

МАТЕМАТИЧЕСКИЙ АНАЛИЗ ЭКОНОМИЧЕСКИХ МОДЕЛЕЙ

Методы оценивания интегрированной дисперсии: проблемы устойчивости
к скачкам в высокочастотных временных рядах

© 2024 г. З.О. Косимов

З.О. Косимов

МГУ имени М.В. Ломоносова; e-mail: zohirsho1@gmail.com

Поступила в редакцию 13.02.2024

Автор благодарит своих научных руководителей Е.Н. Лукаша и А.А. Гущина за полезные замечания и бесценные комментарии к тексту настоящей статьи. Хотелось бы также отметить значение и важность советов Ф.С. Картаева в процессе подготовки статьи к ее публикации.

Аннотация. Интегрированная дисперсия является мерой волатильности процесса в непрерывном времени и используется в финансовой математике как инструмент оптимизации портфеля, прогноза динамики цены финансового актива. Состоятельность оценки интегрированной дисперсии случайного процесса находится в центре внимания настоящей статьи. Основопологающий диффузионный процесс расширен посредством включения компоненты скачков как средства улучшения описательной функции процесса. Именно активность скачков является тем фактором, который обуславливает состоятельность оценки интегрированной дисперсии. Поэтому состоятельность оценки определяется как степень ее устойчивости к скачкам. Рассмотрены два основных метода оценивания интегрированной дисперсии и проанализирована способность соответствующих оценок в нейтрализации эффекта скачков на сходимость. Приведены доводы, указывающие на необходимость дальнейшего исследования эффекта скачков, ссылаясь на работы авторов, заложившие основу анализа интегрированной дисперсии, а также на работы, в которых содержатся основные асимптотические результаты относительно устойчивости оценки интегрированной дисперсии к скачкам. По результатам проведенного анализа выделены направления дальнейшего развития асимптотической теории для анализа состоятельности оценки интегрированной дисперсии.

Ключевые слова: интегрированная дисперсия, диффузионный процесс, состоятельность, скачки, устойчивость к скачкам, асимптотическая теория.

Классификация JEL: C02, C32, C53, C65.

УДК: 519.214.

Для цитирования: **Косимов З.О.** (2024). Методы оценивания интегрированной дисперсии: проблемы устойчивости к скачкам в высокочастотных временных рядах // *Экономика и математические методы*. Т. 60. № 3. С. 107–117. DOI: 10.31857/S0424738824030095

ВВЕДЕНИЕ

В фокусе настоящей статьи стоит такой вероятностный феномен, как дисперсия. Роль и влияние моделирования дисперсии в финансовом мире сегодня трудно переоценить. Диапазон возможных применений моделей оценки дисперсии достаточно широк — начиная с оценки сложных финансовых инструментов до прогнозирования цен финансовых активов и их роли в формировании структуры портфеля. С доступностью больших данных, развитием инфраструктуры для их хранения и обработки, ростом вычислительной мощности приобретают все большую популярность различные непараметрические методы оценивания. Именно доступность больших данных способствовала росту исследований высокочастотных временных рядов, являющихся *объектом* нашего интереса, развитию методов машинного обучения и трансформации моделирования дисперсии в сторону более непараметрических методов. Это стало важной особенностью последних десятилетий в развитии теории количественных финансов. И эта тенденция сохранится в ближайшем будущем, притягивая все больше внимания как практиков, так и теоретиков.

В основе понимания дисперсии лежит видение ее как непрерывного явления, что представлялось бы менее целесообразным в случае меньшего объема доступных данных, соответствующего и более низкой частоте доступных данных. *Предметом* исследования настоящей статьи является оценка интегрированной дисперсии (ОИД), а именно — устойчивое к скачкам оценивание

интегрированной дисперсии, с чем мы связываем понятие *состоятельности* оценки интегрированной дисперсии. Дисперсия рассматривается как интегрированная переменная, что предполагает непрерывность¹ временных рядов.

Состоятельность² оценки интегрированной дисперсии обусловлена активностью скачков, характерных для динамики финансового актива. В нашем случае именно поведение скачков как фактора устойчивости (jump robustness) оценки интегрированной дисперсии является центральным параметром, который определяет состоятельность оценки интегрированной дисперсии.

Цель статьи — обосновать необходимость дальнейшего анализа влияния поведения скачков на состоятельность оценок интегрированной дисперсии, опираясь на основные классы оценок интегрированной дисперсии. Таким образом, обратив внимание на ограничения оценок в форме степени их устойчивости с учетом различной активности скачков, мы определим контуры дальнейшего развития анализа асимптотической теории как инструмента анализа сходимости оценок. Анализ влияния скачков на устойчивость оценки интегрированной дисперсии имеет теоретическую природу, но в заключительной части статьи будут приведены примеры для приложения скачкообразных процессов.

В разд. 1 описывается основополагающий процесс, представляющий *семимартингал*, состоящий из диффузии и скачкообразной компоненты. Определяется интегрированная дисперсия и ее оценка, а также обсуждается эффект влияния скачков на состоятельность ОИД. В разд. 2 выделяются два основных класса ОИД. Показано, что в некоторых случаях эти оценки интегрированной дисперсии неустойчивы к определенным типам активности скачков. В разд. 3 обосновывается необходимость рассмотрения скачков как фактора, влияющего на состоятельность ОИД. Необходимость более детального анализа оценок на их устойчивость основана на доводах авторов, заложивших основу понятия интегрированной дисперсии и изучивших проблему сходимости ОИД на протяжении почти одного десятилетия.

