

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ М. В. ЛОМОНОСОВА

ФАКУЛЬТЕТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ
И КИБЕРНЕТИКИ

На правах рукописи

Черток Андрей Викторович

**Моделирование потоков заявок
на финансовых рынках
с помощью обобщенных процессов риска**

Специальность 01.01.05 – теория вероятностей и математическая статистика

Диссертация на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
д.ф-м.н. Королев Виктор Юрьевич

Москва – 2015

Оглавление

Введение	4
1 Дисбаланс потоков заявок	22
1.1 Описание модели	23
1.1.1 Книга заявок	23
1.1.2 Динамика книги заявок	25
1.2 Процесс дисбаланса количества заявок	27
1.2.1 Постоянные интенсивности потоков заявок	28
1.2.2 Случайные интенсивности потоков заявок	32
1.2.3 Мультипликативная форма случайных интенсивностей	34
1.3 Процесс дисбаланса потоков заявок	35
1.4 Случайные интенсивности. Обобщённые процессы Кокса и процесс OFI.	39
1.5 Асимптотические аппроксимации распределения процесса дисбаланса потоков заявок	46
1.6 Анализ реальных данных	50
1.6.1 Описание данных	50
1.6.2 Оценка параметров	51
1.6.3 Свойства потоков заявок	54

2 Функциональные предельные теоремы для процесса дисбаланса потоков заявок	59
2.1 Пространство Скорохода. Процессы Леви.	61
2.2 Сходимость процесса OFI к процессам Леви	63
2.3 Обобщённые гиперболические процессы Леви как асимптотическая аппроксимация процесса дисбаланса потоков заявок	73
2.4 Финальная настройка модели	79
3 Токсичность потока заявок	82
3.1 Введение	82
3.2 Модель потоков заявок	83
3.3 Профиль мгновенной токсичности потока заявок	84
3.4 Токсичность потока заявок	86
3.4.1 Байесовский подход	87
3.4.2 Квантильный подход	88
3.5 Модели потоков заявок	89
3.5.1 Модель рынка с заявками единичного объёма	89
3.5.2 Модель рынка с экспоненциальными объёмами заявок	93
3.6 Реальные данные	96
3.6.1 Описание данных	96
3.6.2 Процедура оценки параметров	96
3.6.3 Показатели токсичности	97
Заключение	100
Список литературы	103

Введение

Актуальность темы исследования

Распространение высокочастотных данных о транзакциях, котировках и потоках заявок на электронных рынках привело к кардинальному изменению способов обработки данных и техник статистического моделирования процессов, происходящих на финансовых рынках. Динамика рынка на уровне транзакций не может быть описана только в терминах динамики только цены в ее классическом понимании, также нужно принимать во внимание взаимодействие между заявками разных типов, приходящими от покупателей и продавцов. Единственная возможность воспроизвести такое взаимодействие - это моделирование потоков заявок, приходящих на различные уровни так называемой книги заявок.

В последние годы автоматическая (электронная) торговля в значительной степени заменила так называемую торговлю «в яме» на рынке акций. Electronic Communication Network (ECN), такие как Achipelago, Instinet, Brut и Tradebook - электронные системы осуществления сделок купли-продажи биржевых товаров захватили большую долю рынка. В отличие от рынков, где маркет-мейкер или специалист централизованно собирает все заявки на покупку и продажу и представляет ликвидность, устанавливая собственные котировки на покупку и продажу, эти электронные платформы агрегируют все активные заявки на покупку и продажу в книгу заявок и эти заявки становятся доступными всем участникам рынка. В такой системе рыночные заявки мгновенно исполняются по лучшим

Таблица 1: Среднее число заявок за 10 сек. и количество изменений цены за день на самых ликвидных бумагах биржи NYSE

Акция	Ср. число заявок за 10 сек.	Кол-во изм. цены за день
Citigroup	4469	12499
General Electric	2356	7862
General Motors	1275	9016

ценам, представленным в книге заявок в автоматическом режиме. По мере развития ECN-площадок развитые фондовые рынки, такие как NYSE, Nasdaq, Tokyo Stock Exchange, Toronto Stock Exchange, Vancouver Stock Exchange, Euronext (Paris, Amsterdam, Brussels), London Stock Exchange и Moscow Exchange полностью или частично перешли на электронную систему торгов.

В то же самое время частота прихода заявок очень сильно возросла и время исполнения рыночных заявок на таких электронных рынках снизилось с 25 миллисекунд в 2000 г. до 300 микросекунд в 2014. Как показано в Таблице 1, за 10-секундные интервалы времени по каждому из инструментов приходят тысячи заявок, в результате чего лучшие котировки этих инструментов обновляются сотни тысяч раз в день, также как и цены сделок.

В результате такого развития информация об эволюции спроса и предложения, а также цены на рынках акций стала записываться в огромных количествах: эти данные доступны участникам рынка в режиме реального времени, а исследователям - в виде высокочастотных баз данных. Анализ такого рода данных представляется огромным вызовом не только из-за их большого объема и сложной структуры. Эти данные дают нам возможность получить более детальное представление о сложной динамике процессов, отвечающих за ценообразование ([22], [55]).

Большие объемы имеющихся данных, наличие статистических закономерностей в данных и механическая природа процесса исполнения заявок делает элек-

тронной рынок интересным объектом для статистического анализа и стохастического моделирования.

На фундаментальном уровне статистический анализ и моделирование высокочастотных данных могут дать представление о взаимосвязях между потоками заявок, ликвидностью и динамикой цен ([21], [91], [42]), что может заполнить пробел между работами в области анализа рыночной микроструктуры ([17], [82], [44], [48], [78], [84]), которые сосредоточены на моделях, описывающих механизмы образования цены в моделях равновесия, и стохастическими моделями типа «черный ящик», используемыми в финансовом риск-менеджменте, которые представляют цену как экзогенный случайный процесс. На уровне приложений модели высокочастотных данных дают количественные инструменты для маркет-мейкинга ([6]) и алгоритмов оптимального исполнения больших заявок ([4], [16], [81], [3]). Другое очевидное приложение - это разработка статистических моделей для прогнозирования краткосрочного поведения рыночных переменных, таких как цена, торговый объем и потоки заявок.

Исследование высокочастотной динамики рынков также является важным для риск-менеджмента и финансового регулирования. Даже при том, что горизонты, традиционно рассматриваемые риск-менеджерами и регуляторами, больше (как правило, дневные и более), торговые стратегии на разных частотах могут взаимодействовать довольно сложным способом, что приводит к возмущениям на разных временных масштабах - от высокочастотных до менее частотных, что может даже привести к рыночным сбоям, таким как Flash Crash 2010 ([25], [62]).

Микроструктурные модели динамики книги заявок являются сейчас популярной темой для применения стохастических моделей, см., например, [82], [43], [52], [6], [84]. В работе Конта и коллег [34] предложена непрерывная марковская модель для описания динамики книги заявок. В работе [34] книга заявок рассматривается как специальная система массового обслуживания, в которой приходы заявок

и отмены существующих заявок единичных объемов моделируются независимыми пуассоновскими процессами. Такой тип систем массового обслуживания может быть описан в терминах процессов гибели-размножения, чьи состояния есть количество акций на данном ценовом уровне, операция рождения соответствует приходу новой лимитной заявки, а гибель - отмена лимитной заявки либо же ее сведение в сделку. Процессы гибели-размножения являются хорошо изученными статистическими моделями, которые могут быть рассмотрены как специальные примеры более общих *двухсторонних процессов риска*, известных в теории страхования как *процессы риска со случайными премиями*. Математическая модель, представленная в [34], имеет некоторые формальные предположения. С одной стороны, эти предположения дают возможность представить аналитические выражения для некоторых характеристик, связанных с динамикой книги заявок. С другой стороны, эти предположения являются необходимыми для того, чтобы соответствующая модель имела прикладную значимость. Тем не менее, они имеют существенные ограничения и нереалистичны с практической точки зрения. Следовательно, необходимо иметь очевидную интегральную характеристику текущего состояния книги заявок, которая могла бы быть вычислена и изучена без привлечения аппарата систем массового обслуживания.

Такая характеристика - *процесс дисбаланса потоков заявок OFI* - была рассмотрена в 2011 году в работе [30], финальная версия которой [31] была опубликована в 2014 году. Тот же самый процесс независимо был введен и изучался в [28, 76, 77] под названием *обобщенный процесс цены*. Процесс OFI использует всю доступную информацию о постановке, снятии *всех* заявок, а также информацию о сделках, поэтому является более чувствительным к рыночной информации, чем сам процесс цены в ее классическом понимании. Более подробно этот принцип будет рассмотрен далее.

В рамках подхода, развитого в данной работе, лежит идея о том, что момен-

ты событий прихода заявок являются точечным случайным процессом хаоса на временной оси. Более того, предполагается, что этот точечный процесс является нестационарным (неоднородным по времени), поскольку изменения книги заявок в значительной степени подвержены влиянию нестационарных информационных потоков. Как известно, большинство разумных вероятностных моделей нестационарных (неоднородных по времени) хаотических процессов являются *дважды стохастическими процессами Пуассона*, также называемых *процессами Кокса* (см., например, [14, 53]). Эти процессы определяются как пуассоновские процессы со случайными интенсивностями. Обычные процессы Пуассона можно рассматривать как наилучшие модели для стационарных (однородных по времени) хаотических потоков событий. Заметим, что привлекательность процессов Пуассона как моделей однородного дискретного случайного хаоса обусловлена как минимум двумя причинами. Во-первых, процессы Пуассона являются точечными процессами, такими, что интервалы времени между последовательными событиями являются независимыми случайными величинами с одним и тем же экспоненциальным распределением, а как известно, экспоненциальное распределение обладает максимальной дифференциальной энтропией среди всех абсолютно непрерывных распределений с конечным первым моментом, определенных на неотрицательной полуоси, в то время как энтропия является естественной и удобной мерой неопределенности. Во-вторых, точки, образующие пуассоновский процесс, равномерно распределены вдоль оси времени в том смысле, что для любого конечного временного интервала $[t_1, t_2]$, $t_1 < t_2$, условное совместное распределение точек пуассоновского процесса, попадающих в интервал $[t_1, t_2]$ при условии, что число таких точек является фиксированным и составляет, скажем, n , совпадает с совместным распределением порядковых статистик, построенных из независимых выборок размера n из однородного распределения на отрезке $[t_1, t_2]$, в то время как равномерное распределение обладает максимальной дифференциальной эн-

тропией среди всех абсолютно непрерывных распределений, сосредоточенных на конечных интервалах, и очень хорошо соответствует традиционному представлению абсолютно непредсказуемой случайной величины (см, например. [14, 51]).

Финансовые рынки представляют собой примеры сложных открытых стохастических информационных систем, в которых можно выделить два основных источника случайности: внутренний и внешний. Внутренний источник случайности порождает неопределенность, обусловленную различием стратегий очень большого числа участников рынка. Физическим аналогом такой случайности может служить хаотическое тепловое движение частиц в замкнутых системах. Внешний источник случайности – это плохо поддающийся более или менее полному прогнозированию поток новостей политического и экономического характера (в том числе потоки информации с внешних рынков и инструментов), в соответствии с которыми изменяются интересы и стратегии участников рынка. Эти два источника случайности будут учитываться при построении процесса дисбаланса потока заявок и процесса количества заявок как двухсторонних процессов риска, управляемых процессом Кокса, то есть процесса Пуассона со случайной интенсивностью.

Как известно, эмпирические (статистические) распределения приращений (логарифмов) финансовых активов и, в частности, цен акций на сравнительно коротких временных интервалах являются распределениями с тяжелыми хвостами с вершиной заметно острее (то есть с более высоким показателем коэффициента эксцесса), чем у нормальных распределений (островершинные распределения). В то же время, как было указано в некоторых исследованиях, финансовые данные обладают свойством фрактальности, то есть, обладают свойством самоподобия на различных временных горизонтах, см. [80].

Устойчивые процессы Леви были одними из первых моделей, успешно примененных на практике для объяснения свойства острровершинности конечномерных распределений наблюдаемых процессов на финансовых рынках, ровно как и свой-

ства их самоподобия. В соответствии с подходом, основанным на классических предельных теоремах теории вероятностей, отличные от нормальных устойчивые процессы Леви могут быть получены как предельные в функциональных предельных теоремах для случайных блужданий только если элементарные скачки этих блужданий обладают бесконечными дисперсиями. В [33] были доказаны некоторые функциональные предельные теоремы диффузионного типа для динамики книги заявок на ликвидных рынках и было отдельно отмечено, что подход, используемый в статье имеет место только тогда, когда размеры заявок обладают бесконечными дисперсиями. В свою очередь, последнее возможно только тогда, когда вероятности прихода заявок со сколь угодно большими объемами являются строго положительными. К сожалению, последнее условие представляется трудно выполнимым с практической точки зрения. Таким образом, в рамках классического подхода, использовавшегося в [80] и [33], использование устойчивых процессов Леви как адекватных моделей для описания динамики цен финансовых индексов представляется по крайней мере сомнительным.

В финансовой математике эволюция (логарифмов) цен акций и финансовых индексов на малых временных горизонтах часто моделируются случайными блужданиями. Самый простой пример такого подхода - это модель Кокса-Росса-Рубинштейна (см., например, [89]). В то же самое время, наиболее успешные (адекватные) модели динамики (логарифмов) финансовых индексов на больших временных горизонтах являются подчиненными винеровскими процессами (процессами броуновского движения со случайным временем), такие как обобщенные гиперболические процессы, в частности, variance-gamma (VG) процессы и нормальные\ обратные гауссовские процессы (NIG), см. [89]. Подчиненным винеровским процессам соответствуют конечномерные распределения, обладающие свойствами, упомянутыми выше: тяжелые хвосты и островершинность.

В [29, 63] и [51] тяжелые хвосты эмпирических распределений цен индексов

были объяснены с использованием предельных теорем для сумм со *случайным числом* независимых случайных величин как частного случая случайных блужданий со случайными моментами остановки (randomly stopped random walks).

При рассмотрении случайных блужданий схема случайного суммирования является естественным аналогом схемы подчинения более общих случайных процессов. В [51, 66, 67] была предложена модель эволюции неоднородных хаотических случайных процессов, в частности, динамики цен акций и финансовых индексов, с помощью случайных блужданий, порожденных обобщенным дважды стохастическими пуассоновскими процессами (обобщенными процессами Кокса). *Дважды стохастический пуассоновский процесс* (также называемый *процессом Кокса*) - это случайный точечный процесс вида $N_1(\Lambda(t))$, где $N_1(t)$, $t \geq 0$ - однородный пуассоновский процесс с единичной интенсивностью и случайный процесс $\Lambda(t)$, $t \geq 0$, является независимым от процесса $N_1(t)$ и обладает следующими свойствами: $\Lambda(0) = 0$, $\mathsf{P}(\Lambda(t) < \infty) = 1$ для любого $t > 0$, траектории путей $\Lambda(t)$ не убывают и непрерывны справа. Обобщенный процесс Кокса - это случайная сумма независимых одинаково распределенных случайных величин, для которого количество слагаемых следует процессу Кокса. Подобные непрерывные по времени случайные блуждания рассматривались в [54, 58, 94].

В соответствии с подходом, использовавшимся в [51, 66, 67], случайные распределение для обобщенного процесса Кокса со элементарными скачками, обладающими конечными дисперсиями, должны иметь вид дисперсионно-сдвиговых смесей нормальных законов, которые всегда обладают тяжелыми хвостами и остроконечными, если смещающее распределение невырождено. Кроме того, в [64, 65] было показано, что ненормальные устойчивые законы могут возникать как предельные распределения для сумм независимых одинаково распределенных случайных величин с *конечными* дисперсиями, если количество слагаемых в сумме случайно, а распределение количества слагаемых сходится к устойчивому закону

и определено на неотрицательной полуоси. В терминах процессов Кокса последнее условие означает, что конечномерные распределения управляющего процесса $\Lambda(t)$ являются асимптотически устойчивыми. В свою очередь, это означает, что интенсивность потоков заявок существенно нерегулярна, что приводит к хорошо известному эффекту кластеризации финансовых данных.

Степень изученности и проработанности проблемы

Как уже упоминалось ранее, моделированию различных аспектов динамики книги заявок посвящены десятки работ. Далее упоминаются только те работы, которые лежат в русле исследований, которым посвящена диссертация. В работе [34] для описания динамики книги заявок предложена модель типа марковского процесса с непрерывным временем. В этой работе книга заявок рассматривается как специальная система массового обслуживания с независимыми пуссоновскими входящими потоками заявок единичного объема на покупку/продажу и отмену ранее поступивших заявок. Такие системы массового обслуживания традиционно анализируются с помощью аппарата процессов рождения-гибели. Математическая модель книги заявок типа системы массового обслуживания, предложенная в указанной работе, снабжена целым рядом жестких формальных условий. В частности, интенсивности входящих потоков предполагаются постоянными. Эти условия слишком ограничительны и нереалистичны с практической точки зрения. В 2011 г. в работе [30], финальная версия которой [31] опубликована в 2014 г., была предложена удобная интегральная характеристика текущего состояния книги заявок – *процесс дисбаланса потока заявок* (order flow imbalance process, OFI), – которую можно вычислять и анализировать, не используя аппарат теории массового обслуживания.

Точно такая же формальная модель была независимо предложена и исследована в работах [77] и [76] под названием *процесс обобщенной цены*. Процесс OFI

оказывается существенно более чувствительным к информации о рынке, нежели сам процесс цены. Это обусловлено, во-первых, тем, что на рассматриваемых временных масштабах изменения цены являются чрезвычайно редкими событиями по сравнению с изменениями состояния книги заявок, и во-вторых, тем, что процесс OFI учитывает не только изменения лучших цен покупки и продажи, но и события, происходящие глубоко внутри книги заявок.

В данной работе показано, что процесс дисбаланса потоков заявок вполне естественно трактовать как так называемый двусторонний процесс риска – процесс риска со случайными премиями. При этом в качестве формальной математической модели последнего предложено использовать специальный обобщенный дважды стохастический пуассоновский процесс (обобщенный процесс Кокса), что позволяет применить и обобщить аппарат функциональных предельных теорем, имеющих свою историю развития. В работе [61] были доказаны некоторые функциональные предельные теоремы для обобщенных процессов Кокса с квадратично интегрируемыми управляющими процессами $\Lambda(t)$. Однако, класс предельных процессов для обобщенных процессов Кокса, у которых такие управляющие процессы, а скачки имеют конечную дисперсию, довольно узок, поскольку он, в частности, не может содержать никаких устойчивых процессов Леви (помимо винеровского процесса). В книге [59] приведено общее утверждение (следствие VII.3.6) о сходимости суперпозиций семимартингалов со стационарными приращениями, из которого можно вывести функциональную предельную теорему для обобщенных процессов Кокса с управляющими процессами $\Lambda(t)$, математические ожидания которых удовлетворяют условию Липшица $E\Lambda(t) \leq C t$. Однако в таком случае класс предельных процессов для обобщенных процессов Кокса, у которых скачки имеют конечную дисперсию, также не может содержать никаких устойчивых процессов Леви. В работах [73] и [74] доказана функциональная предельная теорема для обобщенных процессов Кокса, у которых скачки имеют конечную дисперсию, а управляющие

процессы $\Lambda(t)$ удовлетворяют условию $\mathbf{E}\Lambda^\delta(t) \leq Ct^{\delta_1}$ с некоторыми $\delta, \delta_1 \in (0, 1]$. В таком случае класс предельных процессов для обобщенных процессов Кокса, у которых скачки имеют конечную дисперсию, может содержать устойчивые процессы Леви, что, в частности, хорошо описывает часто наблюдаемую кластеризацию заявок. Однако в указанных работах рассмотрен только симметричный случай.

Понятие токсичности потоков заявок, рассматриваемое в третьей главе, было введено в работах [35], [36], после чего получило развитие использования индикатора VPIN в качестве метрики оценки токсичности потоков заявок. Однако, позднее способность индикатора VPIN предсказывать токсичную ликвидность на финансовых рынках была подвергнута сомнению (см., например, [5]).

Цели и задачи исследования

- разработать модель потоков заявок со случайными интенсивностями, получить теоретическое и эмпирическое обоснование модели;
- разработать удобный индикатор текущего состояния книги заявок, чувствительный к информации о потоках всех заявок;
- изучить предельные состояния рассматриваемых процессов для построения их асимптотических аппроксимаций;
- на основе процесса дисбаланса потоков заявок разработать математическую модель токсичности потоков заявок, процедуру оценки токсичности потоков заявок в режиме реального времени, решить задачи прогнозирования токсичной ликвидности и ценовых шоков.

Прикладная значимость:

- риск-менеджмент, финансовое регулирование;

- маркет-мейкинг;
- оптимальное исполнение заявок (optimal execution).

Положения, выносимые на защиту

Предмет защиты составляют следующие положения:

1. Разработана новая удобная модель эволюции книги заявок; в качестве удобного индикатора, описывающего динамику потоков заявок, предложен процесс дисбаланса потоков заявок; для описания процесса дисбаланса потоков заявок предложена математическая модель вида двухсторонних процессов риска (процессов риска со случайными премиями); предложено мультипликативное представление интенсивностей потоков заявок, что позволило рассмотреть процесс дисбаланса потоков заявок как специальный обобщенный процесс Кокса;
2. Предложенная модель изучена аналитически: получена теорема переноса, доказаны функциональные предельные теоремы о сходимости процесса дисбаланса потоков заявок в схеме серий к процессам Леви в пространстве Скорохода, доказаны теоремы о сходимости процессов дисбаланса потоков заявок с элементарными скачками, обладающими конечными дисперсиями, к процессам Леви с распределениями, имеющими вид дисперсионно-сдвиговых смесей нормальных одномерных распределений, в частности, к обобщенным гиперболическим процессам Леви;
3. Построена математическая модель токсичности потоков заявок: на основе аналитической модели процесса дисбаланса потоков заявок formalизованы понятия токсичности потоков заявок; разработаны байесовский и квантильный показатели токсичности, рассчитываемые на основе параметров, описывающих потоки *всех* заявок, поступающих на рынок; реализована процедура

оценки показателя токсичности в режиме реального времени; решены задачи прогнозирования токсичной ликвидности и ценовых шоков с помощью предложенного показателя токсичности.

Достоверность результатов обеспечивается корректными доказательствами теорем и подробными описаниями алгоритмов и экспериментов, допускающими воспроизводимость.

Научная новизна диссертации Чертока А.В. заключается в описании модели потоков заявок с неоднородными интенсивностями, введении процесса дисбаланса потоков заявок как интегрального показателя книги заявок, а также рассмотрении его математической модели в виде обобщенных процессов риска, доказательстве функциональных предельных теорем для процесса дисбаланса потоков заявок, а также в разработке аналитического показателя токсичности потоков заявок на основе информации о потоках всех заявок, поступающих на финансовый рынок, и экспериментальном исследовании предложенных подходов на примере важных практических задач. Теоретическая значимость диссертации Чертока А. В. состоит в том, что предложено теоретическое обоснование адекватности мультиплекативной модели интенсивности потоков заявок, доказаны теоремы о сходимости обобщенных процессов Кокса с элементарными скачками, обладающими конечными дисперсиями, к процессам Леви с несимметричными распределениями, имеющими вид дисперсионно-сдвиговых смесей нормальных одномерных распределений, в частности, к обобщенным гиперболическим процессам Леви. Практическая значимость диссертационной работы состоит в том, что разработанный метод вычисления показателя токсичности на основе процесса дисбаланса потоков заявок позволяет своевременно выявлять токсичную ликвидность и предупреждать ценовые шоки в режиме реального времени.

Используемые подходы и методы

В данной работе используются методы многомерного статистического анализа, прямые вероятностные методы и методы характеристических функций.

Апробация работы

Основные результаты диссертации были представлены на следующих конференциях и семинарах:

- XX Международная конференция студентов, аспирантов и молодых ученых «Ломоносов-2013», г. Москва, 8-13 апреля 2013 года.
- «Достижения и перспективы эконометрических исследований в России», г. Казань, 23 июля 2013 г.
- XXI Международная конференция студентов, аспирантов и молодых ученых «Ломоносов-2014», г. Москва, 7-11 апреля 2014 года.
- XXXII International Seminar on Stability Problems for Stochastic Models, Trondheim (Norway), June 16 – 21, 2014.
- ICNAAM-2014 (Rhodes Island, Greece, 22-28 September 2014).
- Научная конференция «Тихоновские чтения», г. Москва, 27-31 октября 2014.
- Спецсеминар кафедры Математической Статистики ВМК МГУ «Теория риска и смежные вопросы».

Публикации

Результаты диссертационной работы опубликованы в 8 печатных изданиях, в том числе высокорейтинговых журналах; 4 работы изданы в журналах, рекомендованных ВАК, 3 - в тезисах докладов.

Публикации из списка ВАК:

1. В. Ю. Королев, А. В. Черток, А. Ю. Корчагин, А. К. Горшенин Вероятностно-статистическое моделирование информационных потоков в сложных финансовых системах на основе высокочастотных данных *Информатика и ее применение*, 7:1 (2013), 12–21
2. В. Ю. Королев, А. Ю. Корчагин, И. А. Соколов, А. В. Черток, О работах в области моделирования информационных потоков в современных высокочастотных финансовых приложениях. *Системы и средства информатики*, 24:4 (2014), 63–85
3. А. В. Черток. О формализации понятия токсичности потока заявок на финансовых рынках. *Информатика и ее применение*, 8:4 (2014), 20–31
4. V.Yu. Korolev, A. V. Chertok, A. Yu. Korchagin, A.I. Zeifman Modeling high-frequency order flow imbalance by functional limit theorems for two-sided risk processes *Applied Mathematics and Computation*, Volume 253, 15 February 2015, Pages 224–241

Прочие публикации:

5. Andrey K. Gorshenin, Victor Yu. Korolev, Alexander I. Zeifman, Sergey Ya. Shorgin, Andrey V. Chertok, Artem I. Evstafyev and Alexander Yu. Korchagin, Modelling stock order flows with non-homogeneous intensities from high-frequency data *AIP Conf. Proc. 1558* , 2394 (2013)
6. Chertok A. V., Korolev V. Yu., and Korchagin A. Yu. Modeling high-frequency non-homogeneous order flows by compound Cox processes. *Available at: arXiv:1410.1900v2 [math.PR]*, January 14, 2014.

7. Alexander I. Zeifman, Victor Yu. Korolev, Andrey V. Chertok and Sergey Ya. Shorgin, On ergodicity bounds for an inhomogeneous birth-death process *AIP Conf. Proc. 1648* , 250011 (2015)
8. A. Chertok, On order flow toxicity. *AIP Conf. Proc. 1648* , 250013 (2015)

Структура и объем

Диссертация состоит из введения, трех глав, заключения и списка цитируемой литературы. Полный объем диссертации составляет 113 страниц. Список литературы содержит 96 наименований.

Краткое содержание диссертации

Во введении обосновывается актуальность темы диссертации, формулируется цель работы, приводится краткая аннотация диссертационной работы по главам. Указывается научная новизна и положения, выносимые на защиту.

