

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н. Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теории функций и стохастического анализа

**КЛАСТЕРИЗАЦИЯ И ЭФФЕКТЫ ПОСЛЕДЕЙСТВИЯ ДЛЯ
КУРСОВ ВАЛЮТНЫХ ПАР**

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

Студентки 4 курса 412 группы
направления 01.03.02 — Прикладная математика и информатика

механико-математического факультета

Лавровой Полины Романовны

Научный руководитель
старший преподаватель

А. Д. Луньков

Заведующий кафедрой
д. ф.-м. н., доцент

С. П. Сидоров

Саратов 2026

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	3
1 Основное содержание работы	4
ЗАКЛЮЧЕНИЕ	16

ВВЕДЕНИЕ

Финансовые рынки, и особенно валютный рынок, характеризуются высокой волатильностью, нелинейной динамикой и наличием экстремальных событий. Круглосуточный режим торгов, глобальный характер операций, чувствительность к макроэкономическим и политическим шокам формируют сложные вероятностные закономерности, которые не могут быть адекватно описаны классическими предположениями о нормальности распределения доходностей и независимости приращений. Эмпирические исследования выявляют «тяжёлые хвосты», кластеризацию волатильности и долговременную память (последствие). Эти явления связаны с фрактальной структурой рынка и требуют применения специализированных методов анализа.

Моделирование волатильности валютных курсов особенно важно, так как она служит ключевой характеристикой риска при ценообразовании производных инструментов, хеджировании валютных рисков и оценке макроэкономической нестабильности.

Цель работы — изучить методы статистического анализа высокочастотных данных и применить их для анализа динамики курсов трёх валютных пар (EUR/RUB, USD/RUB, CNY/RUB), в частности, осуществить проверку наличия долговременных зависимостей в волатильности и провести анализ адекватности стохастической модели тикового процесса для описания реальных рыночных данных.

Для достижения поставленной цели решаются следующие задачи:

- изучить теоретическую модель тикового процесса и модель распределения интервалов между тиками;
- выполнить обработку и дискретизацию эмпирических данных;
- провести корреляционный анализ приращений логарифмических доходностей;
- применить процедуру деволатилизации и оценить показатель долговременной памяти;
- сравнить статистические свойства трёх валютных пар.

1 Основное содержание работы

Теоретические основы моделирования цен

В первом разделе рассматриваются теоретические основы моделирования цен на финансовых рынках. Основное внимание уделяется описанию тикового процесса, распределению интервалов между тиками и применению распределения Парето для моделирования длительностей между сделками. Такой подход позволяет учесть дискретный характер изменения цен и неравномерность поступления новой информации на рынок.

Цена финансового инструмента изменяется не непрерывно, а скачкообразно в случайные моменты времени, которые принято называть тиками. Последовательность таких моментов: $0 = \tau_0 < \tau_1 < \tau_2 < \dots$ задаёт междутиковые интервалы $\Delta_k = \tau_k - \tau_{k-1}$, в течение каждого из которых цена остаётся постоянной. Мгновенное изменение происходит только в момент тика, что делает траекторию цены кусочно-постоянной.

На рынке одновременно существуют две цены: цена продажи (ask) S_t^a и цена покупки (bid) S_t^b , для упрощения анализа их сводят к единой «средней» цене, используя логарифм геометрического среднего:

$$S_t = \sqrt{S_t^a \cdot S_t^b}. \quad (1)$$

Такой подход позволяет свести задачу к изучению одного ценового процесса, который сохраняет основные свойства реальных данных.

Эволюция цены S_t может быть представлена кусочно-постоянным процессом:

$$S_t = S_0 + \sum_{k \geq 1} s_{\tau_k} I(\tau_k \leq t), \quad (2)$$

где $s_{\tau_k} = S_{\tau_k} - S_{\tau_{k-1}}$ — абсолютное изменение цены в момент k -го тика. Для удобства часто переходят к логарифмической цене $H_t = \ln S_t$, а соответствующие приращения

$$h_{\tau_k} = \ln \frac{S_{\tau_k}}{S_{\tau_{k-1}}} \quad (3)$$

называются логарифмическими доходностями. Они измеряют относительное изменение цены, что устраняет зависимость от абсолютного уровня.

