



Орлов Ю.Н., Федоров С. Л.

Моделирование и
статистический анализ
функционалов, заданных на
выборках из
нестационарного
временного ряда

Рекомендуемая форма библиографической ссылки: Орлов Ю.Н., Федоров С. Л. Моделирование и статистический анализ функционалов, заданных на выборках из нестационарного временного ряда // Препринты ИПМ им. М.В.Келдыша. 2014. № 43. 26 с. URL: <http://library.keldysh.ru/preprint.asp?id=2014-43>

**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М.В.Келдыша
Российской академии наук**

Ю.Н. Орлов, С.Л. Федоров

**Моделирование и статистический
анализ функционалов,
заданных на выборках
из нестационарного
временного ряда**

Москва — 2014

Орлов Ю.Н., Федоров С.Л.

Моделирование и статистический анализ функционалов, заданных на выборках из нестационарного временного ряда

Излагается метод генерации пучка траекторий нестационарного временного ряда, эволюция плотности распределения которого моделируется уравнением Фоккера-Планка. Плотность распределения моделируется гистограммой с оптимальным равномерным разбиением на классовые интервалы. На основе этого метода анализируются свойства выборочных статистик, которые представляют собой функционалы, заданные на выборках из нестационарного временного ряда с определенным правилом эволюции его выборочных распределений.

Ключевые слова: нестационарный временной ряд, уравнение Фоккера-Планка, анализ выборочных статистик

Orlov Yu.N., Fedorov S.L.

Modeling and statistical analysis of functional, defining on the sampling from non-stationary time series

The method of sampling analysis for non-stationary time-series with Fokker-Planck evolution equation for its density is considered. The optimal histogram class intervals are used for solving this equation. The properties of sampling statistics are examined.

Key words: non-stationary time series, Fokker-Planck equation, sampling statistical analysis

Работа выполнена при поддержке гранта РФФИ, проект

№ 13-01-00617

Содержание

Введение	3
1. Построение выборки из нестационарного временного ряда	5
2. Статистические свойства исходного ряда	7
3. Тестирование метода генерации пучка нестационарных траекторий	14
4. Генерация траекторий с помощью уравнения эволюции ВПФР	16
5. Статистический анализ функционалов, заданных на траекториях случайного процесса.....	20
Заключение.....	24
Литература	26

Введение

Настоящая работа развивает идеи и подходы, изложенные в статье [1], по генерации нестационарного временного ряда с заданными свойствами эволюции его выборочных распределений. Сама эволюция моделируется уравнением Лиувилля или Фоккера-Планка для выборочных плотностей, как это предложено в монографии [2].

Во многих задачах прикладного статистического анализа существует необходимость тестирования тех или иных индикаторов локального поведения временного ряда с целью оценки вероятности их правильного срабатывания. Как правило, индикаторы представляют собой функционалы от фрагментов траектории случайного процесса. Примеры индикаторов: отношение числа положительных приростов значений временного ряда к числу отрицательных за определенный период времени; угловой коэффициент прямой регрессии для выборки заданной длины; расстояние между выборочными плотностями функции распределения ряда в той или иной норме, дисперсия накопленного размаха за определенный период и т.п.

Чтобы оценить эмпирическую условную вероятность того, что определенный интервал значений индикатора отвечает ожидаемому исследователем поведению ряда в настоящем или будущем, нужно иметь много реализаций изучаемого процесса, тогда как в наличии имеется лишь одна фактически наблюдаемая траектория. На практике берется фрагмент траектории, который представляется достаточно большим, и на нем собирается требуемая статистика по индикатору: число ошибочных срабатываний, число ошибочных несрабатываний и число правильных срабатываний. Однако, если временной ряд нестационарный, то, например, оптимизация длины выборки для получения минимальной ошибки индикатора по некоторому фрагменту прошлой траектории не имеет особого смысла, поскольку оптимум искался для конкретного фрагмента траектории случайного процесса. При другой последовательности тех же самых значений временного ряда возможен иной результат оптимизации. Возникает задача тестирования индикатора на устойчивость относительно различных реализаций случайного процесса, имеющего близкие (с точки зрения исследователя – квазистационарные) выборочные распределения. Для этого требуется сгенерировать пучок возможных траекторий временного ряда, исходящий из заданного текущего состояния, и проверить на нем устойчивость срабатывания индикатора.

Если процесс стационарный, то набор его траекторий может быть получен, исходя из следующих соображений. Рассмотрим равномерно распределенную на конечном отрезке $[a; b]$ случайную величину ξ . Она имеет плотность функции распределения (далее ПФР)

$$f_U(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases} \quad (1)$$

Распределение (1) бывает удобно задать в виде, когда явно выделен центр μ промежутка длины $\omega = |b - a|$:

$$f_U(x; \mu, \omega) = \begin{cases} \frac{1}{\omega}, & \mu - \frac{\omega}{2} \leq x \leq \mu + \frac{\omega}{2} \\ 0, & x \notin [\mu - \omega/2, \mu + \omega/2] \end{cases} \quad (2)$$

Среднее значение этого распределения равно μ , а дисперсия равна $\omega^2/12$. Стандартизованным равномерным распределением является распределение на отрезке $[0; 1]$, ПФР которого в терминах (2) есть $f_U(x; 1/2, 1)$. Распределение (2) переводится в распределение $f_U(y; 1/2, 1)$ посредством линейного преобразования $y = \frac{1}{2} + \frac{x - \mu}{\omega}$.

Равномерное распределение часто используется для того, чтобы определить, является ли выборка $\{x_1, x_2, \dots, x_n\}$ выборкой из распределения с некоторой предполагаемой непрерывной функцией распределения $F(x)$. В основе различных статистических критериев согласия с гипотезой о виде распределения $F(x)$ лежит следующее утверждение (см., напр., [3]). Пусть ξ есть случайная величина с непрерывной функцией распределения (далее ФР) $F(x)$. Тогда случайная величина $\eta = F(\xi)$ имеет ПФР $f_U(y; 1/2, 1)$. Действительно, поскольку

$$P(F(x) \leq y) = \begin{cases} 1, & y > 1 \\ y, & 0 < y \leq 1 \\ 0, & y \leq 0 \end{cases}$$

то плотность этой функции распределения имеет вид $f_U(y; 1/2, 1)$.

На основе этого утверждения строятся алгоритмы генерации временного ряда $\{x_k\}$ с заданной функцией распределения $F(x)$. Для этого генерируется произвольная последовательность чисел $\{y_k\}$, равномерно распределенных на $[0; 1]$, после чего по формуле

$$y_k = F(x_k), \quad x_k = F^{-1}(y_k) \quad (3)$$

находятся элементы ряда $\{x_k\}$. Обращение функции распределения в (3) возможно в силу ее строгой монотонности.

Этот подход можно применить и в том случае, если генеральная совокупность $F(x)$ не известна, а оценивается по наблюдаемому временному ряду. Тогда условие строгой монотонности ФР для де-факто дискретного случайного процесса не выполняется. Тем не менее, удобно считать, что ПФР непрерывна внутри заданного классового интервала. Например, внутри ячейки классового интервала гистограммы можно постулировать равномерное

распределение эмпирической частоты. В этом случае обращение ФР в (3) осуществляется в соответствии с разбиением гистограммы для ПФР. Если постулируется вид распределения внутри классового интервала, то обращение является однозначной процедурой. Так, для равномерного разбиения области гистограммы $[0; 1]$ на n классовых интервалов по заданным вероятностям f_k , $k = 1 \div n$ строится ФР, которая имеет вид

$$F(x) = (nx - j)f_{j+1} + \sum_{k=1}^j f_k, \quad x \in \left[\frac{j}{n}; \frac{j+1}{n} \right], \quad j = 0 \div n - 1. \quad (4)$$

Что касается генерации равномерно распределенного ряда на $[0; 1]$, то для решения этой задачи существует достаточное количество алгоритмов (см. [4]). Для целей настоящей работы конкретный способ генерации не принципиален. Важно, что с его помощью можно построить и нестационарный временной ряд, если к текущему выборочному распределению подойти как к локальной генеральной совокупности, как это предложено в [1, 2]. Разумеется, в нашем подходе генерируется не сам ряд как бесконечная последовательность данных, а всего лишь конечное множество его элементов, т.е. выборка некоторой заданной длины.

1. Построение выборки из нестационарного временного ряда

Сложность анализа и, следовательно, моделирования нестационарного временного ряда с заданными статистическими свойствами его распределения состоит в том, что на практике в каждый момент времени наблюдается только одно значение из предполагаемой генеральной совокупности текущего распределения вероятностей. Поэтому с увеличением периода наблюдения выборочное распределение, построенное по увеличивающемуся количеству данных, не сходится к генеральной совокупности, из которой эти данные получены. Следовательно, анализу доступны только выборочные распределения, и задача анализа состоит не в установлении вида генеральной совокупности, что невозможно, а в нахождении эмпирических правил изменения его выборочной плотности функции распределения (далее ВПФР) на определенном промежутке времени. Эти правила и представляют собой те свойства, которые могут быть приписаны выборочному распределению и протестированы на различных реализациях случайного процесса с близкими распределениями.

