

ВЫСШАЯ ШКОЛА ЭКОНОМИКИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Л.Г. Егорова, И.Ю. Климяк

**ПРИМЕНЕНИЕ ПРОЦЕССОВ ХОУКСА
ДЛЯ ПРОГНОЗИРОВАНИЯ
ФИНАНСОВЫХ РИСКОВ**

Препринт WP7/2017/02

Серия WP7

Математические методы анализа решений
в экономике, бизнесе и политике

Москва
2017

УДК 519.8
ББК 22.176
Е30

Редакторы серии WP7
«Математические методы анализа решений
в экономике, бизнесе и политике»
Ф.Т. Алескеров, В.В. Подиновский, Б.Г. Миркин

Е30 **Егорова, Л. Г., Климяк, И. Ю.**

Применение процессов Хоукса для прогнозирования финансовых рисков [Текст] : препринт WP7/2017/02 / Л. Г. Егорова, И. Ю. Климяк ; Нац. исслед. ун-т «Высшая школа экономики». – М. : Изд. дом Высшей школы экономики, 2017. – (Серия WP7 «Математические методы анализа решений в экономике, бизнесе и политике»). – 36 с. – 20 экз.

Рассматриваются модели прогнозирования потрясений на фондовых рынках, учитывающих эндогенный характер подобных явлений, на основе процессов Хоукса. Интенсивность процессов Хоукса зависит от предыдущих событий, что позволяет моделировать кластеризацию событий. Модели протестированы на данных фондового индекса S&P500 и валютной паре USD/RUB и демонстрируют неплохую прогнозную силу.

УДК 519.8
ББК 22.176

Егорова Л.Г., Международная научно-учебная лаборатория анализа и выбора решений НИУ ВШЭ, департамент математики факультета экономических наук НИУ ВШЭ, Москва, Россия; legorova@hse.ru.

Климяк И.Ю., магистратура факультета экономических наук НИУ ВШЭ, Москва, Россия.

**Препринты Национального исследовательского университета
«Высшая школа экономики» размещаются по адресу: <http://www.hse.ru/org/hse/wp>**

© Егорова Л. Г., 2017
© Климяк И. Ю., 2017
© Оформление. Издательский дом
Высшей школы экономики, 2017

Введение

Биржевой кризис – это резкое изменение (как правило, падение) за продолжительное время биржевых котировок одной или нескольких ценных бумаг или всех (почти всех) ценных бумаг, торгуемых на бирже.

Биржевые кризисы могут быть вызваны внутренними и внешними причинами. Биржа чутко реагирует на финансовые, политические, социальные изменения в стране и мире, даже на природные явления и стихийные бедствия, и биржевой кризис может быть вызван и внешними (неэкономическими) причинами – войнами, стихийными бедствиями, политическими действиями, например, введением эмбарго. Как правило, такие внешние причины мало предсказуемы, но могут оказать сильное влияние на биржу, например, в первый торговый день после терактов 11 сентября 2001 г. индекс Доу-Джонса потерял 7,1%.

Внутренняя нестабильность биржевых процессов обусловлена:

1. Большой сложностью системы. Биржа представляет собой сложную социально-экономическую систему со скрытыми взаимосвязями между элементами. Огромное число участников и их разнородность (по характеристикам, целям, возможностям, информированности и т.п.), быстрота взаимодействия на бирже, рефлексивность как положительная обратная связь (цены зависят от принятых участниками биржи решений, которые они принимают, основываясь на ценах, и т.д.) – все это порождает неопределенность и непредсказуемость будущего состояния биржи для ее участников.

2. Возможной нерациональностью некоторых ее участников. Человек как лицо, принимающее решение, не всегда рационален, часто принимает решения под воздействием эмоций, склонен к стадному поведению и т.д. Многие трейдеры не обладают способностью правильно предсказывать направление движения цен [Odean, 1999; Barber, Odean, 2000], и даже у финансовых аналитиков доля верных прогнозов часто не достигает 50% [Kahneman, 2011; Penikas, Proskurin, 2013].

3. Намеренными действиями по дестабилизации биржи. Причиной биржевого кризиса могут быть сознательные действия его участников, создающих спекулятивный пузырь – когда участники торгов, даже понимая, что цены «оторвались» от разумных, реальных значений, продолжают покупать или продавать в надежде успеть продать раньше, чем цены снизятся

[Devenow, Welch, 1996]. Кроме того, спекулятивные пузыри приводят к обвалам, скорее всего, из-за «стадного» поведения инвесторов [Sornette, 2003].

Хотя биржевые кризисы и могут быть вызваны экзогенным фактором, но неустойчивость биржи гораздо сильнее определяется эндогенными причинами. Поэтому для прогноза потрясений на фондовом рынке необходимо моделировать самовозбуждающиеся следствия поведения инвесторов, когда действия одних участников побуждают действовать других участников рынка. Такой характер самовозбуждения естественным образом наблюдается в сейсмической области: последовательность слабых толчков может вызвать более сильное землетрясение, которое, в свою очередь, может генерировать новые толчки и т.д. То же самое происходит и на фондовом рынке – «стадное» поведение приводит к тому, что обвалы локально самоусиливаются.

В данной работе мы используем модель эпидемической толчковой последовательности (Epidemic-type Aftershock Sequence), чтобы описать и смоделировать возможность появления финансовых крахов как экзогенной, так и эндогенной природы. Модель ETAS была разработана в 1988 г. для прогнозирования землетрясений в Японии [Ogata, 1988]. В модели ETAS для моделирования скорости возникновения землетрясений выше определенного порога используется процесс Хоукса. Интенсивность в процессе Хоукса увеличивается, когда происходит скачок (или шок), после которого скорость убывает как функция времени, прошедшего после прыжка: $\lambda(t | H_t) = \mu(t) + \sum_{t_i < t} \phi(t - t_i)$. Здесь H_t обозначает историю прошедших событий, т.е. даты t_i произошедших к моменту t событий. Поскольку вероятность последующих скачков возрастает после каждого случившегося скачка, процесс Хоукса называется самовозбуждающимся.

Модель ETAS применялась к финансовым данным, например, для моделирования данных о прибылях сделок купли-продажи [Hewlett, 2006], продолжительности между сделками [Bauwens, Hautsch, 2009] или доходности по нескольким индексам [Ait-Sahalia et al., 2015], а также для оценки уровня преступности [Mohler et al., 2011] и распространения красных банановых растений [Balderama et al., 2011].

В этой работе на основе данных фондового рынка (индекса S&P500 и валютной пары USD/RUB) мы предлагаем систему раннего предупреждения (EWS) обвалов на фондовом рынке на основе модели ETAS. Система

EWS тестировалась на данных S&P500 и валютной пары USD/RUB. Мы анализируем эффективность модели с использованием показателя Хансена – Куйпера [Tressa et al., 2012] и сравниваем с наиболее распространенной моделью объяснения кластеризации волатильности на финансовых рынках GARCH.

Результаты оценки подтверждают, что ситуации обвалов финансовых рынков являются самоусиливающимися. Тестируя наши модели по данным S&P500 во время недавнего финансового кризиса, мы обнаружили положительные оценки Хансена – Куйпера. Таким образом, поскольку наша модель использует самовозбуждающееся поведение доходностей акций вблизи финансовых обвалов на рынке, она способна предсказывать вероятности шоков в среднесрочной перспективе.

Структура работы такова: первый раздел посвящен описанию математических моделей точечных процессов, процессов Пуассона, процессов Хоукса как расширения процессов Пуассона и их свойств. Во втором разделе приводятся алгоритмы симуляции процессов Хоукса. В третьем разделе приводятся результаты применения моделей к реальным финансовым данным, приведены оценки параметров моделей и сравнение с наиболее распространенными моделями вида GARCH.

Благодарности. Препринт подготовлен в ходе проведения исследования в рамках Программы фундаментальных исследований Национального исследовательского университета «Высшая школа экономики» (НИУ ВШЭ) и с использованием средств субсидии в рамках государственной поддержки ведущих университетов Российской Федерации «5–100». Авторы благодарны Ф.Т. Алескерову за ценные замечания и помощь.

1. Математическая модель процессов Хоукса

1.1. Точечные процессы

Для анализа финансовых процессов необходимо подсчитывать события с течением времени. Мы будем рассматривать точечные процессы, представляющие, по сути, временной ряд появлений некоторых событий – каждая точка представляет собой «время и/или место события, например, удар молнии, землетрясение или биржевую сделку» [Schoenberg, 2016].