Совокупность пар (τ_k, h_{τ_k}) образует маркированный точечный процесс, который полностью определяет вероятностное поведение цены. Число тиков к моменту t характеризуется счётным процессом N_t .

Эмпирические исследования показывают, что длительности между последовательными тиками обладают «тяжёлыми хвостами», то есть вероятность встретить очень большой интервал убывает по степенному закону. В качестве модели для такого поведения используется распределение Парето с плотностью:

$$f_{\alpha,b}(x) = \begin{cases} \frac{\alpha b^\alpha}{x^{\alpha+1}}, & x \geq b, \\ 0, & x < b, \end{cases} \quad (4)$$

где параметр $\alpha > 0$ является хвостовым индексом, а $b > 0$ — минимальным значением интервала.

Чем меньше значение α , тем тяжелее хвост, и тем выше вероятность появления длительных пауз между сделками. Таким образом, параметр α служит количественной мерой рыночной активности.

Математическая формализация процесса ценообразования

Второй раздел посвящён построению математической модели динамики цен. Основное внимание уделяется переходу от нерегулярных тиковых данных к дискретным временным рядам, анализу распределения логарифмических приращений, методам оценивания параметров «тяжёлых хвостов» и статистической проверке соответствия теоретических распределений эмпирическим данным. Такой подход позволяет количественно описать вероятностные закономерности, характерные для высокочастотных финансовых данных.

Поскольку тиковые данные регистрируются в нерегулярные моменты времени, для применения стандартных статистических методов их необходимо предварительно привести к равномерной временной сетке. Выбирается фиксированный временной шаг Δ (например, 10 минут, 1 час) и определяются моменты наблюдений

$$t_k = k\Delta, \quad k = 0, 1, 2, \dots \quad (5)$$

Тогда дискретизованный ценовой ряд задаётся как

$$S_\Delta = (S_{t_k}). \quad (6)$$

Такую дискретизацию по времени можно выполнить двумя способами: либо взять значения исходного кусочно-постоянного процесса в моментах t_k , либо предварительно сгладить процесс с помощью линейной интерполяции между тиками, а затем выполнить равномерную выборку.

После дискретизации переходят к анализу логарифмических приращений на интервале Δ :

$$h_{t_k}(\Delta) = \ln \frac{S_{t_k}}{S_{t_{k-1}}}.$$

Эти величины интерпретируются как относительные изменения цены за один шаг. При малых Δ распределение $h_{t_k}(\Delta)$ имеет «тяжёлые хвосты» — вероятность больших по модулю значений убывает медленнее, чем у нормального распределения. Асимптотическое поведение хвоста описывается степенным законом:

$$P(h > x) \sim x^{-\alpha} L(x), \quad x \rightarrow \infty,$$

$\alpha > 0$ — хвостовой индекс, а $L(x)$ — медленно меняющаяся функция (например, константа или логарифм).

Для параметрического описания хвостов часто используется распределение Парето. Его параметры оцениваются несколькими методами. Если параметр b известен, оценка максимального правдоподобия имеет вид:

$$\frac{1}{\hat{\alpha}_N} = \frac{1}{N} \sum_{i=1}^N \ln \frac{X_i}{b}.$$

В реальных ситуациях b часто неизвестен. Тогда вводят порог M и рассматривают только наблюдения $X_i \geq M$. Модифицированная оценка получается как

$$\frac{1}{\hat{\alpha}_{N,M}} = \frac{1}{\nu_{N,M}} \sum_{\{i: X_i \geq M\}} \ln \frac{X_i}{M},$$

где $\nu_{N,M}$ — число наблюдений, превышающих M . Широко применяется также оценка Хилла, устойчивая к выбору порога. Для этого упорядочивают выборку по убыванию: $X_{(1)}^* \geq X_{(2)}^* \geq \dots \geq X_{(N)}^*$, выбирают число M наибольших

наблюдений и вычисляют

$$\gamma_{N,M}^* = \frac{1}{M} \sum_{i=1}^M \ln \frac{X_{(i)}^*}{X_{(M)}^*}, \quad \hat{\alpha}_{N,M}^* = \frac{1}{\gamma_{N,M}^*}.$$

Статистический анализ волатильности и зависимостей в валютных парах

В третьем разделе рассматривается статистический анализ волатильности и зависимостей в валютных парах. Исследуются свойства волатильности, её периодическая структура, фрактальные характеристики, корреляционные связи между приращениями, а также методы устранения периодичности и выявления долговременных зависимостей, включая R/S-анализ.