1. ИНТЕГРИРОВАННАЯ ДИСПЕРСИЯ³

1.1. Определение интегрированной дисперсии и ее оценка

Предположим, у нас имеются данные с достаточно высокой частотой. С теоретической точки зрения, возможность использования высокочастотных данных обосновывает целесообразность рассмотрения временного ряда как процесса с непрерывным временем. Непрерывность основывается на наблюдениях, обеспечивающих бесконечную делимость приращений промежутков времени, в результате чего диаметр разделения сетки моментов времени (интервал между приращениями $N = 1/\Delta$ при $N \rightarrow \infty$) стремится к нулю. Таким образом, мы имеем необходимые основания для аппроксимации дисперсии как интегрируемой переменной.

Итак, основополагающий для моделирования процесс следует рассматривать как процесс в непрерывном времени. Не менее фундаментальным для моделируемого процесса компонентом является броуновское движение W_t , позволяющее отразить чистую случайность, — *диффузию*. Такой стохастический процесс называется диффузионным.

Определим традиционный стохастический одномерный процесс без учета скачков:

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s. \quad (1)$$

Процесс X_t является броуновским семимартингалом (или обобщенной диффузией), состоит из локального мартингала M_t , для которого W — стандартное броуновское движение, и согласованного процесса A_t с локально ограниченной вариацией, играющей роль сдвига, т.е. процесс X_t можно записать как $X_t = M_t + A_t$, где $M = \int_0^t \sigma_s dW_s$, $A = \int_0^t a_s ds$, $A_0 = M_0 = 0 \forall t \geq 0$. Семимартингал всегда является согласованным и с ограниченной вариацией.

¹ Непрерывность временных рядов и непараметрический метод оценки их дисперсии подразумевает необходимость наличия больших данных, что равносильно наличию высокочастотных временных рядов.

² Состоятельность оценки в данном случае должна пониматься в классическом смысле; соответствующее определение в контексте интегрированной дисперсии будет приведено далее.

³ Более логичным было бы назвать данный раздел как «Реализованная вариация», так как именно она предлагается в данной работе в качестве состоятельной ОИД. Но так как анализ направлен на оценку интегрированной дисперсии как центральной части квадратической вариации, мы решили обратить внимание читателя именно на этот факт.

Разложение X_t не будет уникальным, так как всегда может прибавить к A_t и вычесть из M_t мартингал с ограниченной вариацией. Процесс σ_s непрерывен справа и имеет предел слева — особенность, из-за которой σ_s локально интегрируема, в частности $\int_0^t \sigma_s^2 ds < \infty$ для всех t .

Замечание 1. Строго говоря, в силу данного предположения диффузионную часть процесса X_t принято обозначать как $\int_0^t \sigma_s dW_s$. Класс непрерывных функций (càdlàg) образует пространство Скорохода.

Процесс стохастической дисперсии $\int_0^t \sigma_s^2 dW_s$ как компонент процесса X_t всегда можно представить в виде локального мартингала с непрерывными траекториями в силу вышеуказанной особенности процесса σ_s . Все непрерывные локальные мартингалы с абсолютно непрерывной квадратической вариацией можно представить в форме процесса стохастической дисперсии. Таким образом, исключение возможности арбитража и моделирование компоненты локального мартингала как процесса стохастической дисперсии с учетом особенностей, изложенных выше, делает процесс вида (1) каноническим для непрерывных процессов в теории количественных финансов.

Следующим фундаментальным понятием является «квадратическая вариация» (QV), развитие теории которой тесно связана с формулой Ито и разложением интеграла Ито на части.

Замечание 2. В обобщенном виде квадратическую вариацию можно представить как $[X, X]_t = X_t^2 - 2 \int_0^t X_{s-} dX_s$. Скачки в процессе X_t можно определить, используя обозначение X_{s-} как $\Delta X_s = X_s - X_{s-}$.

Принимая во внимание, что процесс X_t задан в виде (1), определим квадратическую вариацию этого процесса:

$$[X, X]_t = \langle X^c, X^c \rangle_t. \quad (2)$$

Здесь X^c представляет так называемую непрерывную мартингальную часть процесса X_t . Например, если $X_t = \sigma W_t$, то квадратическая вариация такого процесса равна $\sigma^2 t$, $[X, X]_t = \sigma^2 t$. В этом случае $[X, X]_t$ не является случайной величиной. Когда же процесс X_t задан соотношением (1), квадратическая вариация становится случайной и будет равна

$$[X, X]_t = X^c, \quad X^c_t = \int_0^t \sigma_s^2 ds, \quad (3)$$

что и есть непосредственно интегрированная дисперсия (IV). Так, в (Protter, 2004, р. 66) П. Проттер отмечает важность квадратической вариации в теории семимартингалов (важные и для нас процессы, см. п. 1.2) и в теории финансового риска.

Введем понятие «реализованная вариация» (RV), базирующееся на эмпирических данных как оценку интегрированной вариации:

$$RV_N = \sum_{j=1}^N (X_{t_j} - X_{t_{j-1}})^2. \quad (4)$$

Замечание 3. Определить RV можно и другим образом — предполагая равномерно распределенные данные логарифмированной доходности $r_i = X_{i/N} - X_{(i-1)/N}$ из некоего промежутка, $RV = \sum_{i=1}^N r_i^2$.

С помощью RV мы попробуем состоятельно оценить IV . В силу особенностей реальных данных для временных рядов и вытекающих практических соображений, процесс X_t берется как логарифм начальных данных.