В первой главе вводятся ключевые модели и терминология. В Разделе 1.2 даются необходимые определения и в условиях, когда интенсивности потоков заявок остаются постоянными, вводится двухсторонний процесс риска, адаптирующий определение процесса риска со случайными премиями из теории страхований к специфике задачи, и доказывается, что он является обобщенным процессом Пуассона. В Разделе 1.3 вводится условный неоднородный процесс дисбаланса потока заявок (order flow imbalance, OFI), предлагается мультипликативная форма интенсивностей потоков заявок и, наконец, условие случайной природы внешнего информационного фона и получен процесс дисбаланса потоков заявок в общем виде (безусловный) как специальный обобщенный дважды стохастический процесс Пуассона. Также сформулированы предельные теоремы переноса для одномерных распределений процесса дисбаланса потоков заявок в условиях существования

ния вторых моментов объемов заявок. Глава завершается эмпирическим анализом реальных данных.

Вторая глава посвящена функциональным предельным теоремам для процесса дисбаланса потоков заявок OFI. В Разделе 2.1 второй главы содержится некоторый предварительный материал о пространстве Скорохода и процессах Леви. В Разделе 2.2 доказывается общая функциональная теорема, устанавливающая условия сходимости процессов OFI к процессам Леви в пространстве Скорохода в контексте роста интенсивностей потока заявок. Для этих целей расширены классические результаты, представленные, например, в [59]. В Разделе 2.3 рассматриваются условия сходимости процессов OFI с элементарными скачками (т.е., размерами заявок), обладающими конечными дисперсиями, к процессам Леви с сдвиг-масштабными смесями нормальных одномерных распределений, то есть, к подчиненным винеровским процессам, в частности, к обобщенным гиперболическим процессам Леви.

В третьей главе в рамках рассмотренной модели были введены такие понятия, как мгновенный профиль токсичности, а также байесовский и квантильный показатели токсичности, рассчитываемые на основе параметров, описывающих потоки всех заявок, поступающих на рынок. Эти показатели рассчитываются для двух модельных типов потоков заявок, в первом из которых заявки имеют единичный объем, во втором – объем заявок является случайным и имеющим показательное распределение. Для последней из двух моделей была проведена валидация на реальных данных (фьючерс на индекс РТС) и были построены показатели токсичности в режиме реального времени.

В заключении сформулированы основные полученные результаты, выводы из

проведенных исследований, а также возможные направления для дальнейших исследований.

Благодарности

Автор диссертационной работы выражает огромную благодарность своему научному руководителю д.ф.-м.н. профессору Королеву Виктору Юрьевичу за всестороннюю помощь и поддержку на всех этапах подготовки диссертации, студентам факультета ВМК МГУ им. М.В. Ломоносова Хазеевой Гелане, Полдневу Антону и Николайчук Дарье за помощь в подготовке материала диссертации, профессору Чикагского университета Баласанову Юрие Георгиевичу за ценные идеи и замечания, а также выпускникам факультета ВМК МГУ им. М.В. Ломоносова Корчагину Александру, Горшенину Андрею и Евстафьеву Артему за плодотворную совместную деятельность при написании статей.

Глава 1

Дисбаланс потоков заявок

В данной главе представлена пошаговая процедура построения модели, описывающей динамику потоков заявок на финансовых рынках, а также вводятся ключевые характеристики потоков заявок, различные свойства которых будут изучаться в последующих главах. В Разделе 1 даются необходимые определения при условии, что интенсивности потоков заявок являются постоянными, вводится процесс дисбаланса количества заявок, соответствующий по структуре понятию процесса риска со случайными премиями, известного в теории страхования, а также доказывается, что он является специальным обобщённым пуассоновским процессом. В Разделе 2 вводится условный неоднородный процесс дисбаланса потоков заявок (order flow imbalance, OFI), предлагаются мультипликативная форма представления интенсивностей потоков заявок и условие о случайной природе интенсивности внешнего информационного фона. В завершение получен (безусловный) процесс дисбаланса потоков заявок в общем виде как специальный обобщённый дважды стохастический процесс Пуассона. В завершение представляется анализ реальных данных о потоках заявок на рынке FORTS и проводится проверка гипотез, выдвинутых в рамках предложенной модели.

1.1 Описание модели

1.1.1 Книга заявок

На классических электронных рынках, которые в данный момент принадлежат к числу наиболее активно развивающихся типов рынка, биржевая цена финансового инструмента в ее классическом понимании является результирующей, интегральной характеристикой системы торгов, которая описывается динамикой так называемой *книги заявок* (limit order book), представляющей из себя список всех актуальных на данный момент предложений о покупке и продаже инструмента по различным ценам (см. рис. 1.1). Динамику книги заявок на электронном рынке определяют три типа заявок:

- *лимитная* заявка обозначает желание купить (продать) заданное количество акций по цене не выше (не ниже) заданной, при этом такая заявка немедленно добавляется в книгу заявок;
- *рыночная* заявка обозначает желание купить или продать заданное количество акций по лучшей цене, представленной в книге заявок, после чего немедленно происходит сделка;
- заявка *на отмену* обозначает намерение отменить существующую лимитную заявку, после чего она удаляется из книги заявок.

Разумеется, лимитная заявка может оказаться рыночной, если заявленная в ней цена позволяет немедленно произвести сведение с одной из лимитных заявок на противоположной стороне книги заявок. Участники рынка, присылающие лимитные ордера, являются *поставщиками ликвидности* (liquidity providers), а те, кто присылают рыночные заявки – *потребителями ликвидности* (liquidity takers).

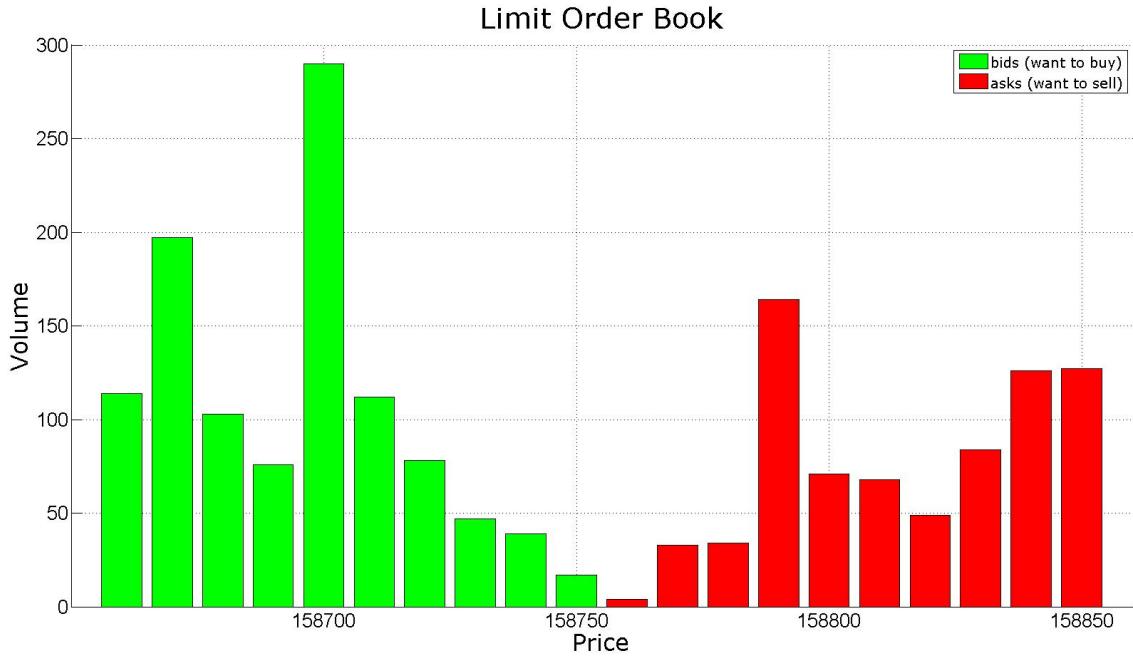


Рис. 1.1: Книга заявок в некоторый момент времени. Высота столбиков равна суммарному объёму лимитных заявок на соответствующем ценовом уровне (зеленый – покупки, красный – продажи)

Более формально, рассмотрим прежде динамику книги заявок на дискретной сетке цен $\Pi = \{1, 2, \dots, M\}$ как процесс с непрерывным временем

$$book(t) \equiv (\mathbf{v}^b, \mathbf{v}^a) \equiv (v_1^b(t), v_2^b(t), \dots, v_M^b(t); v_1^a(t), v_2^a(t), \dots, v_M^a(t)),$$

где $v_p^b(t)$ ($v_p^a(t)$) обозначает количество лимитных заявок на покупку (продажу) с ценой $p \in \Pi$. Так как в один момент не может существовать заявок на покупку и продажу по одной цене (иначе они будут сведены), мы должны потребовать, чтобы $\min(v_p^b(t), v_p^a(t)) = 0$ для всех p и t .

Лучшая цена на продажу $a(t)$ (лучший *ask*) определяется как

$$a(t) = \inf\{p : v_p^a(t) > 0\} \wedge (M + 1),$$

лучшая цена на покупку $b(t)$ (лучший *bid*) определяется как

$$b(t) = \sup\{p : v_p^b(t) > 0\} \vee 0.$$

При этом процесс цены можно, например, определить как

$$p(t) = \frac{a(t) + b(t)}{2},$$

также называемый мидпрайсом, а процесс $s(t) = b(t) - a(t)$ – спредом. Таким образом, процесс цены $p(t)$ является результатом процесса эволюции книги заявок, инициированного потоком заявок трех типов.

1.1.2 Динамика книги заявок

Предположим вначале, что поток информации, поступающей извне, фиксирован. Тогда при фиксированной информации можно считать, что внутренняя случайность является установившимся хаосом. Как показано, например, в [67] и [69], естественными математическими моделями хаотических потоков являются пуассоновские процессы, характеризуемые тем, что интервалы времени между информативными событиями являются независимыми одинаково распределенными случайными величинами с экспоненциальным распределением. Поэтому на первом этапе при построении рассматриваемой модели потоки заявок моделируются с использованием независимых пуассоновских процессов (как это сделано, например, в работах [34], [32]):

- моменты появления лимитных заявок на покупку (продажу) на ценовой уровень, расположенный на расстоянии i от лучшей котировки противоположного типа, образуют пуассоновский процесс с параметром $\lambda_i^+(\lambda_i^-)$ (эмпири-

ческие исследования [96] и [21] показывают, что степенной закон

$$\lambda_i^\pm = \frac{k}{i^\alpha}$$

является хорошей аппроксимацией);

- моменты появления рыночных заявок на покупку (продажу) образуют пуассоновский процесс с параметром $\mu^+(\mu^-)$;
- моменты появления заявок на отмену лимитных заявок на покупку (продажу), находящихся на расстоянии i от лучшей котировки противоположного типа, образуют пуассоновский процесс с интенсивностью $\theta_i^+(\theta_i^-)$.

Вначале мы рассматриваем потоки заявок единичного объема, однако позже все рассуждения могут быть распространены на более общий случай (также см. [58]). Таким образом, $book(t)$ является цепью Маркова с непрерывным временем с пространством состояний $(\mathbb{Z}^+)^{2M}$ и следующими переходами:

$$\begin{aligned} v_i^a(t) &\rightarrow v_i^a(t) + 1 \text{ с интенсивностью } \lambda_{i-b(t)}^- & \text{для } i > b(t), \\ v_i^a(t) &\rightarrow v_i^a(t) - 1 \text{ с интенсивностью } \theta_{i-a(t)}^- & \text{для } i \geq a(t), \\ v_i^a(t) &\rightarrow v_i^a(t) - 1 \text{ с интенсивностью } \mu^+ & \text{для } i = a(t) > 0. \end{aligned}$$

$$\begin{aligned} v_i^b(t) &\rightarrow v_i^b(t) + 1 \text{ с интенсивностью } \lambda_{a(t)-i}^+ & \text{для } i < a(t), \\ v_i^b(t) &\rightarrow v_i^b(t) - 1 \text{ с интенсивностью } \theta_{b(t)-i}^+ & \text{для } i \leq b(t), \\ v_i^b(t) &\rightarrow v_i^b(t) - 1 \text{ с интенсивностью } \mu^- & \text{для } i = b(t) < M + 1. \end{aligned}$$

Введём следующие независимые пуассоновские процессы:

- $L_i^\pm(t)$: потоки лимитных ордеров с интенсивностями λ_i^\pm ;

- $M^\pm(t)$: потоки рыночных ордеров с интенсивностями $\mu^+ \mathbb{I}(\mathbf{v}^\mathbf{a} \neq 0)$ и $\mu^- \mathbb{I}(\mathbf{v}^\mathbf{b} \neq 0)$;
- $C_i^\pm(t)$: потоки заявок на отмену лимитных ордеров с интенсивностями θ_i^\pm ,

а также пуассоновский процесс

$$N(t) = M^+(t) + M^-(t) + \sum_{i=1}^M (L_i^+(t) + L_i^-(t)) + \sum_{i=1}^M (C_i^+(t) + C_i^-(t)),$$

описывающий поток всех заявок, поступающих на рынок.

Процессы $L_i^\pm(t)$, $M^\pm(t)$, $C_i^\pm(t)$ полностью определяют процесс цены $p(t)$ и для него, вообще говоря, могут быть выписаны соответствующие стохастические дифференциальные уравнения (см. [2]), однако дальнейшая его аналитическая интерпретация представляется очень сложной или вообще невозможной даже при довольно сильных допущениях о постоянных и независимых интенсивностях потоков заявок разных типов, что также слабо соотносится с реальностью.

1.2 Процесс дисбаланса количества заявок

В данном разделе мы опишем базовые модели процесса дисбаланса количества заявок (как индикатора состояния книги заявок) с его некоторыми важными свойствами. Эти модели необходимы для построения непрерывной модели процесса дисбаланса потоков заявок. Вначале мы определим условную модель, предполагающую, что интенсивности потоков заявок являются постоянными, а затем расширим эту модель для случайных интенсивностей. В последней части мы предложим мультиплекативную форму случайных интенсивностей.

1.2.1 Постоянные интенсивности потоков заявок

Текущее состояние книги заявок желательно описать с помощью некоторого простого индикатора. Пример такого индикатора – это мидпрайс $p(t)$. Но этот индикатор учитывает только лучшие котировки книги заявок и с теоретической точки зрения описательных статистик играет роль экстремальной порядковой статистики, аккумулирующей информацию только о значениях минимального аска и максимального бида. В то же самое время было бы справедливо учитывать не только изменения лучших котировок, но и постановку/снятие заявок в глубине книги заявок, поскольку каждое такое действие оказывает влияние на текущее распределение сил покупателей и продавцов. Для этих целей мы вводим *процесс дисбаланса количества заявок*.

Напомним, что внешний информационный фон пока предполагается неизменным.

Зафиксируем прежде достаточно небольшой интервал времени $[0, t]$, позволяющий считать, что на таком интервале интенсивности описанных событий постоянны. Пусть по-прежнему $N(t)$, $t \in [0, T]$, – пуассоновский процесс, соответствующий всем событиям в книге заявок и имеющий интенсивность

$$\lambda = \mu^+ + \mu^- + \sum_{i=1}^M (\lambda_i^+ + \lambda_i^-) + \sum_{i=1}^M (\theta_i^+ + \theta_i^-).$$

Представим его как суперпозицию процессов $N^+(t)$ и $N^-(t)$ с интенсивностями соответственно

$$\lambda^+ = \mu^+ + \sum_{i=1}^M \lambda_i^+ + \sum_{i=1}^M \theta_i^-$$

и

$$\lambda^- = \mu^- + \sum_{i=1}^M \lambda_i^- + \sum_{i=1}^M \theta_i^+.$$

Таким образом, $\lambda = \lambda_+ + \lambda_-$, а процессы $N^+(t)$ и $N^-(t)$ характеризуют накоплен-

ную силу покупателей и продавцов соответственно (при этом заметим, что снятие заявок на стороне продавцов в данном случае увеличивает силу покупателей, и наоборот) и являются условно независимыми при фиксированном потоке информации, поступающем извне за время $[0, t]$.

Определим процесс

$$NOI(t) = N^+(t) - N^-(t),$$

который может быть определён как (условный) *процесс дисбаланса количества заявок*. Это определение допускает другую формулировку, которая окажется полезной для дальнейших обобщений моделей с заявки произвольного объёма. А именно, условным *процессом дисбаланса количества заявок* $NOI(t)$ мы также можем назвать процесс, приращения которого на интервале $[0, t] \subseteq [0, T]$ имеют вид

$$NOI(t) - NOI(0) = \sum_{j=1}^{N(t)} X_j, \quad (1.1)$$

где X_1, X_2, \dots – независимые одинаково распределенные величины, такие, что

$$X_j = \begin{cases} +1 & \text{с вероятностью } \frac{\lambda^+}{\lambda^+ + \lambda^-} \\ -1 & \text{с вероятностью } \frac{\lambda^-}{\lambda^+ + \lambda^-}, \end{cases} \quad j = 1, 2, \dots \quad (1.2)$$

причем случайные величины X_1, X_2, \dots стохастически независимы от процесса $NOI(t)$ (в этом можно убедиться, непосредственно выписав характеристическую функцию случайной величины $NOI(t) - NOI(0)$ (это будет сделано позднее в более общей ситуации, см. Лемму 1.1)).

При этом

$$\mathbb{E}X_j = \frac{\lambda^+ - \lambda^-}{\lambda^+ + \lambda^-}, \quad \mathbb{D}X_j = 1 - \left(\frac{\lambda^+ - \lambda^-}{\lambda^+ + \lambda^-}\right)^2 = \frac{4\lambda^+\lambda^-}{(\lambda^+ + \lambda^-)^2},$$

так что

$$\mathbb{E} \sum_{j=1}^{N(t)} X_j = t(\lambda^+ - \lambda^-),$$

$$\mathbb{D} \sum_{j=1}^{N(t)} X_j = t(\lambda^+ + \lambda^-) \left[\left(\frac{\lambda^+ - \lambda^-}{\lambda^+ + \lambda^-} \right)^2 + \frac{4\lambda^+ \lambda^-}{(\lambda^+ + \lambda^-)^2} \right] = t(\lambda^+ + \lambda^-).$$

В дальнейшем без ограничения общности будем считать, что $NOI(0) = 0$. Для удобства временно будем считать, что $T = 1$ и будем рассматривать поведение $NOI(1)$, т.е. приращения процесса NOI за единицу времени.

Если $\lambda = \lambda^+ + \lambda^-$ очень велико, то есть в единицу времени происходит очень много информативных событий, то по центральной предельной теореме для пуссоновских случайных сумм справедливо приближенное соотношение

$$\mathbb{P}(NOI(1) < x) \approx \Phi \left(\frac{x - \lambda^+ + \lambda^-}{\sqrt{\lambda^+ + \lambda^-}} \right), \quad x \in \mathbb{R} \quad (1.3)$$

где $\Phi(x)$ – стандартная нормальная функция распределения

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-z^2/2} dz, \quad x \in \mathbb{R}.$$

При этом с учётом результатов работ [70, 88] точность аппроксимации (1.3) может быть оценена как

$$\sup_{x \in \mathbb{R}} \left| \mathbb{P}(NOI(1) < x) - \Phi \left(\frac{x - \lambda^+ + \lambda^-}{\sqrt{\lambda^+ + \lambda^-}} \right) \right| \leq \frac{0.3031}{\sqrt{\lambda}}, \quad \lambda = \lambda^+ + \lambda^-.$$

Теперь вспомним, что выше внешний поток информации считался фиксированным. Это предположение, в частности, широко используется в большинстве работ по моделированию динамики книги заявок и дает возможность использовать аппарат марковских цепей с непрерывным временем, для которых условие марковости в определенном смысле эквивалентно тому, что распределение вероят-

ностей интервалов времени между информативными событиями является экспоненциальным. В реальной практике это условие не выполняется, как видно из Рис. 1.2. На этом рисунке приведены гистограмма интервалов времени между событиями, произошедшими в течение всего рабочего дня на фьючерсе на индекс РТС 2013.03.05, и график плотности гамма-распределения с параметром формы 0.3642 и соответствующим параметром масштаба. Это распределение хорошо согласуется с гистограммой и заметно отличается от экспоненциального.

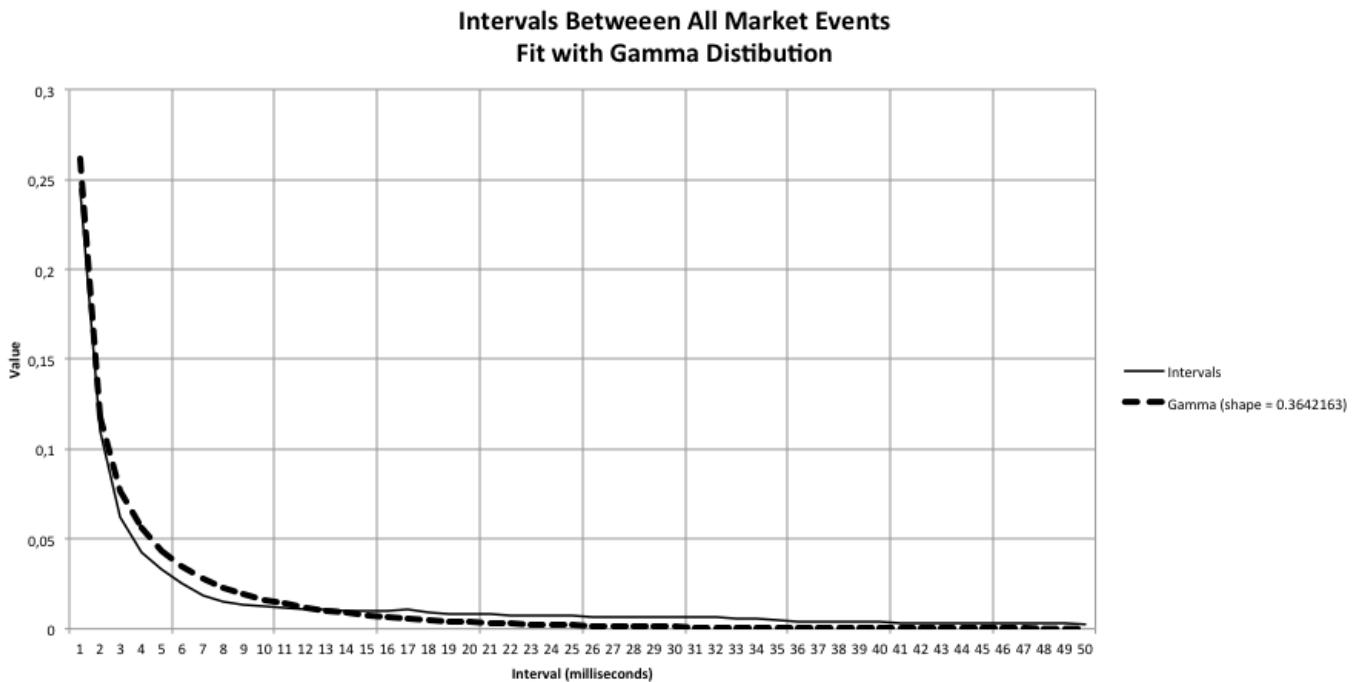


Рис. 1.2: Гистограмма времён между приходами заявок и плотность гамма-распределения

С другой стороны, хорошее согласие распределения вероятностей интервалов времени между событиями, заметное на Рис. 1.2, с указанным выше гамма-распределением подтверждает правильность рассуждений об *условной* марковости рассматриваемых процессов, поскольку, как известно, получение безусловного распределения из условного сводится к смешиванию условного распределения по распределению вероятностей, соответствующему закону распределения параметра, описывающего фиксированное условие. В то же время гамма-распределение

может быть представлено в виде смеси экспоненциальных распределений, только если его параметр формы не превосходит единицы, см. [46]. В той же работе показано, что если параметр формы r гамма-распределения, соответствующего плотности

$$g(x; r, \mu) = \frac{\mu^r}{\Gamma(r)} x^{r-1} e^{-\mu x}, \quad x \geq 0,$$

удовлетворяет условию $0 < r \leq 1$, тогда плотность $g(x; r, \mu)$ может быть представлена в виде

$$g(x; r, \mu) = \int_0^\infty p_\mu(z) z e^{-zx} dz, \quad (1.4)$$

где

$$p_\mu(z) = \frac{(z - \mu)^{-r} \mu^r}{z \Gamma(1 - r) \Gamma(r)} \mathbb{I}(\mu \leq z) \quad (1.5)$$

– плотность распределения Снедекора–Фишера. В работе [77] можно найти также результаты статистического анализа эволюции параметра μ модели (1.4)–(1.5) в течение дня.

Тем не менее, рабочий день, за который накапливались исходные данные для Рис. 1.2, – это слишком большой интервал времени и интенсивности потоков заявок сильно меняются в течение дня, поэтому распределение Снедекора–Фишера (1.4) это усреднённое статистическое распределение интенсивностей потоков заявок, так что практическая ценность этой модели сродни ценности информации о «средней температуре по больнице».

1.2.2 Случайные интенсивности потоков заявок

Для получения более тонких моделей *безусловного* распределения величины $NOI(1)$, в силу непредсказуемости потока внешней информации, следует считать, что λ^+ и λ^- – это некоторые конкретные значения *случайных величин* Λ^+ и Λ^- . Та-

ким образом, для безусловного распределения приращения $Q(1)$ «форсированно» получается модель

$$\begin{aligned} \mathsf{P}(NOI(1) < x) &\approx \int_{\mathbb{R}^+ \times \mathbb{R}^+} \Phi\left(\frac{x - \lambda^+ + \lambda^-}{\sqrt{\lambda^+ + \lambda^-}}\right) d\mathsf{P}(\Lambda^+ < \lambda^+, \Lambda^- < \lambda^-) = \\ &= \mathsf{E}\Phi\left(\frac{x - \Lambda^+ + \Lambda^-}{\sqrt{\Lambda^+ + \Lambda^-}}\right), \quad x \in \mathbb{R}. \end{aligned} \quad (1.6)$$

Эту модель можно статистически исследовать методом скользящего разделения смесей (СРС-методом), используя конечные аппроксимации для смеси (1.6):

$$\mathsf{P}(NOI(1) < x) \approx \sum_{j=1}^k p_j \Phi\left(\frac{x - a_j}{\sigma_j}\right). \quad (1.7)$$

Для этого должны быть оценены параметры $k, p_1, \dots, p_k, a_1, \dots, a_k, \sigma_1, \dots, \sigma_k$. В общем случае распределение $NOI(t)$ меняется с изменением t и, таким образом, параметры $k, p_1, \dots, p_k, a_1, \dots, a_k, \sigma_1, \dots, \sigma_k$ зависят от времени и должны оцениваться в «скользящем» режиме, для чего необходимо применять метод скользящего разделения смесей (СРС-метод), изложенный в [67], где много внимания уделено применению СРС-метода к декомпозиции волатильности в следующем ключе.