Ключевой характеристикой изменчивости ценового процесса служит волатильность. В условно-гауссовских моделях логарифмические доходности представляются в виде $h_n = \sigma \varepsilon_n$, где ε_n — белый шум, а последовательность σ_n описывает текущий уровень волатильности, который может зависеть от предыстории. На практике σ_n часто оценивают выборочным среднеквадратическим отклонением приращений.

$$\bar{\sigma}_n = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (h_k - \bar{h}_n)^2}. \quad (7)$$

Для более детального исследования временной структуры волатильности вводится понятие Δ -волатильности — среднего абсолютного приращения процесса $H_t = \ln(S_t/S_0)$ на интервале $(a, b]$ при шаге дискретизации Δ :

$$\nu_{(a,b]}^{(\delta)}(H; \Delta) = \frac{\sum |H_{t_k} - H_{t_{k-1}}|^{(\delta)}}{\left[\frac{b-a}{\Delta} \right]} \quad (8)$$

На элементарном шаге ν совпадает с модулем логарифмической доходности $|h_k|$. Эмпирически установлено, что зависимость Δ -волатильности от величины шага хорошо описывается степенным законом

$$\tilde{\nu}_T(\Delta) \approx C_T \Delta^H,$$

H — показатель Харста (параметр фрактальности).

- $H = 0.5$ соответствует броуновскому движению (отсутствие памяти),
- $H > 0.5$ указывает на персистентность (долговременную память),
- $H < 0.5$ — на антиперсистентность.

Анализ реальных валютных пар выявляет отчётливую внутрисуточную и недельную неоднородность Δ -волатильности, связанную с активностью азиатских, европейских и американских торговых сессий. Эта периодичность накладывается на фрактальную структуру и затрудняет прямое оценивание долговременных зависимостей.

Чтобы отделить календарные эффекты от фундаментальных свойств процесса, применяется процедура деволатилизации, основанная на переходе к операционному времени. В непрерывном случае θ определяется как накопленная активность: $\theta = \tau(t) = \int_0^t \sigma^2(u) du$. Для дискретных данных локальные дисперсии σ_k^2 оцениваются, например, по часам недели, после чего моменты операционного времени находят из условия:

$$\tau^*(\theta) = \min \left\{ m : \sum_{k=1}^m \sigma_k^2 \geq \theta \right\}. \quad (9)$$

Деволатилизованные доходности

$$h_\theta^* = \sum_{\tau^*(\theta-1) < k \leq \tau^*(\theta)} h_k. \quad (10)$$

После такого преобразования периодическая составляющая исчезает, однако долговременная память, характерная для исходного ряда, сохраняется в полной мере.

Для количественной оценки степени последствия в работе используется метод нормированного размаха (R/S-анализ). По накопленной сумме логарифмических доходностей \tilde{H}_k вычисляется размах отклонений от линейного тренда R_n и выборочное стандартное отклонение S_n . Нормированная статистика $Q_n = R_n/S_n$ подчиняется асимптотическому закону $Q_n \sim c \cdot n^H$, что даёт возможность найти показатель Харста H по наклону линейной зависимости, построенной в осях $\ln Q_n$ от $\ln n$. В отличие от оценки H через

Δ -волатильность, R/S-анализ устойчив к шумам измерений и позволяет не только надёжно разделять процессы с короткой и долгой памятью, но и обнаруживать на графике характерные изломы, соответствующие циклическим компонентам.