Итак, следует построить численную модель нестационарного временного ряда, ВПФР которого имеет определенный тип – не обязательно конкретно параметрический для той или иной функции, а просто эмпирически наблюдаемый для изучаемой случайной величины, причем эволюционирует эта ВПФР в соответствии с правилами, которые следуют из вида модельного эволюционного уравнения.

Из практических соображений удобно считать, что наша случайная величина равномерно ограничена по времени, так что все ее значения

принадлежат отрезку $[0;1]$. Пусть $x(t)$ есть значение случайной величины в дискретный момент времени t , где шаг по времени считается единичным, а $f_T(x,t)$ есть ВПФР выборки длины T с окончанием в момент времени t , т.е. для построения ВПФР взят фрагмент ряда $\{x(t-T+1), \dots, x(t)\}$.

По ВПФР $f_T(x,t)$ можно построить непрерывную функцию распределения $F_T(x,t)$, считая распределение вероятностей в отдельном классовом интервале равномерным, равным частоте попадания в этот интервал наблюдаемых значений, деленной на длину интервала. Тогда соответствующая выборочная функция распределения (далее ВФР) определяется формулой

$$F_T(x,t) = \int_0^x f_T(y,t) dy. \quad (5)$$

Чтобы теперь имитировать процесс, близкий к реальным наблюдениям, например, за динамикой цен на бирже на какой-либо финансовый инструмент (цена акции, индекс набора ценных бумаг, курс валюты и т.п. – далее примеры будут относиться именно к этой области, хотя методика формально может быть применена к любым нестационарным временным рядам), предлагается следующая схема действий.

1. По имеющимся историческим данным строятся выборочные распределения приростов $x(t) = price(t+1) - price(t)$ цен на этот инструмент за тот промежуток времени T , который представляет интерес. Например, строятся распределения приростов цен закрытия в определенном временном интервале для какого-нибудь инструмента за два соседних месяца по скользящей выборке длиной в месяц. Шаг по времени предполагается единичным в соответствии с выбранным тайм-фреймом отсчетов значений ряда. Тем самым в каждый момент времени t определена ВПФР $f_T(x,t)$, допустим, 5-минутных приростов цен закрытия, и соответствующая ей $F_T(x,t)$ согласно формуле (5). В этом рассуждении длина T выражается в единицах 5-минутных интервалов. Имея в виду гистограммное представление ВПФР, фактически будем применять формулу вида (4).

2. Определяется эмпирическая скорость $u(x,t;T,\tau)$ изменения ВПФР по разности между $f_T(x,t)$ и $f_T(x,t+\tau)$ в соответствии с моделью эволюции. Подробнее этот аспект обсуждается в следующем разделе. При этом считаются известными и все промежуточные ВПФР $f_T(x,t+k)$, $k=1,2,\dots,\tau$. Они даны либо как фактически состоявшиеся ВПФР, либо как прогнозные в рамках той или иной модели эволюции ВПФР.

3. Генерируется стационарный равномерно распределенный на $[0;1]$ ряд чисел $\{y_k\}$ длиной τ , равной горизонту прогнозирования. Часто бывает актуальным исследовать изменение ВПФР при сдвиге ее по времени на горизонт, равный длине окна. В примерах часто будет рассматриваться вариант, когда $\tau=T$ (так называемые встык-выборки).

4. Строятся ВФР $F_T(x,t+k)$, $k=1,2,\dots,\tau$ по формуле (4).

5. Ключевой пункт программы – генерация собственно траектории временного ряда в соответствии с меняющейся ВПФР. Пусть t_0 есть начальный момент времени, в который начинает строиться прогноз. Тогда в последующие моменты времени одна из возможных траекторий случайного процесса, для которого ВПФР меняется от $f_T(x, t_0)$ до $f_T(x, t_0 + \tau)$, моделируется по формуле обращения соответствующей локальной по времени функции распределения, движущейся в скользящем окне длины T :

$$y_k = F_T(x_k, t_0 + k). \quad (6)$$

Подчеркнем, что, согласно (6), в каждый момент времени $t_k = t_0 + k$ генерируется только одно значение ряда из распределения $F_T(x, t_k)$. Сама же $F_T(x, t)$ выступает в этот момент времени как генеральная совокупность. Тем самым имитируется процесс наблюдения за динамикой нестационарного временного ряда. Этот метод впервые был описан в [1]. В настоящей работе проводится его тестирование на примере следующей выборки данных: это цены закрытия RTS по 5-минутным интервалам, общая длина массива 40 тыс.

Также дается пример использования генерации пучка нестационарных траекторий с характерным эмпирическим изменением выборочных распределений рассматриваемого ряда для изучения статистических свойств функционалов, определенных на выборочных траекториях. В качестве такого функционала приводится пример несложной торговой стратегии «движения вслед тренду» и делается анализ ее доходности с использованием пучка траекторий. Результаты сравниваются с оценкой доходности той же стратегии по имеющемуся историческому массиву данных.

2. Статистические свойства исходного ряда

Пред-обработка исходного ряда состоит в том, что приросты цен нормируются на отрезок $[0;1]$. Получившиеся в итоге числа и обозначаются далее как $x(t)$, где t есть последовательный номер элемента ряда, т.е. момент времени в единицах выбранного промежутка (5 мин.). Выборки из этого ряда сравниваются одна с другой не как динамические траектории, а как отвечающие этим выборкам плотности функций распределения.

Будем рассматривать две нормы: норму в пространстве непрерывных функций

$$\rho(F, F') = \|F - F'\| = \sup_x |F(x, t) - F'(x, t')| \quad (7)$$

и норму в пространстве суммируемых функций

$$r(f, f') = \|f - f'\| = \sum_{i=1}^n |f(i, t) - f'(i, t')|, \quad (8)$$

где сумма отвечает некоторому способу разбиения гистограммы для ВПФР. Заметим, что норма (7) не зависит от способа разбиения гистограммы.

Для моделирования ВПФР в настоящей работе применяется равномерное разбиение гистограммы с мелкостью разбиения, равной статистической неопределенности в определении получающихся при этом вероятностей. Такое согласованное разбиение было введено в [5]. Число классовых интервалов n в зависимости от длины выборки T и конкретной получающейся формы распределения определяется из неявной формулы

$$\frac{1}{n} = 2\varphi^{-1}[2z(T, n)] = \varepsilon, \quad (9)$$

где

$$\varphi(\varepsilon) = \frac{t_{1-\varepsilon}}{\varepsilon}, \quad z(T, n) = \frac{\sqrt{T}}{\sum_{i=1}^n \sqrt{f_T(i)(1-f_T(i))}} \quad (10)$$

и $t_{1-\varepsilon}$ есть квантиль распределения Стьюдента с $T-1$ степенью свободы.

Для рассматриваемого ряда зависимость числа классовых интервалов от длины выборки приведена на рис. 1.

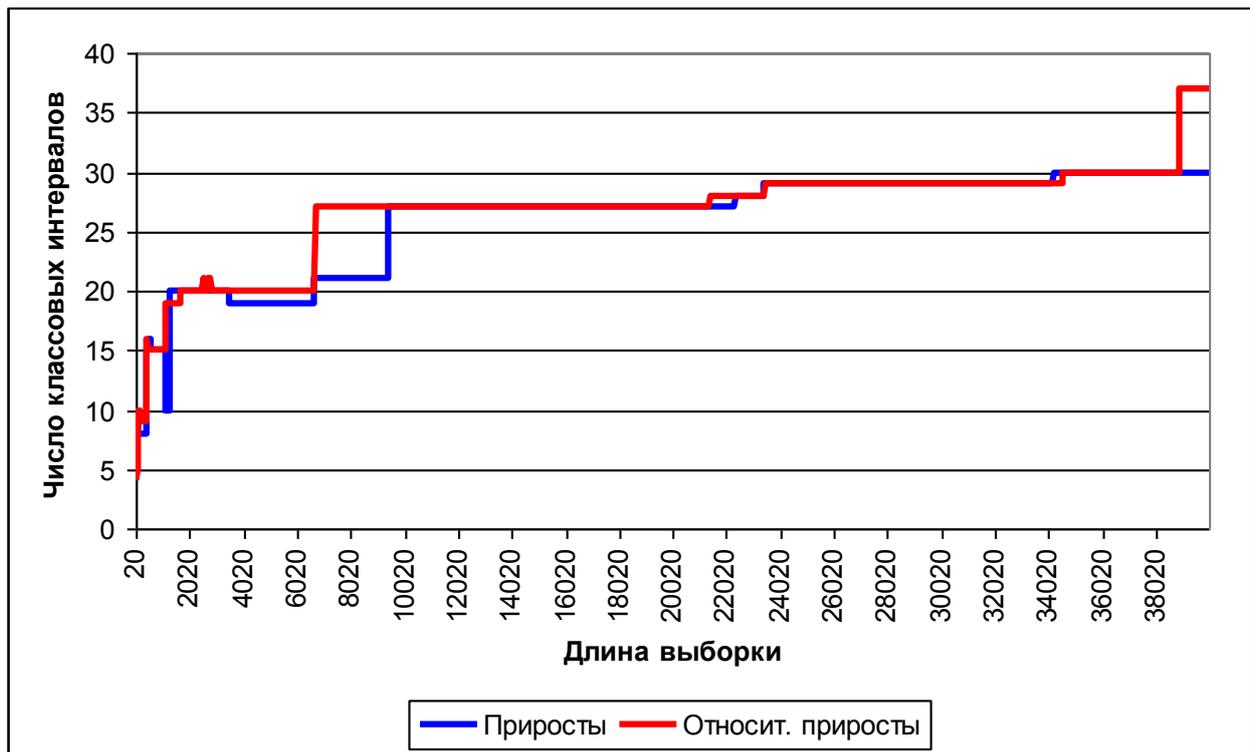


Рис. 1 – Число классовых интервалов в зависимости от длины выборки

Чем больше получается число интервалов в (9), тем, как правило, меньше разброс в случайных величинах. Наименьшее число интервалов требуется для представления равномерного распределения. В свою очередь, чем меньше разброс в случайных величинах, тем более четко можно выделить нестационарность в их последовательности.