Пусть (Ω, \mathcal{F}, P) – вероятностное пространство и случайные моменты времени $\{t_i: 0 = t_0 < t_1 < t_2 < \dots\}$ обозначают моменты наступления

событий. Последовательность $\{t_i\}_{i \in N^*}$, $N^* = N \cup \{0\}$, называется точечным процессом (point process). Процесс $N(t) = N_t = \sum_{i \in N^*} 1_{\{t_i \leq t\}}$, показывающий сколько событий наступило к моменту времени t , называется считающим процессом (counting process), связанным с процессом $\{t_i\}_{i \in N^*}$. Также можно рассматривать процесс $\{\delta t_i\}_{i \in N}$, $\forall i \in N \delta t_i = t_i - t_{i-1}$, который отмечает длительность временных интервалов между последовательными событиями и называется процессом дюраций (duration process), связанным с процессом $\{t_i\}_{i \in N^*}$.

Процесс $N(t)$ является неубывающим, непрерывным справа, может принимать только неотрицательные целочисленные значения (не может быть отрицательных событий) и имеет разрыв слева в каждом событии t_j (когда новое событие повышает значение функции на единицу).

Если, помимо значений времени, точечный процесс содержит дополнительные переменные, то он называется выделенным точечным процессом (т.е. функцией с несколькими входными переменными). Финансовым примером точечного процесса с выделенными точками будет набор точек $\{(t_i, m_i)\}$, который содержит не только время различных рыночных сделок t_i , но и размеры торгов или цену m_i .

Пусть N_t – считающий процесс, заданный на измеримом пространстве (Ω, \mathcal{F}) и адаптированный к естественной фильтрации \mathcal{F}_t , т.е. множество $\{\omega: N_t(\omega) \leq x\}$ принадлежит σ -алгебре \mathcal{F}_t при любом x . Фильтрацией случайного процесса называется семейство σ -алгебр $\{\mathcal{F}_t\}_{t \in T}$, такое что $\forall s \leq t, s, t \in T \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$, естественной фильтрацией случайного процесса N_t называется семейство σ -алгебр $\{\mathcal{F}_t\}_{t \in T} = \sigma\{N_s | s \leq t, s \in T\}$. Условие измеримости гарантирует, что интересующие нас подмножества пространства Ω принадлежат σ -алгебре \mathcal{F}_t и, следовательно, имеют вероятности.

Определим процесс интенсивности следующим образом:

$$\lambda(t|\mathcal{F}_t) = \lim_{\Delta t \rightarrow 0} E \left[\frac{N(t + \Delta t) - N(t)}{\Delta t} | \mathcal{F}_t \right]$$

или эквивалентным образом:

$$\lambda(t|\mathcal{F}_t) = \lim_{\Delta t \rightarrow 0} \frac{P[N(t + \Delta t) - N(t) > 0 | \mathcal{F}_t]}{\Delta t}.$$

В литературе иногда используется термин «плотность», а не «интенсивность». В дальнейшем для удобства опустим упоминание о фильтрации \mathcal{F}_t .

1.2. Процессы Пуассона

В простейшем случае при $P[N(t + \Delta t) - N(t) = 1] = \lambda\Delta t + o(\Delta t)$, $P[N(t + \Delta t) - N(t) > 1] = o(\Delta t)$ получается однородный процесс Пуассона (homogeneous Poisson process), в котором интенсивность не зависит от произошедших событий и длительности, $\{\delta t_i\}_{i \in \mathbb{N}}$ независимы и одинаково распределены по экспоненциальному закону с параметром λ .

Собственно, однородный процесс Пуассона моделирует случайную последовательность событий во времени. Например, физики используют пуассоновские процессы для описания эмиссии радиоактивных частиц, а инженеры-телефонисты применяют их для описания времени поступления вызовов на телефонную станцию. Это естественно, если считать, что наступление рассматриваемых событий вызвано разными атомами (в первом случае) или абонентами (во втором случае), независимыми друг от друга.

Понятно, что однородный процесс Пуассона может применяться только в случаях, когда изучаемые события независимы, но во многих реальных задачах предпосылка о независимости не выполняется и необходимо моделировать более сложную структуру точечного процесса.

Например, интенсивность процесса может быть функцией от времени $\lambda(t)$ – такие процессы называются неоднородными процессами Пуассона (non-homogeneous Poisson process). Например, в [Massey et al., 1996] трафик звонков в сети компании AT&T моделировался с помощью неоднородного процесса Пуассона с интенсивностью в виде кусочно-линейной функции. В [Weinberg et al., 2007] при моделировании поступлений телефонных звонков в кол-центр крупного банка неоднородность процесса Пуассона используется для моделирования различных паттернов звонков в различные дни недели. В [Lee et al., 1991] неоднородный процесс Пуассона применялся для предсказания штормов: авторы моделировали такие события (high-wave events) с помощью функции интенсивности с циклической компонентой вида $\lambda(t) = \exp(\sum_{i=0}^m \alpha_i t^i + \gamma \sin(\omega t + \phi))$.

Дальнейшее усложнение связано с рассмотрением интенсивности процесса не просто как функции, а как случайного процесса. Одна из первых работ на эту тему была опубликована в 1955 г. Д. Коксом. В его честь такие модели называются процессами Кокса (Cox process), или двойными стохастическими процессами (doubly stochastic Poisson process), или смешанными пуассоновскими процессами (mixed Poisson process). В двойном

стохастическом процессе $\lambda(t)$ является стационарным случайным процессом, так что можно считать, что $\lambda(t)$ определен для всех t до того, как рассматривается процесс $N(t)$.

Такого рода процессы используются для моделирования различных паттернов нейронной реакции (spike trains) в нейронной сети [Krumin, Shoham, 2009], возникающих, например, в мозге при обработке информации или при обучении. Для отражения сложной корреляции между процессами были предложены различные алгоритмы трансформации случайного гауссовского процесса с заранее искаженной структурой корреляции в неотрицательный процесс интенсивности, который затем используется для разыгрывания двойного стохастического пуассоновского процесса с требуемой корреляцией. Также двойные стохастические процессы используются в финансовой математике, например, в работе [Lando, 1998] моделируется кредитный риск, и двойной стохастический процесс используется для моделирования процесса перехода контрагента из одной кредитной категории в другую и возможного дефолта контрагента.

1.3. Процессы Хоукса

Финансовые данные часто требуют не только моделирования различной интенсивности появления событий, но и моделирования при этом такого эффекта, как кластеризация, когда появление одного события влияет на вероятность появления последующих.

Такой класс случайных процессов называется «самовозбуждающиеся процессы» (self-exciting process). Поток событий в самовозбуждающихся процессах учитывает влияние произошедших событий, увеличивая вероятность наступления новых событий после произошедшего. Для этого можно представить процесс интенсивности в следующем виде:

$$\lambda(t) = \lambda_0(t) + \int_{-\infty}^t v(t-s) dN_s = \lambda_0(t) + \sum_{t_i < t} v(t-t_i),$$

где $\lambda_0: \mathbb{R} \rightarrow \mathbb{R}_+$ обозначает детерминированную «базовую» интенсивность, а функция $v: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ отражает влияние произошедших событий в моменты времени t_i на текущее значение интенсивности. Функцию $\lambda_0(t)$ называют также скоростью появления внешних событий, а функцию $v(t)$ – ядром.

Алан Хоукс [Hawkes, 1971] предложил использовать экспоненциальную функцию

$$v(t) = \sum_{j=1}^P \alpha_j e^{-\beta_j t},$$

и в этом случае интенсивность будет задаваться как

$$\lambda(t) = \lambda_0(t) + \int_{-\infty}^t \sum_{j=1}^P \alpha_j e^{-\beta_j(t-s)} dN_s = \lambda_0(t) + \sum_{t_i < t} \sum_{j=1}^P \alpha_j e^{-\beta_j(t-t_i)}.$$

Простейшая форма интенсивности с $P = 1$ и постоянной базовой интенсивностью $\lambda_0(t) = \lambda_0$ выглядит так:

$$\lambda(t) = \lambda_0 + \int_{-\infty}^t \alpha e^{-\beta(t-s)} dN_s = \lambda_0 + \sum_{t_i < t} \alpha e^{-\beta(t-t_i)}.$$

Здесь параметр α задает «скачок» интенсивности вероятности появления последующих событий при появлении очередного события, а параметр β определяет скорость снижения интенсивности обратно к базовому уровню с течением времени.