Обозначив $V = \Lambda^+ - \Lambda^-$ и $U = \Lambda^+ + \Lambda^-$, мы видим, что в рамках модели (1.7)

$$\mathsf{P}((U, V) = (\sigma_j^2, a_j)) = p_j, \quad j = 1, \dots, k. \quad (1.8)$$

Кроме того, в этом случае мы можем предположить, что

$$NOI(1) \stackrel{d}{=} X \cdot \sqrt{U} + V \quad (1.9)$$

где случайная величина X имеет стандартное нормальное распределение и яв-

ляется независимой от пары (U, V) . Из представления (1.9) следует, что

$$\mathsf{DNOI}(1) = \mathsf{DV} + \mathsf{EU} = \sum_{j=1}^k p_j (a_j - \bar{a})^2 + \sum_{j=1}^k p_j \sigma_j^2, \quad (1.10)$$

где $\bar{a} = \sum_{j=1}^k p_j a_j$. Первое слагаемое в правой части уравнения (1.10), DV , зависит только от весов p_j и средних a_j , которые в действительности обозначают средние значения *трендов* компонент смеси (1.7). Кроме того, a_j - это среднее изменение координаты в единицу времени, то есть скорость. Это означает, что DV зависит только от локальных трендов, т.е. *динамической компоненты* полной волатильности $\mathsf{DNOI}(1)$.

В то же время, второе слагаемое в правой части уравнения (1.10), EU , зависит только от весов p_j и дисперсий (диффузионных компонент) σ_j^2 . Это означает, что EU зависит только от локальных диффузионных компонент, и потому является *диффузионной* компонентой полной волатильности $\mathsf{DNOI}(1)$.

СРС-метод является *непараметрическим*, потому что методически аналогичен непараметрическим процедурам ядерного оценивания распределений. Он квазинепараметрический, потому что выбор нормальных ядер здесь форсирован и обусловлен центральной предельной теоремой для пуассоновских случайных сумм. Но как любой непараметрический метод, этот метод плох тем, что хорош только для *ретроспективного* анализа. Для *перспективного* анализа (например, прогнозирования), намного удобнее параметрические модели, к построению которых мы и переходим.

1.2.3 Мультипликативная форма случайных интенсивностей

Предположим, что

$$\Lambda^+ = L \cdot \alpha^+, \quad \Lambda^- = L \cdot \alpha^-,$$

где L – неотрицательная случайная величина, имеющая смысл внешнего новостного фона на бирже, а α^+ и α^- – параметры, описывающие тенденции торгов (пока для простоты изложения будем считать параметры α^+ и α^- неслучайными). Ниже мы представим эмпирические подтверждения этого принципа.

Тогда модель (1.6) примет вид

$$\mathbb{P}(NOI(1) < x) \approx \int_0^\infty \Phi\left(\frac{x - \lambda(\alpha^+ - \alpha^-)}{\sqrt{\lambda(\alpha^+ + \alpha^-)}}\right) d\mathbb{P}(L < \lambda), \quad x \in \mathbb{R}. \quad (1.11)$$

Заметим, что модель (1.11) – это хорошо известная *дисперсионно-сдвиговая смесь* (variance-mean mixture) нормальных законов, в которой смещивание производится и по параметру сдвига, и по параметру масштаба, но фактически смесь является однопараметрической. К такому типу смесей относятся, в частности, обобщенные гиперболические законы, включая дисперсионные гамма-распределения (variance-gamma distributions), скошенные распределения Стьюдента, нормальные\\ обратные гауссовские распределения, некоторые устойчивые законы, а также многие другие. Методы исследования и использования таких моделей хорошо известны. Позже мы обсудим асимптотические аргументы в пользу модели (1.6) более детально.

1.3 Процесс дисбаланса потоков заявок

В данном разделе модель, принципиальное устройство которой описано выше, будет адаптирована с учетом нестационарного и стохастического характера интенсивности внешнего информационного потока Λ и параметров α_+ и α_- , описывающих степень реакции на него покупателей и продавцов.

В действительности интенсивности потоков заявок являются нестационарны-

ми, поскольку внешний поток информации, определяющий интенсивность этих событий, сам по себе является нестационарным. В таком случае интенсивности потоков заявок разных типов, во-первых, не могут являться независимыми, а во-вторых, определенным образом зависят от некоторого случайного процесса $\lambda(t)$, определяющего внешний новостной фон. Для формализации этих идей предположим, что интенсивности потоков заявок меняются во времени:

$$\mu^+ = \mu^+(t), \quad \lambda_i^+ = \lambda_i^+(t), \quad \theta_i^+ = \theta_i^+(t),$$

$$\mu^- = \mu^-(t), \quad \lambda_i^- = \lambda_i^-(t), \quad \theta_i^- = \theta_i^-(t),$$

$i = 1, \dots, M$. Для $t \geq 0$ введём положительные функции

$$\lambda^+(t) = \mu^+(t) + \sum_{i=1}^M \lambda_i^+(t) + \sum_{i=1}^M \theta_i^-(t),$$

$$\lambda^-(t) = \mu^-(t) + \sum_{i=1}^M \lambda_i^-(t) + \sum_{i=1}^M \theta_i^+(t),$$

и введём функции

$$\Lambda^+(t) = \int_0^t \lambda^+(\tau) d\tau \quad \text{and} \quad \Lambda^-(t) = \int_0^t \lambda^-(\tau) d\tau, \quad t \geq 0. \quad (1.12)$$

Предположим, что $\Lambda^+(\infty) = \Lambda^-(\infty) = \infty$.

Пусть $N_1^+(t)$ и $N_1^-(t)$ – независимые пуассоновские процессы с единичной интенсивностью. Обозначим

$$N^+(t) = N_1^+(\Lambda^+(t)), \quad N^-(t) = N_1^-(\Lambda^-(t)).$$

Процессы $N^+(t)$ and $N^-(t)$ являются неоднородными процессами Пуассона с мгновенными интенсивностями $\lambda^+(t)$ и $\lambda^-(t)$ соответственно.

Теперь мы избавимся от предположения о единичном объёме заявок и перей-

дём к более общей модели динамики книги заявок, предположив, что заявки могут иметь произвольные объёмы. Пусть X_1^+, X_2^+, \dots – одинаково распределённые положительные случайные величины и X_1^-, X_2^-, \dots – также одинаково распределённые положительные случайные величины. Предположим, что для каждого t , случайные величины $X_1^+, X_2^+, \dots, X_1^-, X_2^-, \dots, N_1^-(t)$ и $N_1^+(t)$ независимы. Определим процесс

$$OFI(t) = \sum_{j=1}^{N^+(t)} X_j^+ - \sum_{j=1}^{N^-(t)} X_j^-. \quad (1.13)$$

Процесс $OFI(t)$ назовём (условным) *процессом дисбаланса потоков заявок*. Этот процесс является интегральной мгновенной характеристикой потоков заявок (и, как следствие, книги заявок) при наличии идеализированного предположения о том, что внешний поток информации не является случайным.

Некоторые эмпирические исследования (см., например, [27, 93]) показывают, что объёмы лимитных заявок в действительности имеют экспоненциальное распределение, что в рамках модели (1.13) в общем случае распределение независимых случайных величин X_j является асимметричным распределением Лапласа (конечная смесь экспоненциальных распределений и распределение симметрично экспоненциальному распределению, определённому на отрицательной полуоси).

Формально процесс $OFI(t)$, предложенный в (1.13), совпадает с хорошо известным в теории страхований *процессом риска со стохастическими премиями*, положительная сумма описывает поток страховых премий, а отрицательная – поток страховых выплат, см., например, [69]. Однако, как это было отмечено в [69], такие модели вряд ли могут считаться адекватными в страховой практике хотя бы потому, что две компоненты $OFI(t)$ считаются независимыми, хотя в реальной жизни страховые случаи не могут возникнуть прежде, чем в страховую фирму будет внесена соответствующая страховая премия. Тем не менее как было также отмечено

в [69], модель (1.13) может быть успешно применена при описании спекулятивной активности на финансовых рынках. В финансовых приложениях процессы типа (1.13) описывают динамику баланса покупателей и продавцов и результирующие риски. Таким образом, в дальнейшем мы будем использовать терминологию, более подходящую для финансовых приложений и называть процессы вида (1.13) *процессами дисбаланса или двусторонними процессами риска*.

В рассматриваемой ситуации, таким образом, при описании динамики процесса дисбаланса потоков заявок $OFI(t)$ «положительные» и «отрицательные» компоненты процесса (1.13) не могут считаться независимыми. Рассматривая зависимость этих компонент, мы сделаем предположение, которое, с одной стороны, является очевидным с аналитической точки зрения и, с другой стороны, соответствует наблюдаемой реальности. А именно, предположим, что мгновенные интенсивности $\lambda^+(t)$ и $\lambda^-(t)$ зависят от *одного и того же процесса* $\ell(t)$, так что имеют место соотношения

$$\lambda^+(t) = a^+(t)\ell(t), \quad \lambda^-(t) = a^-(t)\ell(t) \quad (1.14)$$

с одной и той же неотрицательной функцией $\ell(t)$ и некоторыми неотрицательными функциями $a^+(t)$ and $a^-(t)$, так что

$$\Lambda^+(t) = \int_0^t a^+(\tau)\ell(\tau)d\tau, \quad \Lambda^-(t) = \int_0^t a^-(\tau)\ell(\tau)d\tau.$$

Обозначим

$$L(t) = \int_0^t \ell(\tau)d\tau.$$

Рассматривая функции $a^+(t)$ and $a^-(t)$, мы предполагаем, что они достаточно регулярны, так что выполняются условия теоремы о среднем и для каждого $t \geq 0$

существуют $\tau^+(t) \in [0, t]$ и $\tau^-(t) \in [0, t]$, такие, что

$$\Lambda^+(t) = \int_0^t a^+(\tau) \ell(\tau) d\tau = a^+(\tau^+(t)) \int_0^t \ell(\tau) d\tau = a^+(\tau^+(t)) L(t),$$

$$\Lambda^-(t) = \int_0^t a^-(\tau) \ell(\tau) d\tau = a^-(\tau^-(t)) \int_0^t \ell(\tau) d\tau = a^-(\tau^-(t)) L(t).$$

Заметим, что числа $\tau^+(t)$ and $\tau^-(t)$ могут быть определены неоднозначно. Обозначим

$$\alpha^+(t) = a^+(\tau^+(t)), \quad \alpha^-(t) = a^-(\tau^-(t)).$$

Таким образом, аналогично случаю с дискретным временем мы в конце концов приходим к мультипликативной форме в непрерывной версии

$$\Lambda^+(t) = \alpha^+(t) L(t), \quad \Lambda^-(t) = \alpha^-(t) L(t), \quad t \geq 0. \quad (1.15)$$

Как мы уже упоминали, ниже будут представлены некоторые эмпирические наблюдения, подтверждающие соотношения (1.14) и (1.15).

1.4 Случайные интенсивности.

Обобщённые процессы Кокса и процесс OFI.

Процесс $L(t)$, описанный выше может быть проинтерпретирован как компонента интенсивности торгов, связанная с ажиотажем, обусловленным внешним информационным фоном. В таком случае функции $\alpha^+(t)$ and $\alpha^-(t)$ описывают степень реакции покупателей и продавцов на этот новостной фон.

Рассматривая процесс $L(t)$, мы сделаем следующее предположение. Во-первых, поскольку он описывает активность, вызванную крайне плохо прогнозируемым потоком неопределённых новостей, довольно было бы правильно считать $L(t)$ слу-

чайным процессом. Во-вторых, мы будем предполагать, что $L(t)$, как случайный процесс, удовлетворяет следующим свойствам: $\Lambda(0) = 0$, $\mathbb{P}(\Lambda(t) < \infty) = 1$ для всех $t > 0$, траектории процесса $\Lambda(t)$ являются неубывающими и непрерывными справа. Случайные процессы с этими свойствами называются *случайными мерами*.

Рассмотрим временной интервал $[0, T]$, достаточно малый для того, чтобы средние значения параметров распределения объёмов заявок считались известными на этом интервале. Это предположение не является серьёзным ограничением с практической точки зрения, поскольку эмпирические исследования демонстрируют наличие значимых эффектов памяти у потоков заявок для малых значений T . В общем случае параметры распределений меняются во времени и представляется возможным построить модели, описывающие эволюцию этих параметров в течение торговой сессии.

Итак, мы зафиксируем временной интервал $[0, T]$ и рассмотрим модель двустороннего процесса риска для процесса дисбаланса потоков заявок в общем виде. Пусть X_1^+, X_2^+, \dots – одинаково распределённые положительные случайные величины, соответствующие объёмами приходящих заявок от покупателей, и X_1^-, X_2^-, \dots – одинаково распределённые положительные случайные величины, соответствующие объёмам заявок, приходящих от продавцов. Пусть $N_1^+(t)$ и $N_1^-(t)$ – два стандартных пуассоновских процесса (каждый с единичной интенсивностью). Пусть $L(t)$ – случайная мера. Предположим, что случайные величины $X_1^+, X_2^+, \dots, X_1^-, X_2^-, \dots$ и случайные процессы $N_1^-(t)$, $N_1^+(t)$ и $L(t)$ независимы. Дополнительно будем считать, что траектории процессов $\Lambda^+(t) = \alpha^+(t)L(t)$ и $\Lambda^-(t) = \alpha^-(t)L(t)$ не убывают. Эти предположения соответствуют тому, что $N^+(t) = N_1^+(\Lambda^+(t))$ и $N^-(t) = N_1^-(\Lambda^-(t))$ являются дважды стохастическими пуассоновскими процессами (процессами Кокса). Процесс

$$OFI(t) = \sum_{j=1}^{N^+(t)} X_j^+ - \sum_{j=1}^{N^-(t)} X_j^- \quad (1.16)$$

будем называть обобщённым (безусловным) *процессом дисбаланса потоков заявок*. Процесс $OFI(t)$, определённый таким образом, соответствует специальному двустороннему процессу риска, чьи положительные и отрицательные компоненты не являются независимыми: они связаны с одним и тем же процессом $L(t)$, описывающим общую компоненту интенсивности событий о приходе заявок от покупателей и продавцов..

Данное выше определение предполагает, что процесс $OFI(t)$ в некотором смысле аккумулирует в себе больше информации, чем процесс *процесс цены* $p(t)$, который может быть получен как прореживание процесса $OFI(t)$ - остаются только те события (а точнее, рыночные заявки), которые приводили к увеличению или уменьшению мидпрайса. В этом смысле процесс цены, часто фигурирующий в академических и прикладных исследованиях, является сужением исходного процесса дисбаланса потоков заявок и может изучаться похожими техниками.

Также легко заметить, что в рамках концепции двухсторонних процессов риска, процесс дисбаланса количества заявок NOI является частным случаем более общего процесса дисбаланса потока заявок OFI.

Процесс дисбаланса потоков заявок является намного более чувствительным индикатором (показателем) текущего состояния книги заявок, поскольку в условиях активной деятельности большого количества высокочастотных торговых систем интервалы времени между последовательными изменениями состояний книги заявок обычно так малы, что изменения цены (мидпрайса) $p(t)$ по сравнению с ними являются редкими событиями. Поэтому процесс цены $p(t)$ является намного более грубым показателем, характеризующим книгу заявок и дающим грубое и весьма ограниченное описание динамики рынка. Вместе с тем процесс дисбаланса потоков заявок учитывает не только текущие значения наилучших цен покупки и продажи, но и влияние событий «в глубине» книги заявок и потому меняется существенно быстрее и позволяет интерполировать динамику рынка между из-

менениями цены, в частности, отслеживать ситуации, связанные с токсичностью потоков заявок, то есть чреватые необоснованными трендами в поведении цены (см. рис. 1.3).

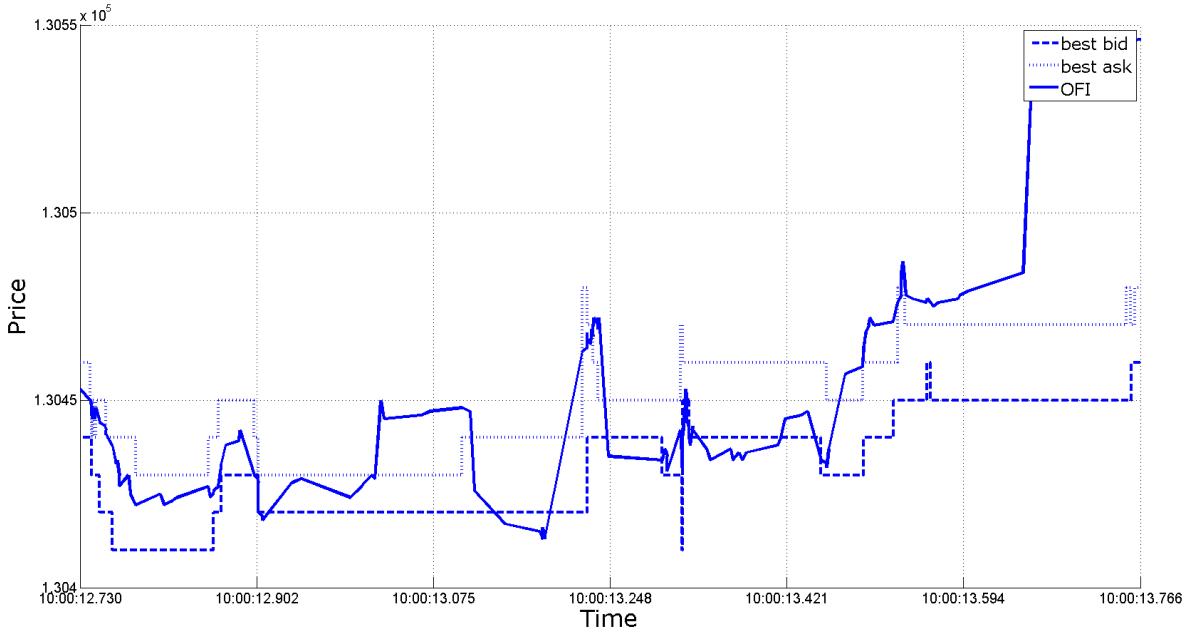


Рис. 1.3: Динамика лучшей цены покупки, лучшей цены продажи и процесса дисбаланса потока заявок $OFI(t)$ в течение 1 секунды с момента 10:00:12,730 01.07.2014 (фьючерс на индекс РТС)

В работе [31] с помощью линейной модели

$$\frac{S(t + \Delta) - S(t)}{\delta} = c \frac{Q(t, t + \Delta)}{D(t)} + \epsilon(t)$$

было показано, что процесс дисбаланса потока заявок $Q(t)$ имеет сильную связь с высокочастотными изменениями цены финансового актива $S(t)$, построенной по ценам сделок, где δ – минимальный шаг цены (тик цены), $\epsilon(t)$ – белый шум и $D(t)$ – мера глубины книги заявок (количество заявок на лучшем биде/аске). Эмпирический анализ высокочастотных данных для американских акций подтверждает наличие линейной связи: коэффициент c оценивается между 0.1 и 1 и оказывается статистически значим в 98% случаев. Наличие такого рода связи позволяет

напрямую исследовать свойства процесса дисбаланса потока заявок и соотносить их со свойствами процесса цены $S(t)$.

Также следует заметить, что предположение о мультипликативности интенсивностей согласуется с другими популярными моделями, использующими хорошо известное свойство кластерности интенсивности прихода заявок. Кроме того, если процесс $\ell(t)$ существенно меняется во времени или осциллирует, то эти осцилляции дают возможность объяснить феномен кластеризации прихода заявок. Мультипликативные конструкции также используются при моделировании интенсивностей потоков заявок с использованием процессов Хоукса [32, 56, 57].

Обозначим характеристические функции случайных величин X_1^+ и X_1^- как $\mathfrak{f}^+(s)$ и $\mathfrak{f}^-(s)$ соответственно, $s \in \mathbb{R}$.

ЛЕММА 1.1. *При выполнении условий для каждого $t \geq 0$ $OFI(t)$ имеет обобщённое смешанное пуссоновское распределение. А именно, для каждого $t \geq 0$*

$$\mathsf{P}(OFI(t) < x) = \mathsf{P}\left(\sum_{j=1}^{N(t)} X_{t,j} < x\right), \quad x \in \mathbb{R},$$

где $N(t) = N_1(\Lambda(t))$, $\Lambda(t) = (\alpha^+(t) + \alpha^-(t))L(t)$, $N_1(t)$ - стандартный пуссоновский процесс, независимый от процесса $L(t)$ и $X_{t,1}, X_{t,2}, \dots$ - одинаково распределённые случайные величины с общей характеристической функцией

$$\mathfrak{f}_t(s) \equiv \mathsf{E}e^{isX_{t,1}} = \frac{\alpha^+(t)\mathfrak{f}^+(s)}{\alpha^+(t) + \alpha^-(t)} + \frac{\alpha^-(t)\mathfrak{f}^-(s)}{\alpha^+(t) + \alpha^-(t)}, \quad s \in \mathbb{R}, \quad (1.17)$$

кроме того, для каждого $t \geq 0$ случайные величины $N_1(t), \Lambda(t), X_{t,1}, X_{t,2}, \dots$ являются независимыми.

ДОКАЗАТЕЛЬСТВО. В силу независимости процессов $N_1^+(t), N_1^-(t), L(t)$

$$\mathsf{P}(OFI(t) < x) = \mathsf{P}\left(\sum_{j=1}^{N^+(t)} X_j^+ - \sum_{j=1}^{N^-(t)} X_j^- < x\right) =$$

$$= \int_0^\infty \mathsf{P} \left(\sum_{j=1}^{N_1^+(\alpha^+(t)\ell)} X_j^+ - \sum_{j=1}^{N_1^-(\alpha^-(t)\ell)} X_j^- < x \right) d\mathsf{P}(L(t) < \ell). \quad (1.18)$$

Рассмотрим характеристическую функцию процесса дисбаланса потока заявок под знаком вероятности в подынтегральном выражении:

$$\begin{aligned} \mathsf{E} \exp \left\{ i s \left(\sum_{j=1}^{N_1^+(\alpha^+(t)\ell)} X_j^+ - \sum_{j=1}^{N_1^-(\alpha^-(t)\ell)} X_j^- \right) \right\} &= \\ &= \exp \left\{ \alpha^+(t)\ell(\mathfrak{f}_t^+(s) - 1) \right\} \cdot \exp \left\{ \alpha^-(t)\ell(\mathfrak{f}_t^-(s) - 1) \right\} = \\ &= \exp \left\{ (\alpha^+(t) + \alpha^-(t))\ell \left(\frac{\alpha^+(t)\mathfrak{f}_t^+(s)}{\alpha^+(t) + \alpha^-(t)} + \frac{\alpha^-(t)\mathfrak{f}_t^-(s)}{\alpha^+(t) + \alpha^-(t)} - 1 \right) \right\}, \end{aligned}$$

то есть для каждого фиксированного $\ell > 0$

$$\sum_{j=1}^{N_1^+(\alpha^+(t)\ell)} X_j^+ - \sum_{j=1}^{N_1^-(\alpha^-(t)\ell)} X_j^- \stackrel{d}{=} \sum_{j=1}^{N_1((\alpha^+(t)+\alpha^-(t))\ell)} X_{t,j},$$

где $X_{t,1}, X_{t,2}, \dots$ - независимые случайные величины с общей характеристической функцией (1.17), независимые от стандартного пуассоновского процесса $N_1(t)$. Следовательно, используя (1.18), для каждого $x \in \mathbb{R}$ получаем

$$\begin{aligned} \mathsf{P}(OFI(t) < x) &= \int_0^\infty \mathsf{P} \left(\sum_{j=1}^{N_1((\alpha^+(t)+\alpha^-(t))\ell)} X_{t,j} < x \right) d\mathsf{P}(L(t) < \ell) = \\ &= \mathsf{P} \left(\sum_{j=1}^{N(t)} X_{t,j} < x \right). \end{aligned}$$

Лемма доказана.

Если бы функции $\alpha^+(t)$ и $\alpha^-(t)$ были постоянными, что может быть достигнуто, когда интервал $[0, t]$ достаточно мал, тогда распределения $X_{t,j}$ не зависят от t и в соответствии с Леммой 1.1 процесс дисбаланса потока заявок $OFI(t)$ может быть рассмотрен как обычный обобщённый процесс Кокса.

Легко видеть, что случайная величина $X_{t,1}$ – это рандомизация:

$$X_{t,1} = \begin{cases} X_1^+ & \text{с вероятностью } \frac{r(t)}{r(t) + 1}, \\ -X_1^- & \text{с вероятностью } \frac{1}{r(t) + 1} \end{cases}$$

где $r(t) = \alpha^+(t)/\alpha^-(t)$, так что

$$\mathbb{E}X_{t,1} = \frac{r(t)\mathbb{E}X_1^+}{r(t) + 1} - \frac{\mathbb{E}X_1^-}{r(t) + 1},$$

$$\mathbb{D}X_{t,1} = \frac{r(t)\mathbb{E}(X_1^+)^2 + \mathbb{E}(X_1^-)^2}{r(t) + 1} - \frac{(r(t)\mathbb{E}X_1^+ + \mathbb{E}X_1^-)^2}{(r(t) + 1)^2}.$$

Можно заключить, что величина $r(t)$ в значительной степени определяет мгновенное поведение процесса OFI. В то же время, из предположений о мультипликативности для достаточно малых значений t следует

$$r(t) = \frac{\alpha^+(t)}{\alpha^-(t)} \approx \frac{\lambda^+(t)}{\lambda^-(t)}.$$

На практике это означает, что для оценки ненаблюдаемого процесса $r(t)$, характеризующего мгновенный дисбаланс между силами покупателей и продавцов нам достаточно оценивать отношение количества заявок, пришедших от покупателей и продавцов в единицу времени. Определим процесс $r(t)$ как *процесс дисбаланса интенсивностей потоков заявок*.

1.5 Асимптотические аппроксимации распределения процесса дисбаланса потоков заявок

Для того, чтобы рассмотреть асимптотические конструкции, которые позволяют формализовать сценарий бесконечного возрастания интенсивностей потоков заявок и рассмотреть асимптотические ("heavy-traffic") аппроксимации одномерных распределений процесса OFI, зафиксируем момент времени t и рассмотрим вспомогательный параметр n . Всюду в дальнейшем предполагается сходимость при $n \rightarrow \infty$, если не указано иное. Таким образом, на основе Леммы 1.1 рассмотрим последовательность обобщённых процессов Кокса в виде

$$Q_n(t) = \sum_{i=1}^{N_1^{(n)}(\Lambda_n(t))} X_{n,i}, \quad t \geq 0, \quad (1.19)$$

где $\{N_1^{(n)}(t), t \geq 0\}_{n \geq 1}$ – это последовательность пуассоновских процессов с единичными интенсивностями; для каждого $n = 1, 2, \dots$ случайные величины $X_{n,1}, X_{n,2}, \dots$ одинаково распределены; для каждого $n \geq 1$ случайные величины $X_{n,1}, X_{n,2}, \dots$ и процессы $N_1^{(n)}(t), t \geq 0$, являются независимыми; для каждого $n = 1, 2, \dots$ процессы $\Lambda_n(t), t \geq 0$, являются управляемыми, то есть неубывающими положительными процессами Леви, независимыми от процесса

$$Z_n(t) = \sum_{i=1}^{N_1^{(n)}(t)} X_{n,i}, \quad t \geq 0, \quad (1.20)$$

и такими, что $\Lambda_n(0) = 0$.

Всюду в дальнейшем для определённости мы полагаем $\sum_{i=1}^0 = 0$. В терминах предыдущих разделов $Z_n(t)$ является условным однородным процессом OFI.