В качестве изучаемого ряда рассмотрены как приросты цен, так и относительные приросты. В целом поведение обоих рядов в статистическом смысле близко, но есть определенные длины выборок, на которых ряд относительных приростов ведет себя заметно более нестационарно, чем ряд самих приростов. Это длины 1 тыс., 6-8 тыс. и 39-40 тыс. Если ряд приростов окажется недостаточно нестационарен, чтобы эволюцию его распределения можно было достоверно выделить на фоне чисто статистического шума, то следует перейти к ряду относительных приростов, который на определенных длинах выборки оказывается более чувствительным.

Если ряд стационарный, то мерой разброса случайной величины служит его дисперсия. Для нестационарного ряда она уже не является той статистикой, которая может иметь критериальные свойства. Для анализа нестационарных рядов в работе [6] был введен так называемый согласованный уровень стационарности (СУС), показывающий характерный уровень шума для квазистационарного участка траектории ряда.

Следуя [6], рассмотрим встык-выборки длины T . Для стационарного ряда случайных величин с независимыми приращениями это будут независимые выборки, выборочные функции распределения которых (ВФР) отличаются одна от другой на величину, оцениваемую в соответствии с критерием Колмогорова-Смирнова [3]. Отличия от этого критерия эмпирической статистики расстояний между ВФР покажут эффекты нестационарности или зависимости выборок.

Задача о принадлежности двух выборок одной генеральной совокупности решается непараметрической статистикой Колмогорова-Смирнова

$$S_T = \sup_x |F_{1,T}(x) - F_{2,T}(x)|, \quad (11)$$

для которой справедлива асимптотическая формула

$$\lim_{T \rightarrow \infty} P \left\{ 0 < \sqrt{\frac{T}{2}} S_T < z \right\} = K(z), \quad (12)$$

где $K(z)$ - табулированная функция Колмогорова [3].

Если для двух ВФР было найдено значение S_T и вычислена величина $z = \sqrt{\frac{T}{2}} S_T$, то величина $1 - K(z)$ приближенно считается равной вероятности того, что $\sqrt{\frac{T}{2}} S_T \geq z$. На практике критерий (12) применяется с учетом требуемого уровня значимости α , который считается известным. Если оказывается, что $1 - K(z) < \alpha$, то выборки следует считать различными на заданном уровне значимости. Некоторая неопределенность вывода состоит в том, что надо априори задать желаемый уровень малости критерия $1 - K(z)$. Какую вероятность следует считать достаточной для того, чтобы признать выборки одинаковыми?

Чтобы практически ответить на данный вопрос, поставим статистический эксперимент по генерации выборок из равномерного распределения. Построим

функцию распределения расстояний между независимыми выборками в смысле нормы (7). На рис. 2 показан временной ряд расстояний в норме C (7) между ВФР для двух выборок длины $T=1000$ данных. Распределение расстояний между этими выборками дано на рис. 3 по результатам 5000 экспериментов по определению расстояния между выборками.

Поскольку все выборки в этом примере взяты из одного распределения, то ошибкой на уровне значимости ε будет непризнание выборок «одинаковыми», если расстояние между ними оказалось больше ε . Смысл же уровня значимости таков, что в среднем в доле случаев, равной уровню значимости, будет сделан ошибочный вывод из статистического эксперимента. Но из рис. 2, в частности, следует, что при сравнении ВФР для выборок в 1000 данных бессмысленно задавать уровень значимости менее 0,01, так как тогда почти всегда придется признавать выборки различными, а на самом деле это не так.

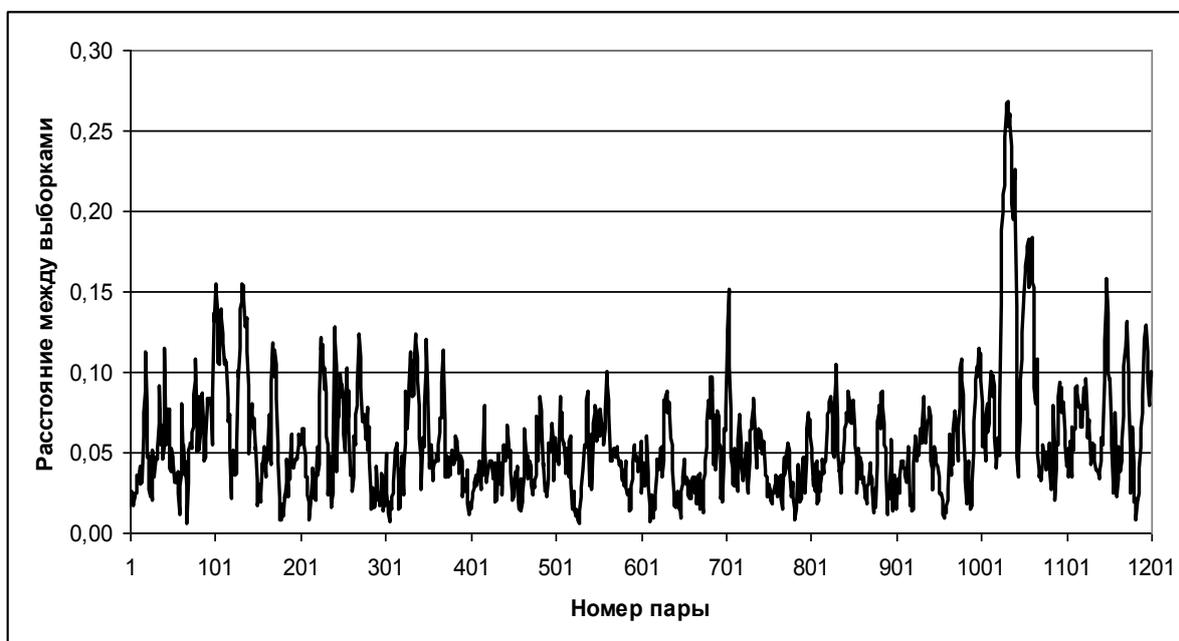


Рис. 2 – Фрагмент временного ряда расстояний между независимыми ВФР из равномерно распределенной совокупности длиной 1000 данных



Рис. 3 – Распределение расстояний между независимыми ВФР

Следовательно, при анализе выборок определенной длины было бы неправильно задавать априори желаемый уровень значимости, так как для заданной длины T выборки лишь при одном значении $\varepsilon = \varepsilon_0(T)$ вероятность превышения значения ε_0 равна значимости используемого для этой цели критерия, и это значение находится из уравнения

$$1 - \varepsilon_0 = K \left(\sqrt{\frac{T}{2}} \varepsilon_0 \right). \quad (13)$$

Решение уравнения (13) единственно, поскольку правая часть как функция ε_0 монотонно возрастает от нуля до единицы, а левая монотонно убывает от единицы до нуля. В частности, $\varepsilon_0(1000) \approx 0,06$. Это означает, что вероятность того, что расстояние между двумя независимыми выборками длиной 1000 данных из одного и того же стационарного распределения будет больше 0,06, равна 0,06. Найденное решение называется согласованным уровнем значимости или согласованным отклонением между ВФР. Это, так сказать, «типичное» отклонение между выборками из одного и того же стационарного распределения.

Для нестационарного ряда также можно построить расстояния (7) между непересекающимися выборками. Пусть $G_T(\rho)$ есть ВФР расстояний между встык-выборками длины T . Тогда согласованным уровнем стационарности называется величина $\rho^*(T)$, являющаяся решением уравнения

$$G_T(\rho) = 1 - \rho. \quad (14)$$

Если расстояние вычисляется в норме (8), то СУС определяется формулой

$$G_T(r) = 1 - r/2. \quad (15)$$

Уравнения (14)-(15) определяют функции $\rho^*(T)$ и $r^*(T)$, которые обладают тем свойством, что при проведении достаточно большого числа экспериментов по вычислению расстояний между двумя выборочными распределениями длины T , сдвинутыми на окно T , в долях ρ^* и r^* случаев будет наблюдаться превышение расстояния, равного ρ^* или r^* . Тем самым СУС можно трактовать как характерное расстояние между распределениями на уровне значимости, не превосходящем этого расстояния.