Для демонстрации роли RV_N приведем еще одно определение квадратической вариации

$$QV = p \lim_{N \rightarrow \infty} \sum_{j=1}^N (X_{t_j} - X_{t_{j-1}})^2. \quad (5)$$

Цель построения ОИД — достичь такую сходимость, когда

$$RV \xrightarrow{u.c.p} QV = IV. \quad (6)$$

Замечание 4. «u.c.p.» (uniform convergence in probability) означает равномерную (по времени на конечных интервалах) сходимость по вероятности, т.е. $\mathbb{P}[d(RV, QV) > \varepsilon] \rightarrow 0$. В литературе можно столкнуться с иными обозначениями для определения RV и QV . Например, в (Aït-Sahalia, Jacod, 2014) введены обозначения $RV = \hat{C}(\Delta_n)_t$, $QV = C_t$, $\Delta^n_t X = X_{i\Delta_n} - X_{(i-1)\Delta_n}$ и, таким образом, $\hat{C}(\Delta_n)_t \xrightarrow{p} [X, X]_t = C_t + \sum_{i=1}^{\lfloor t/\Delta_n \rfloor} (\Delta^n_t X)^2$, где $C_t = \langle X^c, X^c \rangle_t$.

Разделение промежутка $[0, t]$ таково, что $t_0 = 0 < t_1 < \dots < t_N = t$ и $\sup |t_j - t_{j-1}| \rightarrow 0 \quad \forall j$ при $N \rightarrow \infty$, что соответствует предположению о наличии высокочастотных данных при условии регулярности данных, упомянутых выше, а именно — при их равномерном расположении, т.е. исключается случай микроструктурного шума, который сам по себе представляет обособленный и привлекающий растущий академический интерес⁴.

1.2. Процесс X_t с учетом скачков и их последствия для RV

До этого момента процесс был определен как броуновский семимартингал вида $X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s$. Теперь введем в процесс дополнительный компонент скачков Y_t , преобразовав процесс в

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s + Y_t, \quad (7)$$

где Y_t — как пример процесса в непрерывном времени — может быть выражен посредством сложного пуассоновского процесса.

Замечание 5. Сложный пуассоновский процесс Y_t — это стохастический процесс со скачками в непрерывном времени. Моменты скачков ведут себя согласно пуассоновскому процессу N_t с параметром λ , размер которых D_i случаен с некоторой функцией распределения G , не зависящей от N_t . В качестве примера рассмотрим процесс вида $Y_t = \sum_{i=1}^{N_t} D_i$, где (как частный случай) $\{D_i : i \geq 1\}$ независимы и одинаково распределены. В силу предположения независимости и одинакового распределения D_i применимо тождество Вальда для вычисления математического ожидания $\mathbb{E}(Y_t) = \mathbb{E}\left[\left(D_1 + \dots + D_{N_t}\right)\right] = \lambda \mathbb{E}(D)$ и дисперсии $V(Y_t) = \lambda \mathbb{E}(D^2)$.

Основополагающим в финансовой математике, как и для всей теории ценообразования активов, является предположение об отсутствии арбитража как одного из краеугольных условий для моделирования процесса. Отсутствие арбитража, в свою очередь, предполагает, что процесс должен быть семимартингалом, который, в свою очередь, должен являться càdlàg. Данное требование удовлетворяется нашим предположением о свойстве càdlàg.

В свете (7) квадратическая вариация равна

$$QV = \int_0^t \sigma_s^2 ds + \sum_{0 \leq s \leq t} (\Delta X_s)^2, \quad (8)$$

где появилось дополнительное слагаемое $\sum_{0 \leq s \leq t} (\Delta X_s)^2$, отображающее эффект скачков, но X^c , $X_t^c = \int_0^t \sigma_s^2 ds$ все еще является IV .

Замечание 6. Следуя обозначению из замечания 2 и определению процесса X_t как $X_t = M_t + A_t$, квадратическая вариация X_t равна

$$\begin{aligned} [X, X]_t &= \langle X^c, X^c \rangle_t + \sum_{0 \leq s \leq t} (X_s - X_{s-})^2 = \\ &= \langle X^c, X^c \rangle_t + \sum_{0 \leq s \leq t} (M_s - M_{s-})^2 + \sum_{0 \leq s \leq t} (A_s - A_{s-})^2 + 2 \sum_{0 \leq s \leq t} (M_s - M_{s-}) \sum_{0 \leq s \leq t} (A_s - A_{s-}) = \\ &= [M, M]_t + \sum_{0 \leq s \leq t} (A_s - A_{s-})^2 + 2 \sum_{0 \leq s \leq t} (M_s - M_{s-}) \sum_{0 \leq s \leq t} (A_s - A_{s-}). \end{aligned}$$

Если процесс A_s непрерывен, то независимо от процесса M_t , $[X, X]_t = [M, M]_t$ является интегрированной дисперсией.

Так как RV — как оценка QV — теперь должна оценить QV с учетом дополнительного ее компонента, а именно, $\sum_{0 \leq s \leq t} (\Delta X_s)^2$, то данный факт не может не иметь определенных последствий для способности RV самостоятельно и эффективно оценить QV . Основополагающей задачей, таким образом, становится необходимость обособления эффекта скачков от компонента IV и тем самым реабилитация оценки RV . Но как это сделать?

Потенциальное исследование последствий включения в процесс дополнительного компонента скачков должно быть сконцентрировано на качестве оценивания интегрированной дисперсии. При этом тип активности скачков определяет, как они будут влиять на поведение интегрированной

⁴ В некоторой литературе используется термин «in-fill asymptotics» для обозначения сходимости (6), в силу того что на фиксированном промежутке (поэтому «in-fill») предполагается все большая масса данных, $N \rightarrow \infty$.