Ниже мы покажем, что рассуждения, приведённые в [68] и [72] для симмет-

ричного случая могут быть с незначительными изменениями обобщены на случай предельных процессов с несимметричными функциями распределения. Более того, с помощью этого подхода мы получим асимптотические остроконечные распределения с тяжёлыми хвостами даже в случае, когда дисперсии размеров заявок ограничены.

Поскольку везде, где это предполагается, параметр t остаётся зафиксированным, для простоты мы будем писать Q_n , N_n , Λ_n и $X_{n,j}$ вместо $Q_n(t)$, $N_n(t)$, $\Lambda_n(t)$ и $X_{t,j}^{(n)}$ соответственно.

Нам потребуется несколько вспомогательных утверждений из теории случайного суммирования.

Пусть $\{\xi_{n,j}\}_{j \geq 1}$, $n = 1, 2, \dots$ – двумерный массив построчно одинаково распределённых случайных величин. Пусть $\{N_n\}_{n \geq 1}$ – последовательность целочисленных неотрицательных случайных величин, таких, что для каждого $n \geq 1$ случайные величины $N_n, \xi_{n,1}, \xi_{n,2}, \dots$ являются независимыми. Обозначим

$$S_{n,k} = \xi_{n,1} + \dots + \xi_{n,k}.$$

Как было показано в [68, 72], сдвиг-масштабные смеси нормальных законов являются идентифицируемыми, поскольку при каждом фиксированном $\mu \in \mathbb{R}$ and $\sigma > 0$ однопараметрическое семейство распределений $\{\Phi((x - \mu z)/\sigma\sqrt{z}) : z \geq 0\}$ является аддитивно-замкнутым. В [68] было доказано следующее общее утверждение (см. также [72]).

ЛЕММА 1.2. *Предположим, что существует последовательность $\{k_n\}_{n \geq 1}$ натуральных чисел и конечные числа $\mu \in \mathbb{R}$ и $\sigma > 0$ такие, что*

$$\mathbb{P}(S_{n,k_n} < x) \implies \Phi\left(\frac{x - \mu}{\sigma}\right). \quad (1.21)$$

Предположим, что $N_n \rightarrow \infty$ по вероятности. Тогда распределение случайных сумм слабо сходится к некоторой функции $F(x)$:

$$\mathsf{P}(S_{n,N_n} < x) \implies F(x),$$

тогда и только тогда, когда существует функция распределения $A(x)$ такая, что $A(0) = 0$,

$$F(x) = \int_0^\infty \Phi\left(\frac{x - \mu z}{\sigma\sqrt{z}}\right) dA(z), \quad (1.22)$$

и

$$\mathsf{P}(N_n < xk_n) \implies A(x). \quad (1.23)$$

Условие 1.21 выполнено в следующем достаточно общем случае. Предположим, что случайные величины $\xi_{n,j}$ обладают конечными дисперсиями. Также предположим, что $\xi_{n,j}$ представимы в виде

$$\xi_{n,j} = \xi_{n,j}^* + \mu_n,$$

где $\mu_n \in \mathbb{R}$ и $\xi_{n,j}^*$ является случайной величиной, для которой выполняется $\mathsf{E}\xi_{n,j}^* = 0$, $\mathsf{D}\xi_{n,j}^* = \sigma_n^2 < \infty$, так что $\mathsf{E}\xi_{n,1} = \mu_n$ и $\mathsf{D}\xi_{n,1} = \sigma_n^2$. Предположим, что $\mu_n k_n \rightarrow \mu$ и $k_n \sigma_n^2 \rightarrow \sigma^2$ as $n \rightarrow \infty$. Тогда в силу хорошо известного результата о необходимых и достаточных условиях нормальной сходимости распределений сумм независимых случайных величин с конечными дисперсиями в предельной схеме двумерного массива (см., например, [50]), мы можем видеть, что (6) выполнено тогда и только тогда, когда выполнено условие Линдеберга: для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} k_n \mathsf{E}(\xi_{n,1}^*)^2 \mathbb{I}(|\xi_{n,1}^*| \geq \varepsilon) = 0,$$

(здесь $\mathbb{I}(B)$ – индикатор множества B).

ЛЕММА 1.3. Пусть $N_n = N_1^{(n)}(\Lambda_n)$, $n \geq 1$, где $\{N_1^{(n)}(t), t \geq 0\}$, $n = 1, 2, \dots$ являются стандартными пуассоновскими процессами и Λ_n , $n = 1, 2, \dots$ – положительные случайные величины такие, что для каждого $n \geq 1$ случайная величина Λ_n является независимой от процесса $N_1^{(n)}(t)$. Тогда

$$\mathbb{P}(N_n < k_n x) \implies A(x)$$

для некоторой бесконечно возрастающей последовательности k_n вещественных чисел и некоторой функции распределения $A(x)$ тогда и только тогда, когда

$$\mathbb{P}(\Lambda_n < k_n x) \implies A(x).$$

Доказательство см. в [51].

Из Лемм 1.2 и 1.3 следует

ТЕОРЕМА 1.1. Предположим, что существует бесконечно возрастающая последовательность $\{k_n\}_{n \geq 1}$ натуральных чисел и конечные числа $\mu \in \mathbb{R}$ и $\sigma > 0$ такие, что случайные обобщены заявок $X_{n,j}$ удовлетворяют условию

$$\mathbb{P}(X_{n,1} + \dots + X_{n,k_n} < x) \implies \Phi\left(\frac{x - \mu}{\sigma}\right).$$

Сходимость

$$\mathbb{P}(OFI_n < x) \implies F(x) \tag{1.24}$$

с некоторой функцией распределения $F(x)$ имеет место тогда и только тогда, когда существует функция распределения $A(x)$ такая, что $A(0) = 0$, функция

распределения $F(x)$ представима в виде (1.21) и

$$\mathbb{P}(\Lambda_n < xk_n) \implies A(x). \quad (1.25)$$

1.6 Анализ реальных данных

1.6.1 Описание данных

Рассматриваются данные о потоках всех заявок (лимитных, рыночных и заявок на отмену) на первые $d = 5$ уровней книги заявок фьючерса на индекс РТС. Эти данные дают доступ к самой детальной информации о рыночных торгах, в отличие от данных о сделках и котировках (TAQ, Trades and Quotes), которые часто используются для анализа высокочастотных данных и состоящих из цен и объёмов сделок (что соответствует только рыночным заявкам в потоке всех заявок), а также информации о цене и объёме лучших котировок на покупку и продажу (т. е. только первый уровень книги заявок) с проставленными моментами времени. Мы не рассматриваем первые и последние 5 минут торгов (с 10:00 до 10:05 и с 18:40 до 18:45), поскольку эти периоды характеризуются аномальными всплесками волатильности и слабо описываются в рамках предложенной модели.

В Таблице 1.1 приведён пример данных о потоке заявок для фьючерса на индекс РТС - самого ликвидного инструмента фьючерсного рынка Московской Биржи - и о том, как выглядел срез книги заявок на первые пять уровней после прихода соответствующей заявки. Заметим, что на рынке FORTS присутствует всего два типа заявок: лимитные (L) и заявки на отмену (C), а механизм рыночных заявок участники рынка реализуют самостоятельно (отправляя лимитные заявки с ценами, гарантирующими их моментальное исполнение). Тем не менее имеется

Таблица 1.1: Описание формата данных о потоках всех заявок (фьючерс на индекс РТС).

Time	T	Dir	Price	Vol	b_1	a_1	v_1^b	v_1^a	v_2^b	v_2^a	v_3^b	v_3^a	v_4^b	v_4^a	v_5^b	v_5^a
10:02:36.444	L	B	130020	2	130040	130050	2	4	22	23	54	22	81	31	759	20
10:02:36.445	L	S	130070	1	130040	130050	2	4	22	23	55	22	81	31	759	20
10:02:36.465	C	B	130040	1	130040	130050	1	4	22	23	55	22	81	31	759	20
10:02:36.473	L	B	130050	3	130040	130050	1	1	22	23	55	22	81	31	759	20

возможность оценить параметры потоков рыночных заявок в рамках предлагаемой модели, рассматривая для этого потоки лимитных заявок, которые приводили к сделкам.

Для тестирования некоторых из вышеизложенных концепций были выбраны высокочастотные данные для самого ликвидного инструмента фьючерсного рынка биржи ММВБ-РТС – фьючерса на индекс РТС. Биржа распространяет информацию о полном потоке обезличенных заявок участников рынка, что, в частности, позволяет провести анализ процессов $N^+(t)$ и $N^-(t)$ в рамках описанной модели процесса дисбаланса потока заявок.

Для анализа были выбраны данные за первые три часа торгов данным инструментом за дневную сессию 11 ноября 2012 года, график цены, построенный по сделкам, изображен на Рис. 1.4.

1.6.2 Оценка параметров

В таблицах ниже приведены оценки на интенсивности (первая таблица) и средние объемы заявок (вторая таблица) за 20-секундные интервалы в различные моменты времени в течение дня.



Рис. 1.4: График цены.

Таблица 1.2: Оценки параметров, связанных с интенсивностями потоков заявок.

Время	Q	n_μ	n_{λ_1}	n_{λ_2}	n_{λ_3}	n_{λ_4}	n_{λ_5}	n_{θ_1}	n_{θ_2}	n_{θ_3}	n_{θ_4}	n_{θ_5}
10:10:00.049	9	20	15	32	18	29	8	27	12	26	10	9
10:30:00.061	10	9	2	12	16	5	4	3	12	9	5	2
13:00:00.017	19	20	15	32	18	29	8	27	12	26	10	9
16:00:00.020	12	20	38	113	46	11	14	17	60	35	18	15
18:30:00.021	6	31	63	76	25	18	33	57	44	17	12	14

При работе с реальными данными имеется возможность наблюдать лишь усреднённые интенсивности

$$\lambda_*^+(t, h) = \frac{1}{h} \int_{t-h}^t \lambda^+(\tau) d\tau \quad \text{и} \quad \lambda_*^-(t, h) = \frac{1}{h} \int_{t-h}^t \lambda^-(\tau) d\tau$$

для различных h . На Рис. 1.5 изображены графики мгновенных интенсивностей $\lambda_*^+(t, h)$ и $\lambda_*^-(t, h)$, вычисленных как количество заявок соответствующего типа за последнюю минуту ($h = 1$ минута). Для лучшей визуализации вместо $\lambda_*^-(t, h)$ на этом рисунке отображён $-\lambda_*^-(t, h)$. Можно видеть, что её график является почти

Таблица 1.3: Оценки параметров, связанных с объёмами заявок.

Время	V^μ	V^{λ_1}	V^{λ_2}	V^{λ_3}	V^{λ_4}	V^{λ_5}	V^{θ_1}	V^{θ_2}	V^{θ_3}	V^{θ_4}	V^{θ_5}
10:10:00.049	4.66	4.24	2.86	2.81	3.88	6.22	3.00	2.75	3.31	4.46	5.27
10:30:00.061	4.43	3.93	2.71	3.17	3.68	4.15	3.08	3.36	3.78	4.16	3.23
13:00:00.017	4.08	4.02	2.90	2.39	3.05	3.48	2.81	2.56	2.46	4.04	3.78
16:00:00.020	3.02	2.72	2.08	2.28	2.57	2.14	2.49	2.01	2.05	2.11	2.12
18:30:00.021	3.73	3.71	2.48	2.67	3.17	5.75	2.80	3.23	2.76	3.52	4.90

зеркальным отражением графика $\lambda_*^+(t, h)$ (т.е. эти графики являются почти симметричными относительно оси x). Этот график подтверждает мультипликативное представление интенсивностей (1.14) с $\ell(t)$, представляющей из себя реализованную траекторию мгновенной интенсивности внешнего информационного потока. В таком контексте процесс $\ell(t)$ может быть представлен как фактор *общего ажио-тажса* на рынке, влияющего на интенсивности как покупателей, так и продавцов.

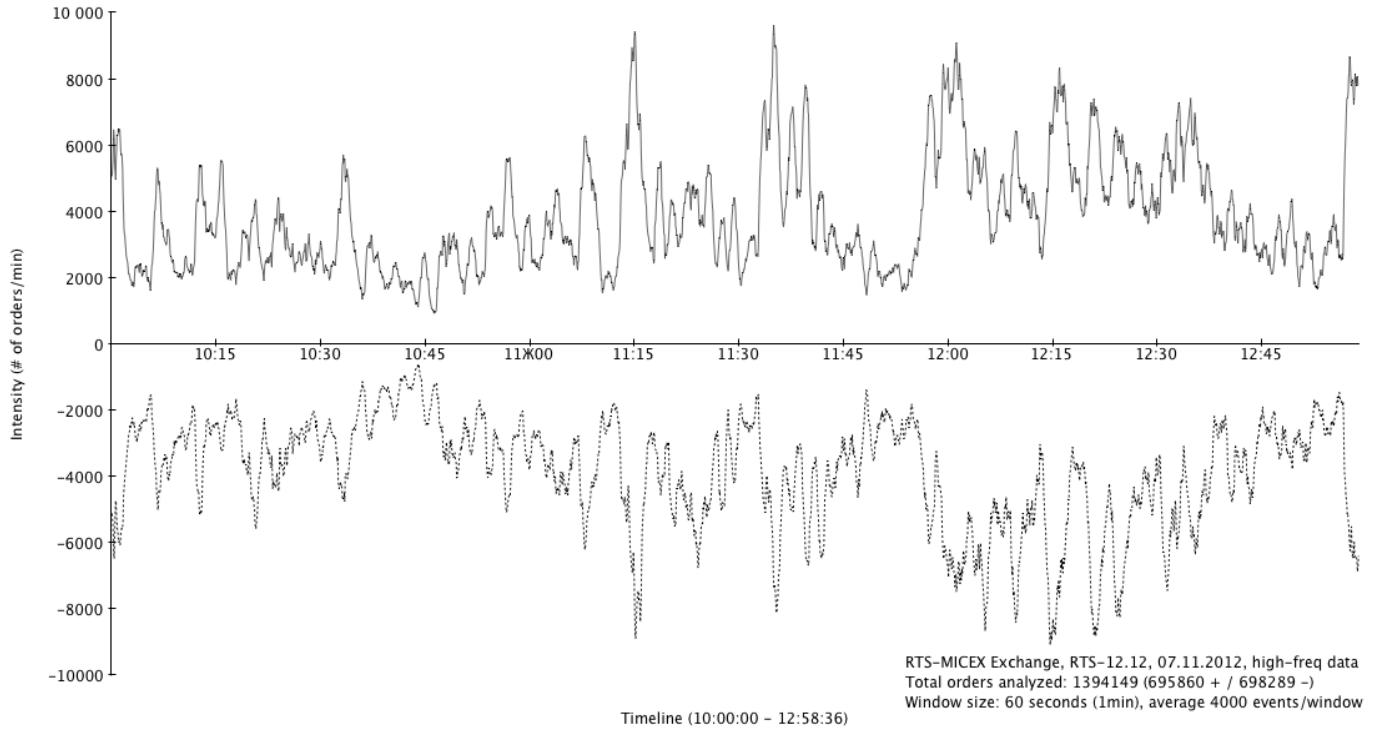


Рис. 1.5: Графики мгновенных интенсивностей потоков заявок покупателей $\lambda_*^+(t)$ (сплошная линия) и продавцов $\lambda_*^-(t)$ (пунктирная линия), размер скользящего окна $w = 60$ сек.

На Рис. 1.6 изображены графики тех же процессов, но совмещенные в положительной полуплоскости. Наблюдаемые расхождения графиков означают локальное

преобладание покупателей над продавцами ($\lambda_*^+(t) > \lambda_*^-(t)$) или продавцов над покупателями ($\lambda_*^-(t) > \lambda_*^+(t)$), согласованные же их падение или рост соответствуют общему падению или росту интенсивности торгов без особенной борьбы между покупателями и продавцами, и, как следствие, изменений цены.

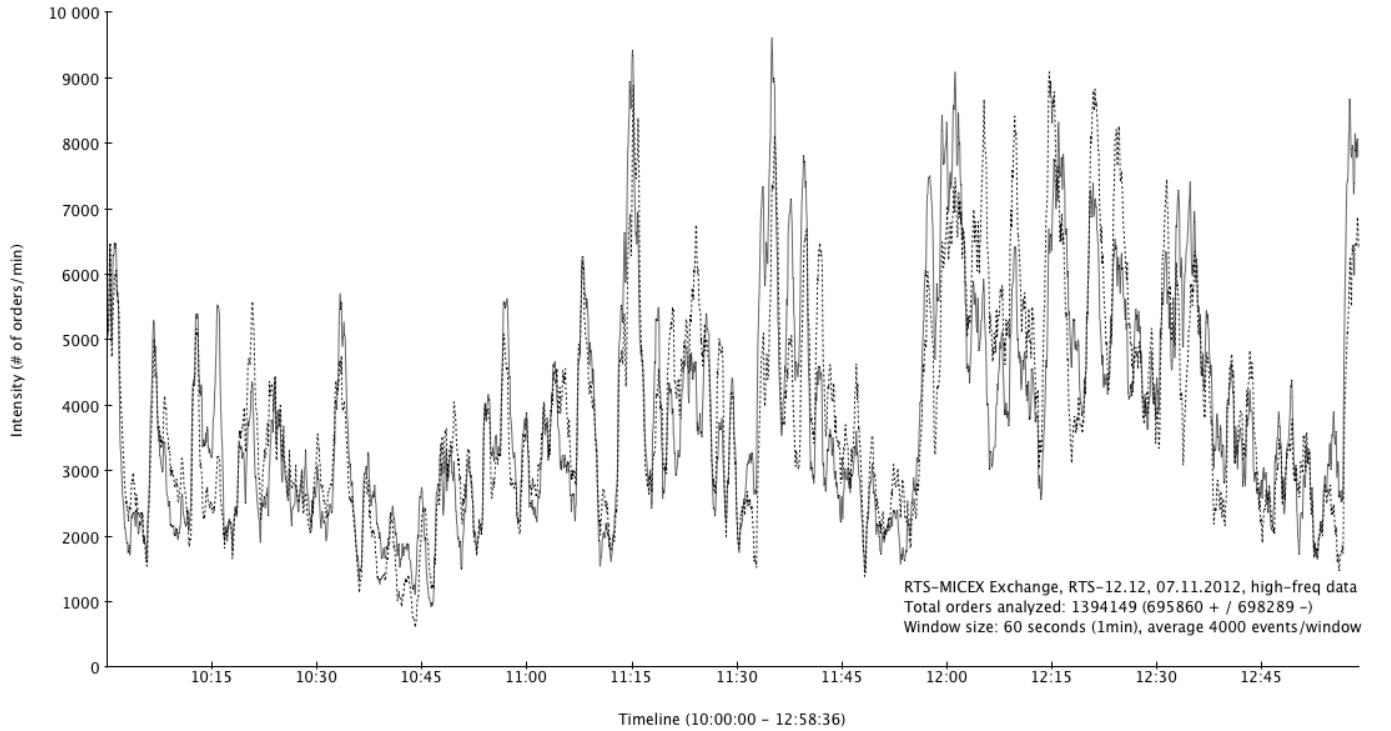


Рис. 1.6: Графики мгновенных интенсивностей потоков заявок покупателей $\lambda_*^+(t)$ (сплошная линия) и продавцов $\lambda_*^-(t)$ (пунктирная линия), размер скользящего окна $w = 60$ сек.

1.6.3 Свойства потоков заявок

Прежде всего заметим, что эмпирические статистические свойства интенсивностей $\lambda_*^+(t, h)$ и $\lambda_*^-(t, h)$ существенно различаются для разных значений h , поскольку на разных временных горизонтах рыночная микроструктура подвержена влиянию разных типов участников рынка (высокочастотные трейдеры для $h < 60$ сек, внутридневные трейдеры для горизонта минут и часов и классические инвесторы для больших значений h).

Таблица 1.4: Оценки и статистики, связанные с $\alpha_*(t, h)$

stat	$h = 1$ сек	$h = 2$ сек	$h = 4$ сек	$h = 8$ сек	$h = 16$ сек	$h = 32$ сек	$h = 64$ сек
$\mathbb{E}(\alpha)$	2.262607	3.003977	3.784542	4.548348	5.293362	6.026013	6.754775
$\mathbb{D}(\alpha)$	0.989517	0.932258	0.832570	0.744913	0.671429	0.611302	0.555113
KS	0.000000	0.214000	0.584200	0.694200	0.388300	0.434700	0.685700
SW				0.028320	0.004415	0.010580	0.010950

Свойство мультипликативности

Рассмотрим процесс дисбаланса с усреднёнными интенсивностями

$$r_*(t, h) = \frac{1}{h} \int_{t-h}^t r_*(\tau) d\tau$$

и процессы

$$\alpha_*(t, h) = \ln \sqrt{\lambda_*^+(t, h) * \lambda_*^-(t, h)},$$

$$\beta_*(t, h) = \ln \sqrt{\frac{\lambda_*^+(t, h)}{\lambda_*^-(t, h)}} = \frac{1}{2} \ln r_*(t, h).$$

Мы используем тесты Колмогорова-Смирнова (KS) и Шапиро-Вилкас (SW) для тестирования соответствия этих процессов нормальному распределению.

В Таблице 1.4 представлены результаты расчёта различных характеристик, связанных с эмпирическим наблюдениями $\alpha_*(t, h)$ при разных значениях h . В первой и второй строках таблицы представлены оценки на матожидание и дисперсию $\alpha_*(t, h)$. В следующих строках представлены значения p -статистик с 5%-уровнем значимости для тестов на соответствие нормальному распределению – тест Колмогорова-Смирнова (KS, третья строка) и Шапиро-Уилка (SW, четвёртая строка).

Аналогичные расчёты для $\beta_*(t, h)$ представлены в Таблице 1.5.

Таблица 1.5: Оценки и статистики, связанные с $\beta_*(t, h)$

stat	$h = 1$ сек	$h = 2$ сек	$h = 4$ сек	$h = 8$ сек	$h = 16$ сек	$h = 32$ сек	$h = 64$ сек
$\mathbb{E}(\beta)$	-0.001902	-0.001567	-0.001858	-0.001499	-0.001156	-0.000612	-0.001030
$\mathbb{D}(\beta)$	0.314539	0.238152	0.168382	0.120034	0.086369	0.062167	0.043387
KS	0.000000	0.000000	0.000144	0.290600	0.500900	0.782200	0.670900
SW				0.000000	0.000019	0.071840	0.148700

Таким образом, установлена лог-нормальность $r_*^+(t, h)$ для $h > 32$ сек. Также обнаружено, что $r_*(t, h)$ не обладает значимым автокорреляционным свойством для любых h , тогда как $\alpha_*(t, h)$ обладает значимой автокорреляцией (≥ 0.1) и имеет нормальное распределение для размера скользящего окна $h > 8$ сек. Эти эмпирические свойства также подтверждают адекватность мультипликативной формы представления интенсивностей заявок, предложенной в данном разделе.

На Рис. 1.7 изображен график процесса $r(t)$, из которого хорошо видно преобладение покупателей над продавцами, что отразилось в поведении цены на протяжении наблюдаемого периода. Другой особенностью графика является наличие уровня поддержки $r = 0.6$, что может означать наличие крупного покупателя, который сдерживал натиск продавцов при достижении данного уровня дисбаланса сил (это также согласуется с поведением цены в эти моменты времени).

Корреляции интенсивностей

Наблюдаются существенные корреляции между значениями интенсивностей автоматных потоков заявок (см Рис. 1.8), что ещё раз подтверждает идею зависимости каждого из потоков от некоторого общего фактора.

Распределение количества заявок

Разобьём торговую сессию на непересекающиеся 15-секундные интервалы времени. На Рис. 1.9 изображены гистограммы количества заявок от покупателей (слева) и продавцов (справа) за 15-секундные интервалы времени. Легко видеть,

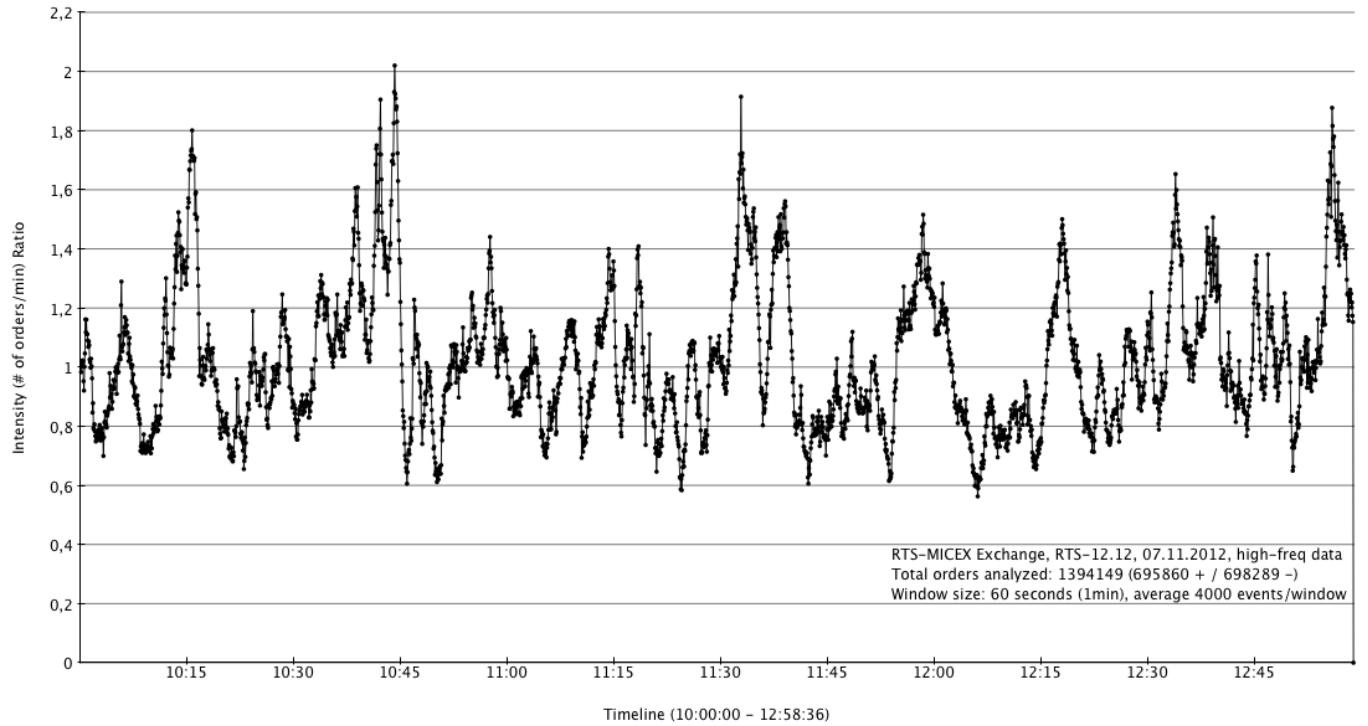


Рис. 1.7: Отношение мгновенных интенсивностей покупателей и продавцов $r(t) = \frac{\lambda^+(t)}{\lambda^-(t)}$, размер скользящего окна $h = 60$ сек.