Если выбрана норма $L1$, и в ней оказалось, что $r^* \leq 2\varepsilon$, где ε определяется в (9) как уровень естественного шума в оценке ВПФР по гистограмме, то на достижимом уровне значимости ряд следует считать стационарным. Если же $r^* > 2\varepsilon$, то на выборках этих длин ряд имеет нестационарное поведение.

Для нормы C уровнем шума является СУС $\varepsilon_0(T)$; если $\rho(T) \leq \varepsilon_0(T)$, то ряд стационарный, а если $\rho(T) > \varepsilon_0(T)$, то нестационарный.

Наилучшая длина выборки, на которой нестационарность проявляется наиболее значимо, определяется максимумом индекса нестационарности $J(T)$ ряда, который определяется как отношение доли расстояний, не превосходящих фактический СУС, построенный по имеющимся эмпирическим данным, к доле расстояний, не превосходящих уровень шума. В норме $L1$ имеем

$$J(T) = \frac{G_T(r^*)}{G_T(2\varepsilon)} = \frac{1 - r^*(T)/2}{G_T(2\varepsilon)}, \quad (16)$$

а в норме C

$$J(T) = \frac{G_T(\rho^*)}{G_T(\varepsilon_0(T))} = \frac{1 - \rho^*(T)}{G_T(\varepsilon_0(T))}, \quad (17)$$

Если $J(T) \leq 1$, ряд считается стационарным, а если $J(T) > 1$, то ряд нестационарный.

Зависимость СУС (16) от длины выборки для ряда приростов цен приведена на рис. 4 в сравнении с удвоенной точностью определения эмпирических вероятностей, а индекс нестационарности – на рис. 5.

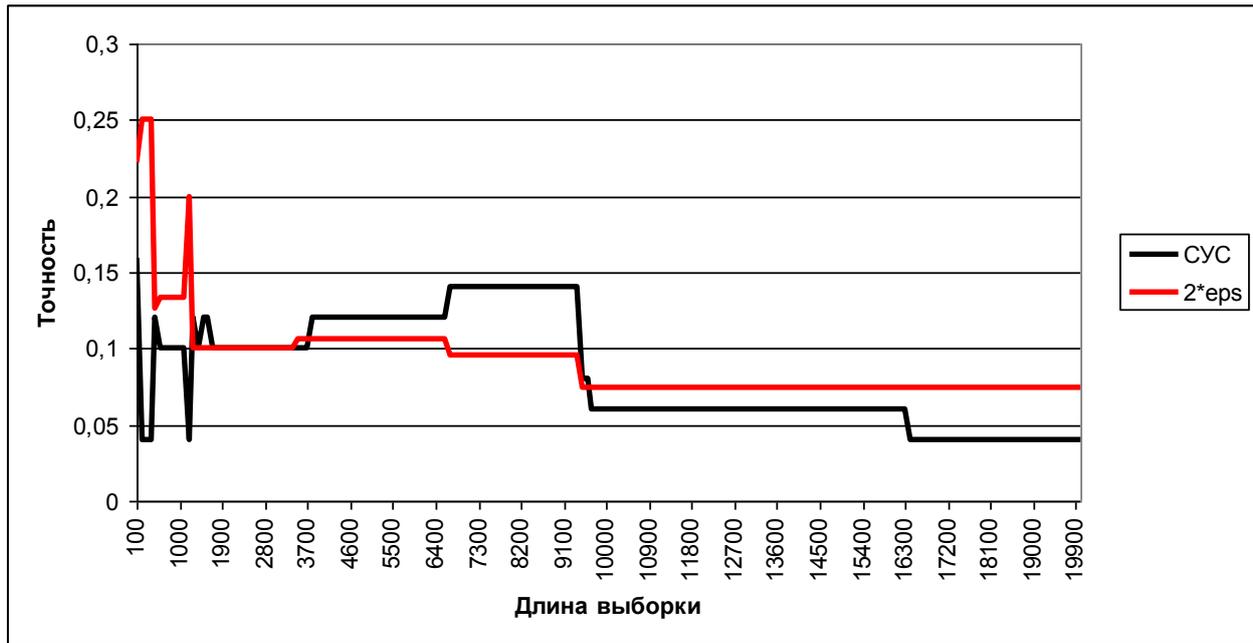


Рис. 4 – СУС в сравнении с точностью оценки вероятностей



Рис. 5 – Индекс нестационарности исходного ряда приростов цен

Из графика рис. 5 следует, что на длинах выборки меньше 7 тыс. и больше 9 тыс. нестационарные эффекты не могут быть выявлены, так как они находятся ниже уровня стационарного шума в оценке вероятностей. Этот анализ уточняет качественные выводы, сделанные выше на основании одного лишь разбиения гистограммы. Далее анализируются приросты цен в норме (8).

3. Тестирование метода генерации пучка нестационарных траекторий

Задавая различные равномерно распределенные ряды $\{y_k\}_j$, где $j=1, \dots, N$ есть номер генерации, можно получить пучок траекторий, ассоциированных с ВПФР текущих выборок.

Пусть в каждый момент времени $t \in [t_0 + 1; t_0 + T]$ по выборке длины T , отсчитанной назад от момента времени t , построены ВПФР $f_T(x, t)$. и соответствующие ВФР $F_T(x, t)$. Если в каждый момент $t > t_0$ сгенерировать N значений случайной величины, распределенной с ВФР $F_T(x, t)$, то получим N траекторий нестационарной случайной величины, т.е. N реализаций выборки из временного ряда на промежутке $[t_0 + 1; t_0 + T]$. Начальная точка для всех этих реализаций общая – это $x(t_0)$.

Каждая из траекторий пучка порождает на отрезке $[t_0 + 1; t_0 + T]$ ВПФР $\tilde{f}_T(\{y\}_j; x, t_0 + T)$, отличную, вообще говоря, от наблюдаемой $f_T(x, t_0 + T)$. По совокупности сгенерированных траекторий можно оценить, насколько значимо отклонение этих распределений в норме L1 одно от другого и от фактической ВПФР $f_T(x, t_0 + T)$. Предположительно все эти выборочные траектории являются реализациями некоторого неизвестного нестационарного распределения вероятностей, однако строгого доказательства этого предположения быть не может, ибо ни с увеличением окна выборки T , ни с увеличением числа траекторий N выборочные распределения не должны сходиться к этому распределению. Причем несмотря на то, что в каждый момент времени генерируется случайное число из одного и того же распределения, в силу нестационарности этого распределения итоговые сгенерированные выборки в окне $[t_0 + 1; t_0 + T]$ не обязаны принадлежать одному и тому же распределению. Если окажется, что выборки принадлежат разным распределениям, то такой способ генерации следует признать неадекватным.

Поскольку возможна лишь практическая проверка применимости описанного метода генерации нестационарной выборки с выборочным распределением, меняющимся определенным (т.е. фактически наблюдаемым) образом, то необходимо удостовериться в том, что: а) СУС попарных расстояний между сгенерированными выборками примерно равен $2/n$, где n есть число классовых интервалов разбиения гистограммы; б) СУС попарных расстояний между сгенерированной и фактической ВПФР $f_T(x, t_0 + T)$ также примерно равен $2/n$; в) СУС расстояний между сгенерированной и фактической исходной ВПФР $f_T(x, t_0)$ примерно равен расстоянию между $f_T(x, t_0 + T)$ и $f_T(x, t_0)$.

Если какое-либо из трех перечисленных условий не выполнено, то данный временной ряд не может быть смоделирован этим методом.

На рис. 6 приведен пучок траекторий (красных), сгенерированных описанным методом для ряда приростов цен закрытия 5-минутных интервалов

для индекса RTS. Зеленой линией выделена фактически состоявшаяся траектория. В начальный момент времени t_0 из точки $x(t_0)$ выпускается пучок траекторий в соответствии с распределениями, которые известны в момент t_k по выборке длины $T = 8000$, как описано выше.