В предположении, что процесс является стационарным (в широком смысле), подсчитаем его математическое ожидание $E[\lambda(t)]$:

$$\begin{aligned} E[\lambda(t)] &= E \left[\lambda_0(t) + \int_{-\infty}^t v(t-s) dN_s \right] = \lambda_0 + E \left[\int_{-\infty}^t v(t-s) \lambda(s) ds \right] = \\ &= \lambda_0 + E[\lambda(s)] \int_{-\infty}^t v(t-s) ds = \frac{\lambda_0}{1 - \int_0^{\infty} v(u) du} = \frac{\lambda_0}{1 - \sum_{j=1}^P \frac{\alpha_j}{\beta_j}} \end{aligned}$$

Поэтому, для того чтобы случайный процесс интенсивности Хоукса был стационарным (в широком смысле), необходимо выполнение условия

$$\sum_{j=1}^P \frac{\alpha_j}{\beta_j} < 1.$$

Важную роль в процессах Хоукса играет коэффициент ветвления n – он представляет собой среднее число дочерних событий первого поколе-

ния от инициировавшего их события. Этот коэффициент, в случае экспоненциального ядра в процессе Хоукса, равен

$$n = \sum_{j=1}^P \frac{\alpha_j}{\beta_j}.$$

Для оценки параметров процесса Хоукса [Ogata, 1978; Ozaki, 1979; Voogaard van den, 1986] можно рекурсивно вывести функцию правдоподобия в виде

$$\begin{aligned} \ln L(\{t_i\}_{i=1,\dots,n}) &= t_n - \int_0^{t_n} \lambda_0(s) ds - \sum_{i=1}^n \sum_{j=1}^P \frac{\alpha_j}{\beta_j} (1 - e^{-\beta_j(t_n - t_i)}) + \\ &+ \sum_{i=1}^n \ln \left[\lambda_0(t_i) + \sum_{j=1}^P \alpha_j R_j(i) \right], \\ \forall j, R_j(1) &= 0, R_j(i) = e^{-\beta_j(t_i - t_{i-1})} (1 + R_j(i-1)). \end{aligned}$$

В [Ogata, 1978] доказано, что для одномерного процесса Хоукса оценки, полученные с помощью такой функции правдоподобия, состоятельны, асимптотически эффективны и асимптотически нормальны.

Легко можно обобщить такую конструкцию на многомерный случай. Пусть $\{(t^m)_i\}_{i \in N^*}, m = 1, \dots, M$ – M -мерный точечный процесс и $N_t = (N_t^1, \dots, N_t^M)$ – связанный с ним считающий процесс. Тогда многомерный процесс Хоукса с интенсивностями $\lambda^m, m = 1, \dots, M$ определяется как

$$\lambda^m(t) = \lambda_0^m(t) + \sum_{n=1}^M \int_0^t \sum_{j=1}^P \alpha_j^{mn} e^{-\beta_j^{mn}(t-s)} dN_s^n$$

и в простейшей форме с $P = 1$ и постоянной базовой интенсивностью

$$\lambda^m(t) = \lambda_0^m + \sum_{n=1}^M \sum_{t_i^n < t} \alpha^{mn} e^{-\beta^{mn}(t-t_i^n)}.$$

Достаточным условием для стационарности в этом случае является условие того, что спектральный радиус (т.е. максимальное по модулю собственное число матрицы) равно

$$\Gamma = \int_0^\infty G(u) du = \left(\frac{\alpha^{mn}}{\beta^{mn}} \right)_{m,n=1,\dots,M}, \text{ где } G(t) = (\alpha^{mn} e^{-\beta^{mn}t})_{m,n=1,\dots,M}.$$

Для стационарного процесса математическое ожидание будет равно

$$E[\lambda(t)] = \left(1 - \int_0^{\infty} G(u)du\right)^{-1} \lambda_0.$$

Так же, как и для одномерного случая, можно рекурсивно вывести функцию правдоподобия и оценить параметры.

Процессы Хоукса позволяют моделировать кластеризованность моментов наступления событий. Так, в работе [Hewlett, 2006] анализировались моменты выставления заявок на покупку (Buy) и продажу (Sell) ценных бумаг с помощью двумерного процесса Хоукса с интенсивностями

$$\begin{aligned} \lambda^B(t) &= \lambda_0^B + \int_0^t \alpha^{BB} e^{-\beta^{BB}(t-u)} dN_u^B + \int_0^t \alpha^{BS} e^{-\beta^{BS}(t-u)} dN_u^S, \\ \lambda^S(t) &= \lambda_0^S + \int_0^t \alpha^{SB} e^{-\beta^{SB}(t-u)} dN_u^B + \int_0^t \alpha^{SS} e^{-\beta^{SS}(t-u)} dN_u^S. \end{aligned}$$

Таким образом, введенные интенсивности учитывают эффект как самовозбуждения (повышения вероятности выставления заявок при выставлении заявки такого же типа), так и перекрестного возбуждения (повышения вероятности выставления заявок противоположного типа). Hewlett предположил, что потоки симметричны и $\lambda_0^B = \lambda_0^S = \lambda_0$, $\alpha^{BB} = \alpha^{SS} = \alpha^{self}$, $\beta^{BB} = \beta^{SS} = \beta^{self}$, $\alpha^{SB} = \alpha^{BS} = \alpha^{cross}$, $\beta^{SB} = \beta^{BS} = \beta^{cross}$. Автор оценил параметры такой модели на данных по покупкам/продажам на валютном рынке (два месяца торгов на валютной паре евро/польский злотый) и получил следующие оценки:

$$\lambda_0 = 0,0033, \alpha^{cross} = 0, \alpha^{self} = 0,0169, \beta^{self} = 0,0286,$$

т.е. взаимного влияния ордера на покупку и продажу не оказывают, зато появление ордера на покупку/продажу вызывает повышение вероятности появления ордера на покупку/продажу. Эта модель была существенным образом расширена в [Large, 2007], где возникающие на бирже заявки были классифицированы в десять категорий.

Исследование, аналогичное [Hewlett, 2006], но без ограничения симметричности параметров, было проведено в [Toke, 2011] на данных торгов акций VNPP.PA в течение 14 торговых дней. Автор оценил параметры и выяснил, что параметры, отвечающие за перекрестное возбуждение, действительно на порядок ниже остальных параметров.

Также автор исследовал модель взаимосвязи двух типов ордеров – рыночных и лимитных, чтобы понять, насколько вероятно появление ордера одного типа сразу после появления ордера другого типа.

Процессы Хоукса используются также для моделирования цен [Vasuy et al., 2011]: цена представляется в виде разности двух процессов Хоукса $p(t) = N^1(t) - N^2(t)$, интенсивность каждого из которых учитывает только перекрестное влияние. Таким образом, авторы моделируют эффект “mean reversion”, отражающий концепцию возврата цены к среднему уровню при отклонении в большую или меньшую сторону от средней. В [Ait-Sahalia et al., 2015] представлена модель динамики доходностей активов с периодами кризисов и «заражением», когда событие в одном регионе вызывает увеличение интенсивности наступления следующего события как в самом регионе, так и в других регионах.

1.4. Модель ETAS

Модель ETAS основана на самовозбуждающемся точечном процессе Хоукса и представляет собой модель ветвления, в которой каждое событие может инициировать последующие события, которые также, в свою очередь, могут инициировать последующие события.

Рассмотрим процесс с событиями $(t_1, m_1), \dots, (t_n, m_n)$, где t_i определяет время, а m_i – амплитуду события i . Пусть $H_t = \{(t_i, m_i) : t_i < t\}$ представляет всю историю событий вплоть до момента времени t . Условная интенсивность процесса Хоукса определяется выражением

$$\lambda(t | \theta; H_t) = \mu_0 + \sum_{i:t_i < t} \phi(t - t_i, m_i),$$

где $\mu_0 > 0$ и $\phi(s - t_i, m_i) > 0$ всякий раз, когда $s > 0$ и 0 – в иных случаях. Теперь интенсивность состоит из постоянного члена μ_0 и самовозбуждающейся функции $\phi(s)$, которая зависит не только от времени, прошедшего с тех скачков, которые произошли до t , но и их размера.

Используем ядра, предложенные Огатай в 1988 г.:

$$\phi_{pow}(t - t_i, m_i) = \frac{K_0}{(t - t_i + c)^p} e^{\alpha(m_i - M_0)}; \quad (1)$$

$$\phi_{exp}(t - t_i, m_i) = K_0 e^{-\beta(t-t_i) + \alpha(m_i - M_0)}, \quad (2)$$

где K_0 – максимальная интенсивность функции [Ogata, 1988].

Вероятность возникновения события, инициирующего последующее событие, затухает в соответствии со степенным законом в уравнении (1), или по экспоненциальному закону в (2). Параметры p и β определяют, насколько быстро вероятность инициирования последующих событий затухает с соответствующим временем, прошедшим после события. Когда p и β большие, вероятность того, что событие вызывает другое событие, быстро снижается.

В формулах (1) и (2) M_0 обозначает минимальную величину событий, которые моделируются и маркируются как событие. Когда $\alpha \neq 0$, интенсивность, с которой последующие события возбуждаются наступившим событием, зависит от размера события. Предполагается $\alpha > 0$, потому что более крупные события должны вызывать больше событий, чем более мелкие события. Следовательно, вероятность возбуждения последующих событий увеличивается с размером избыточной величины события ($x = m_i - M_0$). Чем больше α , тем сильнее влияние размера событий на процесс.

