та дисперсію $\xi(S, K)$, а, також, порівняємо її результати з загальновідомою стратегією розбиття заявки на дрібніші та виконання їх з деякою періодичністю. Досліджувати результати стратегій будемо за допомогою імітаційного моделювання.

4.1 Стратегія розбиття заявки на заявки мінімального обсягу

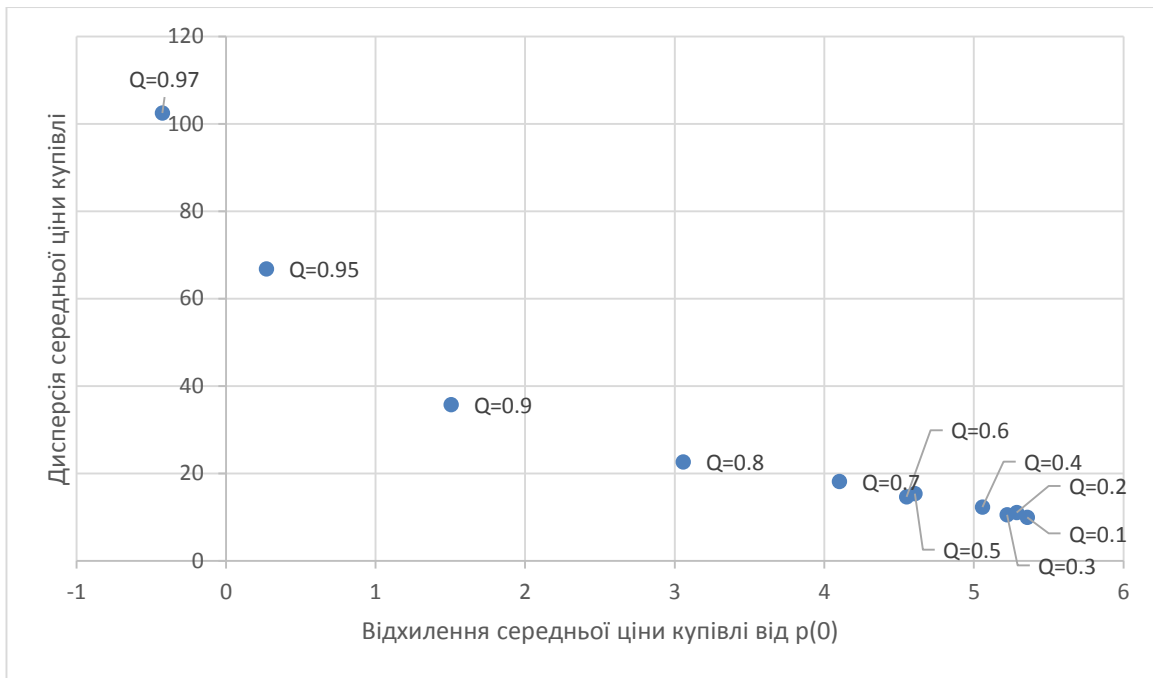
Розглянемо стратегію, яка полягає у розбитті заявки на послідовність заявок мінімального обсягу та виконання їх з деякою періодичністю. Проведемо імітаційне моделювання для купівлі 100 контрактів з періодичністю в 1 купівлю мінімально можливого обсягу ф'ючерсних контрактів на кожні N транзакцій.



Ефективність стратегії розбитті заявки на послідовність заявок мінімального обсягу в залежності від частоти купівлі

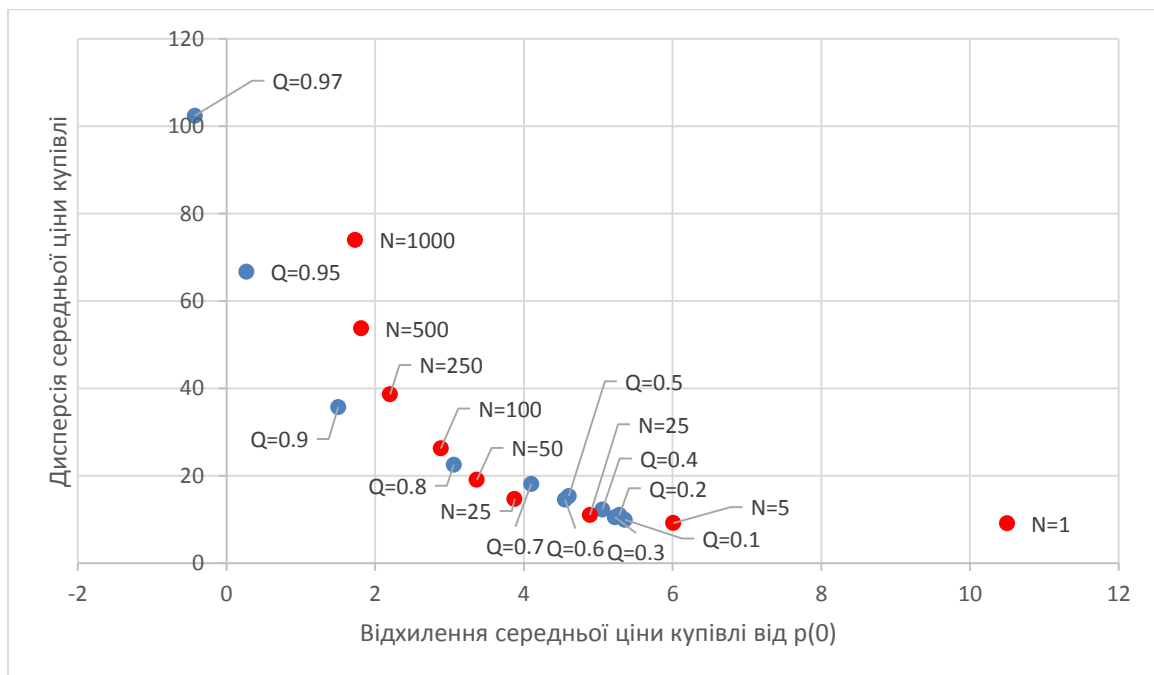
4.2 Стратегія на основі марковської моделі

Розглянемо стратегію на основі аналітичної формули, отриманої у розділі 3.2. Будемо проводити купівлі мінімального обсягу, в кожен момент часу, коли різниця між ціною кращих цін купівлі та продажу рівна одному крокові ціни та $p^{up}(n, p) > Q$, де Q – деяке заздалегідь визначене число від 0 до 1.