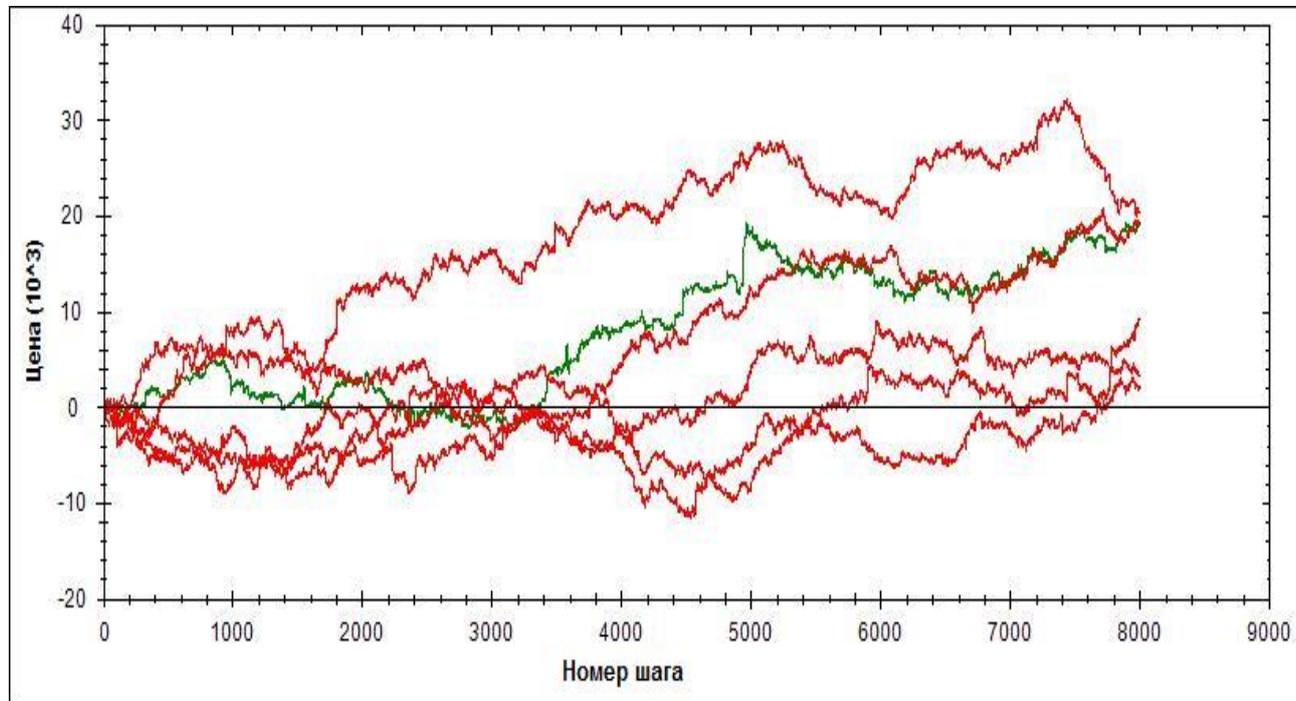


Рис. 6 – Пучок траекторий нестационарного процесса в окне $T = 8000$

В рассмотренном примере оказалось, что согласованное расстояние между сгенерированными распределениями по выборкам длины T на промежутке $[t_0 + 1, t_0 + T]$, равно $r^* = 0,09$, что близко к величине 2ε (рис. 4). Такой же величине равно и согласованное расстояние между сгенерированными выборками и фактически наблюдаемой ВПФР $f_T(x, t_0 + T)$. Расстояние между фактическими распределениями, разделенными временным промежутком длины T , в данном случае равно $r_{fact} = 0,15$. Этой же величине равен и СУС между исходной и сгенерированной выборками. Заметим, что СУС встык-выборок временного ряда приращений индекса RTS равен на этой длине величине 0,14.

Таким образом, можно считать, что эксперимент генерации нестационарного ряда в заданном классе распределений проведен корректно: сгенерированные траектории имеют уровень стационарности, значительно меньший, чем расстояние между встык-выборками, и в то же время больший, чем для стационарного процесса. Это служит указанием на то, что отклонения между фактическими распределениями, разделенными временным промежутком T , скорее всего, вызваны именно нестационарностью, а не статистическими флуктуациями. Кроме того, описанный метод генерации траектории действительно дает нестационарную выборку, поскольку конечное состояние отличается от начального на величину, заметно большую, чем 2ε .

4. Генерация траекторий с помощью уравнения эволюции ВПФР

При моделировании временных рядов следует помнить, что «затравочный» равномерно распределенный ряд $\{y_k\}$ порождает стационарный ряд величин x_k с независимыми приращениями согласно (3). Поэтому такой способ моделирования некоторых процессов, для которых свойственно порождать траекторию с заметно ненулевой автокорреляцией, как это часто имеет место на финансовых рынках, не может считаться адекватным. Желательно, чтобы корреляционные свойства ряда сохранялись и на сгенерированных выборках. При этом корреляция может быть связана не собственно с зависимостью случайных величин, а с нестационарностью распределения. Тогда для моделирования связей между элементами отдельной траектории можно использовать эмпирический закон эволюции выборочных распределений.

Свойства сгенерированного пучка траекторий можно более детально изучить в рамках определенной модели эволюции ВПФР. Следуя методике, описанной выше в разделе 1, можно построить численный алгоритм моделирования нестационарного временного ряда с определенными непараметрическими свойствами его ВПФР. Эти свойства выражаются в том, что задается некоторое эмпирическое кинетическое уравнение – Лиувилля, Фоккера-Планка или иное модельное уравнение, по которому предположительно эволюционирует ВПФР рассматриваемого ряда.

Пусть, в частности, эволюция определяется уравнением Лиувилля

$$\frac{\partial f}{\partial t} + \frac{\partial(uf)}{\partial x} = 0. \quad (18)$$

Выбираем равномерное разбиение гистограммы на классовые интервалы и считаем шаг по времени единичным. Тогда, опуская далее у ВПФР указание на длину выборки и обозначая через $f(i,t)$ значение ВПФР в i -ом классовом интервале, получаем для эмпирической скорости выражение [7]

$$u(i+1,t) = \frac{u(i,t)f(i,t) + f(i,t) - f(i,t+1)}{f(i+1,t)}. \quad (19)$$

Обозначим

$$F(i,t) = \sum_{k=1}^i f(k,t). \quad (20)$$

Тогда

$$u(1,t) = 0; \quad u(i+1,t) = \frac{F(i,t) - F(i,t+1)}{f(i+1,t)}. \quad (21)$$

В ячейках, где в знаменателе формулы (21) $f(i,t) = 0$, полагаем скорость равной нулю.

Единичным шагом естественно считать шаг по времени, с которым наблюдаются значения временного ряда. Зафиксировав длину выборки T в терминах этих единичных шагов, можно построить изменение ВПФР при

смещении по времени на длину окна выборки. Из формулы (21) следует, что скорость $U(i+1, t)$ изменения ВПФР при смещении на T шагов равна умноженной на T средневзвешенной скорости смещения на 1 шаг:

$$U(i+1, t) = \frac{F(i, t) - F(i, t+T)}{f(i+1, t)} = \frac{1}{f(i+1, t)} \sum_{k=0}^{T-1} u(i+1, t+k) f(i+1, t+k). \quad (22)$$

Формула (22) показывает, как можно построить траекторию временного ряда с заданным профилем скорости изменения ВПФР в классе ε -стационарных распределений, т.е. таких, для которых при сдвиге на ширину окна выборки выполняется условие

$$\sum_{i=1}^n |f(i, t+T) - f(i, t)| \leq \varepsilon, \quad (23)$$

где n есть число классовых интервалов гистограммы, а ε представляет характерное расстояние среди наблюдаемых попарных расстояний между распределениями по непересекающимся выборкам длины T , т.е. СУС встык-выборок. Для выполнения условия (23) на уровне значимости ε можно положить скорость изменения ВПФР на шаге длины 1 равной

$$u(i+1, t+k) = \frac{U(i+1, t)}{(1+\varepsilon)T}, \quad k=0, 1, \dots, T-1. \quad (24)$$

Тем самым на всем интервале времени длиной T можно смоделировать промежуточные распределения, являющиеся решениями модельного уравнения Лиувилля. Более того, если есть модель изменения эмпирической скорости, то можно построить и пучок прогнозных траекторий будущего случайного нестационарного процесса.

Однако модель генерации временного ряда через ПФР в соответствии с уравнением Лиувилля имеет существенный недостаток, который состоит в том, что если исходный ряд стационарный, но коррелированный, то стационарность распределения на пучке траекторий воспроизводится, а корреляция – нет. Это следует из того, что, как показано в [7], изменение выборочной дисперсии в силу уравнения Лиувилля определяется формулой

$$\frac{d\sigma^2}{dt} = 2\text{cov}_{x,u}(t), \quad \sigma^2(t) = \int (x - \bar{x}(t))^2 f(x, t) dx. \quad (25)$$

Отсюда вытекает, что в стационарном случае, когда $d\sigma^2/dt=0$, ковариация координаты и скорости (т.е. ковариация соседних значений ряда по выборке заданной длины) также равна нулю, что исключает возможность моделировать коррелированные ряды. Равенство нулю корреляции по выборке понимается в смысле подтверждения соответствующей статистической гипотезы на уровне значимости не меньшем, чем уровень стационарности ВПФР.