Процесс будет стационарен, если коэффициент ветвления n меньше 1. Если $n \geq 1$, то количество порождаемых событий будет расти со временем до бесконечности, и процесс будет взрывным. Условия стационарности процесса Хоукса с ядрами (1) и (2) выглядят следующим образом [Ogata, 1988]:

$$n = \int_0^{\infty} \phi_{pow}(t - t_i, m_i) dt = \frac{K_0}{p} < 1 \text{ для (1),}$$

$$n = \int_0^{\infty} \phi_{exp}(t - t_i, m_i) dt = \frac{K_0}{\beta} < 1 \text{ для (2).}$$

Оценим параметры $\theta = \{\mu, K_0, \alpha, c, p\}$ моделей по максимальному правдоподобию. Логарифмическая функция правдоподобия модели:

$$\log L(\theta) = \sum_{i=1}^N \log \lambda(t_i | \theta; H_t) - \int_0^T \lambda(t_i | \theta; H_t) dt,$$

где $\lambda(t_i | \theta; H_t)$ – условная интенсивность, а t_i – время события в интервале $[0, T]$. Для быстрого нахождения максимума функции правдоподобия пользуемся методом прямого поиска симплекса Nelder – Mead. Этот метод значительно ускоряет время вычисления. В связи с большой сложностью точной оценки параметров процесса Хоукса мы использовали разные ме-

тоды оценки и алгоритмы оптимизации и протестировали нашу методику на известных данных.

Вероятность возникновения события для процесса Хоукса с условной интенсивностью $\lambda(t|\theta; H_t)$ между t_{n-1} и t_n определяется формулой

$$\begin{aligned} Pr(N(t_n) - N(t_{n-1}) > 0) &= 1 - (Pr(N(t_n) - N(t_{n-1}) = 0)) \\ &= 1 - F(t > t_n - t_{n-1}) \\ &= 1 - \exp\left(-\int_{t_{n-1}}^{t_n} \lambda(t|\theta; H_t) dt\right). \end{aligned} \quad (3)$$

Таким образом, оценив параметры реального процесса с помощью метода максимума правдоподобия, можно получить условную интенсивность. Используя условную интенсивность, определенную оценочными параметрами, можно посчитать вероятность появления события во временном промежутке с помощью (3). Эти вероятностные прогнозы составляют основу системы раннего предупреждения обвалов (EWS).

2. Симуляция процессов Хоукса

Задача моделирования процесса Хоукса очень интересна потому, что, зная истинные значения параметров процесса, мы можем изменять процесс оценки параметров и видеть поведение функции правдоподобия.

Для имитации на компьютере в [Ogata, 1981] был предложен метод, основанный на «процедуре прореживания» (thinning procedure) [Lewis, Shedler, 1979]. С ее помощью можно моделировать случайный процесс с ограниченной интенсивностью. После разыгрывания первого события интенсивность увеличивается на α , и далее события разыгрываются с интенсивностью $\lambda^* = \lambda_0 + \alpha$, но могут быть «прорежены», т.е. новое событие может быть пропущено с вероятностью $\frac{\lambda(t)}{\lambda^*}$. Если новое событие не было пропущено, то максимальная интенсивность λ^* снова увеличивается на α , и процедура продолжается. Этот же метод используется и для моделирования неоднородного процесса Пуассона.

Еще один метод был предложен в [Ozaki, 1979], в котором необходимо численно решать трансцендентное уравнение для определения времени следующего события. Впоследствии алгоритм существенно доработан и улучшен в [Simon, 2016].

Алгоритм основывается на поиске времени u ($k + 1$)-го события при наличии времен наступления k предыдущих событий. В каждой итерации цикла генерируется U – случайная величина, распределенная равномерно на промежутке $(0,1)$. Для генерации используется соотношение, предложенное Озаки в 1979 г.:

$$\log U + \int_{t_k}^u \left(\mu + \sum_{i=1}^k \phi(t - t_i | \theta) \right) dt = 0.$$

В этом уравнении известны все параметры, кроме u . Например, для простейшего вида интенсивности процесса Хоукса в виде $\lambda(t) = \mu(t) + \sum_{t_i < t} \alpha e^{-\beta(t-t_i)}$, это уравнение принимает вид

$$\log U + \mu(u - t_k) - \frac{\alpha}{\beta} \left(\sum_{i=1}^k e^{-\beta(u-t_i)} - \sum_{i=1}^k e^{-\beta(t_k-t_i)} \right) = 0.$$

Обозначим через $Y(k)$ часть этого выражения:

$$Y(k) = \left(- \sum_{i=1}^k e^{-\beta(u-t_i)} + \sum_{i=1}^k e^{-\beta(t_k-t_i)} \right).$$

В [Simon, 2016] было доказано, что

$$Y(k) = S(k)(1 - e^{-\beta(u-t_k)}),$$

где $S(k) = e^{-\beta(t_k-t_{k-1})}S(k-1) + 1$.

Таким образом, начальное уравнение решается с помощью рекурсивного соотношения

$$\log U + \mu(u - t_k) - \frac{\alpha}{\beta} S(k)(1 - e^{-\beta(u-t_k)}) = 0,$$

где $S(1) = 1$ и $S(k) = e^{-\beta(t_k-t_{k-1})}S(k-1) + 1$.

Для решения вышенаписанного уравнения используется следующая последовательность шагов [Ozaki, 1979]:

1. Обозначаем $t_k = -\log U / \mu$.

2. Обновляем значения: $u_{i+1} = u_i - \frac{f(u_i)}{f'(u_i)}$,

где $f(u) = \log U + \mu(u - t_k) - \frac{\alpha}{\beta} S(k)(1 - e^{-\beta(u-t_k)})$;

$$f'(u) = \mu + \alpha S(k)(e^{-\beta(u-t_k)}).$$

3. Остановить при условии $|u_{i+1} - u_i| \leq \epsilon^1$.

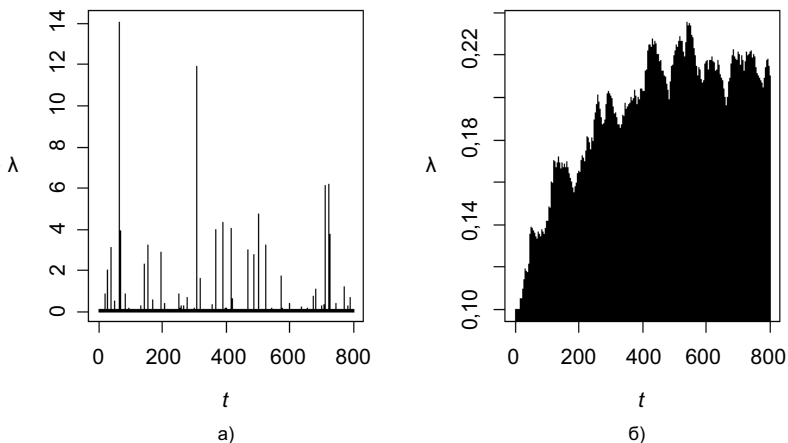


Рис. 1. Симуляция процесса с параметрами: а) $\mu = 0,1$, $\alpha = 5$, $\beta = 10$;
б) $\mu = 0,1$, $\alpha = 0,05$, $\beta = 0,1$

Параметр μ определяет интенсивность экзогенных событий (приблизительное количество событий в единицу времени), параметры α и β определяют процесс кластеризации и поведение функции интенсивности между событиями. При одинаковом коэффициенте ветвления n и μ уменьшение остальных параметров приводит к уменьшению кластеризации (рис. 1). Каждый из процессов имеет коэффициент ветвления $n = 0,5$, при этом коэффициенты α и β отличаются в 100 раз. Наибольшая разница между процессами заключается в том, что при больших значениях β интенсивность между событиями затухает очень резко, что сильно снижает вероятность появления событий-потомков.

3. Применение процессов Хоукса для прогноза среднесрочных обвалов

3.1. Построение модели EWS на основе данных индекса S&P500

Мы рассматриваем данные индекса S&P500 за период с 3 января 1950 г. по 7 апреля 2017 г. Для калибровки наших моделей используем

¹ Изначально заданная величина точности.

промежутков до 10 октября 2008 г., данные последующих девяти лет используем для выборочной оценки моделей. Набор данных состоит из суточных значений $R = \frac{p_t - p_{t-1}}{p_{t-1}}$, где p_t обозначает цену индекса в точке t .

На рис. 2 показана стоимость индекса S&P500 и доходности по нему. Серьезные обвалы и отрицательные доходности можно заметить вблизи известных исторических периодов: «черный понедельник» (1987) и спад фондового рынка 2002 г. после «доткомовского пузыря» (1997–2000). Более того, на рис. 2 видна кластеризация экстремальных доходностей, т.е. спокойные периоды с небольшими изменениями цен чередуются с «турбулентными» периодами с большими изменениями цен. Эта функция кластеризации может быть связана со «стадным» поведением инвесторов и эндогенным ростом нестабильности на финансовом рынке.