дисперсии — центрального параметра динамики цены финансового актива. Следует отметить, что под качеством ОИД должна пониматься состоятельность RV . Качество ОИД представлено двумя фундаментальными для способности оценки характеристиками: скоростью, с которой RV стремится к IV , а также законом распределения ошибки оценивания IV посредством RV .

2. МЕТОДЫ ОЦЕНИВАНИЯ ИНТЕГРИРОВАННОЙ ВОЛАТИЛЬНОСТИ

Для анализа мы используем данные достаточно высокой частоты, которые распределены равномерно.

Существует несколько практических ограничений для предположения равномерности. Например, они характерны для высокочастотных временных рядов и подтверждают их эмпирические факты, указывающие на неравномерность предположения равномерности в ее полной мере. Одно из наиболее частых явлений — микроструктурный шум, источником которого могут быть такие скачки, как разница цен «bid» и «ask», а также кластеризация данных, возникающая в определенные периоды времени, или же устоявшиеся типы поведения в распределении данных (например, более высокая волатильность в начале и в конце дня, так называемая U-образная волатильность). Это сказывается на стабильности измеряемого параметра и представляет отдельную область исследования, приобретающую популярность с ростом доступности больших данных, увеличивающейся вычислительной мощности.

Изложим кратко еще раз задачу, стоящую перед нами с введением обозначений согласно замечанию 4. Пусть оценка $RV = \hat{C}(\Delta_n)_t$ определяется на некотором временном промежутке $[0, T]$, состоящем из интервалов разбиения $\Delta_n \rightarrow 0$ при $n \rightarrow \infty$. Посредством RV мы стремимся оценить $QV = C_t$, — при том что центральным для нашей оценки является непосредственно интегрированная дисперсия, $\int_0^t \sigma_s^2 ds$ — параметр, описывающий динамику изменчивости процесса. Поэтому при наличии скачков в основополагающем процессе X_t предстоит обособить $\int_0^t \sigma_s^2 ds$ от компоненты скачков $\sum_{i=1}^{[T/\Delta_n]} (\Delta_i^n X)^2$ (см. (8)), что равносильно достижению устойчивости оценки к скачкам.

Рассмотрим пример. Пусть процесс имеет форму $X_t = X_0 + \int_0^t \sigma_s dW_s + Y_t$, с несложным компонентом Y_t , описывающим активность скачков. Разделим промежуток $[0, T]$ на интервалы $((i-1)\Delta_n, i\Delta_n]$. Если данный интервал включает скачок размера J , то $\Delta_i^n X$ из этого интервала будет близок к J , отличен от нуля и не зависеть от n . Предполагая конечное число скачков, например N_T на промежутке $[0, T]$ — для достаточно большого n среди всех $((i-1)\Delta_n, i\Delta_n]$. И тогда именно N_T число разбиений будут содержать скачки J , а все остальные интервалы из $[0, T]$ будут свободны от скачков. Для выявления скачков и выделения их из оценки (и тем самым, обособляя их эффект) с целью состоятельного оценивания интегрированной дисперсии определим последовательность u_n , стремящуюся к нулю, и тем самым исключим приращения, превышающие порог u_n . Таким образом, наша оценка RV имеет вид

$$\hat{C}(\Delta_n, u_n)_t = \sum_{i=1}^{[T/\Delta_n]} (\Delta_i^n X)^2 1_{\{|\Delta_i^n X| \leq u_n\}}. \quad (9)$$

Выбор u_n должен быть достаточно осторожным, т.е. скорость, с которой u_n приближается к 0, не должна исключать приращения, не включающего скачки. Данная идея представляет основу метода порогового усечения (truncated (threshold) realised variance) как одного из классов оценки RV , который будет обсужден в следующих разделах.

Таким образом, при $\int_0^t \sigma_s^2 ds = C_t$ и, принимая во внимание процесс $\hat{C}(\Delta_n, u_n)_t$, предельное вероятностное распределение ошибки ОИД посредством оценки из класса пороговых вариаций вида

$$(\hat{C}(\Delta_n, u_n)_t - C_t) / \sqrt{\Delta_n}, \quad (10)$$

будет гауссовским распределением с дисперсией распределения ошибки, имеющей наиболее простой вид в случае конечной активности скачков.

Замечание 7. $\mathbb{E} \left(\left((\hat{C}(\Delta_n, u_n)_t - C_t) / \sqrt{\Delta_n} \mid \mathcal{F} \right)^2 \right) = 2 \int_0^t \sigma_s^4 ds$, и поэтому

$$\mathcal{L} \left((\hat{C}(\Delta_n, u_n)_t - C_t) / \left\{ \sqrt{\Delta_n} \sqrt{2 \int_0^t \sigma_s^4 ds} \right\} \mid \mathcal{F} \right) \rightarrow \mathcal{N}(0, 1).$$

Нормальное распределение скачков будет иметь более сложный вид в случае бесконечной активности скачков, что и является фокусом анализа сходимости оценки. Конечно, помимо

активности скачков, на предельное распределение ошибки будут также влиять и другие факторы (например, поведение дисперсии исследуемого процесса, зависимость дисперсии по отношению к броуновскому движению и др.).

Описав суть задачи, представим два основных класса оценок реализованной вариации.