Модель, в которой правильно учитывается корреляция, использует уравнение Фоккера-Планка с определенным образом подобранным коэффициентом диффузии. В работе [8] показано, что если ВПФР удовлетворяет уравнению

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial x}(uf) - \frac{D(t)}{2} \frac{\partial^2 f}{\partial x^2} = 0, \quad (26)$$

то эволюция выборочной дисперсии дается уравнением

$$\frac{d\sigma^2}{dt} = D + 2\text{cov}_{x,u}(t). \quad (27)$$

В [8] также доказано утверждение о том, что оценка величины D по выборке в соответствии с дискретным аналогом дисперсии имеет вид

$$\begin{aligned} D(t) &= \sigma^2(t+1) - \sigma^2(t) - 2\text{cov}_{x,u}(t) = \\ &= \frac{1}{T} \sum_{k=t-T+1}^t (x(k) - x(k+1))^2 - \frac{1}{T^2} (x(t+1) - x(t-T+1))^2 \end{aligned} \quad (28)$$

и при $T > 1$ последнее выражение строго положительно. Тогда уравнение Фоккера-Планка может быть корректно использовано для генерации случайной траектории нестационарного процесса с сохранением присущих ему автокорреляционных свойств (в данном случае с единичным лагом).

Таким образом, если использовать разностную схему решения уравнения (26), в которой снос $u(x,t)$ и диффузия $D(t)$ определены непосредственно по временному ряду (в этом и состоит смысл понятия эмпирического уравнения эволюции), например, взятые с предыдущего слоя по времени, ПФР $f(x,t)$ можно будет использовать для генерации временного ряда с присущей ему автокорреляцией с лагом $l=1$. Для сохранения автокорреляций с большим лагом следует использовать уравнение Фоккера-Планка большей размерности, когда аргументы ПФР – не только координаты (значения случайной величины), но и скорости (приращения случайной величины), ускорения и т.д.

В настоящей работе мы ограничимся рассмотрением уравнения (26). Его решение на следующем временном слое имеет вид

$$\begin{aligned} f(x,t+1) &= f(x,t)(1+u(x,t-1)) - f(x+1,t)u(x+1,t-1) + \\ &+ \frac{D(t-1)}{2} (f(x+2,t) - 2f(x+1,t) + f(x,t)). \end{aligned} \quad (29)$$

Здесь, однако, следует учесть, что явные схемы типа (29) при решении уравнений диффузионного типа неустойчивы, в связи с чем они имеют малый горизонт прогнозирования. Для повышения устойчивости при прогнозировании на длину окна выборки мы далее использовали схему, в которой, во-первых, каждый классовый интервал разбит еще на 10 ячеек (с учетом равномерности в них, по построению, выборочной плотности), и, во-вторых, аппроксимация второй производной делается в лево-разностном шаблоне, в котором значение функции в ячейке x берется со следующего шага по времени. Описанная процедура приводит к разностному уравнению

$$\begin{aligned} f(x,t+1) &= f(x,t)(1+u(x,t-1)) - f(x+1,t)u(x+1,t-1) + \\ &+ \frac{D(t-1)}{2} (2f(x-1,t) - f(x,t+1) - f(x-2,t)). \end{aligned} \quad (30)$$

Разрешая его относительно $f(x, t + 1)$, получаем схему расчета:

$$f(x, t + 1) \cdot (1 + D/2) = f(x, t)(1 + u(x, t - 1)) - f(x + 1, t)u(x + 1, t - 1) + \frac{D}{2}(2f(x - 1, t) - f(x - 2, t)). \quad (31)$$

Временной ряд, который получается последовательным применением формулы (31) на заданном горизонте τ по алгоритму, описанному в п.1 (формула (6)), обладает как типичной для данного инструмента эволюцией его ВПФР, так и автокорреляцией первого порядка. Это – новое важное следствие из модели эволюции ВПФР, использующей уравнение Фоккера-Планка.

На рис. 7 показан пучок траекторий в том же окне, что и в примере выше (рис. 6), но только функции распределения для генерации случайных значений ряда взяты не с реальной траектории, а из решения уравнения (31).

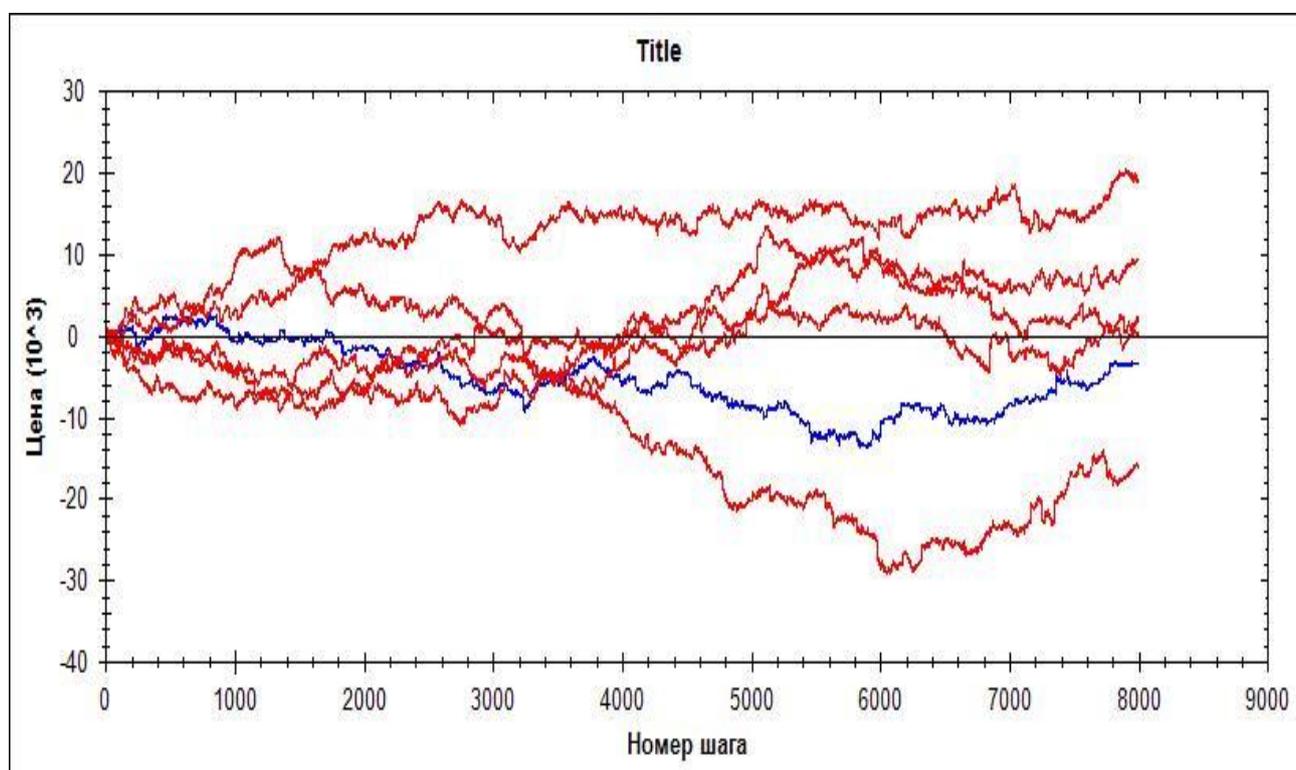


Рис. 7 – Пучок траекторий в окне $T = 8000$ для уравнения Фоккера-Планка

Отметим прежде всего, что расстояние между последними двумя встык-выборками, трактуемыми как «опорная история» и «тестируемый факт», составило 0,16. Среднее же расстояние между прогнозной ВПФР, полученной как решение уравнения Фоккера-Планка (31), и «историей» равно 0,17, т.е. отклонение модельных траекторий от «истории» примерно равно расстоянию между историей и «тестируемым фактом».

Далее, решение уравнения (31) находится на расстоянии 0,11 от распределения в прогнозируемом окне, т.е. от «тестируемого факта», что примерно равно величине 2ε . Это означает, что расстояние между модельной и фактической ВПФР получилось близким к уровню статистической

неопределенности самих распределений, как и должно быть для двух выборок из одного и того же распределения. Важно, что это расстояние заметно меньше, чем указанное выше расстояние 0,16 между «опорной историей» и «тестируемым фактом», ибо именно этому расстоянию было бы равно расстояние между прогнозом и фактом, если бы прогноз строился в предположении стационарности исходного распределения.

Таким образом, использование уравнения Фоккера-Планка позволило, во-первых, сгенерировать ряд с теми статистическими свойствами его выборочного распределения, которые для него характерны, и, во-вторых, спрогнозировать выборочное распределение в среднем несколько точнее, чем наивный прогноз, считающий распределение стационарным.

Отметим также, что согласованное расстояние между сгенерированными траекториями по выборкам длины T на промежутке $[t_0 + 1, t_0 + T]$, оказалось равным $r^* = 0,08$, что примерно равно величине 2ε , как и должно быть для выборок из одного распределения.

Тем самым численный эксперимент показал, что кинетическая модель эволюции распределения действительно отвечает определенным тенденциям изменения ВПФР, а не просто дает некоторое случайное ее изменение.

Следовательно, описанный метод может быть использован для тестирования различных функционалов, заданных на траектории случайного процесса, причем не на фрагменте предыстории, как это обычно делается, а на определенном горизонте в будущем. Пример такого тестирования приведен в следующем разделе работы.