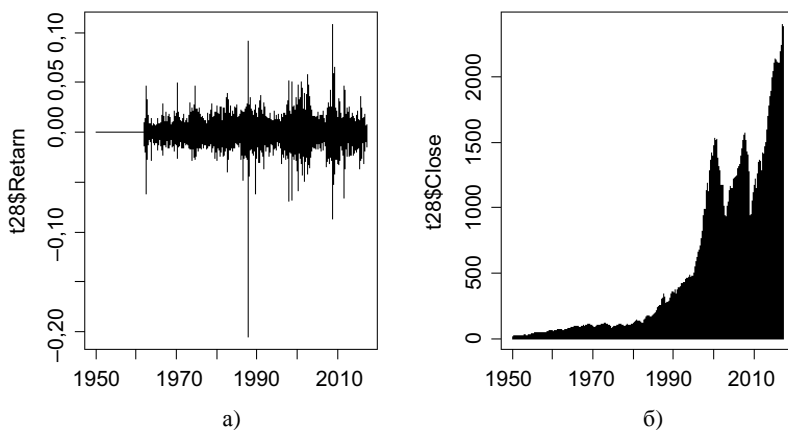


Рис. 2. а) Доходность индекса S&P500; б) Цена индекса S&P500

Мы применяем модели ETAS к 95% и 99% квантилям экстремальных отрицательных доходностей. Эти ограничения соответствуют 160 и 846 событиям за период времени 16 926 торговых дней соответственно. Выборку следует ограничивать по причине того, что количество событий (отрицательная дневная прибыль) огромно, и, как следствие, такие процессы имеют высокую интенсивность внешних событий и очень низкую вероятность возникновения потомков. Такая ситуация приводит к тому, что рассмотрение кластеризации невозможно.

Коэффициент ветвления n , т.е. ожидаемое число потомков событий во всех моделях Хоукса, находится в пределах 75–80%. Это означает, что около 20% событий произошло из-за собственного возбуждения. Интенсивность фоновых событий колеблется в пределах от 0,79 до 3,25. В моделях ETAS коэффициент ветвления гораздо ниже: 15–35%.

Поэтому мы можем утверждать, что многие мощные движения цены индекса S&P500 вызваны предыдущими движениями индекса, что ожидаемо, поскольку кластеризация и повышение вероятности нового пика при наличии старого – хорошо известное свойство доходности акций. Это подтверждает ожидание того, что обвалы являются локально самоусиливающимися и растущими эндогенно, поскольку события провоцируют появления новых событий.

Таблица 1. Оценки параметров моделей для различных фильтраций (для выборок из 1% и 5% экстремальных отрицательных доходностей фондового индекса)

Оценки параметров моделей	Процесс Хоукса с ядром $\phi(t - t_i) = \alpha e^{-\beta(t-t_i)}$		Модель ETAS с ядром (1)	
	1%	5%	1%	5%
μ	0,79	3,25	0,49	5,71
α	6,51	7,87	0,14	0,48
β	8,66	9,97		
K_0			1,86	0,02
c			0,91	136,46
p			4,97	0,13
$\log L$	– 1375,65	– 4600,88	–9,67	440,08
AIC	2757,3	9207,75	29,33	–870,19

Исходя из данных табл. 1, модели ETAS обладают более высоким потенциалом в сравнении с процессами Хоукса. Важно заметить, что с увеличением количества событий качество оценки параметров моделей Хоукса снижается, а ETAS, наоборот, увеличивается.

Таким образом, наиболее оптимальной моделью является ETAS на 5%-й выборке экстремальных отрицательных доходностей. Также рассматриваем процесс Хоукса на 1%-й выборке.

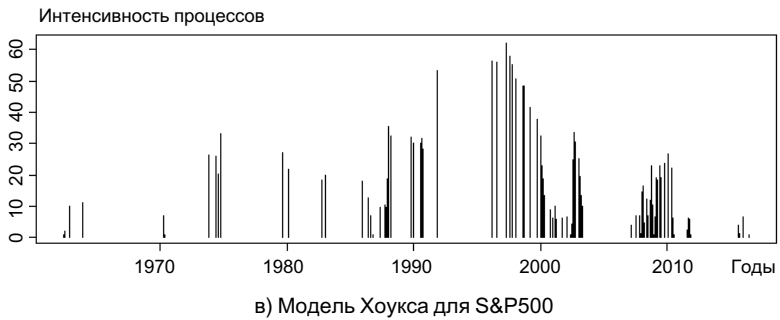
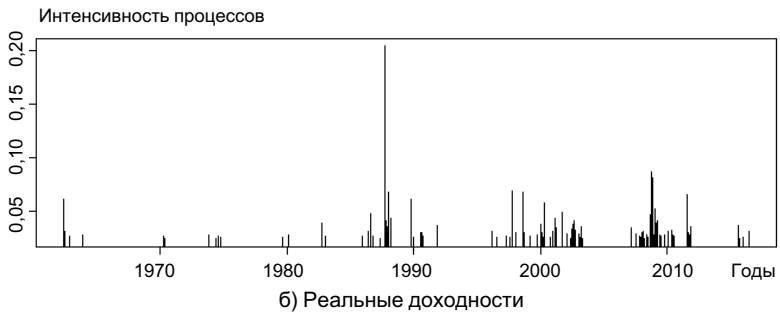
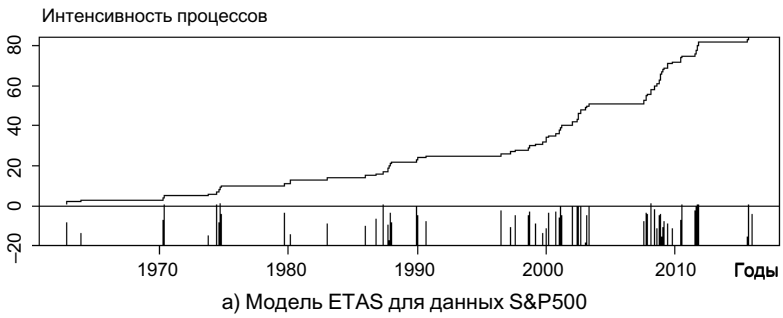
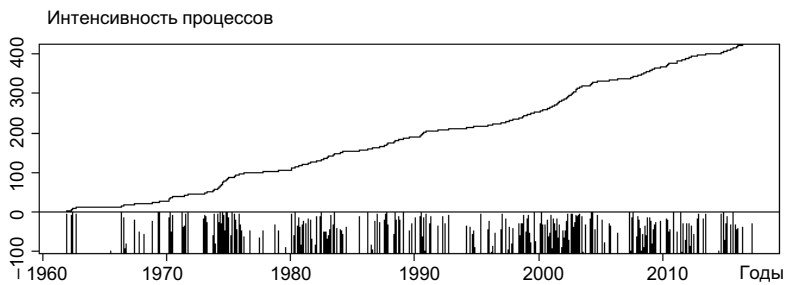
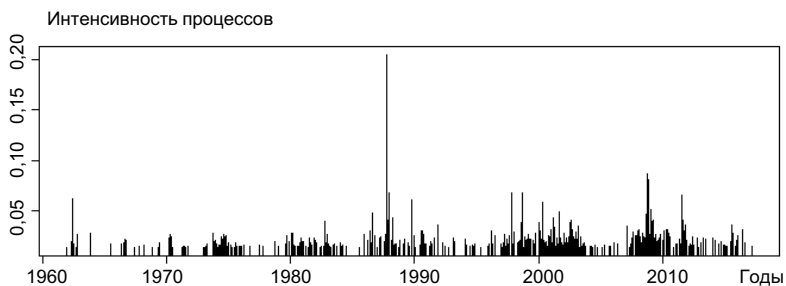


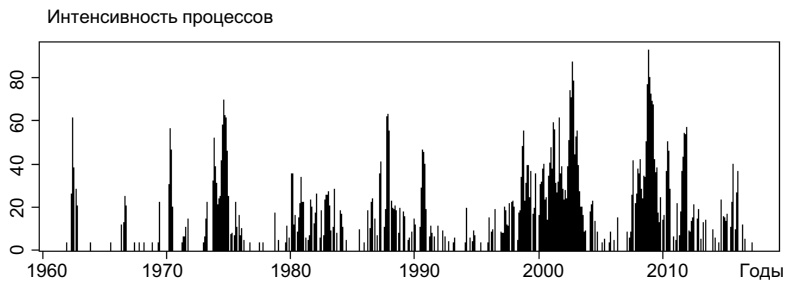
Рис. 3. Результаты расчетов, основанных на 1%-й выборке отрицательных доходностей индекса S&P500



а) Модель ETAS для данных S&P500



б) Реальные доходности



в) Модель Хоукса для S&P500

Рис. 4. Результаты расчетов, основанных на 5%-й выборке отрицательных доходностей индекса S&P 500

Оцененная интенсивность показывает большие всплески вокруг известных периодов обвалов «черный понедельник» (1987 г.) и «спад фондового рынка 2002 г.» после «доткомовского пузыря» (1997–2000 гг.). Как и ожидалось, скорость, с которой происходят события, велика вокруг обвалов прибылей, что отражает увеличение вероятности возникновения новых потомков.