Практическая часть

В практической части исследования были использованы тиковые данные валютных пар USD/RUB, EUR/RUB и CNY/RUB, полученные с платформы Finam за период с 1 января 2023 года по 11 июня 2024 года. Основные характеристики исходных данных представлены в таблице 1

Валютная пара	Число тиков	Ср. цена, руб.	Диапазон цен, руб.
USD/RUB	7 916 505	87,34	67,09 – 102,35
EUR/RUB	4 677 302	93,75	72,44 – 111,42
CNY/RUB	18 547 285	12,30	9,87 – 14,03

Таблица 1 – Характеристика исходных тиковых данных

После загрузки и очистки данных были сформированы унифицированные тиковые ряды. Наибольшей интенсивностью торгового потока характеризуется пара CNY/RUB — в среднем 52 099 тиков в день, что в 2,4 раза превышает аналогичный показатель для USD/RUB и в 4 раза для EUR/RUB. Это свидетельствует о переориентации российского валютного рынка на китайский юань в рассматриваемый период.

Для анализа временной структуры рыночной активности были рассчитаны последовательные интервалы между соседними тиками (в секундах) отдельно для каждого торгового дня, что позволило исключить ночные и выходные паузы. Объём выборки составил 2,95 млн интервалов для USD/RUB, 2,32 млн для EUR/RUB и 4,96 млн для CNY/RUB. Параметры распределения Парето оценивались двумя методами — методом максимального правдоподобия (ММП) и методом Хилла. Параметр масштаба b определялся как минимальный интервал в выборке. Оценка α вычислялась по всей совокупности интервалов на основе метода максимального правдоподобия. Оценка Хилла α строилась по 10% наибольших интервалов, что позволяет сосредоточиться непосредственно на хвосте распределения. Проверка гипотезы о соответствии хвоста закону Парето проводилась с помощью критерия Колмогорова–Смирнова с параметрическим бутстрэпом (500 репликаций). Результаты представлены в таблице 2.

Таблица 2 – Результаты оценивания параметров распределения Парето для междутиковых интервалов

Критерий / Параметр	USD/RUB	EUR/RUB	CNY/RUB
N интервалов	2 952 508	2 323 631	4 956 043
Порог X_M^* , сек	10,0	11,0	6,0
α (ММП)	1,851	2,106	1,801
α (Хилл)	1,488	1,490	1,525
D_N (KS)	0,202	0,458	0,205
p -value	0,002	0,002	0,002
H_0	отвергается	отвергается	отвергается

Полученные оценки параметра формы α находятся в диапазоне от 1,49 до 2,11, что соответствует распределению с конечным математическим ожиданием; для некоторых оценок дисперсия не существует, что типично для распределений с тяжёлыми хвостами. Формальное отвержение гипотезы Парето по критерию Колмогорова–Смирнова ($p = 0,002$ для всех трёх пар) обусловлено прежде всего огромным объёмом хвостовых подвыборок (от 328 до 585 тыс. наблюдений): при таком N критерий реагирует на любые, даже незначительные отклонения от теоретической модели. Согласованность оценок обоими методами подтверждает степенную природу хвоста.

На следующем этапе были построены синтетические тиковые процессы методом обратных функций. Интервалы между тиками генерировались из распределения Парето с оценёнными параметрами, логарифмические доходности — из нормального распределения. Для каждой пары сгенерировано 5 000 синтетических тиков. Вся генерация реализована с применением самостоятельно написанных функций обратных преобразований, без использования встроенных библиотечных генераторов.

После этого реальные и синтетические ряды были дискретизированы по шести временным шагам: 1 минута, 5 минут, 20 минут, 1 час, 4 часа и 1 день. По дискретизированным рядам вычислялись логарифмические доходности и оценивались их хвостовые свойства с помощью метода Хилла. Ключевые

результаты представлены в таблице 3.