5. Статистический анализ функционалов, заданных на траекториях случайного процесса

Рассмотрим возможность использования решения уравнения Фоккера-Планка для изучения статистик, заданных на траектории случайного процесса. Пусть на выборке длины T задан некоторый функционал $\Psi\{x(t - T + 1), \dots, x(t)\}$. Требуется определить, во-первых, его статистические свойства на выборках, отвечающих данной модели эволюции ВПФР, и, во-вторых, устойчивость функционала при изменении параметров уравнения эволюции или при разладке динамики ВПФР.

Первая задача решается следующим образом. Пусть выбран интересующий нас фрагмент временного ряда и на нем построен пучок виртуальных траекторий числом N . Обозначим Ψ_j значение функционала на j -ой траектории. Его статистические свойства полностью определяются распределением $\Phi(\Psi)$, которое строится по имеющимся N значениям. В частности, можно определить среднее, дисперсию, нормированное среднее (коэффициент Шарпа)

$$\bar{\Psi} = \frac{1}{N} \sum_{j=1}^N \Psi_j, \quad \sigma_{\Psi}^2 = \frac{1}{N} \sum_{j=1}^N (\Psi_j - \bar{\Psi})^2, \quad S_{\Psi} = \frac{\bar{\Psi}}{\sigma_{\Psi}}, \quad (32)$$

а также моду, медиану, меньшие квантили и тому подобные величины.

Решение второй задачи требует предварительной кластеризации фрагментов временного ряда по типичным состояниям и определения характерных изменений скорости (коэффициента сноса) с тем, чтобы можно было ввести разумный коэффициент модуляции в формуле (26). В частности, изменяя скорость в крайних значениях, можно проверить устойчивость функционала на изменение волатильности. Влияние корреляции между элементами ряда на значение функционала изучается варьированием коэффициента диффузии.

Например, если требуется смоделировать изменение вероятности срабатывания индикатора в определенной области значений случайной величины, то для этого можно провести «модуляцию» скорости в (26) с помощью произвольной подходящей функции $\lambda(i, t)$, т.е. рассматривать вместо $u(i, t + k)$ величину $v(i, t + k) = u(i, t + k)\lambda(i, t + k)$. Величина $\lambda(i, t)$ только должна быть такой, чтобы на следующем шаге по времени расчетная вероятность, получаемая по формуле (31), была бы неотрицательной.

Применительно к биржевым рядам можно выделить три основных кластера, к которым относят ВПФР соответствующих фрагментов выборок: тренд вверх, тренд вниз, боковое движение цены. Если требуется протестировать поведение функционала при смене кластерной принадлежности фрагмента траектории ряда, то берутся два эталонных представителя ВПФР каждого кластера и генерируются соответствующие два фрагмента ряда, которые затем соединяются встык, и через которых в скользящем окне заданной ширины T пропускается изучаемый функционал.

К задаче смешанного типа относится задача оптимизации функционала, заданного на траектории. Предположим, что определено окно, в котором задан функционал (в нашем примере это окно длиной 8000), и имеется определенный исторический отрезок траектории, на котором этот функционал тестируется. Тестирование предполагает определение оптимального значения некоторого параметра функционала, чтобы его значение на тестируемом отрезке было максимальным. Это – типичная задача, возникающая при разработке торговых стратегий на бирже.

Покажем на простом примере, что тестирование функционала, называемого «доходностью торговой стратегии», на случайном пучке траекторий, типичных для выбранного инструмента, гораздо более отвечает реальности, чем тестирование пусть и на достаточно длинной, но единственной исторической реализации.

Рассмотрим стратегию «следования за трендом» для торговли внутри одного торгового дня. Соответствующий отрезок 5-минутных данных составляет окно в 144 точки. Стратегия состоит в следующем. Имеющаяся историческая траектория делится на фрагменты длиной 144 точек. Внутри

каждого фрагмента по выборке в скользящем окне 20 точек (фрагменты и окна могут быть и других длин, для нашего примера это не принципиально) определяется наличие (или отсутствие) тренда. Если текущее состояние определено как трендовое, то в этот момент делается вход в соответствии с направлением тренда. Выход из позиции осуществляется по достижении заданного уровня прибыли – например, удвоенной комиссии, либо по стоп-сигналу минимизации убытка, либо по окончании фрагмента. Параметром стратегии является уровень стоп-сигнала.

Для определенности трендом вверх в момент времени t будем называть такое состояние, в котором разность между ценами $p(t) - p(t-199)$ лежит в последнем квартиле распределения таких отклонений на всем промежутке времени. Аналогично определим тренд вниз. Уровень тренда не считается здесь оптимизируемым параметром. Наша цель имеет иллюстративный, а не прикладной характер в части оптимизации параметров стратегии, так что задача формулируется в максимально упрощенном виде. В частности, не проводится анализ наилучшей длины выборки и наилучшего квантиля распределения приростов цен за этот период времени, т.е. таких значений, при которых условная вероятность продолжения того же направления движения цены, что и знак тренда, в зависимости от горизонта планирования максимальна.

При заданном уровне комиссии $q=0,002$ на промежутке 32 тыс. данных оказалось, что оптимальным является уровень стоп-сигнала, равный 0,09. Соответствующий график движения доходности как функции времени приведен на рис. 8. Как видно, такая тривиальная стратегия не представляет особого интереса с практической точки зрения, но, как говорилось выше, на этом примере мы демонстрируем применение методики тестирования функционала доходности на случайной траектории.

Доля положительных сделок в области оптимизации параметра составила ровно половину 0,5; средняя доходность на положительную сделку равна 0,0048, средний убыток на отрицательную сделку равен -0,0043.

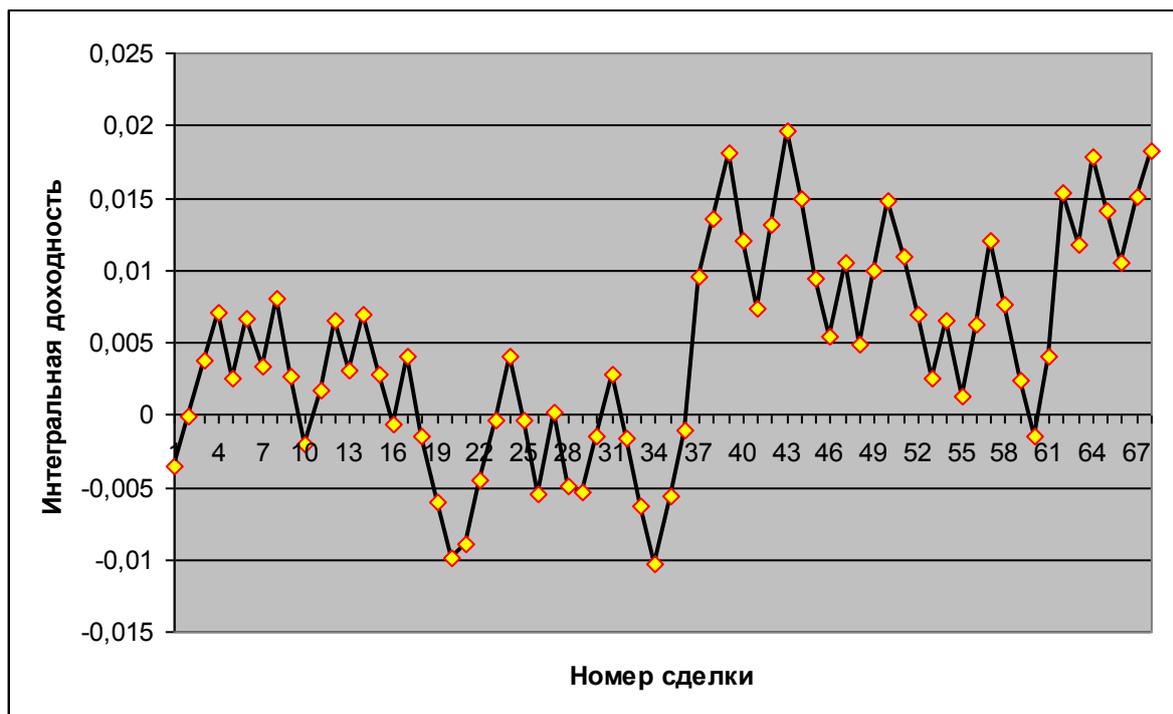


Рис. 8 – График движения доходности торговой системы в области оптимизации параметров

Рассмотрим, какую доходность можно ожидать за период $T = 8000$. Для этого выберем произвольный фрагмент ряда длиной 8000 из области, где проводилась оптимизация, протестируем на нем торговую систему и определим значение функционала доходности. В результате было получено распределение доходности, показанное на рис. 9 и отмеченное легендой «факт».