Рассчитаем вероятность появления события в период следующих пяти дней, начиная с сегодняшнего. Для этого используется формула (3). Если вероятность превышает значение 0,5, считаем, что события предсказываются положительно. Ниже приведена таблица, показывающая попадания и промахи системы EWS с ядром (1) для выборки 1% максимальных отрицательных доходностей в период с 10 октября 2008 г. до 7 апреля 2017 г.

Таблица 2. Попадания и промахи системы EWS с ядром Хоукса для выборки 1% максимальных отрицательных доходностей в период с 10 октября 2008 г. по 7 апреля 2017 г.

Предсказание	Наблюдение		
	Событие произошло	Отсутствие события	Итого предсказанных
Событие произошло	25	44	69
Отсутствие события	30	2034	2064
Итого наблюдаемых	55	2078	

Большое количество промахов системы (30) при условии того, что событие наблюдалось, объясняется тем, что система может предсказывать только иммигрантские события. События, произошедшие из-за внешних шоков, система предсказать не может. В действительности, по данному индексу количество внезапных обвалов составляет 30. Тем не менее система генерирует небольшое количество ошибок второго рода (69).

Для оценки прогнозной силы модели воспользуемся скорингом KSS [Tressa et al., 2012].

KSS (Hanssen Kuiper Skill Score) – это показатель, принимающий значения от -1 до 1 , где 0 означает отсутствие прогнозирования, а отрицательные значения интерпретируются как «перевернутое прогнозирование»², 1 – полное прогнозирование.

² То есть можно добиться положительного значения скоринга, поменяв местами исходы.

$$KSS_{hawkes\ 1\%} = \frac{25 * 2034 - 44 * 30}{(25 + 30)(44 + 2034)} = 0,43.$$

На рис. 5 изображены вероятности вместе с доходностями фондового индекса для наглядности.

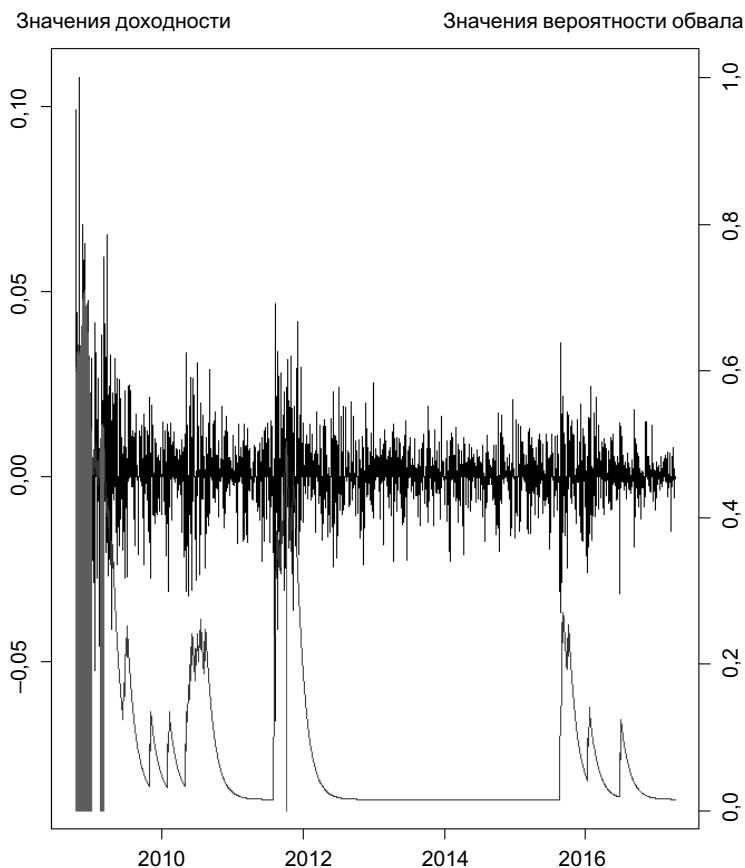


Рис. 5. Доходность индекса S&P500 (ломаная линия в центре) и вероятность появления обвала в течение следующих пяти дней, рассчитанная системой EWS на основе процесса Хоукса для 1% фильтрации событий (нижняя линия, пики означают появление события). Вертикальные линии отмечают даты, когда модель показывает вероятность обвала выше 0,5 – это начало периода и вторая половина 2011 г.

Таким образом, из 55 событий на данном промежутке было предсказано верно 25, 44 предсказано неверно, 30 не предсказано. Сравним KSS с моделью ETAS на 5%-й выборке на том же периоде.

Таблица 3. Попадания и промахи системы ETAS с ядром (14) для выборки 5% максимальных отрицательных доходностей в период с 10 октября 2008 г. по 7 апреля 2017 г.

Предсказание	Наблюдение		
	Событие произошло	Отсутствие события	Итого предсказанных
Событие произошло	99	282	381
Отсутствие события	74	1678	1842
Итого наблюдаемых	173	1960	

Следовательно, из 173 произошедших событий верно предсказаны 99, в ошибку первого рода вошли 74 события, второго – 282.

Столь высокие ошибки первого и второго рода связаны с тем, что не все события иммигрантские. Часть из них имеет внешнее происхождение.

$$KSS_{ETAS5\%} = \frac{99 * 1678 - 282 * 74}{(99 + 74)(282 + 1678)} = 0,42.$$

Как показывают коэффициенты Хансенна – Куйпера, прогнозная сила моделей практически одинакова.

Как показывают рис. 5 и 6, система предсказания EWS действительно улавливает периоды кластеризации отрицательных доходностей и может предсказывать обвалы индекса в пятидневном интервале.

3.2. Построение модели EWS на основе данных USD/RUB

Мы рассматриваем данные Московской биржи по валютной паре USD/RUB за период с 26 января 1999 г. по 10 апреля 2017 г. Для калибровки моделей используется весь временной ряд. Набор данных состоит из суточных значений $R = \frac{p_t - p_{t-1}}{p_{t-1}}$, где p_t обозначает цену индекса в точке t .

На рис. 7 показаны цена пары и дневная доходность по ней. Серьезные обвалы и отрицательные доходности можно заметить вблизи известных исторических периодов: дефолта 1998 г. и снижения цены на нефть 2014 г.

Значения доходности

Значения вероятности обвала

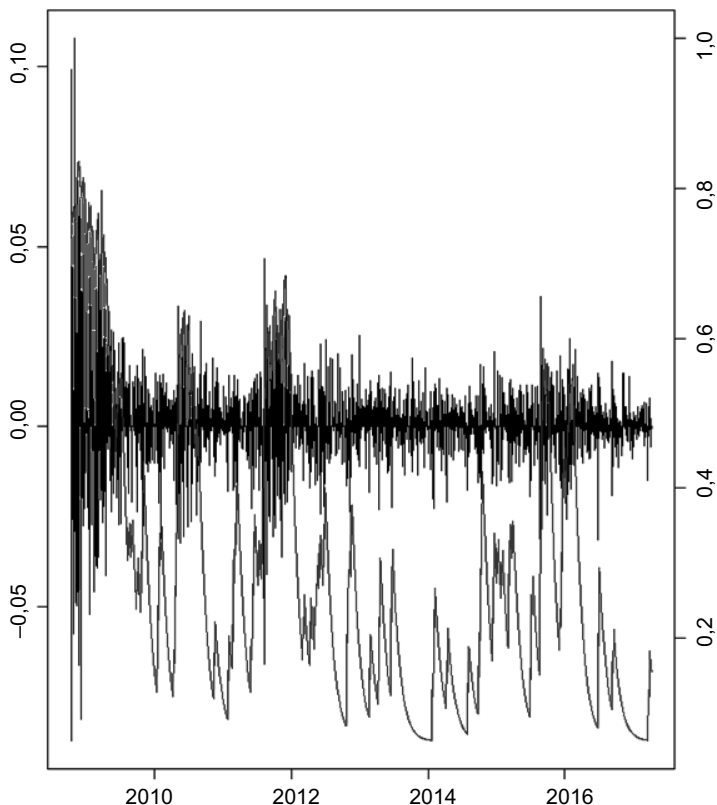


Рис. 6. Доходность индекса S&P500 (ломаная линия в центре) и вероятность появления обвала в течение следующих пяти дней, рассчитанная системой EWS на основе процесса Хоукса для 5%-й фильтрации событий (нижняя линия, пики означают появление события). Вертикальные линии отмечают даты, когда модель показывает вероятность обвала выше 0,5 – это начало периода, середина 2010 г. и вторая половина 2011 г., а также конец 2015 г. и начало 2016 г.