Таблица 3 – Хвостовые индексы логарифмических доходностей

Пара	1 мин	1 час	1 день	синт.
USD/RUB	2,38	2,62	3,78	4,32
EUR/RUB	2,22	2,83	3,51	4,20
CNY/RUB	2,46	2,56	3,06	3,97

Из таблицы (3) видны два ключевых результата. Во-первых, хвостовые индексы реальных доходностей существенно ниже синтетических: для пары USD/RUB синтетический индекс превышает реальный на 81%, для EUR/RUB — на 89%, для CNY/RUB — на 61% (на минутном горизонте). Это означает, что реальные валютные доходности обладают значительно более тяжёлыми хвостами, чем предполагает нормальное распределение, и простая нормальная модель существенно недооценивает вероятность экстремальных движений курсов. Во-вторых, с ростом шага дискретизации Δ хвостовой индекс реальных доходностей возрастает: для USD/RUB — с 2,38 при $\Delta = 1$ мин до 3,78 при $\Delta = 1$ день, для EUR/RUB — с 2,22 до 3,51, для CNY/RUB — с 2,46 до 3,06. Это подтверждает, что при укрупнении временного масштаба распределение доходностей постепенно приближается к нормальному.

Далее анализировалась Δ -волатильность — средняя абсолютная доходность $MAD(\Delta)$. Для каждого шага Δ рассчитывалось значение $MAD(\Delta)$, после чего оценивалась степенная зависимость $MAD(\Delta) \sim \Delta^{H_1}$ методом наименьших квадратов в логарифмических координатах. Результаты представлены в таблице 4.

Таблица 4 – Оценки показателя масштабирования H_1 и коэффициента детерминации R^2

Пара	H_1	R^2
USD/RUB	0,463	0,998
EUR/RUB	0,445	0,996
CNY/RUB	0,466	0,999

Зависимость $\ln MAD(\Delta)$ от $\ln \Delta$ оказалась близкой к линейной для всех трёх валютных пар, что подтверждается высокими значениями R^2 (выше

0,99). Это свидетельствует о выполнении степенного закона масштабирования.

Оценки показателя масштабирования H_1 составили 0,463 для USD/RUB, 0,445 для EUR/RUB и 0,466 для CNY/RUB. Полученные значения несколько ниже 0,5, что указывает на более медленный рост средней абсолютной доходности по сравнению с моделью броуновского движения. Другими словами, при укрупнении временного шага волатильность возрастает, но слабее, чем в случае чисто случайного процесса.

Дополнительно выявлена выраженная календарная неоднородность: максимальная торговая активность и волатильность наблюдаются в 10–11 часов московского времени, минимальная — в ночные часы и по субботам. По дням недели наиболее высокие значения характерны для четверга–пятницы.

Таким образом, Δ -волатильность рублёвых валютных пар подчиняется степенному закону с показателем, близким к 0,5, но немного ниже него, и характеризуется выраженной внутрисуточной и недельной периодичностью.

Для устранения календарной периодичности был осуществлён переход к операционному времени. По 20-минутным доходностям для каждого часа каждого дня недели оценивались локальные дисперсии $\sigma^2(d, h)$. После нормировки на среднее значение величины в ночные часы опускались до $\sim 0,005$, а в пиковые часы достигали 24,5 и выше, что наглядно демонстрирует неравномерность внутрисуточной активности.

Накопленная активность вычислялась последовательным суммированием нормированных дисперсий. Как только она достигала целого значения θ , все попавшие в этот интервал доходности суммировались, образуя деволатилизованное приращение h_θ^* . Основные параметры полученных рядов приведены в таблице 5.

Пара	$N_{\text{исх}}$	$N_{\text{опер}}$	$\mathbb{E}[h^*]$	$\text{std}[h^*]$
USD/RUB	13 407	11 669	≈ 0	0,0017
EUR/RUB	13 088	8 043	≈ 0	0,0020
CNY/RUB	13 156	13 244	≈ 0	0,0014

Таблица 5 – Характеристики деволатилизованных рядов

Средние значения приращений остались нулевыми, стандартные отклонения близки к исходным, то есть общий масштаб вариации сохранился, а

временная структура стала более однородной.

Ключевой результат дало сравнение автокорреляций абсолютных доходностей до и после процедуры (таблица 6).