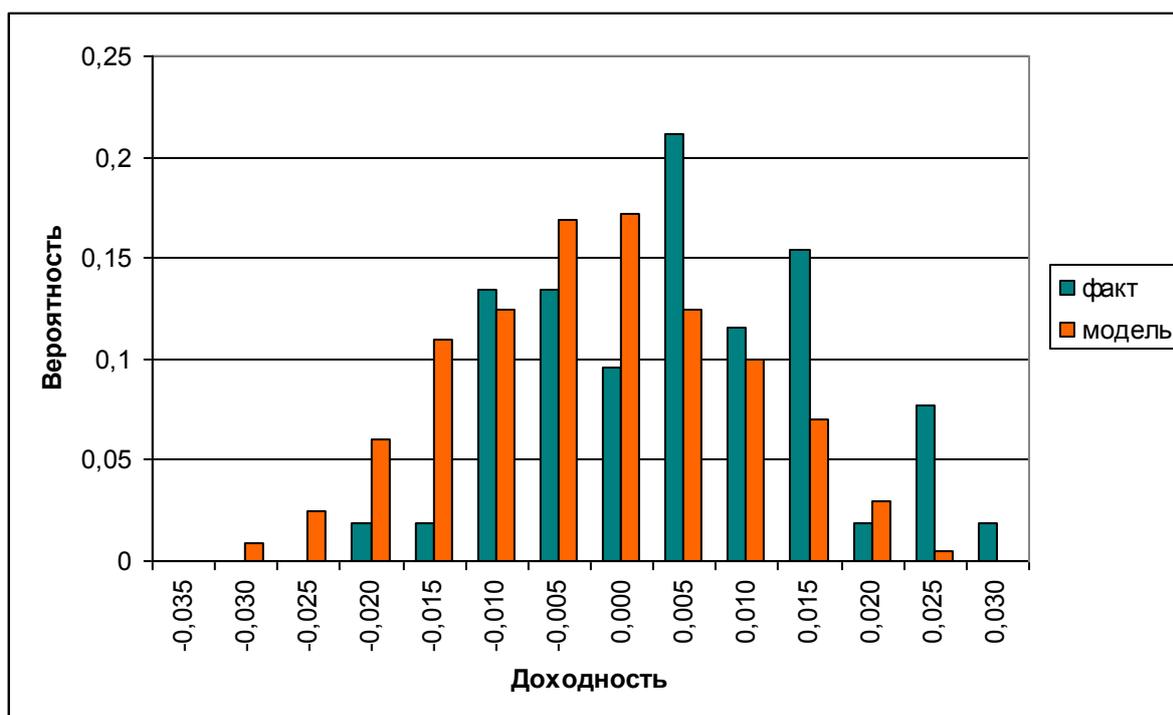


Рис. 9 – Распределение доходности в окне $T = 8000$

Среднее значение доходности за этот период составило $\bar{E} = 0,0022$ при среднеквадратичном отклонении $\sigma_E = 0,011$, так что коэффициент Шарпа равен $S_T = \frac{\bar{E}}{\sigma_E} = 0,20$. В промежутке $\bar{E} \pm \sigma_E$ находится 58 % ожидаемых доходностей. В частности, за последний промежуток оптимизации 24 ÷ 32 тыс. точек доходность оказалась равной 0,007.

Запустим теперь эту оптимизированную стратегию на новом участке, где оптимизации не проводилось, т.е. на промежутке 32 ÷ 40 тыс. Обнаруживаем, что доходность системы на этом участке составила $-0,004$, т.е. было получено существенно иное значение, чем математическое ожидание из области оптимизации. Из распределения доходности (рис. 9, «факт») находим, что вероятность события «доходность на промежутке 8000 не более $-0,004$ » составляет приблизительно 0,30. Возможно, что нам просто не очень повезло с «реальным» включением? На самом деле это не так. Чтобы составить более реальное представление о возможностях этой торговой модели, сгенерируем на промежутке 32 ÷ 40 тыс. 200 траекторий с тем же распределением приращений по методике, описанной в п.3. На каждой из них запустим торговую систему и определим доходность. Соответствующее распределение также показано на рис. 9 (легенда «модель»). По этому распределению средняя доходность равна $\bar{E}' = -0,005$, а среднеквадратичное отклонение равно $\sigma_{E'} = 0,013$, так что коэффициент Шарпа равен $S_T' = -0,38$. Вероятность того, что доходность будет лежать в промежутке $\bar{E}' \pm \sigma_{E'}$, равна 0,80. Тем самым показано, что доходность на уровне $-0,004$ за 8000 шагов для этой модели вполне обычна. Вероятность же того, что доходность будет более 0,07, оказывается примерно равной 0,10. Такой результат получился вследствие того, что оптимизация была проведена на одной случайной траектории, и параметр, оптимальный для нее, не будет таковым на любой другой траектории с тем же распределением. Следовательно, более корректно проводить оптимизацию и тестирование не на одной исторической реализации, а на пучке траекторий с характерным для данного ряда нестационарным распределением.

Заключение

В работе описан метод генерации траекторий нестационарных временных рядов, имитирующих временной ряд, возникающий на практике в процессе наблюдения случайной величины, распределение которой не является стационарным. Моделирование ряда основано на решении эмпирического кинетического уравнения Фоккера-Планка, что позволяет задавать скользящее среднее, дисперсию, размах и аналогичные величины изменяющимися во времени в соответствии с требованиями статистического эксперимента, причем выборочные распределения остаются в классе распределений, характерных для

данного ряда. Подчеркнем, что класс распределений выделяется непараметрическими методами.

Метод позволяет тестировать индикаторы-предикторы изменения какого-либо свойства временного ряда в широком диапазоне изменения его средних выборочных характеристик. В настоящее время для такого тестирования используются исторические данные за достаточно большой промежуток времени, но надо помнить, что при этом берется единственная реализация случайного процесса. Можно предположить, что такой промежуток действительно содержит все нужные для теста ситуации, и одной прогонки по длинной истории будет достаточно. По этому поводу следует заметить, что тест на прошлых данных, например, функционала доходности торговой системы почти никогда не подтверждается в будущем с требуемой точностью, и весьма часто подтверждается с точностью до наоборот. Это связано в первую очередь с тем, что в будущем реализуется произвольная траектория из ансамбля возможных путей, а в прошлом есть только одна. Этой одной траектории достаточно, если только процесс эргодический. Поскольку это последнее требование не выполняется (во всяком случае на финансовых рынках), то тестирования в прошлом всегда недостаточно.

Разумеется, описанный в работе метод не моделирует сам финансовый инструмент, поскольку, возможно, на действительную траекторию влияют не учитываемые в модели внешние факторы. К преимуществам метода следует отнести то, что он позволяет провести стресс-тест на работоспособность индикатора в пределах, контролируемых исследователем. Исторический же ряд данных не предоставляет таких возможностей. Кроме того, для квалифицированного тестирования ряд прошлых данных требует предварительного выявления интересных ситуаций, кластеризации их, определения ошибок при кластеризации, что весьма трудоемко и не дает полного представления об имеющихся локальных паттернах ряда. Рассмотренный в работе практический пример оптимизации торговой стратегии показывает причину ошибочности выводов, получаемых при тестировании определенного функционала на одной случайной траектории.

Таким образом, численный код, генерирующий по фрагменту траектории нестационарного ряда ансамбль его нестационарных же реализаций, представляет не только теоретический интерес в качестве новой методики, но и имеет практическую важность в связи с тем, что он более корректно позволяет определить вероятностные характеристики практически используемых функционалов, заданных на случайных траекториях.

Литература

1. Босов А.Д., Кальметьев Р.Ш., Орлов Ю.Н. Моделирование нестационарного временного ряда с заданными свойствами выборочного распределения // Мат. Мод., 2014. № 3. С. 97-107.
2. Орлов Ю.Н., Осминин К.П. Нестационарные временные ряды: методы прогнозирования с примерами анализа финансовых и сырьевых рынков. – М.: Editorial URSS, Книжный дом «ЛИБРОКОМ», 2011. – 384 с.
3. Гнеденко Б.В. Курс теории вероятностей. – М.: Физматлит, 1961. – 406 с.
4. Прохоров С.А. Математическое описание и моделирование случайных процессов. – Уральск: Самарский научный центр РАН, 2001.
5. Орлов Ю.Н. Оптимальное разбиение гистограммы для оценивания выборочной плотности функции распределения нестационарного временного ряда / Препринты ИПМ им. М.В. Келдыша РАН, № 53, 2011. – 20 с.
URL: <http://library.keldysh.ru/preprint.asp?id=2013-14>
6. Орлов Ю.Н., Шагов Д.О. Индикативные статистики для нестационарных временных рядов // Препринты ИПМ им. М.В. Келдыша РАН, № 53, 2011. – 20 с. URL: <http://library.keldysh.ru/preprint.asp?id=2011-53>
7. Босов А.Д., Орлов Ю.Н. Кинетико-гидродинамический подход к прогнозированию нестационарных временных рядов на основе уравнения Фоккера-Планка // Труды МФТИ, 2012. Т. 3. № 4. С. 134-140.
8. Босов А.Д., Орлов Ю.Н. Эмпирическое уравнение Фоккера-Планка для моделирования нестационарных временных рядов // Препринты ИПМ им. М.В. Келдыша РАН. 2013. № 3. 30 с.
URL: <http://library.keldysh.ru/preprint.asp?id=2013-3>