Более того, на рисунке видна кластеризация экстремальных доходностей, т.е. спокойные периоды с небольшими изменениями цен чередуются с «турбулентными» периодами с большими изменениями цен. Стадное поведение инвесторов и возможное предсказание сильных движений индекса будет рассмотрено далее.

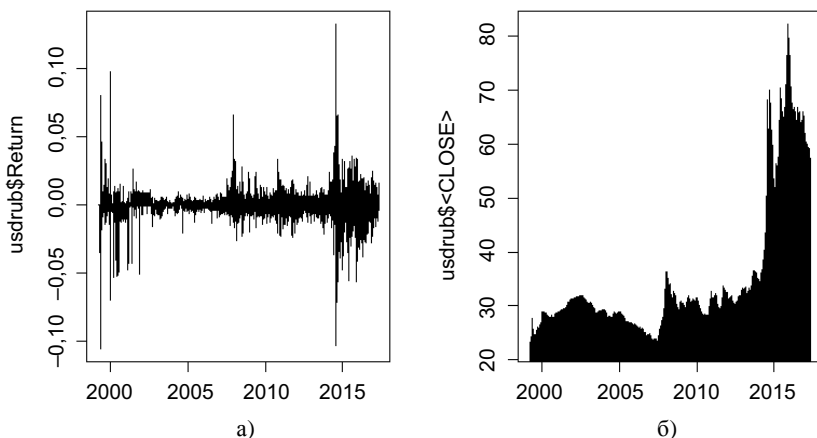


Рис. 7. а) Доходность валютной пары USD/RUB; б) Цена валютной пары USD/RUB

Мы применяем модели Хоукса и ETAS к 95% и 99% квантилям экстремальных отрицательных доходностей. Эти ограничения соответствуют 53 и 277 событиям за период времени 5543 торговых дней соответственно. Ограничение выборки – вынужденный шаг для того, чтобы кластеризацию отрицательных доходностей было возможно рассмотреть. Без фильтрации большое количество событий (около 30% всех торговых дней) приводит к тому, что процессы имеют высокую интенсивность внешних событий и низкую вероятность возникновения иммигрантских событий.

Коэффициент ветвления n в процессах Хоукса ожидаемо выше и находится в пределах 87–95%. Это означает, что около 10% событий произошло из-за собственного возбуждения. Интенсивность фоновых событий колеблется в пределах от 0,23 до 2,12. Уже на данном примере видно, что с увеличением выборки интенсивность фоновых событий растет очень быстро, коэффициент ветвления падает. Если не проводить фильтрацию вообще, то мы получим процесс с огромным μ , низким коэффициентом ветвления и кластеризацию рассмотреть не сможем.

Аналогично поведению коэффициентов моделей для S&P500 качество оценки параметров модели ETAS возрастает с увеличением количества событий.

Таблица 4. Оценки параметров моделей для различных фильтраций (для выборок из 1% и 5% экстремальных отрицательных доходностей по валютной паре USD/RUB)

Оценки параметров моделей	Процесс Хоукса с ядром $\phi(t - t_i) = \alpha e^{-\beta(t-t_i)}$		ETAS с ядром (1)	
	1%	5%	1%	5%
μ	0,23	2,12	0,27	1,16
α	5,13	7,16	0	0,69
β	5,39	8,18		
K_0			0,26	0,55
c			0,37	0,55
p			2,99	3,045
$\log L$	-384,77	-5854,85	20,92	196,125
AIC	775,54	11715,7	-31,85	-382,245

Как видно из рис. 8 и 9, модели, основанные на 1%-й выборке доходностей, имеют интенсивность 0 в промежутке между 2003 и 2014 г. Это свидетельствует о том, что модель не очень корректно отображает интенсивность на половине временного интервала. Более того, отсутствие интенсивности в этом периоде будет означать невозможность подсчета вероятности обвала в пятидневном интервале.

Из рис. 10 следует, что для 1%-й выборки доходностей все прогнозные вероятности обвалов оказались ниже 0,5. Это связано с малым количеством событий (55) на промежутке в 17 лет и отсутствием самовозбуждения процесса на протяжении с 2003 по 2014 г. Таким образом, интенсивность затухает настолько быстро, что потомки не успевают образоваться. Следовательно, мы не можем использовать кластеризацию событий для прогноза обвалов.

Рассмотрим результаты расчетов для фильтрации 5% экстремальных отрицательных доходностей.

Таким образом, рис. 11 демонстрирует прогнозы системы вместе с реальными доходностями. Вероятности обвалов начали расти вместе с увеличением волатильности на рынке в конце 2014 г. Практически все время до этого (кроме одного дня в 2001 г.) система не давала прогнозов событий.

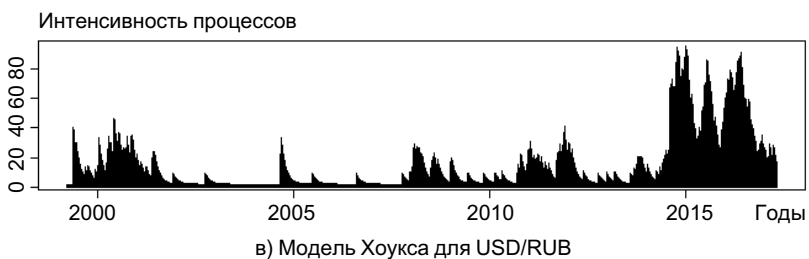
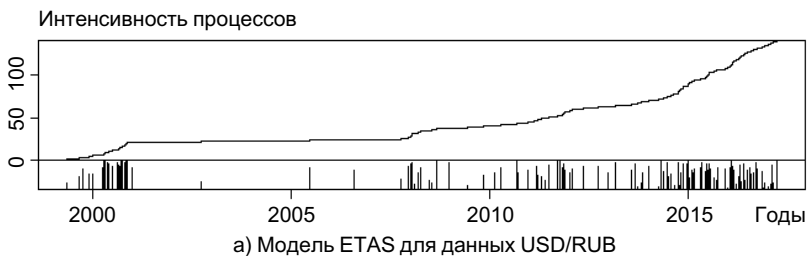


Рис. 8. Результаты расчетов, основанных на 5%-й выборке отрицательных доходностей валютной пары USD/RUB



Рис. 9. Результаты расчетов, основанных на 1%-й выборке отрицательных доходностей валютной пары USD/RUB

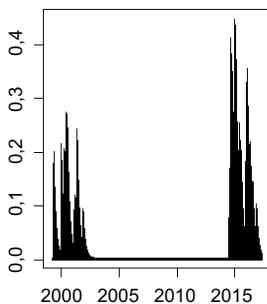


Рис. 10. Вероятности возникновения обвалов в течение следующих пяти дней для фильтрации 1% данных валютной пары USD/RUB

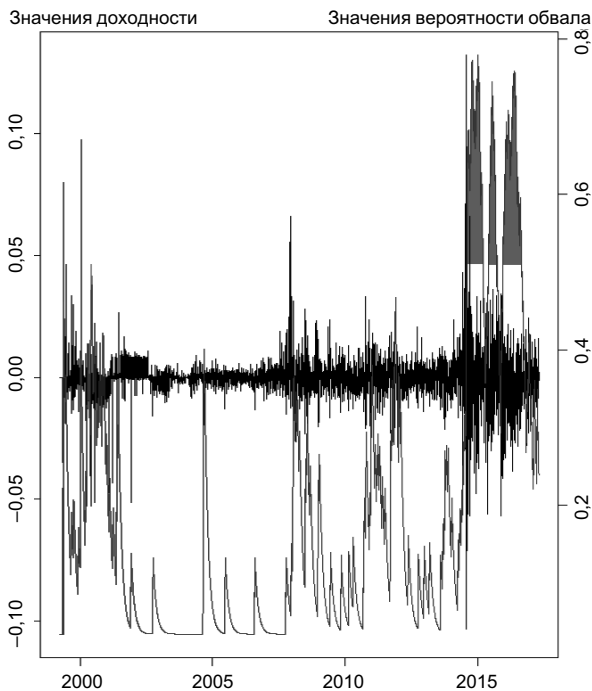


Рис. 11. Доходность валютной пары USD/RUB (ломаная линия в центре) и вероятность появления обвала в течение следующих пяти дней, рассчитанная системой EWS на основе процесса Хоукса для 5%-м фильтрации событий (нижняя линия, пики означают появление события).

Закрашенная область соответствует датам, когда модель показывает вероятность обвала выше 0,5 – это 2015 и 2016 гг.