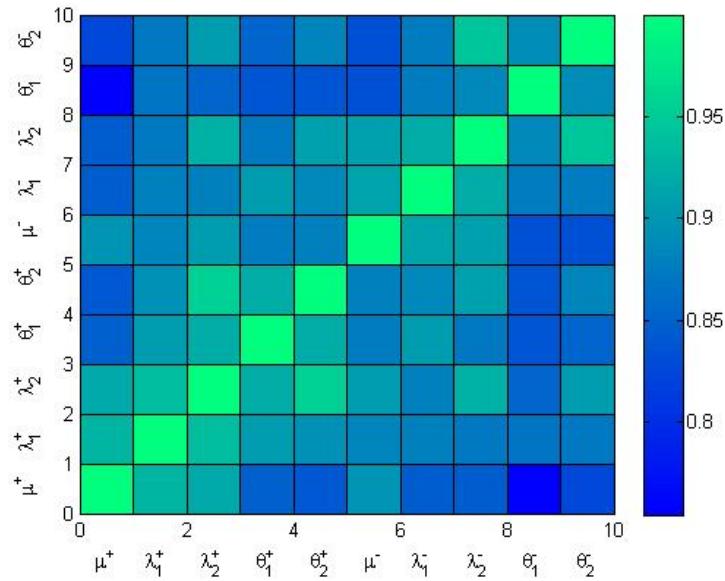


Рис. 1.8: Корреляции интенсивностей.

что эти гистограммы неплохо аппроксимируются графиком плотности гамма-распределения с соответствующими параметрами. Это является хорошим свиде-

тельством в пользу выбора процессов Кокса для описания соответствующих считающих процессов. Обоснование такого хорошего соответствия даётся в следующей главе о предельных функциональных теоремах для процесса OFI: если ожидаемая интенсивность потоков заявок велика, тогда асимптотическое распределение накопленной интенсивности совпадает со смешанным пуассоновским.

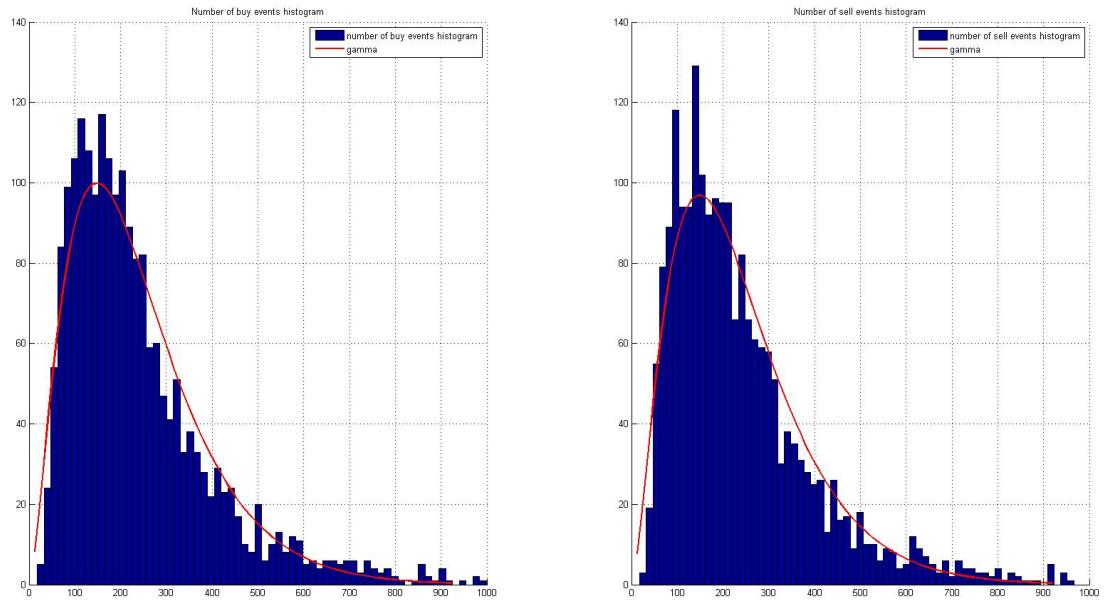


Рис. 1.9: Гистограмма количества заявок на покупку (слева) и продажу (справа) в течение 15-секундных интервалов, фьючерс на индекс РТС, дневная сессия 2014.07.01).

Глава 2

Функциональные предельные теоремы для процесса дисбаланса потоков заявок

В работе [74] были доказаны некоторые предельные функциональные теоремы, устанавливающие сходимость случайных блужданий, порождённых обобщёнными процессами Кокса со скачками, обладающими конечными дисперсиями, к процессам Леви с симметричными распределениями, включая симметричные строго устойчивые процессы Леви. В данной главе мы расширим эти результаты на несимметричный случай и применим их для описания динамики процесса дисбаланса потока заявок, являющегося интегральной характеристикой динамики книги заявок.

Функциональные предельные теоремы устанавливают вполне естественную связь между случайными блужданиями и подчинёнными винеровскими процессами. Операция подчинения даёт хорошее объяснение наличию тяжёлых хвостов распределений приращений (логарифмов) цен акций и финансовых индексов.

сов. Функциональные предельные теоремы для обобщённых процессов Кокса доказаны в [74] и устанавливают связь между формальными микроструктурными моделями, имеющими форму непрерывных случайных блужданий, порождённых обобщёнными процессами Кокса и популярным макромоделями подчинённых винеровских процессов, включающих в себя обобщённые гиперболические процессы, variance-gamma процессы и т.д. Практическая значимость таких моделей обусловлена случайной природой интенсивностей хаотических потоков информативных событий в больших информационных финансовых системах и, в частности, при описании деятельности высокочастотных торговых систем. Использование высокочастотных финансовых данных, доступных благодаря электронным системам торговли, даёт возможным верифицировать модели, упомянутые выше и связать их с процессом образования цены, являющейся результатом эволюции книги заявок.

В данной главе представлено дальнейшее развитие моделей и техник, предложенных в работах [77], [76], [28]. Предлагается удобная и достаточно реалистичная модель для процесса дисбаланса потока заявок на основе двухсторонних процессов риска. Для этой задачи мы используем подходы, предложенные в [74].

Таким образом, материал данной главы посвящён функциональным предельным теоремам для процесса дисбаланса потоков заявок OFI. Раздел 2.1 содержит некоторый предварительный материал о пространстве Скорохода и процессах Леви. В Разделе 2.2 доказывается общая функциональная теорема, устанавливающая условия сходимости процессов OFI к процессам Леви в пространстве Скорохода в контексте роста интенсивностей потока заявок. Для этих целей мы немного расширяем классические результаты, представленные, например, в [59]. В Разделе 2.3 мы рассматриваем условия сходимости процессов OFI с элементарными скачками (т.е., размерами заявок), обладающими конечными дисперсиями, к процессам Леви с сдвиг-масштабными смесями нормальных одномерных распределений, то

есть, к подчинённым винеровским процессам, в частности, к обобщённым гиперболическим процессам Леви.

Условия, приведённые выше, дают возможность использовать хорошо развитый аналитический аппарат обобщённых процессов Кокса для изучения асимптотического поведения процесса OFI при росте интенсивности потоков заявок. Это даёт возможность описать возможные асимптотические аппроксимации процесса дисбаланса потока заявок OFI. Для этого нам потребуются некоторые вспомогательные определения и результаты.

2.1 Пространство Скорохода. Процессы Леви.

Пусть $D = D[0, 1]$ – пространство вещественных функций, определённых на интервале $[0, 1]$, непрерывных справа и имеющих конечный предел слева в каждой точке.

Пусть \mathcal{F} - класс строго возрастающих отображений отрезка $[0, 1]$ на себя. Пусть f - неубывающая функция, определённая на отрезке $[0, 1]$ и $f(0) = 0, f(1) = 1$.

Определим

$$\|f\| = \sup_{s \neq t} |\log [(f(t) - f(s))/(t - s)]|$$

Если $\|f\| < \infty$, тогда функция f является непрерывной и строго возрастающей и, таким образом, принадлежит \mathcal{F} .

Определим расстояние $d_0(x, y)$ на множестве $D[0, 1]$ как наибольшую нижнюю границу множества положительных чисел ϵ , для которых класс \mathcal{F} содержит функцию f , такую что $\|f\| \leq \epsilon$ и $\sup_t |x(t) - y(f(t))| \leq \epsilon$.

Пространство $D[0, 1]$ является полным относительно расстояния d_0 . Метрическое пространство $\mathcal{D} = (D[0, 1], d_0)$ называется *пространством Скорохода*. Мы будем рассматривать стохастические процессы как \mathcal{D} -измеримые случайные вели-

чины.

Пусть X, X_1, X_2, \dots являются \mathcal{D} -измеримыми случайными величинами. Пусть T_X - подмножество отрезка $[0, 1]$, такое что $0 \in T_X$, $1 \in T_X$ и если $0 < t < 1$, то $t \in T_X$ тогда и только тогда, когда $\mathsf{P}(X(t) \neq X(t-)) = 0$. Нижеследующая теорема устанавливает достаточные условия для слабой сходимости случайного процесса в \mathcal{D} (обозначаемой далее как \implies и рассматриваемой в пределе при $n \rightarrow \infty$).

ТЕОРЕМА 2.1. *Пусть $(X_n(t_1), \dots, X_n(t_k)) \implies (X(t_1), \dots, X(t_k))$ для любого натурального k и t_1, \dots, t_k принадлежат T_X . Пусть $\mathsf{P}(X(1) \neq X(1-)) = 0$ и пусть существует неубывающая непрерывная функция F , определённая на отрезке $[0, 1]$, такая, что для любого $\epsilon > 0$*

$$\mathsf{P}(|X_n(t) - X_n(t_1)| \geq \epsilon, |X_n(t_2) - X_n(t)| \geq \epsilon) \leq \epsilon^{-2\nu} [F(t_2) - F(t_1)]^{2\gamma} \quad (2.1)$$

для $t_1 \leq t \leq t_2$ and $n \geq 1$, где $\nu \geq 0$, $\gamma > 1/2$. Тогда $X_n \implies X$.

ДОКАЗАТЕЛЬСТВО Теоремы 2.1 приводится, например, в [18].

Всюду в дальнейшем символ $\stackrel{d}{=}$ будет обозначать совпадение распределений.

Процессом Леви мы будем называть случайный процесс $X(t)$, $t \geq 0$, обладающий следующими свойствами:

- $X(0) = 0$ почти всюду;
- $X(t)$ - процесс с независимыми приращениями, то есть для любых $N \geq 1$ и t_0, t_1, \dots, t_N ($0 \leq t_0 \leq t_1 \leq \dots \leq t_N$) случайные величины $X(t_0), X(t_1) - X(t_0), \dots, X(t_N) - X(t_{N-1})$ независимы в совокупности;
- $X(t)$ - однородный процесс, то есть $X(t+h) - X(t) \stackrel{d}{=} X(s+h) - X(s)$ для любых $s, t, h > 0$;

- процесс $X(t)$ является стохастически непрерывным, то есть для любых $t \geq 0$ и $\epsilon > 0$

$$\lim_{s \rightarrow t} \mathsf{P}(|X(t) - X(s)| > \epsilon) = 0;$$

- траектории процесса $X(t)$ являются непрерывными справа и имеют конечные пределы слева в каждой точке.

Пусть $\psi_t(s)$ - характеристическая функция случайной величины $X(t)$ ($\psi_t(s) = \mathsf{E} e^{isX(t)}$, $s \in \mathbb{R}$). Следующее утверждение описывает хорошо известное свойство процессов Леви.

ЛЕММА 2.1. *Пусть $X = X(t)$, $t \geq 0$ - случайный процесс Леви. Для любого $t > 0$ характеристическая функция случайной величины $X(t)$ является безгранично делимой и представима в виде*

$$\psi_t(s) = [\psi_1(s)]^t = \left[\mathsf{E} e^{isX(1)} \right]^t, \quad s \in \mathbb{R}. \quad (2.2)$$

Другими словами, пусть Y - произвольная безгранично делимая случайная величина, тогда семейство безгранично делимых распределений с характеристическими функциями вида $[\mathsf{E} e^{isY}]^t$ полностью определяют конечномерные распределения некоторого процесса Леви $X(t)$, $t \geq 0$, при этом $X(1) \stackrel{d}{=} Y$.

Свойства процессов Леви детально описаны [15, 85]. Книги [13, 86] и обзорная статья [45] рассматривают приложения процессов Леви для моделирования динамики цен акций и финансовых индексов.

2.2 Сходимость процесса ОФИ к процессам Леви

Без потери общности всюду в дальнейшем мы будем рассматривать случайные процессы, определённые на отрезке $0 \leq t \leq 1$. Фактически это означает, что мы

будем изучать поведение обобщённого двухстороннего процесса риска, следовательно и процесса OFI, на конечном временном интервале. Равенство правой границы единице может быть обеспечено соответствующим выбором единицы измерения времени. Другими словами, мы рассматриваем процесс OFI в пространстве Скорохода \mathcal{D} .

Для того, чтобы рассмотреть асимптотические конструкции, которые позволяют формализовать сценарий бесконечного возрастания интенсивностей потоков заявок и рассмотреть асимптотические («heavy-traffic») аппроксимации одномерных распределений процесса OFI, зафиксируем момент времени t и рассмотрим вспомогательный параметр n . Всюду в дальнейшем предполагается сходимость при $n \rightarrow \infty$, если не указано иное. Таким образом, рассмотрим последовательность обобщённых процессов Кокса в виде

$$Q_n(t) = \sum_{i=1}^{N_1^{(n)}(\Lambda_n(t))} X_{n,i}, \quad t \geq 0, \quad (2.3)$$

где $\{N_1^{(n)}(t), t \geq 0\}_{n \geq 1}$ – это последовательность пуассоновских процессов с единичными интенсивностями; для каждого $n = 1, 2, \dots$ случайные величины $X_{n,1}, X_{n,2}, \dots$ одинаково распределены; для каждого $n \geq 1$ случайные величины $X_{n,1}, X_{n,2}, \dots$ и процессы $N_1^{(n)}(t), t \geq 0$, являются независимыми; для каждого $n = 1, 2, \dots$ процессы $\Lambda_n(t), t \geq 0$, являются управляемыми, то есть неубывающими положительными процессами Леви, независимыми от процесса

$$Z_n(t) = \sum_{i=1}^{N_1^{(n)}(t)} X_{n,i}, \quad t \geq 0, \quad (2.4)$$

и такими, что $\Lambda_n(0) = 0$ и существуют такие $\delta \in (0, 1]$, $\delta_1 \in (0, 1]$ и константы $C_n \in (0, \infty)$, обеспечивающие для любых $t \in (0, 1]$ выполнение неравенства

$$\mathbb{E} \Lambda_n^\delta(t) \leq (C_n t)^{\delta_1}. \quad (2.5)$$

Всюду в дальнейшем для определённости мы полагаем $\sum_{i=1}^0 = 0$. В терминах предыдущих разделов $Z_n(t)$ является условным однородным процессом OFI.

Ниже мы покажем, что рассуждения, приведённые в [74] для симметричного случая могут быть с незначительными изменениями обобщены на случай предельных процессов с несимметричными функциями распределения.

Из (2.3) и (2.4) несложно увидеть, что $Q_n(t) = Z_n(\Lambda_n(t))$. Поскольку для любого $n \geq 1$ $Z_n(t)$ и $\Lambda_n(t)$ являются независимыми процессами Леви и, кроме того, $\Lambda_n(t)$ является управляющим процессом, то суперпозиция $Q_n(t) = Z_n(\Lambda_n(t))$ также является процессом Леви (см., например, Теорему 3.1.1 в [61]). Таким образом, верно следующее утверждение:

ЛЕММА 2.2. Для любых $0 \leq t_1 < t_2 < \infty$ и любого $n \geq 1$ мы имеем $Q_n(t_2) - Q_n(t_1) \stackrel{d}{=} Q_n(t_2 - t_1)$.

Определим $a_n = \mathbb{E}X_{n,1}$ и предположим, что

$$0 < m_n^\beta \equiv \mathbb{E}|X_{n,1}|^\beta < \infty \quad (2.6)$$

для некоторого $\beta \in [1, 2]$.

ЗАМЕЧАНИЕ 3. В контексте рассматриваемой задачи о моделировании динамики потока заявок мы можем считать, что

$$\mathbb{E}e^{isX_{n,1}} = \frac{\alpha_n^+ \mathfrak{f}_n^+(s)}{\alpha_n^+ + \alpha_n^-} + \frac{\alpha_n^- \mathfrak{f}_n^+(-s)}{\alpha_n^+ + \alpha_n^-}, \quad s \in \mathbb{R},$$

где $\mathfrak{f}_n^+(s)$ и $\mathfrak{f}_n^+(s)$ - это характеристические функции случайных величин $X_{n,1}^+$ и $X_{n,1}^-$, тогда

$$a_n \equiv \mathbb{E}X_{n,1} = \frac{\alpha_n^+ \mathbb{E}X_{n,1}^+}{\alpha_n^+ + \alpha_n^-} - \frac{\alpha_n^- \mathbb{E}X_{n,1}^-}{\alpha_n^+ + \alpha_n^-}$$

и

$$m_n^\beta \equiv \mathbb{E}|X_{n,1}|^\beta = \frac{\alpha_n^+ \mathbb{E}|X_{n,1}^+|^\beta}{\alpha_n^+ + \alpha_n^-} + \frac{\alpha_n^- \mathbb{E}|X_{n,1}^-|^\beta}{\alpha_n^+ + \alpha_n^-},$$

а условие (2.6) обеспечивается условиями

$$0 < \mathbb{E}|X_{n,1}^+|^\beta < \infty, \quad 0 < \mathbb{E}|X_{n,1}^-|^\beta < \infty.$$

ЛЕММА 2.3. *Пусть $Q_n(t)$ является обобщённым процессом Кокса (2.3), удовлетворяющим условиям (2.5) and (2.6). Тогда для любого $t \in [0, 1]$ и любого $\epsilon > 0$ имеем $\mathbb{P}(|Q_n(t)| \geq \epsilon) \leq \epsilon^{-\beta\delta} m_n^{\beta\delta} \cdot (C_n t)^{\delta_1}$.*

ДОКАЗАТЕЛЬСТВО. Поскольку распределения одномерных процессов Кокса (2.3) являются смесями пуассоновских распределений, то

$$\begin{aligned} \mathbb{P}(|Q_n(t)| \geq \epsilon) &= \sum_{k=0}^{\infty} \mathbb{P}(N_1^{(n)}(\Lambda_n(t)) = k) \mathbb{P}\left(\left|\sum_{j=1}^k X_{n,j}\right| \geq \epsilon\right) = \\ &= \int_0^{\infty} \left[\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \mathbb{P}\left(\left|\sum_{j=1}^k X_{n,j}\right| \geq \epsilon\right) \right] d\mathbb{P}(\Lambda_n(t) < \lambda). \end{aligned} \quad (2.7)$$

Изменение порядка суммирования и интегрирования становится возможным благодаря очевидной равномерной сходимости соответствующего ряда. Последовательно применим к уравнению (2.7) неравенства Маркова и Йенсена с $\delta \in (0, 1]$, а также с учётом (2.5) и (2.6) получаем

$$\mathbb{P}(|Q_n(t)| \geq \epsilon) \leq \frac{1}{\epsilon^{\beta\delta}} \int_0^{\infty} \left[\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \left(\mathbb{E} \left| \sum_{j=1}^k X_{n,j} \right|^\beta \right)^\delta \right] d\mathbb{P}(\Lambda_n(t) < \lambda), \quad (2.8)$$

поскольку для $\delta \in (0, 1]$ функция $f(x) = x^\delta$ является выпуклой при $x \geq 0$. Легко видеть, что $\mathbb{E} \left| \sum_{j=1}^k X_{n,j} \right|^\beta \leq \sum_{j=1}^k \mathbb{E} |X_{n,j}|^\beta = km_n^\beta$ для $1 \leq \beta \leq 2$. Следовательно, продолжая цепочку для (2.8) с учётом неравенства Йенсена для выпуклых

функций и условия (2.5), получаем

$$\begin{aligned} \mathsf{P}(|Q_n(t)| \geq \epsilon) &\leq \frac{m_n^{\beta\delta}}{\epsilon^{\beta\delta}} \int_0^\infty \mathsf{E}[N_1^{(1)}(\lambda)]^\delta d\mathsf{P}(\Lambda_n(t) < \lambda) \leq \\ &\leq \frac{m_n^{\beta\delta}}{\epsilon^{\beta\delta}} \int_0^\infty [\mathsf{E}N_1^{(1)}(\lambda)]^\delta d\mathsf{P}(\Lambda_n(t) < \lambda) = \frac{m_n^{\beta\delta}}{\epsilon^{\beta\delta}} \cdot \mathsf{E}\Lambda_n^\delta(t) \leq \frac{m_n^{\beta\delta}}{\epsilon^{\beta\delta}} \cdot (C_n t)^{\delta_1}. \end{aligned}$$

Лемма доказана.

Для установления слабой сходимости случайного процесса $Q_n(t)$ в пространстве Скорохода \mathcal{D} прежде всего требуется найти предельное распределение случайной величины $Q_n(t)$ для каждого $t > 0$. Символ \xrightarrow{d} будет обозначать сходимость по распределению, т.е. поточечную сходимость функций распределения во всех точках непрерывности предельной функции распределения.

Пусть $t = 1$. Обозначим $N_n = N_1^{(n)}(\Lambda_n(1))$. Предположим, что для некоторого $k_n \in \mathbb{N}$ имеет место сходимость

$$\mathsf{P}(X_{n,1} + \dots + X_{n,k_n} < x) \xrightarrow{d} H(x), \quad (2.9)$$

где $H(x)$ - некоторая функция распределения безгранично делимой случайной величины.

Также предположим, что

$$\mathsf{P}(\Lambda_n(1) < k_n x) \xrightarrow{d} \mathsf{P}(U < x), \quad (2.10)$$

где U – неотрицательная случайная величина такая, что её распределение не вырождается в нуле. Заметим, что поскольку $\Lambda_n(t)$ – процесс Леви, то случайная величина U также является безгранично делимой будучи слабым пределом безгранично делимых случайных величин.

ЛЕММА 2.4. Пусть $N_n = N_1^{(n)}(\Lambda_n)$, $n \geq 1$, где $\{N_1^{(n)(t)}, t \geq 0\}$, $n = 1, 2, \dots$

это стандартные процессы Пуассона и Λ_n , $n = 1, 2, \dots$ - положительные случайные величины, такие, что для каждого $n \geq 1$ случайная величина Λ_n является независимой от процесса $N_1^{(n)}(t)$. Тогда для некоторой бесконечно возрастающей последовательности k_n вещественных чисел и некоторой функции распределения $A(x)$ сходимость

$$\mathbb{P}(N_n < k_n x) \xrightarrow{d} A(x)$$

имеет место быть тогда и только тогда, когда

$$\mathbb{P}(\Lambda_n < k_n x) \xrightarrow{d} A(x).$$

Доказательство может быть найдено, например в [51].

Из Леммы 2.4 следует, что сходимость (2.10) эквивалентна сходимости

$$\mathbb{P}(N_n < k_n x) \xrightarrow{d} \mathbb{P}(U < x). \quad (2.11)$$

По теореме переноса Гнеденко-Фахима [49] из условий (2.9) и (2.11) следует, что

$$Q_n(1) = X_{n,1} + \dots + X_{n,N_n} \xrightarrow{d} Q, \quad (2.12)$$

где Q - случайная величина с характеристической функцией

$$\mathfrak{f}(s) = \int_0^\infty (h(s))^u d\mathbb{P}(U < u), \quad (2.13)$$

$h(s)$ является характеристической функцией, соответствующей функции распределения $H(x)$. Заметим, что $H(x)$ может не удовлетворять условию $H(-x) = 1 - H(x)$, то есть она может быть несимметричной.

Пусть Y - безгранично делимая случайная величина с функцией распределения $H(x)$. Поскольку обе случайные величины Y и U являются безгранично

делимыми, мы можем определить независимые процессы Леви $Y(t)$ и $U(t)$, $t \geq 0$, такие, что $Y(1) \stackrel{d}{=} Y$ и $U(1) \stackrel{d}{=} Y$. Тогда с учётом Леммы 2.1 легко установить, что $\mathbf{f}(s) = \mathbf{E}e^{isQ} = \mathbf{E}\exp\{isY(U(1))\}$, $s \in \mathbb{R}$, то есть $Q \stackrel{d}{=} Y(U(1))$. Кроме того, повторяя рассуждения из [61] (см. Теорему 3.3.1), мы можем легко видеть, что случайная величина Q является безгранично делимой, следовательно, можно определить процесс Леви $Q(t)$, $t \geq 0$, такой, что $Q(1) \stackrel{d}{=} Q$. Из Леммы 2.1 и из сказанного выше следует, что мы можем рассматривать процесс $Q(t)$ как суперпозицию $Q(t) \stackrel{d}{=} Y(U(t))$.

Поскольку в соответствии с (2.12) мы имеем $Q_n(1) = \sum_{i=1}^{N_n} X_{n,i} \Rightarrow Q(1)$ и $Q_n(t)$ и $Q(t)$ являются процессами Леви, тогда, используя (2.2) заключаем, что для любого $t > 0$

$$Q_n(t) = \sum_{i=1}^{N_n(\Lambda_n(t))} X_{n,i} \xrightarrow{d} Q(t). \quad (2.14)$$

Поскольку $Q_n(t)$ и $Q(t)$, $0 \leq t \leq 1$ - процессы Леви, то почти всюду траектории этих процессов будут лежать в пространстве Скорохода \mathcal{D} .

Далее мы попробуем установить дополнительные условия для слабой сходимости обобщённых процессов $Q_n(t)$ к процессу Леви $Q(t)$ в пространстве \mathcal{D} . Мы последовательно рассмотрим каждое из условий Теоремы 2.1.

Во-первых, без потери общности пусть $0 \leq t_1 < t_2 < \dots < t_k \leq 1$. Сходимость $(Q_n(t_1), \dots, Q_n(t_k)) \xrightarrow{d} (Q(t_1), \dots, Q(t_k))$ эквивалентна сходимости

$$\begin{aligned} (Q_n(t_1), Q_n(t_2) - Q_n(t_1), \dots, Q_n(t_k) - Q_n(t_{k-1})) &\xrightarrow{d} \\ &\xrightarrow{d} (Q(t_1), Q(t_2) - Q(t_1), \dots, Q(t_k) - Q(t_{k-1})), \end{aligned} \quad (2.15)$$

поскольку линейное преобразование $(x_1, x_2, \dots, x_{k-1}, x_k) \mapsto (x_1, x_2 - x_1, \dots, x_k - x_{k-1})$ из \mathbb{R}^k в \mathbb{R}^k является взаимнооднозначным и непрерывным в обе стороны. Но сходимость (2.15) следует из (2.14) и того факта, что $Q_n(t)$ and $Q(t)$ являются

процессам Леви.