2.1. Класс многостепенных вариаций

Наиболее популярным классом оценок интегрированной дисперсии является класс многостепенных вариаций (multipower variation, MPV). Асимптотическая теория для данного класса оценок RV была разработана группой ученых под руководством Оле Барндорфф-Нелсона (Ole Barndorff-Nielsen). Наиболее фундаментальными в развитии асимптотической теории для данного класса оценок являются работы (Barndorff-Nielsen, Shephard, Winkel, 2006; Barndorff-Nielsen et al., 2006), наиболее цитируемыми работами, где исследован этот класс оценок, — (Barndorff-Nielsen, Shephard, 2002, 2004). Значительный вклад в развитие предельных теорем для стохастических процессов внесла работа выдающихся ученых в области теории вероятностей (Jacod, Shiryaev, 2003).

Определим многостепенную вариацию общего вида

$$MPV_N(m; r) = d_{m,r} \left(\frac{N}{N-m+1} \right) N^{0,5r-1} \sum_{i=1}^{N-m+1} |\Delta X_i|^{r/m} \dots |\Delta X_{i+m-1}|^{r/m}, \quad (11)$$

где ΔX_i — гауссовские независимые одинаково распределенные случайные величины; $d_{m,r}$ — константа, зависящая от параметров m и r ; $(N / (N - m + 1)) N^{0,5r-1}$ — корректирующий фактор, необходимый в силу конечности выборки данных. Таким образом, $MPV_N(m; r)$ как оценка сходится к интегрированной дисперсии в случае конечной активности скачков:

$$MPV(m; r) \xrightarrow{u.c.p.} \int_0^1 \sigma_u^r du. \quad (12)$$

Так, например, $MPV_N(1; 2) = RV_N$ (реализованная вариация — realised variation), $MPV_N(2; 2) = BV_N$ (двухстепенная вариация — bipower variation), $MPV_N(3; 2) = TPV_N$ (трехстепенная вариация — tripower variation) и т.д.

Далее на примере BV_N отметим определенные особенности данной оценки как частного случая многостепенных вариаций и тем самым определим вектор потенциального направления для исследования состоятельности ОИД этого класса:

$$BV_N = 0,5\pi \left(\frac{N}{N-1} \right) \sum_{i=1}^{N-1} |\Delta X_i| |\Delta X_{i+1}|. \quad (13)$$

Замечание 8. $\Delta X_i, \Delta X_{i+1} \sim \text{iid } \mathcal{N}(0, \sigma^2 / N)$ и поэтому $\mathbb{E}[|\Delta X_i| |\Delta X_{i+1}|] = 2\sigma^2 / \pi N$. См. также (Barndorff-Nielsen, Shephard, 2006).

Для случая конечной активности скачков оценка BV_N является устойчивой к скачкам (т.е. состоятельной) — поскольку мы предполагаем бесконечно делимые приращения, у каждых двух смежных таких приращений может быть только один скачок (в силу предположения конечной активности скачков!), который будет нейтрализован в силу перемножения приращения, содержащего скачок, и смежного приращения (диффузии), не имеющего скачка, порядка $1/\sqrt{N}$. При $N \rightarrow \infty$ эффект скачка будет асимптотически устранен. На практике, однако, эти приращения представляют некий фиксированный промежуток и поэтому размер такого промежутка имеет определенный предел. По этой причине приращения не могут достаточно быстро приближаться к нулю (так как предпосылка $N \rightarrow \infty$ является теоретической), что является причиной чаще всего завышенной оценки BV_N . Вследствие этого для определения оптимального размера приращения обычно руководствуются общепринятыми на практике установками, связанными с ликвидностью и особенностями микроструктурного шума. Более того, нейтрализация *эффекта скачков* за счет стремления к нулю со скоростью $1/\sqrt{N}$ будет недостаточной для применения центральной предельной теоремы в случае непрерывного процесса X_t , содержащего скачки даже конечной активности (Andersen, Dobrev, Schaumburg, 2012, p. 77). Недостаточность асимптотической теории, основанной на применении предельной теоремы для непрерывного процесса, содержащего скачки различного типа, является поводом для развития и применения иных многостепенных вариаций.

Даже в случае конечной активности скачков RV_N уже не будет состоятельной ОИД, тогда как BV_N является состоятельной оценкой интегрированной дисперсии в случае конечной активности скачков. Однако BV_N не предоставляет возможности применять предельную теорему в случае как конечной, так и бесконечной активности скачков. Наименьшую предельную вариацию в данном

классе оценок имеет TPV_N , что позволяет ей более гибко оперировать асимптотической теорией в случае наличия определенных типов активности скачков.

2.2. Класс пороговых вариаций⁵

В начале разд. 2 был приведен пример нейтрализации эффекта скачков с использованием метода пороговых вариаций. Далее, обобщим определение данного класса оценок реализованной вариации (TV) для многомерного процесса X_l , где $j, l = 1, \dots, m$:

$$\hat{C}(\Delta_n u_n)_T^{jl} = \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \Delta_i^n X^j \Delta_i^n X^l 1_{\{|\Delta_i^n X| \leq u_n\}}. \quad (15)$$

Как было отмечено раньше, важным для данного класса оценок является выбор стремящейся к нулю последовательности u_n так, чтобы непрерывная часть приращений была сохранена тогда, как скачки должны быть исключены (усечены) при соблюдении определенных условий.

Замечание 9. Одним из таких условий станет выбор порогового значения $u_n \asymp \Delta_n^\omega$, где $\omega \in (0; 0,5)$ такого, что $u_n / \sqrt{\Delta_n \log(1/\Delta_n)} \rightarrow \infty$ и $u_n \rightarrow 0$ — как условия непрерывности согласно модулю непрерывности Леви.

Для иллюстрации оценок данного класса приведем оценки, предложенные в (Andersen, Dobrev, Schaumburg, 2012, p. 78)⁶. Авторы предлагают две оценки интегрированной дисперсии.