Таблица 5. Попадания и промахи системы ETAS с ядром (1) для выборки 5% максимальных отрицательных доходностей валютной пары USD/RUB

Предсказание	Наблюдение		
	Событие произошло	Отсутствие события	Итого предсказанных
Событие произошло	112	400	512
Отсутствие события	165	4866	5031
Итого наблюдаемых	277	5266	

Ситуация с ошибками первого и второго рода полностью аналогична ситуации с индексом S&P500. Тем не менее скоринг KSS значительно отличается:

$$KSS_{ETAS5\%} = \frac{112 * 4866 - 400 * 165}{(112 + 165)(400 + 4866)} = 0,33.$$

Это произошло по причине того, что для валютной пары временной промежуток для оценки параметров был гораздо меньше, а прогноз осуществлялся для всего временного промежутка, а не его части.

3.3. Сравнение с другими моделями

В описанных выше моделях KSS имеет наибольшее значение для индекса S&P500 и меньшее для валютной пары USD/RUB. Исследователи ставили аналогичную задачу: создать модель раннего оповещения о крахах на рынке, используя модель GARCH [Gresnigt et al., 2014]. Поскольку тестирование модели происходило на данных S&P500, мы сравниваем результаты скоринга KSS для моделей ETAS с полученными в другом исследовании и делаем вывод о лучшей модели [Gresnigt et al., 2014].

В приведенном исследовании $KSS_{GARCH} \in (0,37; 0,46)$ в зависимости от фильтрации данных. В нашем исследовании для S&P500 $KSS_{ETAS} \in (0,42; 0,43)$.

Равенство показателя Хансена – Куйпера единице соответствует идеальной предсказательной силе модели, а у описанных выше моделей показатели разнятся в зависимости от фильтрации данных и не дотягивают до 0,5. Для разных фильтраций модели GARCH и ETAS попеременно показывают лучшие результаты. Более того, GARCH не полностью учитывает самовозбуждающуюся природу обвалов [Gresnigt et al., 2014].

Таким образом, в целом можно судить об эквивалентности моделей для данных S&P500.

Заключение

В данной работе рассматриваются модели прогнозирования потрясений на фондовых рынках, учитывающих кластеризацию доходностей во время обвалов рынка. Все модели основаны на самовозбуждающемся процессе Хоукса, интенсивность которого определяется произошедшими в прошлом событиями (экзогенными или эндогенными). Модели протестированы на данных фондового индекса S&P500 и валютной паре USD/RUB. Все модели имеют неплохую прогнозную силу и сравнимы по эффективности с широко распространенными эконометрическими моделями. Модели подобного вида могут быть использованы для оценки финансовых рисков и в качестве элемента системы поддержки принятия решений трейдером на бирже.

Литература

Aït-Sahalia Y., Cacho-Diaz J., Laeven R.J.A. Modeling financial contagion using mutually exciting jump processes // *Journal of Financial Economics*. 2015. Vol. 117 (3). P. 585–606.

Bacry E., Delattre S., Hoffmann M., Muzy J.F. Modeling microstructure noise with mutually exciting point processes. 2011. arXiv:1101.3422v1.

Balderama E. et al. Application of branching models in the study of invasive species // *Journal of the American Statistical Association*. 2012. Vol. 107 (498). P. 467–476.

Barber B., Odean T. Trading is hazardous to your wealth: The common stock investment performance of individual investors // *Journal of Finance*. 2000. Vol. 55 (2). P. 773–806.

Bauwens L., Hautsch N. Modelling financial high frequency data using point processes // *Handbook of Financial Time Series*. 2009. P. 953–979.

Boogaard H. van den Maximum Likelihood Estimations in a Nonlinear Self-Exciting Point Process Model // *Biol. Cybern.* 1986. No. 55. P. 219–225.

Devenow A., Welch I. Rational herding in financial economics // *European Economic Review*. 1996. No. 40.

Gresnigt F., Kole E., Franses P.H. Interpreting financial market crashes as earthquakes: A new early warning system for medium term crashes // *Journal of Banking & Finance*. 2015. No. 56. P. 123–139.

Hawkes A.G. Spectra of Some Self-Exciting and Mutually Exciting Point Processes // *Biometrika*. 1971 (Apr.) Vol. 58. No. 1. P. 83–90.

Hewlett P. Clustering of order arrivals, price impact and trade path optimisation // *Workshop on Financial Modeling with Jump processes, Ecole Polytechnique*, 2006. P. 6–8.

Kahneman D. *Thinking, fast and slow*. N.Y.: Penguin, 2011.

Krumin M., Shoham S. Generation of Spike Trains with Controlled Auto- and Cross-Correlation Functions // *Neural Computation*. 2009. No. 21 (6). P. 1642–1664.

Lando D. On Cox processes and credit risky securities // *Review of Derivatives Research*. 1998. No. 2 (2–3). P. 99–120.

Lee S., Wilson J.R., Crawford M.M. Modeling and Simulation of a Nonhomogeneous Poisson Process having Cyclical Behavior // *Communications in Statistics – Simulation and Computation*. No. 20. Vol. 2–3. P. 777–809.

Lewis P.A.W., Shedler G.S. Simulation of Nonhomogeneous Poisson processes by thinning // *Naval Research Logistics Quarterly*. 1979. No. 26 (3). P. 403–413.

Massey W.A., Parker G.A., Whitt W. Estimating the parameters of a non-homogeneous Poisson process with linear rate // *Telecommunication Systems*. 1996. No. 5. P. 361–388.

Mohler G.O. et al. Self-exciting point process modeling of crime // *Journal of the American Statistical Association*. 2011. Vol. 106 (493). P. 100–108.

Odean T. Do investors trade too much? // *American Economic Review*. 1999. Vol. 89. No. 5. P. 1279–1298.

Ogata Y. The asymptotic behavior of maximum likelihood estimators for stationary point processes // *Annals of the Institute of Statistical Mathematics*. 1978. No. 30 (1). P. 243–261.

Ogata Y. On Lewis' simulation method for point processes // *IEEE Transactions on Information Theory*. 1981. No. 27 (1). P. 23–31.

Ogata Y. Statistical models for earthquake occurrences and residual analysis for point processes // *Journal of the American Statistical Association*. 1988. No. 83 (401). P. 9–27.

Ozaki T. Maximum Likelihood Estimation of Hawkes' Self-Exciting Point Process // *Ann. Inst. Statist. Math.*, 31, Part B. 1979. P. 145–155.

Penikas H., Proskurin S. How Well Do Analysts Predict Stock Prices? Evidence from Russia // *Working papers by NRU Higher School of Economics*. Series FE "Financial Economics", WP BRP 18/FE/2013.

Schoenberg F.P. Introduction to point processes // *Wiley Encyclopedia of Operations Research and Management Science*. 2010.

Simon G. Hawkes processes in finance: a review with simulations. University of Oregon, 2016.

Sornette D. Critical market crashes // *Physics Reports*. 2003. Vol. 378 (1). P. 1–98.

Toke I.M. "Market making" in an order book model and its impact on the bid-ask spread // *Econophysics of Order-Driven Markets*. New Economic Windows / F. Abergel, B. Chakrabarti, A. Chakraborti, M. Mitra (eds). Springer-Verlag Milan, 2011. P. 49–64.

Tressa L. Fowler, Tara L. Jensen, Barbara G. Brown. Introduction to Forecast Verification / National Center for Atmospheric Research Thesis. 2012.

Weinberg J., Brown L.D., Stroud J.R. Bayesian Forecasting of an Inhomogeneous Poisson Process with Applications to Call Center Data // Journal of the American Statistical Association. 2007. No. 102. Vol. 480. P. 1185–1198.

Egorova Lyudmila, Klymyuk Igor.

Hawkes processes for forecasting financial risk [Text] : Working paper WP7/2017/02 / L. Egorova, I. Klymyuk ; National Research University Higher School of Economics. – Moscow : Higher School of Economics Publ. House, 2017. – (Series WP7 “Mathematical methods for decision making in economics, business and politics”). – 36 p. – 20 copies.

We consider models of forecasting shocks in stock markets that take into account the endogenous nature of such phenomena on the basis of Hawkes processes. The intensity of the Hawkes processes depends on previous events, that allow modeling the clustering effect. The models were tested on the S&P500 stock index and the USD/RUB currency pair and are capable to predict some financial crashes.

Препринт WP7/2017/02

Серия WP7

Математические методы анализа решений
в экономике, бизнесе и политике

Егорова Л.Г., Климюк И.Ю.

**Применение процессов Хоукса
для прогнозирования финансовых рисков**

Зав. редакцией оперативного выпуска *А.В. Заиченко*
Технический редактор *Ю.Н. Петрина*

Отпечатано в типографии
Национального исследовательского университета
«Высшая школа экономики» с представленного оригинал-макета
Формат 60×84 ¹/₁₆. Тираж 20 экз. Уч.-изд. л. 2,3
Усл. печ. л. 2,1. Заказ № . Изд. № 2059

Национальный исследовательский университет
«Высшая школа экономики»
125319, Москва, Кочновский проезд, 3
Типография Национального исследовательского университета
«Высшая школа экономики»