Ефективність стратегії на основі розділу 3.2 в залежності від мінімальної імовірності зростання ціни для здійснення купівлі

4.3 Порівняння стратегій



Порівняння ефективності стратегій. Червоні точки – стратегія з розділу 4.1, сині – з розділу 4.2

З графіку можна зробити висновок, що для $Q > 0.8$ результати стратегії на основі марковської моделі стають оптимальними за Парето. Конкретне значення параметра Q можна підібрати в залежності від кількості контрактів, яку необхідно придбати та цільової функції.

Висновки

В рамках даного дослідження показано суперечливість підходів до моделювання динаміки ринків з неперервним двостороннім аукціоном на основі припущення про експоненційний та гамма-розподіл випадкових величин, які дорівнюють проміжкам часу між подіями одного типу.

Розроблено марковську модель для моделювання динаміки книги заявок у дискретному часі та отримано аналітичну формулу для оцінки ймовірностей зростання та спадання ціни на основі поточного стану книги заявок. Дана формула стала основою стратегії купівлі ф'ючерсних контрактів. Дослідження властивостей цієї стратегії за допомогою імітаційного моделювання показало її високу ефективність.

Також результатом даного дослідження є програмне забезпечення для дослідження і моделювання динаміки фінансових ринків з неперервним двостороннім аукціоном.

Незважаючи на отримані результати, тема, що досліджувалася, має широкі перспективи подальших досліджень. Основними задачами є подальше вивчення властивостей динаміки книги заявок (знаходження формул чи алгоритмів для розрахунку імовірності виконання лімітних заявок, врахування динаміки фінансових інструментів з високою кореляцією з інструментом, динаміка якого моделюється за допомогою марковської моделі), розробка стратегій для інших задач (маркет-мейкінгу, високочастотної торгівлі), використання інших методів моделювання динаміки ринку (деякі здобутки у цьому напрямку наявні у [14]).

Література

1. Dunis C. L. et al. Statistical arbitrage and high-frequency data with an application to Eurostoxx 50 equities //Liverpool Business School, Working paper. – 2010.
2. Какойченко А.І. Розробка алгоритму перевірки стратегій статистичного арбітражу//XXII International Conference Problems of decision making under uncertainties (PDMU-2013): Матеріали наук. праць міжнародної конференції. – Ялта, 2013.- С. 82
3. Какойченко А.І. Моделювання ціноутворення фінансових інструментів з неперервним двостороннім аукціоном//VII Міжнародна наукова конференція імені І. І. Ляшка "Обчислювальна та прикладна математика" (СМ-2014): Матеріали наук. праць міжнародної конференції. – Київ, 2014.- С. 52.
4. Parlour C. A. Price dynamics in limit order markets //Review of Financial Studies. – 1998. – Т. 11. – №. 4. – С. 789-816.
5. Gouriéroux C., Jasiak J., Le Fol G. Intra-day market activity //Journal of financial markets. – 1999. – Т. 2. – №. 3. – С. 193-226.
6. Avellaneda M., Stoikov S. High-frequency trading in a limit order book //Quantitative Finance. – 2008. – Т. 8. – №. 3. – С. 217-224.
7. Korolev V. Y. et al. Probability and statistical modeling of information flows in complex financial systems based on high-frequency data // Informatics and its Applications. – 2013. – Т. 7. – №. 1. – С. 12-21.
8. Harris L. E., Panchapagesan V. The information content of the limit order book: evidence from NYSE specialist trading decisions //Journal of Financial Markets. – 2005. – Т. 8. – №. 1. – С. 25-67.
9. Cont R., Stoikov S., Talreja R. A stochastic model for order book dynamics //Operations research. – 2010. – Т. 58. – №. 3. – С. 549-563.
10. Minka, Thomas P. (2002) "Estimating a Gamma distribution". <http://research.microsoft.com/en-us/um/people/minka/papers/minka-gamma.pdf>
11. Biais B., Hillion P., Spatt C. An empirical analysis of the limit order book and the order flow in the Paris Bourse //the Journal of Finance. – 1995. – Т. 50. – №. 5. – С. 1655-1689.
12. Cont R., De Larrard A. Price dynamics in a Markovian limit order market //SIAM Journal on Financial Mathematics. – 2013. – Т. 4. – №. 1. – С. 1-25.

Додатки

Додаток А. Програма для статистичного дослідження потоків транзакцій (мовою R)

Допоміжні функції

```
HistExt<-function(x,minX,maxX,cnt)
{
  r<-matrix(0,cnt,1);
  for(i in 1:length(x))
  {
    ind=floor((x[i]-minX)/(maxX-minX)*cnt)+1;
    #print(ind);
    if(ind >= cnt)
    {
      ind<-cnt;
    }
    if(ind<=0)
    {
      ind<-1;
    }
    r[ind]<-r[ind]+1;
  }
  return(r);
}
```

```
GetTimes<-function(x,beg,N)
```

```
{
```