1. *MinRV* включает два смежных наблюдения, которые в предельном случае никогда не могут содержать два скачка — инструмент для минимизации эффекта скачка посредством умножения компонента, содержащего скачок (например, $|\Delta X_i|$), на компонент, содержащий диффузию, $|\Delta X_{i+1}|$ для определенного значения i :

$$MinRV = \frac{\pi}{\pi - 2} \left(\frac{N}{N-1} \right) \sum_{i=1}^{N-1} \min(|\Delta X_i|, |\Delta X_{i+1}|)^2. \quad (16)$$

2. *MedRV* состоит из трех смежных наблюдений, посредством чего авторы стремятся уменьшить вариацию, предоставляемую TPV_N — как оценки с наименьшей вариацией в классе многостепенных вариаций⁷:

$$MedRV = \frac{\pi}{6 - 4\sqrt{3} + \pi} \left(\frac{N}{N-2} \right) \sum_{i=2}^{N-1} \text{med}(|\Delta X_i|, |\Delta X_{i+1}|, |\Delta X_{i+2}|)^2. \quad (17)$$

Так, исключение скачков конечной активности происходит посредством *MinRV*, тогда как минимальная вариация достигается с помощью *MedRV*. Авторы назвали данный метод усечением ближайшей окрестности (nearest neighbour truncation). Важно отметить, что активность скачков в данном случае предполагается конечной.

Сравнивая состоятельность различных оценок реализованной вариации, следует отметить, что краеугольным является достижение оценкой компромисса между двумя противоречащими друг другу результатами. Так, при достижении оценкой лучшей эффективности, необходимо учесть, что такое достижение постоянно находится в определенном конфликте со степенью устойчивости данной оценки к скачкам. Именно баланс между двумя этими характеристиками и определяет уровень гибкости асимптотической теории, подразумеваемой той или иной оценкой.

3. СКАЧКИ КАК ВЕКТОР РАЗВИТИЯ АНАЛИЗА СОСТОЯТЕЛЬНОСТИ ОЦЕНКИ

Рассмотрим процессы, состоящие из двух компонентов — диффузионного и скачкообразного. Ранее было введено понятие «*состоятельность оценки*» для интегрированной дисперсии данного процесса. Именно скачки являются фактором, обуславливающим состоятельность оценки, точнее, — скорости ее сходимости и предельного закона распределения ошибки. Поэтому состоятельная ОИД должна пониматься как оценка, устойчивая к скачкам. Насколько она устойчива

⁵ Класс пороговых вариаций также называют классом усечений, так как усечение происходит посредством установленного порога.

⁶ Иногда в литературе данный метод приводится как отдельный от класса пороговых вариаций, наряду с оценками Фурье, оценками, основанными на ранжировании и иными. В своем классическом виде метод пороговых вариаций предполагает непосредственно анализ последовательности u_n — как правила для усечения скачков.

⁷ Выражения перед корректирующими факторами в случаях (15) и (16) необходимы для несмещенности предлагаемых авторами оценок.

к скачкам, т.е. каково влияние компонента скачков на скорость сходимости оценки и распределение ее ошибки, исследуется с помощью асимптотической теории, главным инструментом которой являются предельные теоремы.

Несмотря на огромную работу, проделанную различными авторами в отношении развития асимптотической теории для ряда оценок интегрированной дисперсии, можно привести некоторые соображения для мотивации дальнейшего развития анализа устойчивости оценки интегрированной дисперсии.

Так, в работе (Aït-Sahalia, Jacod, 2014, p. 200), посвященной высокочастотной финансовой эконометрике (high frequency financial econometrics), авторы уделяют много внимания анализу квадратической вариации и ее оценке посредством реализованной вариации, делая акцент на практической стороне работы с использованием данных высокой частоты. Описав особенности обоих классов ОИД (многостепенные и пороговые вариации) и применяя метод усечения ближайшей окрестности, описанный выше, авторы делают заключение: «Оценивая интегрированную дисперсию менее точно, чем другие методы, метод усечения ближайшей окрестности достаточно мало теряет в точности оценивании интегрированной дисперсии. Несмотря на это, результат использования данного метода может оказаться потенциально лучшим для процессов с бесконечной активностью скачков, которые являются *основным ограничением* применимости многостепенных и пороговых вариаций».

Под лучшим результатом подразумевается устойчивость оценки к различной активности скачков, включая бесконечную, с чем мы связываем высший уровень гибкости, т.е. лучшую устойчивость оценки интегрированной дисперсии.

В работе (Barndorff-Nielsen, Shephard, 2007, p. 39) отмечается, что согласно (Barndorff-Nielsen, Shephard, Winkel, 2006) даже конечная активность скачков влияет на сходимость BV_N к интегрированной дисперсии, т.е. скачки некоторым образом воздействуют на скорость сходимости оценки, что, в свою очередь, ограничивает применение асимптотической теории в случае скачков, тогда как такое воздействие отсутствует для сходимости многостепенных вариаций высшего порядка, в частности, когда $m \geq 3$ для $MPV(m; r)$. Таким образом, этот результат открывает определенную перспективу получения новых результатов о сходимости многостепенных вариаций высшего порядка при наличии скачков различной активности. Интересно, что в дополнение к результатам (Barndorff-Nielsen, Shephard, Winkel, 2006) в работе (Woerner, 2004) доказывается, что свойства предела по вероятности для BV_N не меняются в некоторых случаях бесконечной активности скачков. Авторы этих работ подчеркивают необходимость дополнительных усилий для получения окончательных результатов.