Пара	$\rho(h , 1)$	$\rho(h^* , 1)$	$\rho(h , 50)$	$\rho(h^* , 50)$
USD/RUB	0,292	0,129	0,056	0,047
EUR/RUB	0,213	0,153	0,046	0,025
CNY/RUB	0,276	0,097	0,074	0,061

Таблица 6 – Автокорреляции абсолютных доходностей до и после деволатилизации

На малых лагах корреляции заметно снизились: например, у USD/RUB значение $\rho(|h|, 1)$ уменьшилось с 0,29 до 0,13. Это подтверждает, что значительная часть коротких корреляций была вызвана именно суточной ритмикой. Однако на лаге 50 и далее корреляции остаются статистически значимыми, то есть долговременная память сохранилась. Таким образом, деволатилизация устранила календарную цикличность, но не разрушила медленно затухающие корреляции, что даёт основание перейти к формальной оценке показателя Харста методом R/S-анализа.

Для окончательной оценки долговременной памяти был проведён R/S-анализ деволатилизованных рядов. Для каждого окна n вычислялась статистика нормированного размаха $Q_n = R_n/S_n$, после чего показатель Харста H_2 определялся как наклон линейной регрессии $\ln Q_n$ на $\ln n$. Значимость отличия H_2 от 0,5 проверялась методом Монте-Карло: для каждого ряда генерировалось 1000 суррогатных последовательностей белого шума той же длины, и по ним строилось эмпирическое распределение H_2 . Результаты представлены в таблице 7.

Пара	$N_{\text{опер}}$	H_2	$H_{0.95}$ суррогатов	p-value	Значимость
USD/RUB	11 669	0,560	0,562	0,066	нет
EUR/RUB	8 043	0,538	0,570	0,484	нет
CNY/RUB	13 244	0,571	0,561	0,008	да

Таблица 7 – Показатели Харста по R/S-анализу и результаты проверки значимости

Для CNY/RUB зафиксирована статистически значимая персистентность: $H_2 = 0,571$ при p-value = 0,008. Это означает, что даже после устранения календарных циклов ряд сохраняет положительную долговременную зависимость. У USD/RUB $H_2 = 0,560$ находится на границе значимости (p = 0,066)

— можно говорить о слабой тенденции к персистентности, но на 5%-м уровне она не подтверждается. Пара EUR/RUB ведёт себя практически как случайное блуждание: $H_2 = 0,538$ при $p = 0,484$. Синтетические ряды из-за малой длины не позволили получить устойчивых оценок H_2 .

Таким образом, значимая долговременная память обнаружена только у пары CNY/RUB, тогда как USD/RUB и EUR/RUB не демонстрируют устойчивого отклонения от модели случайного блуждания. Этот результат согласуется с ранее выявленной спецификой юаня и подчёркивает неоднородность статистических свойств рублёвых валютных пар.

ЗАКЛЮЧЕНИЕ

Проведённый анализ высокочастотных данных валютных пар USD/RUB, EUR/RUB и CNY/RUB подтвердил степенной характер распределения межтиковых интервалов с оценками хвостового индекса α в диапазоне 1,49–2,11. Формально гипотеза о точном соответствии распределению Парето отвергается критерием Колмогорова - Смирнова, что объясняется сверхбольшими объёмами выборки. Логарифмические доходности на малых временных масштабах также обладают тяжёлыми хвостами ($\alpha \approx 2,2-2,5$), что свидетельствует о повышенной вероятности экстремальных изменений курса по сравнению с нормальным распределением.

Установлено, что Δ -волатильность подчиняется степенному закону с показателем $H_1 \approx 0,45-0,47$, то есть средняя амплитуда колебаний возрастает с масштабом медленнее, чем в броуновской модели. Деволатилизация устраняет календарную цикличность, но не устраняет полностью медленно затухающие корреляции абсолютных доходностей, что указывает на сохранение внутренней зависимости в динамике волатильности.

По результатам R/S -анализа статистически значимая долговременная память подтверждена только для пары CNY/RUB ($H_2 = 0,571$, $p = 0,008$); для USD/RUB и EUR/RUB устойчивое отклонение от случайного блуждания не выявлено. Тем самым показана неоднородность статистических свойств рублёвых валютных пар и необходимость дифференцированного подхода к их моделированию.

В работе предложена и реализована стохастическая модель тикового процесса, разработаны алгоритмы дискретизации и деволатилизации, а также выполнен полный эмпирический анализ на реальных и синтетических данных. Полученные результаты могут быть использованы при построении моделей волатильности, оценке рыночных рисков и прогнозировании динамики валютных курсов.