Во-вторых, необходимо проверить условие $\mathsf{P}(Q(1) \neq Q(1-)) = 0$. Это условие выполняется тогда и только тогда, когда $\lim_{t \rightarrow 1-} \mathsf{P}(|Q(1) - Q(t)| > \epsilon) = 0$ для любого $\epsilon > 0$ (см. (15.16) в [18]). Рассмотрим $\mathsf{P}(|Q(1) - Q(t)| > \epsilon)$. Поскольку $Q(t)$ - процесс Леви, то $Q(1) - Q(t) \stackrel{d}{=} Q(1-t)$ по Лемме 2.2. Таким образом, $\mathsf{P}(|Q(1) - Q(t)| > \epsilon) = \mathsf{P}(|Q(1-t)| > \epsilon)$. Для каждого $\epsilon > 0$ и каждого $t \in [0, 1]$ существует $\epsilon_t \in [\epsilon/2, \epsilon]$ такое что точки $\pm\epsilon_t$ являются точками непрерывности функции распределения случайной величины $Q(1-t)$. Поскольку $Q_n(t) \xrightarrow{d} Q(t)$ для любого $t \in [0, 1]$, то $\mathsf{P}(|Q(1-t)| > \epsilon_t) = \lim_{n \rightarrow \infty} \mathsf{P}(|Q_n(1-t)| > \epsilon_t)$. Таким образом, для любого $\epsilon > 0$ и любого $t \in [0, 1]$

$$\mathsf{P}(|Q(1-t)| > \epsilon) \leq \mathsf{P}(|Q(1-t)| > \epsilon_t) = \lim_{n \rightarrow \infty} \mathsf{P}(|Q_n(1-t)| > \epsilon_t). \quad (2.16)$$

Из (2.16) с учётом (2.5) и Леммы 2.3 для любого $\delta \in (0, 1]$ получаем

$$\begin{aligned} \mathsf{P}(|Q(1-t)| > \epsilon) &\leq \sup_n \mathsf{P}(|Q_n(1-t)| > \epsilon_t) \leq \\ &\leq \sup_n (\epsilon_t^{-\beta} m_n^\beta)^\delta (C_n |1-t|)^{\delta_1} \leq (2^{\beta\delta} \epsilon^{-\beta\delta} |1-t|)^{\delta_1} \sup_n m_n^{\beta\delta} C_n^{\delta_1}. \end{aligned} \quad (2.17)$$

Поэтому если

$$K \equiv \sup_n C_n^{\delta_1/\delta} m_n^\beta < \infty, \quad (2.18)$$

то из (2.17) следует $\lim_{t \rightarrow 1-} \mathsf{P}(|Q(1) - Q(t)| > \epsilon) \leq 4(K\epsilon^{-\beta})^\delta \lim_{t \rightarrow 1-} |1-t|^{\delta_1} = 0$.

В-третьих, проверим выполнение условия (2.1), если выполняются (2.5) и (2.18). Как было замечено ранее, $Q_n(t)$ - процесс Леви и потому имеет независимые приращения. Поэтому

$$\mathsf{P}(|Q_n(t) - Q_n(t_1)| \geq \epsilon, |Q_n(t_2) - Q_n(t)| \geq \epsilon) =$$

$$= \mathsf{P}(|Q_n(t) - Q_n(t_1)| \geq \epsilon) \cdot \mathsf{P}(|Q_n(t_2) - Q_n(t)| \geq \epsilon). \quad (2.19)$$

Рассмотрим первый множитель правой части уравнения (2.19). По Лемме 2.2 $Q_n(t) - Q_n(t_1) \stackrel{d}{=} Q_n(t - t_1)$. С учётом (2.18) по Лемме 2.3 получаем

$$\mathsf{P}(|Q_n(t) - Q_n(t_1)| \geq \epsilon) = \mathsf{P}(|Q_n(t - t_1)| \geq \epsilon) \leq (K\epsilon^{-\beta})^\delta |t - t_1|^{\delta_1}. \quad (2.20)$$

Для второго множителя в правой части уравнения (2.19) аналогичным образом

$$\mathsf{P}(|Q_n(t_2) - Q_n(t)| \geq \epsilon) = \mathsf{P}(|Q_n(t_2 - t)| \geq \epsilon) \leq (K\epsilon^{-\beta})^\delta |t_2 - t|^{\delta_1}. \quad (2.21)$$

Итого, из (2.20) и (2.21) следует

$$\mathsf{P}(|Q_n(t) - Q_n(t_1)| \geq \epsilon, |Q_n(t_2) - Q_n(t)| \geq \epsilon) \leq (K\epsilon^{-\beta})^{2\delta} [(t - t_1)(t_2 - t)]^{\delta_1} \quad (2.22)$$

Легко видеть, что для любых $t_1 \leq t \leq t_2$ выполняется $(t - t_1)(t_2 - t) \leq \frac{1}{4}(t_2 - t_1)^2$.

Подставляя эту оценку в (2.22), получаем $\mathsf{P}(|Q_n(t) - Q_n(t_1)| \geq \epsilon, |Q_n(t_2) - Q_n(t)| \geq \epsilon) \leq \epsilon^{-2\beta\delta} [\frac{1}{2}K(t_2 - t_1)]^{2\delta_1}$. Таким образом, при выполнении условий (2.5) и (2.18) автоматически выполняется условие (2.1) с $F(t) \equiv \frac{1}{2}Kt$, $\nu = \beta\delta$ и $\gamma = \delta_1$.

Подводя итоги этих рассуждений, связанных с выполнением необходимых условий Теоремы 2.1, получаем следующее утверждение

ТЕОРЕМА 2.2. *Пусть процесс дисбаланса потока заявок $Q_n(t)$ (см. (2.3)) управляемся неубывающим положительным процессом Леви $\Lambda_n(t)$, удовлетворяющим условиям (2.5) и (2.10) с некоторыми $\delta, \delta_1 \in (0, 1]$ и $k_n \in \mathbb{N}$. Предположим, что случайные величины $\{X_{n,j}\}_{j \geq 1}$, $n = 1, 2, \dots$, (случайные объёмы заявок) удовлетворяют условию (2.9) с теми же значениями k_n , а также удовлетворяют условию (2.6) с некоторым $\beta \in [1, 2]$. Также предположим, что выполнено*

условие (2.18). Тогда процесс $Q_n(t)$ слабо сходится в пространстве Скорохода \mathcal{D} к процессу Леви $Q(t)$ такому, что

$$\mathsf{E} \exp\{isQ(1)\} = \int_0^\infty (h(s))^u d\mathsf{P}(U < u), \quad s \in \mathbb{R}, \quad (2.23)$$

где $h(s)$ - характеристическая функция, соответствующая функции распределения $H(x)$ в условии (2.9).

Следует заметить, что в Теореме 2.2 речь идёт о хорошо изученной слабой сходимости семимартингалов с независимыми и стационарными приращениями, см., например, [59]. Однако, случай суперпозиции процессов, рассматриваемый в данной работе, даёт возможность ослабить условия, требуемые для такой сходимости в общем случае, скажем, в Следствии VII.3.6 в [59], где предполагается (в нашей терминологии) $\delta = \delta_1 = 1$.

Некоторые следствия этого результата, имеющие дело с симметричными предельными распределениями, рассмотрены в [74]. В частности, было показано, что симметричные устойчивые процессы Леви могут возникать как предельные для обобщённых процессов Кокса даже в случаях, когда дисперсии элементарных приращений обобщённого процесса Кокса конечны. Как уже было сказано, в большинстве прикладных задач нет причин отвергать это предположение. В частности, это относится к моделированию динамики книги заявок, где объёмы заявок ограничены. Поэтому далее мы сконцентрируемся на случае конечных дисперсий и рассмотрим условия сходимости процесса дисбаланса потоков заявок к некоторым известным моделям, в частности, к обобщённым гиперболическим процессам Леви.

2.3 Обобщённые гиперболические процессы Леви как асимптотическая аппроксимация процесса дисбаланса потоков заявок

Обозначим $\sigma_n^2 = \text{DX}_{n,1}$. Из классической теории предельных теорем хорошо известно, что если при некоторых $a \in \mathbb{R}$, $0 < \sigma^2 < \infty$ и некотором $\epsilon > 0$ выполняются условия (при $n \rightarrow \infty$)

$$k_n a_n \longrightarrow a, \quad k_n \sigma_n^2 \longrightarrow \sigma^2 \quad \text{and} \quad k_n \mathbb{E}(X_{n,1} - a_n)^2 \mathbb{I}(|X_{n,1} - a_n| \geq \epsilon) \longrightarrow 0, \quad (2.24)$$

то имеет место сходимость (2.9) с $H(x) \equiv \Phi(\sigma^{-1}(x - a))$. В таком случае функция распределения $F(x)$ предельной случайной величины $Q(1)$ в Теореме 2.2 является сдвиг-масштабной смесью нормальных законов. Относительно недавно было показано, что нормальные сдвиг-масштабные смеси возникают как предельные в предельных теоремах для случайных сумм независимых одинаково распределённых случайных величин [68, 71, 72]. А именно, пусть $\{\xi_{n,j}\}_{j \geq 1}$, $n = 1, 2, \dots$, - двумерный массив построчно (для каждого фиксированного n) независимых и одинаково распределённых случайных величин. Пусть $\{\nu_n\}_{n \geq 1}$ - последовательность целых неотрицательных случайных величин, таких что для каждого $n \geq 1$ случайные величины $\nu_n, \xi_{n,1}, \xi_{n,2}, \dots$ независимы. Обозначим $S_{n,k} = \xi_{n,1} + \dots + \xi_{n,k}$. Нижеследующая теорема доказана в [68].

ТЕОРЕМА 2.3. *Предположим, что существуют: последовательности $\{k_n\}_{n \geq 1}$ натуральных чисел и конечные числа $\alpha \in \mathbb{R}$ и $\sigma > 0$, такие, что*

$$\mathbb{P}(S_{n,k_n} < x) \xrightarrow{d} \Phi\left(\frac{x - \alpha}{\sigma}\right). \quad (2.25)$$

Предположим также, что $\nu_n \rightarrow \infty$ по вероятности. Тогда функции рас-

пределения случайных сумм S_{n,ν_n} сходятся к некоторой функции распределения $F(x)$:

$$\mathbb{P}(S_{n,\nu_n} < x) \xrightarrow{d} F(x),$$

тогда и только тогда, когда существует функция распределения $A(x)$, такая, что $A(0) = 0$,

$$F(x) = \int_0^\infty \Phi\left(\frac{x - \alpha z}{\sigma\sqrt{z}}\right) dA(z),$$

и

$$\mathbb{P}(\nu_n < xk_n) \xrightarrow{d} A(x).$$

Из Теоремы 2.3 и Леммы 2.4 вытекает следующая

ТЕОРЕМА 2.4. Пусть процессы дисбаланса потоков заявок $Q_n(t)$ (см. (2.3)) управляются неубывающими положительными процессами Леви $\Lambda_n(t)$, удовлетворяющими условию (2.5) с некоторыми $\delta, \delta_1 \in (0, 1]$. Предположим, что обёмы заявок $\{X_{n,j}\}_{j \geq 1}$, $n = 1, 2, \dots$, удовлетворяют условиям (2.24) с некоторыми $k_n \in \mathbb{N}$. Также предположим, что выполняется условие (2.18) с $\beta = 2$. Тогда процессы $Q_n(t)$ слабо сходятся в пространстве Скорохода \mathcal{D} к процессу Леви $Q(t)$ тогда и только тогда, когда существует неотрицательная случайная величина U , такая, что

$$\mathbb{P}(Q(1) < x) = \int_0^\infty \Phi\left(\frac{x - au}{\sigma\sqrt{u}}\right) d\mathbb{P}(U < u), \quad x \in \mathbb{R}, \quad (2.26)$$

и выполняется условие (2.10) с теми же самыми k_n .

Класс распределений вида (2.26) неоднократно рассматривался Оле Барндорфф-Нильсеном и его коллегами [8–10] для того, чтобы ввести *обобщённые гиперболические распределения* и изучить их свойства.

Класс нормальных сдвиг-масштабных смесей (refeq:th2) очень широк. Например, он содержит обобщённые гиперболические законы с обобщёнными обратными гауссовскими смешивающими распределениями, в частности, (a) симметричные и несимметричные (скошенные) распределения Стьюдента (включая распределение Коши), которым в (2.26) соответствуют обратные гамма-распределения в качестве смешивающих; (b) variance-gamma (VG) распределения (включая симметричные и несимметричные распределения Лапласа), которым в (2.26) соответствуют гамма-распределения в качестве смешивающих; (c) нормальные\\обратные гауссовские распределения, которым в (2.26) соответствует обратное гауссовское распределение в качестве смешивающего и много других. Вместе с обобщёнными гиперболическими распределениями класс нормально сдвиг-масштабных смесей включает в себя симметричные строго устойчивые законы с $\mu = 0$ и строго устойчивые смешанные распределения, определённые на положительной полуоси, обобщённые экспоненциальные, степенные распределения и многие другие.

Обобщённые гиперболические распределения демонстрируют высокую адекватность при описании статистических закономерностей в поведении характеристик различных сложных систем, в частности, турбулентных систем и финансовых рынков. В данный момент имеется большое множество публикаций, описывающих модели на базе обобщённых гиперболических распределений. Лишь некоторые из них: [8, 9, 11, 12, 26, 37–40, 79, 83, 89]. Поэтому ниже мы сконцентрируем наше внимание на функциональных предельных теоремах, устанавливающих сходимость процессов дисбаланса потоков заявок OFI к обобщённым гиперболическим процессам Леви.

Конечно же, хорошая описательная способность обобщённых гиперболических моделей связана с наличием достаточно большого количества параметров, которые описывают эти модели и, как следствие, могут быть точно настроены под конкретный изучаемый процесс. Но в действительности было бы разумнее объяс-

нить этот феномен с помощью функциональных предельных теорем, из которых будет следовать возможность использования обобщённых гиперболических процессов Леви как очевидных «heavy-traffic» *асимптотических* аппроксимаций.

Пусть $p_{GIG}(x; \nu, \mu, \lambda)$ - плотность *обобщённого обратного гауссовского распределения*:

$$p_{GIG}(x; \nu, \mu, \lambda) = \frac{\lambda^{\nu/2}}{2\mu^{\nu/2}K_{\nu}(\sqrt{\mu\lambda})} \cdot x^{\nu-1} \cdot \exp\left\{-\frac{1}{2}\left(\frac{\mu}{x} + \lambda x\right)\right\}, \quad x > 0.$$

Здесь $\nu \in \mathbb{R}$,

$$\mu > 0, \quad \lambda \geq 0, \quad \text{если } \nu < 0,$$

$$\mu > 0, \quad \lambda > 0, \quad \text{если } \nu = 0,$$

$$\mu \geq 0, \quad \lambda > 0, \quad \text{если } \nu > 0,$$

$K_{\nu}(z)$ - модифицированная функция Бесселя третьего рода с индексом ν ,

$$K_{\nu}(z) = \frac{1}{2} \int_0^{\infty} y^{\nu-1} \exp\left\{-\frac{z}{2}\left(y + \frac{1}{y}\right)\right\} dy, \quad z \in \mathbb{C}, \quad \text{Re } z > 0.$$

Соответствующая функция распределения будет обозначаться как $P_{GIG}(x; \nu, \mu, \lambda)$,

$$P_{GIG}(x; \nu, \mu, \lambda) = \int_0^x p_{GIG}(z; \nu, \mu, \lambda) dz, \quad x \geq 0,$$

и $P_{GIG}(x; \nu, \mu, \lambda) = 0$, $x < 0$. Согласно [87], обобщённое обратное гауссовское распределение было введено в 1946 by Этьеном Хальфеном (Étienne Halphen), который использовал его для описания месячных объёмов воды, проходящих через гидроэлектростанции. В работе [87] обобщённое обратное гауссовское распределение было названо *распределением Хальфена* (*Halphen distribution*). В 1973 это распределение было заново открыто Гербертом Сихелом (Herbert Sichel) [90], который использовал его как смешивающий закон в специальных смешанных пуассон-

новских распределениях (*Sichel distributions*, см., например, [69]) как дискретные распределения с тяжёлыми хвостами. В 1977 эти распределения было ещё раз открыты Оле Барндорфф-Нильсоном [7, 8], который, в частности, использовал их для описания распределения объёмов частиц.

Класс обобщённых обратных гауссовских распределений довольно богат и содержит, в частности, как распределения с экспоненциально убывающими хвостами (гамма-распределения с $(\mu = 0, \nu > 0)$), так и распределения со степенным убыванием хвостов (обратные гамма-распределения $\lambda = 0, \nu < 0$), обратные гауссовские распределения $(\nu = -\frac{1}{2})$ и их предельный случай при $\lambda \rightarrow 0$ – распределение Леви (устойчивое распределение с характеристической экспонентой, равной $\frac{1}{2}$).

В 1977–78 Оле Барндорфф-Нильсен [7, 8] ввёл класс *обобщённых гиперболических распределений* как специальный класс нормальных сдвиг-масштабных смесей. Для удобства мы будем использовать более простую параметризацию. Пусть $\alpha \in \mathbb{R}, \sigma > 0$. Функция обобщённого гиперболического распределения с параметрами $\alpha, \sigma, \nu, \mu, \lambda$ по определению

$$P_{GH}(x; \alpha, \sigma, \nu, \mu, \lambda) = \int_0^\infty \Phi\left(\frac{x - \alpha z}{\sigma \sqrt{z}}\right) p_{GIG}(z; \nu, \mu, \lambda) dz, \quad x \in \mathbb{R}. \quad (2.27)$$

Заметим, что в (2.27) параметры сдвига и масштаба при смешивании разделены, но поскольку эти параметры напрямую связаны в (2.27), то это фактически однопараметрическая смесь. Для параметризации обобщённых гиперболических распределений существуют различные методики, см., например, [8, 9, 11, 12, 38, 39, 83, 89]. Однако, плотность $p_{GH}(x; \alpha, \sigma, \nu, \mu, \lambda)$ обобщённого гиперболического распределения не может быть выражена в элементарных функциях и имеет вид

$$p_{GH}(x; \alpha, \sigma, \nu, \mu, \lambda) = \int_0^\infty \frac{1}{\sigma \sqrt{z}} \varphi\left(\frac{x - \alpha z}{\sigma \sqrt{z}}\right) p_{GIG}(z; \nu, \mu, \lambda) dz =$$

$$= \int_0^\infty \frac{1}{\sigma \sqrt{2\pi} z} \exp \left\{ -\frac{(x - \alpha z)^2}{2\sigma^2 z} \right\} \frac{\lambda^{\nu/2} z^{\nu-1}}{2\mu^{\nu/2} K_\nu(\sqrt{\mu\lambda})} \exp \left\{ -\frac{1}{2} \left(\frac{\mu}{z} + \lambda z \right) \right\} dz, \quad x \in \mathbb{R},$$

который может быть в дальнейшем упрощён с использованием модифицированных функций Бесселя третьего рода..

Из Теоремы 2.3 мы легко получаем следующее

УТВЕРЖДЕНИЕ 2.1. *Предположим, что существуют: последовательность $\{k_n\}_{n \geq 1}$ натуральных чисел, конечные числа $\alpha \in \mathbb{R}$ и $\sigma > 0$, такие, что имеет место сходимость (2.25). Предположим, что $\nu_n \rightarrow \infty$ по вероятности. Тогда распределение случайной суммы S_{ν_n} сходится к обобщенному гиперболическому распределению*

$$\mathbb{P}(S_{n, \nu_n} < x) \xrightarrow{d} P_{GH}(x; \alpha, \sigma, \nu, \mu, \lambda)$$

тогда и только тогда, когда

$$\mathbb{P}(\nu_n < x k_n) \xrightarrow{d} P_{GIG}(x; \nu, \mu, \lambda)$$

Из Теоремы 2.4 и Утверждения 2.1 с учётом эквивалентности соотношений (2.10) и (2.11) мы легко получаем следующий результат о сходимости процесса ОФИ к обобщённым гиперболическим процессам Леви.

ТЕОРЕМА 2.5. *Пусть процессы дисбаланса потока заявок $Q_n(t)$ (см. (2.3)) управляются неубывающими положительными процессами Леви $\Lambda_n(t)$, удовлетворяющими условию (2.5) с некоторыми $\delta, \delta_1 \in (0, 1]$. Предположим, что случайные обзёмы заявок $\{X_{n,j}\}_{j \geq 1}$, $n = 1, 2, \dots$, удовлетворяют условиям (2.24) с некоторыми $k_n \in \mathbb{N}$ and some $a \in \mathbb{R}$ и $\sigma > 0$. Также предположим, что выполняется условие (2.18) с $\beta = 2$. Тогда процессы $Q_n(t)$ слабо сходятся в пространстве*

Скорохода \mathcal{D} к обобщённому гиперболическому процессу Леви $Q(t)$, такому, что

$$\mathsf{P}(Q(1) < x) = P_{GH}(x; a, \sigma, \nu, \mu, \lambda)$$

тогда и только тогда, когда

$$\mathsf{P}(\Lambda_n(1) < k_n x) \xrightarrow{d} P_{GIG}(x; \nu, \mu, \lambda)$$

с некоторыми k_n , ν , μ and λ .

Для завершения данной главы мы должны заметить, что Теоремы 2.2–2.5, представленные выше, могут выступать в качестве очевидного объяснения высокой адекватности обобщённых гиперболических процессов Леви как моделей, объясняющих динамику процесса дисбаланса потока заявок. Кроме того, они напрямую связывают субординатор в представлении обобщённых гиперболических процессов Леви как подчинённых винеровских процессов с интенсивностями потоков заявок, определяемых процессом общего новостного фона. Поскольку последний является слабо предсказуемым, описание типа его распределения требует гораздо больше параметров (как минимум три в семействе обобщённых обратно-гауссовских законов).

2.4 Финальная настройка модели

В данном разделе мы вернёмся к определению общего процесса дисбаланса потоков заявок, введённого в Главе 1. В дополнение к обозначениям, введённым выше, рассмотрим *процесс дисбаланса (II)* как

$$r(t) = \frac{\alpha^+(t)}{\alpha^-(t)}.$$

Пусть $\Lambda^*(t) = \Lambda^-(t) = \alpha^-(t)\Lambda^*(t)$. Тогда, очевидно, $\Lambda^+(t) = r(t)\Lambda^*(t)$,

$$\Lambda(t) = (1 + r(t))\Lambda^* \quad (2.28)$$

и если $r(t) = r = \text{const}$, тогда

$$\mathsf{E} X_{n,1} = \frac{r}{1+r} \mathsf{E} X_{n,1}^+ + \frac{1}{1+r} \mathsf{E} X_{n,1}^-, \quad \mathsf{E} |X_{n,1}|^\beta = \frac{r}{1+r} \mathsf{E} |X_{n,1}^+|^\beta + \frac{1}{1+r} \mathsf{E} |X_{n,1}^-|^\beta. \quad (2.29)$$

Это означает, что вместо использования параметров α^+ и α^- можно повторить все рассуждения, данные выше, в терминах процесса дисбаланса потоков заявок $r(t)$.

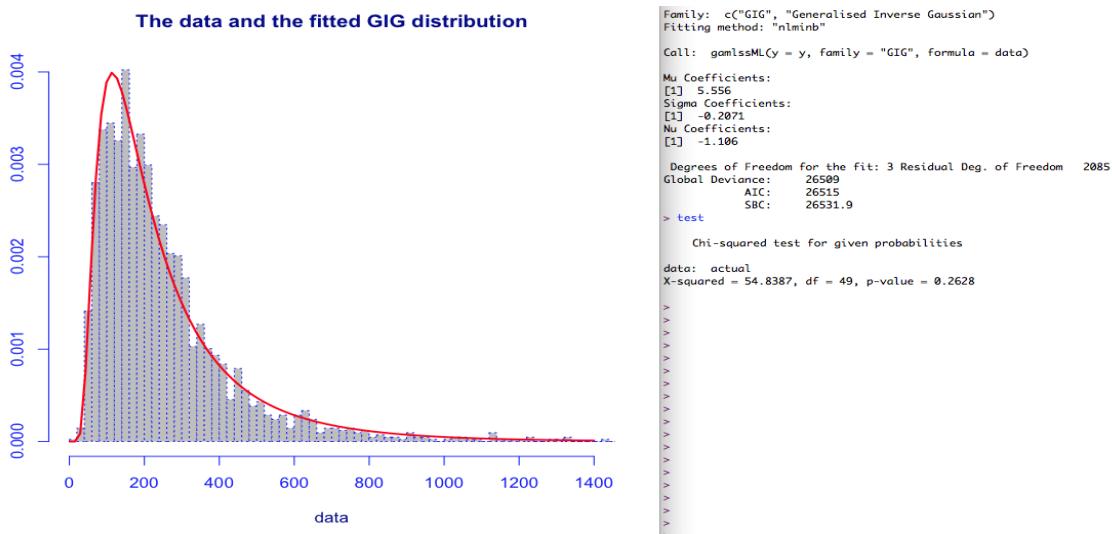


Рис. 2.1: Гистограмма количества заявок на покупку за 15-секундные интервалы времени и плотность подходящего GIG-распределения. Фьючерс на индекс РТС, дневная сессия 2014.07.01).

Как можно было видеть раньше, если $r(t) = r = \text{const}$, то процесс OFI может быть успешно аппроксимирован с помощью обобщённых гиперболических процессов Леви с некоторыми параметрами $a, \sigma, \nu, \mu, \lambda$. Тем не менее, если процесс $r(t)$ считается случайным, то в финальной версии модели предполагается, что коэф-

фициенты $a, \sigma, \nu, \mu, \lambda$ должны зависеть от $r(t)$, так что финальная модель представляется в виде обобщённого гиперболического процесса Леви со случайными параметрами. Потому в соответствии с (2.29) параметры $a(t)$ и $\sigma(t)$ аккумулируют информацию о текущем балансе между размерами заявок и интенсивностями покупателей и продавцов, в то время как в соответствии с (2.28) параметры $\nu(t)$, $\mu(t)$, $\lambda(t)$ зависят только от процесса $r(t)$. Прогнозирование динамики статистических характеристик такого процесса и соответствующих рисков сводится к анализу траектории точки в пятимерном параметрическом пространстве. Для этих целей могут использоваться многомерные авторегрессионные модели.

Глава 3

Токсичность потока заявок

3.1 Введение

Активное развитие электронной торговли на финансовых рынках выявило необходимость анализа биржевых высокочастотных данных для более глубокого понимания рыночной микроструктуры, на которую оказали огромное влияние компании, занимающиеся автоматизированным высокочастотным трейдингом (они формируют до 70–80% дневного оборота на ведущих мировых площадках). Эти высокочастотные системы, как правило, являются маркет-мейкерами – поставщиками ликвидности посредством размещения пассивных (лимитных) заявок на различных уровнях электронной книги заявок. Поставщик ликвидности, выставивший пассивную заявку, не имеет возможности влиять на время её исполнения (разумеется, кроме как снять заявку). Маркет-мейкеры зачастую не прогнозируют в явном виде динамику рынка, а используют шумовую составляющую рыночных движений. Степень эффективности деятельности маркет-мейкеров связана с контролем риска оказаться с большим количеством купленных или проданных контрактов, что напрямую зависит от их способности контролировать эффект неблагоприятного отбора (adverse selection) в отношении пассивных заявок.