В (Eberlein, 2010), описывая особенности и потенциал использования диффузионных процессов, автор утверждает, что диффузионные процессы сами по себе неспособны дать достаточно адекватной картины для некоторых эмпирически установленных фактов, если рассматривать свойства законов распределения этих процессов без использования скачков — одной из причин растущей роли скачков во многих областях финансов (ibid., p. 1894). Автор (Eberlein, 2010) считает, что растущей роли скачков способствуют развитие теории семимартингалов и растущая вычислительная мощь. Начиная со стандартного процесса Пуассона, он вводит меру Леви и заканчивает обсуждение перечислением возможностей и вариантов использования процессов Леви для определения скачков и их включения в семимартингальный процесс. Таким образом, именно добавление скачков в основополагающий диффузионный процесс должно улучшить описательную функцию диффузионных процессов. Значимость теории семимартингалов для современной теории финансов также подчеркивает (Shiryayev, 1999, p. 294–313).

На рис. 1–2 представлены иллюстрации скачкообразных процессов с бесконечной активностью и конечной вариацией, а также с бесконечной активностью и бесконечной вариацией, смоделированные как процессы Леви.

Как следует из эмпирических наблюдений (Christensen, 2014, p. 2), самих по себе диффузионных процессов недостаточно, в частности, для описания цен финансовых активов. Согласно гипотезе эффективного рынка новости должны мгновенно отражаться в изменениях цен финансовых активов. Поэтому скачки являются естественным и привлекательным дополнительным компонентом для моделирования эффекта таких новостей.

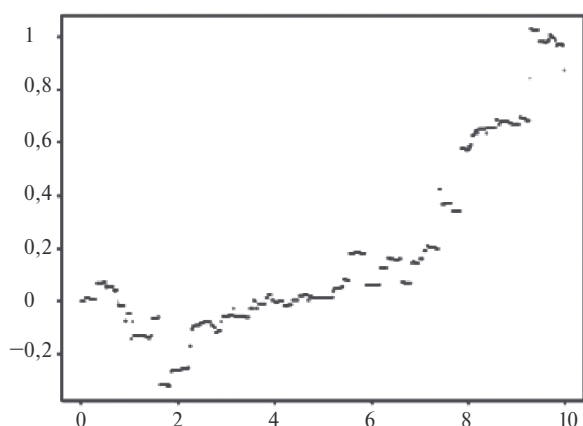


Рис. 1. Скачкообразный процесс с бесконечной активностью и конечной вариацией (вариация гамма)

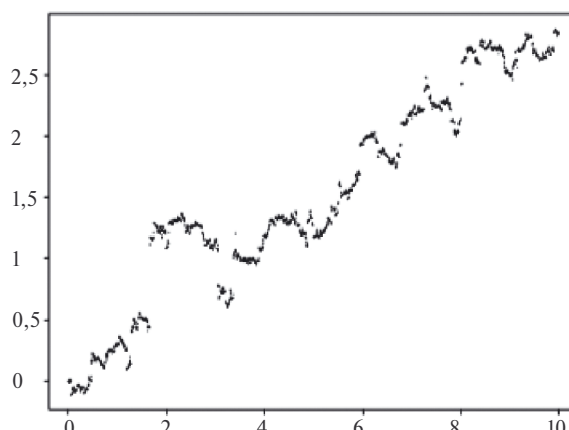


Рис. 2. Скачкообразный процесс с бесконечной активностью и бесконечной вариацией (нормальный обратный гауссовский процесс)

На рис. 1–2 оба процесса представляют собой предельный случай семейства гиперболических функций. Симуляция произведена в программном пакете R.

Для класса пороговых вариаций⁸ в работе (Mancini, 2009, р. 271–276) рассматриваются случаи конечной и бесконечной активности скачков (моделируется посредством процесса Леви). Сравнивая результаты с оценками из класса многостепенных вариаций, автор утверждает, полагаясь на результаты симуляций, что пороговые оценки, развитые ею, не менее эффективны, а в случае когда частота доступных данных ограничена, — даже более эффективны. Более того, пороговая оценка является эффективной даже при неравноудаленных (по времени) эмпирических данных, а также в случае меняющегося порога — вне зависимости от динамики дисперсии. В работе (Mancini, 2012, р. 444) в отношении оценки интегрированной дисперсии обосновывается необходимость дополнительных исследований, посвященных скорости сходимости оценки в случае бесконечной активности скачков. Существуют LM-тесты. Отсутствуют результаты применимости данных тестов локально (то есть не предельные случаи). Более того, необходимо разработать непараметрические тесты для обнаружения скачков для спот-дисперсии.

ЗАКЛЮЧЕНИЕ

Все большая доступность больших данных, растущие возможности их обработки, новые результаты теории семимартингалов расширяют горизонты развития непараметрических методов, каким и является метод оценивания интегрированной дисперсии на основе диффузионно-скачкообразных процессов. ОИД должна пониматься как один из важнейших параметров прогноза динамики цены финансового актива, в частности сложного финансового инструмента, дериватива.

В своей описательной функции диффузионные процессы имеют ряд ограничений. Причина, по которой в процесс включен дополнительный компонент — скачки, приобретающие в силу новых практических возможностей и теоретических усовершенствований все большую роль в финансовой математике. Так, именно *устойчивость* оценки интегрированной дисперсии к скачкам и определяет степень ее состоятельности. Данная степень устойчивости показывает, насколько применима асимптотическая теория, где ее главными условиями являются скорость сходимости оценки и предельный закон распределения ошибки оценки.