Практики, как правило, описывают принцип неблагоприятного отбора как «естественную тенденцию слишком быстрого исполнения пассивных заявок в тех ситуациях, когда они должны исполняться медленно, и наоборот: исполняться слишком медленно в тех ситуациях, когда они должны исполниться быстро» ([60]). Эта интуитивная формулировка согласуется с ранними микроструктурными моделями рынка [35, 47, 78], в которых информированные трейдеры получают преимущество над неинформированными участниками рынка. Поток заявок считается токсичным, когда происходит эффект неблагоприятного отбора маркет-мейкеров, поставляющих ликвидность.

В работе [36] предложена эмпирическая процедура оценки токсичности потока заявок на основе анализа информации о *сделках*. В данной главе рассматривается более точный подход к измерению токсичности рынка, использующий всю доступную информацию о *потоке заявок* (не только сами сделки, но также и постановки/снятия заявок) на основе аналитической модели процесса дисбаланса потока заявок, рассмотренной ранее в работах [28, 76].

3.2 Модель потоков заявок

В Главе 2 в качестве математической модели эволюции процесса дисбаланса потока заявок было предложено использовать двусторонний процесс риска – специальный обобщенный (compound) пуассоновский процесс. Следуя этому подходу, зафиксируем малый интервал времени $[0; T]$, в течение которого параметры распределений, описывающих объёмы заявок, и интенсивности потоков заявок одного типа остаются постоянными и известными. Для $t \in [0, T]$ пусть $N^+(t)$ и $N^-(t)$ – количества заявок, пришедших от покупателей и продавцов соответственно в течение интервала времени $[0, t]$ – независимые пуассоновские процессы с интенсивностями $\lambda^+ > 0$ и $\lambda^- > 0$ ($\mathbb{E}N^+(t) = \lambda^+t$, $N^+(0) = 0$, $\mathbb{E}N^-(t) = \lambda^-t$, $N^-(0) = 0$).

Пусть X_i^+ и X_i^- , $i = 1, 2, \dots$, – объёмы заявок, поступающих от покупателей и продавцов соответственно – две независимые последовательности независимых и одинаково в каждой последовательности распределенных случайных величин с функциями распределения $G(x)$ и $F(x)$ соответственно, независимых от пуассонских процессов $N^+(t)$ и $N^-(t)$. Положим

$$Q^+(t) = \sum_{i=1}^{N^+(t)} X_i^+, \quad Q^-(t) = \sum_{j=1}^{N^-(t)} X_j^-$$

и определим процесс *дисбаланса потока заявок* $Q(t)$ как

$$Q(t) = Q^+(t) - Q^-(t).$$

3.3 Профиль мгновенной токсичности потока заявок

Как уже было сказано выше, поток заявок считается токсичным, когда он оказывается неблагоприятным для маркет-мейкеров, предоставляющих ликвидность в книге заявок. В работе [36] предложена процедура оценки токсичности потока заявок на основе анализа информации об интенсивности и направлении *сделок* (направление сделки определяется в зависимости от того, кто являлся её инициатором – покупатель или продавец). В данном разделе будет предложен более точный подход к измерению токсичности потока заявок, использующий всю доступную информацию о заявках (не только сами сделки, но также и постановки/снятия заявок).

Чтобы формализовать понятие токсичности потока заявок, для начала рассмотрим процесс дисбаланса потока заявок $Q(t) = Q^+(t) - Q^-(t)$ в предположении,

что $\mathbb{E}Q(t) > 0$, т. е.

$$\lambda^+ \mathbb{E}X_1^+ > \lambda^- \mathbb{E}X_1^-,$$

что означает преимущество покупателей над продавцами в рамках интервала $[0, T]$. Предположим, что $Q(0) = 0$.

Для $u > 0$ рассмотрим вероятность

$$\varphi_{\pm}(u, T) = \mathbb{P}(\inf_{0 < t \leq T} Q(t) \geq -u),$$

т. е. вероятность того, что траектория процесса $Q(t)$ в течение интервала времени $[0, T]$ целиком будет находиться не ниже уровня $-u$, а также аналогичную предельную вероятность на бесконечном интервале времени:

$$\varphi_{\pm}(u) = \mathbb{P}(\inf_{t > 0} (Q^+(t) - Q^-(t)) \geq -u) = \lim_{T \rightarrow \infty} \varphi_{\pm}(u, T).$$

Вероятность $\varphi_{\pm}(u)$ описывает вероятность того, что при *положительном* тренде процесс дисбаланса никогда не достигнет *отрицательного* уровня $-u$ при условии, что параметры потока заявок $(\lambda^+, \lambda^-, G(x)$ и $F(x)$) остаются неизменными.

ОПРЕДЕЛЕНИЕ 1. *Функцию $\varphi_{\pm}(u)$ будем называть **профилем мгновенной токсичности** потока заявок.*

Введенная таким образом характеристика – профиль мгновенной токсичности потока заявок – формально совпадает с *вероятностью неразорения* в классической модели коллективного риска со случайными премиями, рассматривавшейся, например, в работах [23, 24, 92]. В некоторых источниках, см., в частности, [69], справедливо отмечено, что интерпретация этого показателя именно как вероятности физического разорения страховой компании некорректна, поскольку изначальное предположение о неизменности основных параметров потоков страховых премий и страховых выплат в течение бесконечного интервала времени заведомо

не выполняется. Тем не менее эта характеристика является удобным показателем текущей эффективности работы страховой компании и имеет смысл некоей оценки качества текущего состояния параметров страховой деятельности. Точно так же профиль мгновенной токсичности потока заявок является удобно интерпретируемым показателем неустойчивости текущего состояния потоков заявок.

Из работ [23, 24] следует

ЛЕММА 3.1. *Функция профиля мгновенной токсичности потока заявок $\varphi_{\pm}(u)$ удовлетворяет интегральному уравнению*

$$(\lambda^+ + \lambda^-)\varphi_{\pm}(u) = \lambda^- \int_0^u \varphi_{\pm}(u-v)dF(v) + \lambda^+ \int_0^{\infty} \varphi_{\pm}(u+v)dG(v).$$

Если R – решение характеристического уравнения

$$\lambda^+(\mathbb{E}e^{-RX_1^+} - 1) + \lambda^-(\mathbb{E}e^{RX_1^-} - 1) = 0,$$

то

$$\varphi_{\pm}(u) = \frac{e^{-Ru}}{\mathbb{E}\{e^{-RQ(t)}|\tau < \infty\}},$$

при этом $\varphi_{\pm}(u) \geq 1 - e^{-Ru}$.

3.4 Токсичность потока заявок

Профиль токсичности представляет собой *функцию*, аргументом которой является уровень u . Это затрудняет сравнение токсичности потоков заявок на разных участков рынка, поскольку, вообще говоря, в множестве функций нельзя ввести отношение порядка. Поэтому хотелось бы иметь некий интегральный показатель токсичности, выражаемый одним числом. Для построения такого показателя можно воспользоваться одним из двух подходов.

3.4.1 Байесовский подход

Выделим некий «характеристический» уровень u_0 , пересечение которого может иметь серьезные последствия. Пусть $w(x)$ – некоторая плотность распределения вероятностей, обладающая свойствами

$$\int_0^\infty w(x)dx = 1, \quad \int_0^\infty xw(x)dx = u_0. \quad (3.1)$$

ОПРЕДЕЛЕНИЕ 2. *Байесовским показателем мгновенной токсичности потока заявок $\theta_\pm^{(w)}$ называется величина*

$$\theta_\pm^{(w)} = \theta_\pm^{(w)}(u_0) = \int_0^\infty \varphi_\pm(u)w(u)du.$$

По сути показатель мгновенной токсичности потока заявок θ_\pm есть математическое ожидание «случайного» профиля мгновенной токсичности $\varphi_\pm(U)$, где U – неотрицательная случайная величина с плотностью распределения $w(x)$ и имеющая математическое ожидание u_0 .

В случае, когда $\mathbf{E}Q(t) < 0$, т. е. $\lambda^+ \mathbf{E}X_1^+ < \lambda^- \mathbf{E}X_1^-$, что означает преимущество продавцов над покупателями на интервале $[0, T]$, вместо $\varphi_\pm(u)$ будем рассматривать вероятность

$$\varphi_\mp(u) = \mathbf{P}(\sup_{t>0} (Q^+(t) - Q^-(t)) \leq u) = \lim_{T \rightarrow \infty} \varphi_\mp(u, T),$$

которая описывает вероятность того, что при *отрицательном* тренде траектория процесса $Q(t)$ не превысит *положительный* уровень u при условии, что параметры потока заявок (λ^+ , λ^- , $G(x)$ и $F(x)$) остаются неизменными.

В таком случае в качестве байесовского показателя **мгновенной токсичности**

сти потока заявок θ возьмем величину

$$\theta_{\mp}^{(w)} = \theta_{\mp}^{(w)}(u_0) = \int_0^{\infty} \varphi_{\mp}(u)w(u)du.$$

3.4.2 Квантильный подход

При условии $\mathbf{E}Q(t) > 0$ на промежутке $[0, T]$ зафиксируем некоторое $0 < \alpha < 1$.

ОПРЕДЕЛЕНИЕ 3. *Квантильным α -показателем мгновенной токсичности потока заявок называется такое минимальное значение q_{\pm} , при котором*

$$\varphi_{\pm}(q_{\pm}) \geq \alpha$$

Таким образом, при наличии положительного тренда у процесса $Q(t)$ квантильный α -показатель мгновенной токсичности – это настолько минимальное значение q_{\pm} , что вероятность того, что траектория процесса $Q(t)$ на интервале $[0, T]$ целиком пройдёт выше уровня $-q_{\pm}$ больше или равна α . Чем больше значение q_{\pm} , тем более токсичен поток заявок от покупателей.

По аналогии с предыдущим пунктом при наличии у процесса $Q(t)$ отрицательного тренда (т. е. при условии $\lambda^+ \mathbf{E}X_1^+ < \lambda^- \mathbf{E}X_1^-$) α -квантильный показатель мгновенной токсичности q_{\mp} определяется из уравнения

$$\varphi_{\mp}(q_{\mp}) \geq \alpha.$$

Чем больше значение q_{\mp} , тем более токсичен поток заявок от продавцов на интервале $[0, T]$.

3.5 Модели потоков заявок

В некоторых случаях удаётся напрямую вычислить профиль мгновенной токсичности потока заявок. Аналоги моделей, приведённых ниже, рассмотрены в работе [23] в рамках модели Крамера–Лундберга со стохастическими премиями.

3.5.1 Модель рынка с заявками единичного объёма

Рассмотрим простейшую модель рынка, где потоки заявок имеют единичный объём, т. е.

$$\mathbb{P}(X_i^+ = 1) = \mathbb{P}(X_i^- = 1) = 1.$$

В этом случае

$$Q(t) = \sum_{i=1}^{N^+(t)} 1 - \sum_{i=1}^{N^-(t)} 1 = N^+(t) - N^-(t).$$

Несмотря на очевидно идеальный характер такого примера, он имеет реальный практический смысл, поскольку при этом становится возможным учитывать чистые интенсивности потоков заявок и отслеживать влияние их соотношения (дисбаланса интенсивностей потоков заявок) на токсичность ситуации. Более того, в таком случае рассматриваемый процесс дисбаланса потоков заявок $Q(t)$ является простейшим процессом рождения и гибели, различные характеристики которого можно исследовать специально разработанными для этого методами.

Если $\lambda_+ > \lambda_-$, то покупатели преобладают над продавцами и характеристическое уравнение имеет вид

$$\lambda^+ [e^{-R} - 1] + \lambda^- [e^R - 1] = 0,$$

откуда $e^R = \frac{\lambda^+}{\lambda^-}$ или $e^R = 1$. По Лемме 3.1 для $u > 0$ имеем

$$\varphi_{\pm}(u) \geqslant 1 - \left(\frac{\lambda^-}{\lambda^+}\right)^u, \quad \varphi_{\pm}(\infty) = 1.$$

Равенство $\varphi_{\pm}(u) = \varphi_{\pm}([u])$ очевидно. Для целых u интегральное уравнение переходит в разностное

$$\lambda_1 \varphi_{\pm}(u+1) - (\lambda^- + \lambda^+) \varphi_{\pm}(u) + \lambda \varphi_{\pm}(u-1) = 0, \quad (3.2)$$

откуда $\varphi_{\pm}(u) = C_1 + C_2 \left(\frac{\lambda^-}{\lambda^+}\right)^u$, $C_1 = \varphi(\infty) = 1$. Константу C_2 найдём при подстановке в уравнение (3.2) $u = 0$:

$$(\lambda^- + \lambda^+) \varphi(0) = \lambda \varphi(1), \quad C_2 = -\frac{\lambda^-}{\lambda^+},$$

откуда получаем, что для $u > 0$ профиль мгновенной токсичности имеет вид

$$\varphi_{\pm}(u) = \varphi_{\pm}([u]) = 1 - \left(\frac{\lambda^-}{\lambda^+}\right)^{[u]+1}.$$

Байесовский показатель мгновенной токсичности

Коль скоро исследователь может сам назначать уровень u_0 , относительно которого будут рассчитываться характеристики токсичности потока заявок, будем рассматривать u_0 на множестве натуральных чисел, а в качестве функции $w(u)$ можно выбрать функцию плотности вероятности распределения Пуассона (относительно считающей меры), также определённую на множестве натуральных чисел. Для

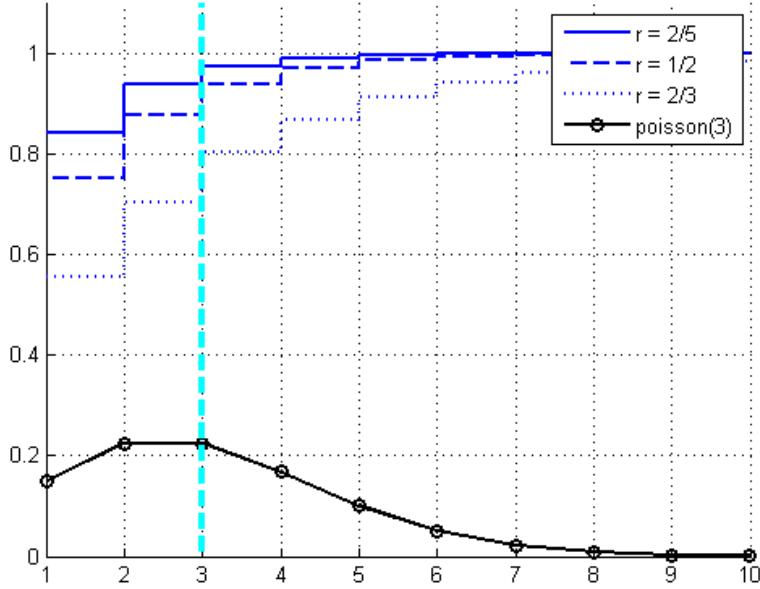


Рис. 3.1: Функция профиля токсичности потока заявок $\varphi_{\pm}(u)$ в модели рынка с единичными потоками заявок для разных значений $r = \frac{\lambda^-}{\lambda^+}$. Чёрным цветом нарисована функция $w(u)$ – плотность (относительно считающей меры) пуассоновского распределения со средним $u_0 = 3$

удобства обозначим $r = \frac{\lambda^-}{\lambda^+}$, при этом $r < 1$. В таком случае

$$\begin{aligned} \theta_{\pm}^{(w)}(u_0) &= \int_{\mathbb{N}} \varphi_{\pm}(u) w(u) du = \sum_{k=0}^{\infty} (1 - r^{k+1}) \frac{u_0^k e^{-u_0}}{k!} = 1 - r e^{-u_0} \sum_{k=0}^{\infty} \frac{(r u_0)^k}{k!} = \\ &= 1 - r e^{u_0(r-1)}. \end{aligned}$$

На рис. 3.2а изображён график токсичности в зависимости от значений $r = \frac{\lambda^-}{\lambda^+}$ для фиксированного значения $u_0 = 3$ в условиях положительного тренда ($\lambda^+ > \lambda^-$).

Чем меньше значение r , тем токсичнее рынок. И напротив: рынок нетоксичен, когда $\lambda^+ = \lambda^-$, то есть наблюдается баланс между покупателями и продавцами.

Квантильный показатель мгновенной токсичности

Для заданного $\alpha \in (0, 1)$ квантильный показатель мгновенной токсичности – это такое минимальное $q_{\pm} \in \mathbb{N}$, при котором

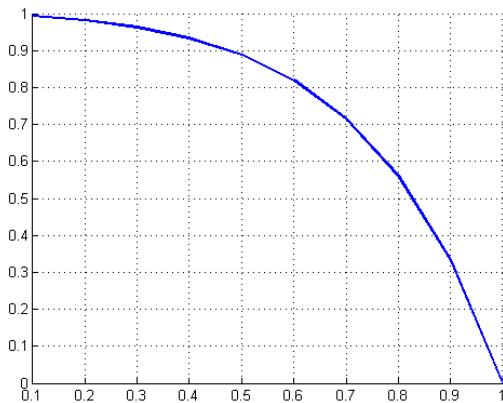
$$\varphi_{\pm}(q_{\pm}) = 1 - r^{q_{\pm}+1} \geq \alpha,$$

откуда

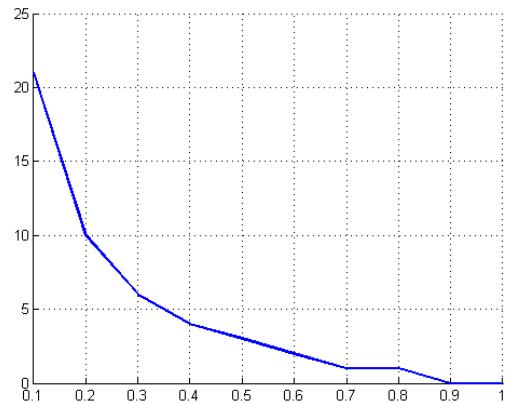
$$q_{\pm}(\alpha) = \left\lceil \frac{\ln(1 - \alpha)}{\ln r} - 1 \right\rceil$$

Заметим, что при $r = 1$ α -квантильный показатель мгновенной токсичности не определён и в таком случае полагается равным нулю.

На рис. 3.2b на график нанесены различные значения квантильного показателя мгновенной токсичности потока заявок в зависимости от значения $r = \frac{\lambda^-}{\lambda^+}$ на промежутке $[0, T]$. Токсичность покупателей максимальна при малых значениях r и близка к нулю при наличии баланса между покупателями и продавцами. Монотонность обоих графиков по r подтверждает обоснованность использования



(a) Байесовский подход



(b) Квантильный подход

Рис. 3.2: Графики показателей токсичности в зависимости от значения $r = \frac{\lambda^-}{\lambda^+}$.

θ_{\pm} и q_{\pm} в качестве показателей токсичности потока заявок в случае модели рынка с заявками единичного объёма.

3.5.2 Модель рынка с экспоненциальными объёмами заявок

Пусть объёмы заявок покупателей и продавцов имеют экспоненциальное распределение, т. е.

$$G(t) = 1 - e^{-bt} \text{ и } F(t) = 1 - e^{-at}.$$

В случае, когда покупатели преобладают над продавцами, т. е. $\lambda^+/b > \lambda^-/a$, характеристическое уравнение имеет вид

$$\lambda^+ \left[\frac{b}{b+R} - 1 \right] + \lambda^- \left[\frac{a}{a-R} - 1 \right] = 0,$$

откуда $R = (\lambda^+a - \lambda^-b)/(\lambda^+ + \lambda^-)$ или 0, а профиль мгновенной токсичности потока заявок ([23])

$$\varphi_{\pm}(u) = \frac{(a+b)\lambda^-}{(\lambda^+ + \lambda^-)a} \exp \left(-\frac{\lambda^+a - \lambda^-b}{\lambda^+ + \lambda^-} u \right).$$

В случае, когда продавцы преобладают над покупателями

$$\varphi_{\mp}(u) = \frac{(a+b)\lambda^+}{(\lambda^+ + \lambda^-)b} \exp \left(-\frac{\lambda^-b - \lambda^+a}{\lambda^+ + \lambda^-} u \right).$$

Байесовский показатель мгновенной токсичности

На множестве функций $w(u)$, удовлетворяющих условиям (3.1), рассмотрим функции, удовлетворяющие также условию

$$\int_0^\infty x^2 w(x) dx - \left(\int_0^\infty x w(x) dx \right)^2 = 1, \quad (3.3)$$

то есть обеспечивающие единичную дисперсию соответствующей случайной величины, имеющей функцию $w(u)$ в качестве плотности своего распределения.

Для вычисления байесовского показателя мгновенной токсичности возьмём в качестве $w(u)$ плотность гамма-распределения

$$w(u) = u^{k-1} \frac{e^{-u/\theta}}{\theta^k \Gamma(k)},$$

где $\Gamma(k)$ – гамма-функция Эйлера:

$$\Gamma(k) = \int_0^{+\infty} t^{k-1} e^{-t} dt.$$

Поскольку математическое ожидание и дисперсия случайной величины U , имеющей гамма-распределение, равны $k\theta$ и $k\theta^2$ соответственно, то с учётом условий (3.1) и (3.3) значения k и θ определяются из уравнений $k\theta = u_0$ и $k\theta^2 = 1$, откуда $k = u_0^2$ и $\theta = u_0^{-1}$.

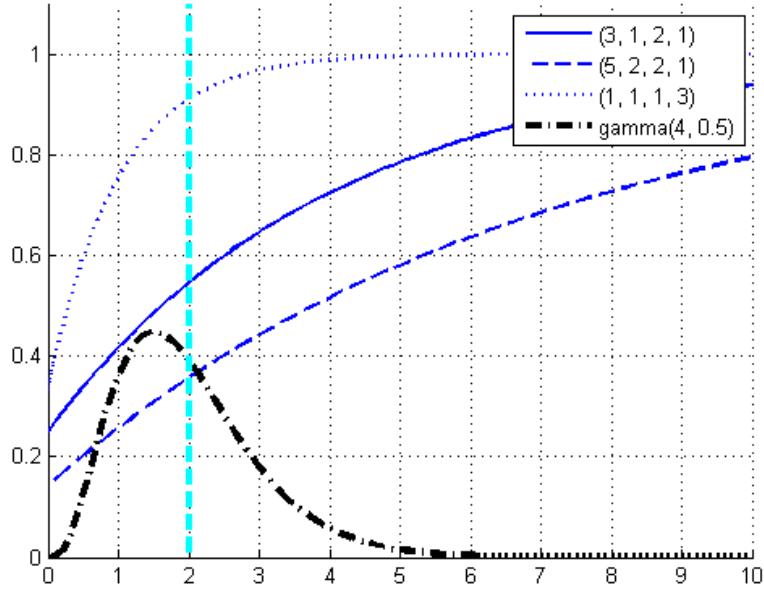


Рис. 3.3: Функция профиля токсичности $\varphi_\pm(u)$ в модели рынка с экспоненциальными объёмами заявок для разных наборов $(\lambda^+, \lambda^-, b, a)$. Чёрным цветом показана весовая функция $w(u)$ – плотность гамма-распределения $\Gamma(u_0^2, u_0^{-1})$ при $u_0 = 2$

Для удобства обозначим

$$\beta = \frac{(a+b)\lambda^-}{(\lambda^+ + \lambda^-)a} \text{ и } \gamma = \frac{\lambda^+ a - \lambda^- b}{\lambda^+ + \lambda^-} > 0. \quad (3.4)$$

Байесовский показатель мгновенной токсичности равен

$$\begin{aligned} \theta_{\pm}^{(w)}(u_0) &= \int_0^{\infty} \varphi_{\pm}(u) w(u) du = \int_0^{\infty} (1 - \beta e^{-\gamma u}) u^{k-1} \frac{e^{-u/\theta}}{\theta^k \Gamma(k)} du = \\ &= 1 - \frac{\beta}{\theta^k \Gamma(k)} \int_0^{\infty} e^{-(\gamma + \theta^{-1})u} u^{k-1} du = \left[t = \frac{\theta\gamma + 1}{\theta} u \right] = \\ &= 1 - \frac{\beta}{(\theta\gamma + 1)^k \Gamma(k)} \int_0^{\infty} e^{-t} \frac{\theta^{k-1}}{(\theta\gamma + 1)^{k-1}} t^{k-1} \frac{\theta}{\theta\gamma + 1} dt = \\ &= 1 - \frac{\beta}{(\theta\gamma + 1)^k}. \end{aligned}$$

После подстановки k и θ получаем значение показателя

$$\theta_{\pm}(u_0) = 1 - \frac{\beta}{(\gamma u_0^{-1} + 1)^{u_0^2}}.$$

Квантильный показатель мгновенной токсичности

Для заданного $\alpha \in (0, 1)$ квантильный показатель мгновенной токсичности – это такое минимальное q_{\pm} , при котором

$$\varphi_{\pm}(q_{\pm}) = 1 - \beta e^{-\gamma q_{\pm}} \geq \alpha.$$

Так как функция φ_{\pm} является непрерывной по q_{\pm} , то данное неравенство может быть обращено в равенство, откуда получаем

$$q_{\pm}(\alpha) = \frac{\ln \beta - \ln(1 - \alpha)}{\gamma}.$$

Заметим, что байесовский и квантильный показатели токсичности являются монотонными по каждой из величин β и γ .

3.6 Реальные данные

В данном разделе описывается структура данных о потоках заявок, на базе которых можно провести валидацию модели, предложенной в разделе 6.2. Далее оценим параметры потока заявок $\lambda^+, \lambda^-, b, a$ в режиме скользящего окна и рассчитаем функции профиля мгновенной токсичности, а также показатели мгновенной токсичности потока заявок $\theta(t)$ и $q(t)$ в режиме реального времени и проанализируем адекватность полученных характеристик.

3.6.1 Описание данных

Как и в Главе 1, рассматриваются данные о потоках всех заявок (лимитных, рыночных и заявок на отмену) на первые $d = 5$ уровней книги заявок фьючерса на индекс РТС за период с 1 по 30 июля 2014 г.

3.6.2 Процедура оценки параметров

Разобьём один из рассматриваемых торговых дней (1 июля 2014 г.) на временные интервалы с шагом $\tau = 15$ секунд. При этом исключим интервалы времени в первые пять минут торгов (с 10:00 до 10:05), а также в последние пять минут торгов (с 18:40 до 18:45), поскольку они характеризуются аномальными всплесками волатильности, слабо поддающейся анализу в рамках представленной модели. Внутри каждого τ -интервала проведём оценку параметров λ^+, λ^-, b и a согласно модели рынка с экспоненциальными объёмами заявок. Результат оценки параметров в режиме реального времени изображён на рис. 3.4.

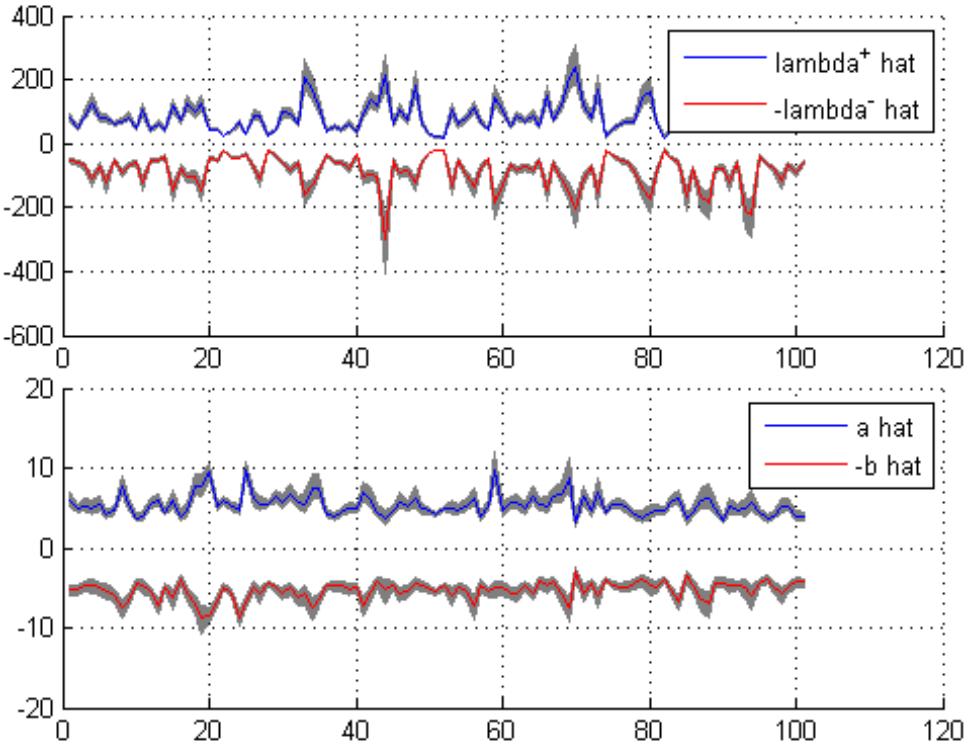


Рис. 3.4: Оценка параметров $\lambda^+, \lambda^-, b, a$ в режиме реального времени (серый цвет – интервалы доверия), ось x – номер соответствующего τ -интервала (фьючерс на индекс РТС, дневная сессия 2014.07.01)

3.6.3 Показатели токсичности

На основе оценок для $\lambda^+, \lambda^-, b, a$ можно вычислить β и γ , а затем построить графики показателей мгновенной токсичности потока заявок $\theta(u_0)$ и $q(\alpha)$ для фиксированных u_0 и α в режиме реального времени (см. рис. 3.5) и идентифицировать участки, на которых деятельность покупателей или продавцов была токсичной. Прикладные исследования демонстрируют достаточную значимость данного показателя для своевременной идентификации участков неблагоприятного отбора маркет-мейкеров.

Таким образом, в данном разделе рассмотрена микроструктурная модель рынка, в которой потоки заявок моделируются пуассоновскими процессами с постоянными интенсивностями (такая аппроксимация возможна на небольших временных

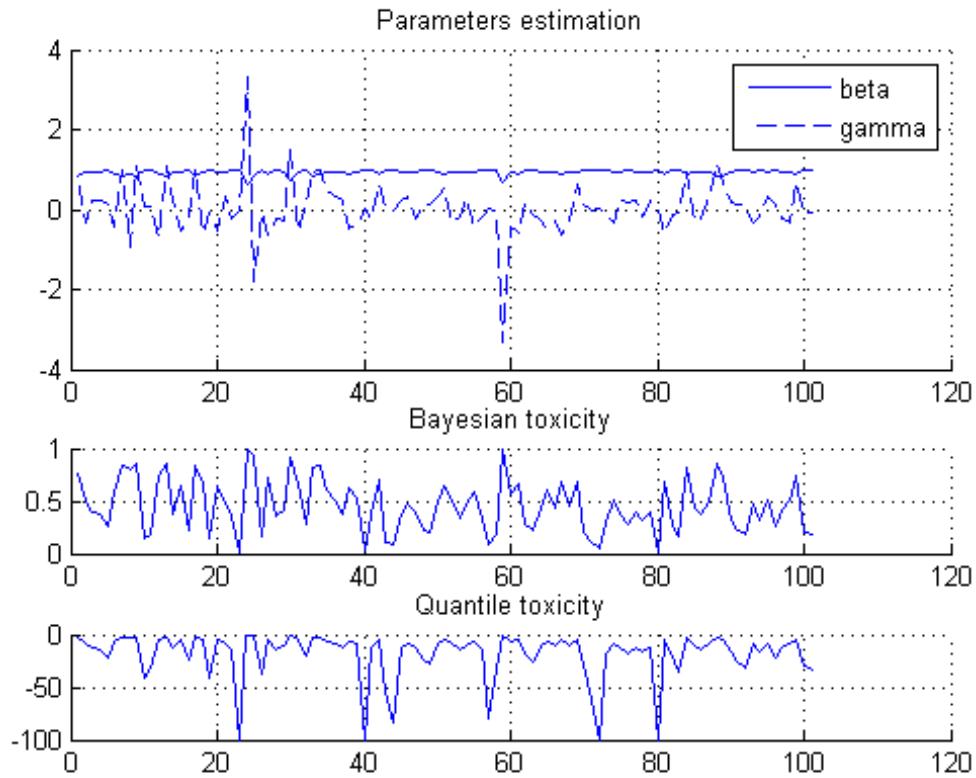


Рис. 3.5: Графики β , γ , рассчитанных по формулам (3.4), байесовского и квантильного показателей токсичности в режиме реального времени, ось x – номер соответствующего τ -интервала (фьючерс на индекс РТС, дневная сессия 2014.07.01)

ных интервалах). В качестве интегрального индикатора текущего состояния книги заявок применялся дисбаланс потока заявок (order flow imbalance), который использует не только текущие значения наилучших цен покупки и продажи, но и влияние событий «в глубине» книги заявок и потому меняется существенно быстрее и позволяет интерполировать динамику рынка между изменениями цены, в частности, отслеживать ситуации, связанные с токсичностью потока заявок. В рамках рассмотренной модели были введены такие понятия, как мгновенный профиль токсичности, а также байесовский и квантильный показатели токсичности, рассчитываемые на основе параметров, описывающих потоки всех заявок. Эти показатели рассчитываются для двух модельных типов потоков заявок, в первом из которых заявки имеют единичный объём, во втором – объём заявок является слу-

чайным и имеющим показательное распределение. Для последней из двух моделей была проведена валидация на реальных данных (фьючерс на индекс РТС) и были построены показатели токсичности в режиме реального времени. Предложенная методика расчёта показателей токсичности на основе информации о потоках всех заявок является перспективной и может быть распространена на модели рынка с неоднородными интенсивностями потоков заявок.

Заключение

Данная работа посвящена разработке модели неоднородных и нестационарных потоков заявок в специализированных системах обслуживания с использованием обобщённых процессов риска, а также её теоретическому и статистическому исследованию.

Основные результаты и выводы В работе показано, что процесс дисбаланса потоков заявок вполне естественно трактовать как так называемый двусторонний процесс риска - процесс риска со случайными премиями. При этом в качестве формальной математической модели последнего предложено использовать специальный обобщенный дважды стохастический пуассоновский процесс (обобщенный процесс Кокса). При обосновании указанной модели определяющую роль играет предложенное мультипликативное представление накопленных интенсивностей, согласно которому интенсивности потоков заявок на покупку и продажу с точностью до неслучайного множителя пропрциональны одному и тому же случайному процессу, характеризующему общий ажиотаж на рынке. Это модельное представление получило в диссертации как теоретическое, так и статистическое обоснование.

Основным теоретическим результатом диссертации являются предельные теоремы для процесса дисбаланса потоков заявок. В работе были получены теоремы переноса для одномерных распределений и функциональные предельные теоремы

в реалистичных условиях существования вторых моментов элементарных скачков. В прежних работах были доказаны функциональные предельные теоремы для обобщенных процессов Кокса, у которых скачки имеют конечную дисперсию, а управляющие процессы удовлетворяют условию типа Гельдера с показателями, меньшими единицы. В данной работе вышеупомянутые результаты распространены на несимметричный случай и приведен критерий сходимости обобщенных процессов Кокса и, в частности, двухсторонних процессов риска, описывающих процесс дисбаланса потоков заявок, к обобщенным гиперболическим процессам Леви.

К числу основных результатов диссертации следует отнести предложенную формализацию понятия понятия токсичности потоков заявок. В рамках предложенного подхода, основанного на формальной тождественности понятия токсичности потока заявок и понятия вероятности неразорения в процессах риска со случайными премиями был разработан и реализован индикатор токсичности потока заявок, способный предсказывать ценовые шоки и токсичную ликвидность. Также был предложен соответствующий метод вычисления такого индикатора и проведено его тестирование на финансовых данных. Данный индикатор может быть внедрён в системы риск-менеджмента различных финансовых институтов, а также может быть успешно использован финансовыми регуляторами для контроля качества финансовой экосистемы.

Дальнейшие перспективы развития исследований

В данной работе используется математическая модель двухсторонних процессов риска, использование которой может быть распространено на широкий спектр задач, включая анализ поведения высоконагруженных телекоммуникационных систем и различных систем массового обслуживания. Теоретические результаты, полученные в данной работе могут быть распространены на многомерный случай, поскольку интенсивности потоков заявок на фиксированном финансовом активе

также зависят от динамики интенсивностей потоков заявок на связанных активах и потому представляется возможным аналитически описать часть динамики внешнего информационного фона $\Lambda(t)$, выступающего в качестве управляющего процесса в определении процесса дисбаланса потоков заявок, а также учесть наблюдаемые кросс-корреляционные эффекты между активами. Показатель токсичности, рассмотренный в третьей главе, может быть очевидным образом адаптирован для модели потоков заявок с неоднородными интенсивностями, а также по сравнению с существующими подходами более тонко настроен для моделей потоков заявок с распределениями объёмов заявок, отличными от экспоненциальных.

Литература

- [1] *Abergel F., Jedidi A.* A mathematical approach to order book modelling // *Econophysics of Order-Driven Markets* / Abergel F., Chakrabarti B. K., Chakraborti A., Mitra M. (Eds.) *Econophysics of Order-Driven Markets*. – New York: Springer, 2011. P. 93–108.
- [2] *Abergel F., Jedidi A.* A mathematical approach to order book modelling // *Econophysics of Order-Driven Markets* / Abergel F., Chakrabarti B. K., Chakraborti A., Mitra M. (Eds.) *Econophysics of Order-Driven Markets*. – New York: Springer, 2011. P. 93–108.
- [3] *Alfonsi A., Schied A., and Schulz A.* Optimal execution strategies in limit order books with general shape function // *Quantitative Finance*, 10 (2010), pp. 143–157.
- [4] *Almgren R. and Chriss N.* Optimal execution of portfolio transactions, *Journal of Risk*, 3 (2000), pp. 5–39.
- [5] *Andersen T. G. and Bondarenko O.* Reflecting on the VPIN Dispute (August 4, 2013) // *Journal of Financial Markets*, Vol. 17, pp. 53-64, 2014.
- [6] *M. Avellaneda and Stoikov S.*, High-frequency trading in a limit order book // *Quantitative Finance*, 8 (2008), pp. 217–224.

- [7] *Barndorff-Nielsen O. E.* Exponentially decreasing distributions for the logarithm of particle size // Proc. Roy. Soc. London. Ser. A, 1977. Vol. A(353). P. 401–419.
- [8] *Barndorff-Nielsen O. E.* Hyperbolic distributions and distributions of hyperbolae // Scand. J. Statist., 1978. Vol. 5. P. 151–157.
- [9] *Barndorff-Nielsen O. E.* Models for non-Gaussian variation, with applications to turbulence // Proc. Roy. Soc. London. Ser. A, 1979. Vol. A(368). P. 501–520.
- [10] *Barndorff-Nielsen O. E., Kent J., Sørensen M.* Normal variance-mean mixtures and z-distributions // International Statistical Review, 1982. Vol. 50. No. 2. P. 145–159.
- [11] *Barndorff-Nielsen O. E., Blæsild P., Schmiegel J.* A parsimonious and universal description of turbulent velocity increments // European Physical Journal, 2004. Vol. B. 41. P. 345–363.
- [12] *Barndorff-Nielsen O. E.* Processes of normal inverse Gaussian type // Finance and Stochastics, 1998. Vol. 2. P. 41–18.
- [13] *Barndorff-Nielsen O. E., Mikosch T., Resnick S. I.* Lévy Processes: Theory and Applications. – Boston: Birkhäuser, 2001.
- [14] *Bening V., Korolev V.* Generalized Poisson Models and Their Applications in Insurance and Finance. – Utrecht, VSP, 2002.
- [15] *Bertoin J.* Lévy Processes. Cambridge Tracts in Mathematics, Vol. 121. – Cambridge: Cambridge University Press, 1996.
- [16] *D. BERTSIMAS AND A. LO*, Optimal control of execution costs, Journal of Financial Markets, 1 (1998), pp. 1–50.

- [17] *B. BIAIS, L. GLOSTEN, AND C. SPATT*, Market microstructure: A survey of microfoundations, empirical results and policy implications, *Journal of Financial Markets*, 8 (2005), pp. 217–264.
- [18] *Billingsley P.* Convergence of Probability Measures. – New York: Wiley, 1968.
- [19] Bouchaud, J.-P., M. Mezard, M. Potters. 2002. Statistical properties of stock order books: Empirical results and models. *Quant. Finance* 2, 251–256.
- [20] *Bouchaud J.-P., Mezard M., Potters M.* Statistical properties of stock order books: Empirical results and models // *Quant. Finance*, 2002. Vol. 2. P. 251–256.
- [21] *J.-P. Bouchaud, M. Mezard , and M. Potters* Statistical properties of stock order books: empirical results and models, *Quantitative Finance*, 2 (2002), pp. 251–256.
- [22] *J.-P. Bouchaud, D. Farmer, and F. Lillo* How markets slowly digest changes in supply and demand, in *Handbook of Financial Markets: Dynamics and Evolution*, T. Hens and K. Schenk-Hoppe, eds., Elsevier:Academic Press, 2008.
- [23] *Boykov A.* Cramer-Lundberg model with stochastic premiums, *Theory of Probability and its Applications*, 47:3, (2002), 549–553.
- [24] *Бойков А. В.* Стохастические модели капитала страховой компании и оценивание вероятности неразорения. – Дис. канд. физ.-матем. наук. Москва, Математический институт им. В. А. Стеклова РАН, 2003, 83 с.
- [25] *CFTC AND SEC* Findings regarding the market events of May 6, 2010 // Report to the joint advisory committee on emerging regulatory issues, CFTC, 2010.
- [26] *Carr P. P., Madan D. B., Chang E. C.* The Variance Gamma process and option pricing // *European Finance Review*, 1998. Vol. 2. P. 79–105.

- [27] *Chakraborti A., Toke I., Patriarca M., Abrergel F.* Empirical facts and agent-based models // arXiv preprint arXiv:0909.1974, 2009.
- [28] *A. Chertok, V. Korolev, A. Korchagin and S. Shorgin.* Modeling High-Frequency Non-Homogeneous Order Flows by Compound Cox Processes // Available at: arXiv:1410.1900v2 [math.PR], 2014.
- [29] *Clark P. K.* A subordinated stochastic process model with finite variance for speculative prices // *Econometrica*, 1973. Vol. 41. P. 135–155.
- [30] *Cont R., Kukanov A. and Stoikov S.* The price impact of order book events // March 01, 2011. Available at SSRN: <http://ssrn.com/abstract=1712822>
- [31] *Cont R., Kukanov A. and Stoikov S.* The price impact of order book events // *Journal of Financial Econometrics* 12(1), pp. 47-88.
- [32] *Cont R., de Larrard A.* Price dynamics in a Markovian limit order market. – Working paper, 2011. Available at: <http://ssrn.com/abstract=1735338>
- [33] *Cont R., de Larrard A.* Order book dynamics in liquid markets: limit theorems and diffusion approximations // Working paper. Laboratoire de Probabilites et Modeles Aleatoires CNRS, Universite Pierre et Marie Curie (Paris VI) August 2011. Revised February 2012 (hal-00672274, version 2 – 1 October 2012).
- [34] *Cont R., Stoikov S., Talreja R.* A stochastic model for order book dynamics // *Operations Research*, 2010. Vol. 58. No. 3. P. 549–563.
- [35] *Easley D. and M. O'Hara.* 1992. Time and the process of security price adjustment // *Journal of Finance*, 47: 576-605.
- [36] *Easley D., Lopez de Prado, Marcos and O'Hara, Maureen.* Flow Toxicity and

Liquidity in a High Frequency World (February 20, 2012). *Review of Financial Studies*, Vol. 25, No. 5, pp. 1457–1493, 2012.

- [37] *Eberlein E.* Application of Generalized Hyperbolic Lévy Motions to Finance. – Freiburg: Universität Freiburg, Institut für Mathematische Stochastic, 1999. Preprint No. 64.
- [38] *Eberlein E., Keller U.* Hyperbolic Distributions in Finance // Bernoulli, 1995. Vol. 1, No. 3. P. 281–299.
- [39] *Eberlein E., Keller U., Prause K.* New insights into smile, mispricing and value at risk: the hyperbolic model // *Journal of Business*, 1998. Vol. 71. P. 371–405.
- [40] *Eberlein E., Prause K.* The Generalized Hyperbolic Model: Financial Derivatives and Risk Measures. – Freiburg: Universität Freiburg, Institut für Mathematische Stochastic, 1998. Preprint No. 56.
- [41] *Embrechts P., Maejima M.* Selfsimilar Processes. – Princeton: Princeton University Press, 2002.
- [42] *Farmer J. D., Gillemot L., Lillo F., Mike S., and Sen A.*, What really causes large price changes? // *Quantitative Finance*, 4 (2004), pp. 383–397.
- [43] *Foucault T.* Order flow composition and trading costs in a dynamic limit order market // *Journal of Financial Markets*, 1999. Vol. 2. P. 99–134.
- [44] *Foucault T., Kadan O., and Kandel E.*, Limit order book as a market for liquidity // *Review of Financial Studies*, 18 (2005), pp. 1171–1217.
- [45] *Geman H.* Pure jump Lévy processes for asset price modelling // *Journal of Banking and Finance*, 2002. Vol. 26. No. 7. P. 1297–1316.

- [46] *Gleser L. J.* The Gamma Distribution as a Mixture of Exponential Distributions: Technical Report # 87-28. – West Lafayette: Purdue University, 1987. 6 p.
- [47] *Glosten, L. R. and P. Milgrom.* 1985. Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders // Journal of Financial Economics, 14: 71-100.
- [48] *L. GLOSTEN*, Is the limit order book inevitable? // Journal of Finance, 49 (1994), pp. 1127–1161.
- [49] *Gnedenko B. V., Fahim H.* On a transfer theorem // Soviet Math. Dokl., 1969. Vol. 187. No. 1. P. 15–17.
- [50] *Gnedenko B. V., Kolmogorov A. N.* Limit Distributions for Sums of Independent Random Variables. – New York: Addison–Wesley, 1954.
- [51] *Gnedenko B. V., Korolev V. Yu.* Random Summation: Limit Theorems and Applications. – Boca Raton: CRC Press, 1996.
- [52] *Goettler R., Parlour C., Rajan U.* Equilibrium in a dynamic limit order market // Journal of Finance, 2005, **60** 2149–2192.
- [53] *Grandell J.* Doubly Stochastic Poisson Processes. Lecture Notes Mathematics, Vol. 529. – Berlin–Heidelberg–New York: Springer, 1976.
- [54] *Granovsky B. L., Zeifman A. I.* The decay function of nonhomogeneous birth and death processes, with application to mean-field models // Stochastic Process. Appl., 1997, Vol. 72. P. 105–120.
- [55] *J. Hasbrouck.* Empirical Market Microstructure. – Oxford University Press, 2007.
- [56] *Hawkes A. G.* Spectra of some self-exciting and mutually exciting point processes // Biometrika, 1971. Vol. 58. No. 1. P. 83–90.

- [57] *Hewlett P.* Clustering of order arrivals, price impact and trade path optimization // Working paper. University of Oxford, 2006.
- [58] *Huang H., Kercheval A. N.* A generalized birth–death stochastic model for high-frequency order book dynamics // Quantitative Finance 12.4 (2012): 547-557.
- [59] *Jacod J., Shiryaev A. N.* Limit theorems for stochastic processes. 2nd edition. Volume 288 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. – Berlin: Springer-Verlag, Berlin, 2003.
- [60] *Jeria, D. and G. Sofianos.* Passive Orders and Natural Adverse Selection // Street Smart, 33 (September 4, 2008).
- [61] *Kashcheev D. E.* Modeling the Dynamics of Financial Time Series and Evaluation of Derivative Securities. PhD Thesis. – Tver: Tver State University, 2001 (in Russian).
- [62] *Kirilenko A., Kyle A., Samadi M., and Tuzun T.*, The flash crash: The impact of high frequency trading on an electronic market // Available at SSRN: <http://ssrn.com/abstract=1686004>, 2010.
- [63] *Korolev V. Yu.* Convergence of random sequences with the independent random indices. I // Theory Probab. Appl., 1994. Vol. 39. No. 2. P. 282–297.
- [64] *Korolev V. Yu.* On convergence of the distributions of random sums of independent random variables to stable laws // Theory Probab. Appl., 1997. Vol. 42. No. 4. P. 818–820.
- [65] *Korolev V. Yu.* On convergence of the distributions of compound Cox processes to stable laws // Theory Probab. Appl., 1998. Vol. 43. No. 4. P. 786–792.

- [66] *Korolev V. Yu.* Asymptotic properties of extrema of compound Cox processes and their application to some problems of financial mathematics // Theory Probab. Appl., 2000. Vol. 45. No. 1, pp. 182–194.
- [67] *Korolev V. Yu.* Probabilistic and Statistical Methods For the Decomposition of Volatility of Chaotic Processes. – Moscow: Moscow State University Publishing House, 2011.
- [68] *Korolev V. Yu.* Generalized hyperbolic laws as limit distributions for random sums // Theory of Probability and Its Applications, 2013. Vol. 58. No. 1. P. 117–132.
- [69] *Королев В. Ю., Бенинг В. Е., Шоргин С. Я.* Математические основы теории риска. 2-е издание, переработанное и дополненное. – Москва: ФИЗМАТЛИТ, 2011. 620 С.
- [70] *Korolev V., Shevtsova I.* An improvement of the Berry–Esseen inequality with applications to Poisson and mixed Poisson random sums // Scandinavian Actuarial Journal, 2012. No. 2. P. 81–105. Available online since 04 June 2010.
- [71] *Korolev V. Yu., Sokolov I. A.* Skew Student distributions, variance gamma distributions and their generalizations as asymptotic approximations // Informatics and Its Applications, 2012. Vol. 6. No. 1. P. 2–10.
- [72] *Zaks L. M., Korolev V. Yu.* Generalized variance gamma distributions as limit laws for random sums // Informatics and Its Applications, 2013. Vol. 7. No. 1. P. 105–115.
- [73] *В. Ю. Королев, Л. М. Закс, А. И. Зейфман.* О сходимости случайных блужданий, порожденных обобщенными процессами Кокса, к процессам Леви // Информатика и ее применения, 2013. Т. 7. Вып. 2. С. 84–91.

- [74] *Korolev V. Yu., Zaks L. M., Zeifman A. I.* On convergence of random walks generated by compound Cox processes to Lévy processes // Statistics and Probability Letters, 2013. Vol. 83. No. 10. P. 2432–2438.
- [75] *Korolev V., Chertok A. and Zeifman A.* Functional Limit Theorems for Order Flow Imbalance Process, 2014. Working paper.
- [76] *Korolev V. Yu., Chertok A. V., Korchagin A. Yu., Gorshenin A. K.* Probabilistic and statistical modeling of information flows in complex financial systems from high-frequency data // Informatics and Its Applications, 2013. Vol. 7. No. 1. P. 12–21.
- [77] *Gorshenin A., Doynikov A., Korolev V., Kuzmin V.* Statistical properties of the dynamics of order books: empirical results // Applied Problems in Theory of Probabilities and Mathematical Statistics Related to Modeling of Information Systems: Abstracts of VI International Workshop. – Moscow.: IPI RAS, 2012. P. 31–51.
- [78] *Kyle, A. S.* 1985. Continuous Auctions and Insider Trading // Econometrica, 53: 1315–1335.
- [79] *Madan D. B., Seneta E.* The variance gamma (V.G.) model for share market return // Journal of Business, 1990. Vol. 63. P. 511–524.
- [80] *Mandelbrot B. B.* Fractals and Scaling in Finance. Discontinuity, Concentration, Risk. – Berlin–Heidelberg: Springer, 1997.
- [81] *A. OBIZHAEVA AND J. WANG*, Optimal trading strategy and supply/demand dynamics, working paper, MIT, 2006.
- [82] *Parlour Ch. A.* Price dynamics in limit order markets // Review of Financial Studies, 1998. Vol. 11. No. 4. P. 789–816.

- [83] *Prause K.* Modeling Financial Data Using Generalized Hyperbolic Distributions. – Freiburg: Universität Freiburg, Institut für Mathematische Stochastic, 1997. Preprint No. 48.
- [84] *Rosu I.* A dynamic model of the limit order book // Rev. Financial Stud., 2009. Vol. 22. P. 4601–4641.
- [85] *Sato K.* Lévy Processes and Infinitely Divisible Distributions. – Cambridge: Cambridge University Press, 1999.
- [86] *Schoutens W.* Lévy Processes in Finance: Pricing Financial Derivatives. – New York: Wiley, 2003.
- [87] *Seshadri V.* Halphen's laws // Kotz, S., Read, C. B., Banks, D. L. (Eds.). Encyclopedia of Statistical Sciences, Update Volume 1. – New York: Wiley, 1997. P. 302–306.
- [88] *Shevtsova I. G.* On the accuracy of the normal approximation to Poisson random sums // Theory Probab. Appl., 2013. Vol. 58. No. 1. P. 152–176.
- [89] *Shiryayev A. N.* Essentials of Stochastic Finance: Facts, Models, Theory. – Singapore: World Scientific. 1999.
- [90] *Sichel H. S.* Statistical evaluation of diamondiferous deposits // Journal of South Afr. Inst. Min. Metall., 1973. Vol. 76. P. 235–243.
- [91] *Smith E., Farmer J. D., GILLEMET L., and Krishnamurthy S.* Statistical theory of the continuous double auction, Quantitative Finance, 3 (2003), pp. 481–514.
- [92] *Темнов Г. О.* Математические модели риска и случайного притока взносов в страховании. – Дис. канд. физ.-матем. наук, С.-Петербург, Санкт-

Петербургский государственный архитектурно-строительный университет,
2004, 102 с.

- [93] *Toke I. M.* Market making in an order book model and its impact on the spread / In: *Econophysics of Order-Driven Markets*. New York – Berlin: Springer Verlag, 2011, 49–64.
- [94] *Zeifman A. I.* Upper and lower bounds on the rate of convergence for nonhomogeneous birth and death processes // *Stochastic Process. Appl.*, 1995. Vol. 59. P. 157–173.
- [95] *Zolotarev V. M.* *One-Dimensional Stable Distributions*. – Providence, R.I.: American Mathematical Society, 1986.
- [96] *Zovko I., Farmer J. D.* The power of patience; A behavioral regularity in limit order placement // *Quant. Finance*, 2002, Vol. 2. P. 387–392.