В настоящей работе приведен ряд доводов в пользу необходимости дальнейших исследований влияния активности скачков на состоятельность ОИД. Данные доводы имеют в большей степени теоретический характер. Результаты, достигнутые асимптотической теорией для конечной активности скачков и для некоторых типов бесконечной активности скачков, позволяют очертить контуры необходимого исследования и наметить направление развития анализа асимптотической теории для оценок интегрированной дисперсии.

⁸ Теоретическая основа данного класса была заложена в (Mancini, 2001).

СПИСОК ЛИТЕРАТУРЫ / REFERENCES

- Andersen T.G., Dobrev D., Schaumburg E. (2012). Jump robust volatility estimation using nearest neighbour truncation. *Journal of Econometrics*, 169, 75–93.
- Aït-Sahalia Y., Jacod J. (2014). *High frequency financial econometrics*. Princeton: Princeton University Press.
- Barndorff-Nielsen O.E., Graversen S.E., Jacod J., Podolskij M., Shephard N. (2006). A central limit theorem for realized power and bipower variations of continuous semimartingales. In: *From stochastic calculus to mathematical finance: The Shiryaev festschrift*. Y. Kabanov, R. Lipster (eds.), 33–68. Berlin, Heidelberg: Springer–Verlag.
- Barndorff-Nielsen O.E., Shephard N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. Part 2. *Journal of the Royal Statistical Society Series B — Statistical Methodology*, 64, 253–280.
- Barndorff-Nielsen O.E., Shephard N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, 2, 1, 1–37.
- Barndorff-Nielsen O.E., Shephard N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, 4, 1, 1–30.
- Barndorff-Nielsen O.E., Shephard N. (2007). Variation, jumps, market frictions and high frequency data in financial econometrics. In: R. Blundell, T. Persson, W. Newey (eds.), *Advances in Economics and Econometrics. Theory and Applications*. Ninth World Congress. Cambridge: Cambridge University Press.
- Barndorff-Nielsen O.E., Shephard N., Winkel M. (2006). Limit theorems for multipower variation in the presence of jumps in financial econometrics. *Stochastic Processes and their Applications*, 116, 796–806.
- Christensen K. (2016). High frequency data econometrics. PhD course, Aarhus University. Available at: https://econ.au.dk/fileadmin/site_files/filer_oekonomi/subsites/creates/Diverse_2016/PhD_High-Frequency/Slides_day_2.pdf
- Eberlein E. (2010). Jump processes. In: *Encyclopaedia of quantitative finance by Rama Cont*. In 4 vols., 1850–1869. Chichester: John Wiley & Sons.
- Jacod J., Shiryaev A.N. (2003). *Limit theorems for stochastic processes*. 2nd ed. N.Y.: Springer-Verlag.
- Mancini C. (2001). Disentangling the jumps of the diffusion in a geometric jumping Brownian motion. *Giornale dell'Istituto Italiano degli Attuari*, 64, 19–47.
- Mancini C. (2009). Non parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36, 270–296.
- Mancini C. (2012). Jumps. *Handbook of volatility models and their applications*. L. Bauwens, C. Hafner, S. Laurent (eds.). New Jersey: Wiley and Sons, 403–445.
- Protter P. (2004). *Stochastic integration and differential equations*. N.Y.: Springer–Verlag.
- Shiryaev A.N. (1999). *Essential of stochastic finance: Facts, models and theory*. Singapore: World Scientific.
- Shiryaev A.N. (1999). *Essential of stochastic finance: Facts, models, theory*. Advanced Series on Statistical Science & Applied Probability. 1st ed. Singapore: Scientific Pub. Co Inc.
- Woerner H.C.J. (2004). Power and multipower variance: Inference for high frequency data. In: Shiryaev A.I. *Handbook of Stochastic Finance*. Chapter 12, 343–364. N.Y.: Springer–Verlag.

Methods for estimating integrated variance: Jump robustness issues in high frequency time series

© 2024 Z.O. Kosymov

Z.O. Kosymov,

Lomonosov Moscow State University, Moscow, Russia; e-mail: zohirsho1@gmail.com

Received 13.02.2024

The author is grateful to his supervisors — Lukas E.H. and Gushchin A.A — for their notes and valuable comments for the text of this article. It is also worth noting the importance of advice by Kartaev F.S. while preparing the article for publication.

Abstract. Integrated variance is a volatility measure of a process in continuous time, which is applicable in financial mathematics as a means to optimize a portfolio, project dynamics of a price of a financial asset. The consistency of an estimator for integrated variance of a random process is at the core of this article. A fundamental diffusion process is extended by adding a jump component as a means of improving the descriptive function of the process. It is activity of jumps that is a factor subject to which is the consistency of the estimator for integrated variance. Due to this fact, consistency is defined as an extent to which the estimator at hand is jump robust. Two main methods for estimating integrated variance are considered and the capacity of corresponding estimators to withstand the effect of jumps while converging is briefly analyzed. The arguments indicating a necessity for further research of the effect of jumps with reference to works of the authors who have established a ground for analysis of integrated variance and those works containing main asymptotic results for consistency of integrated variance estimators are elaborated. Based on this, avenue for further research and development of asymptotic theory for consistency of an estimator for integrated variance is identified.

Keywords: integrated variance, diffusion process, consistency, jumps, jump robustness, asymptotic theory.

JEL Classification: C02, C32, C53, C65.

UDC: 519.214.

For reference: **Kosymov Z.O.** (2024). Methods for estimating integrated variance: Jump robustness issues in high frequency time series. *Economics and Mathematical Methods*, 60, 3, 107–117. DOI: 10.31857/S0424738824030095 (in Russian).