



I T H E A



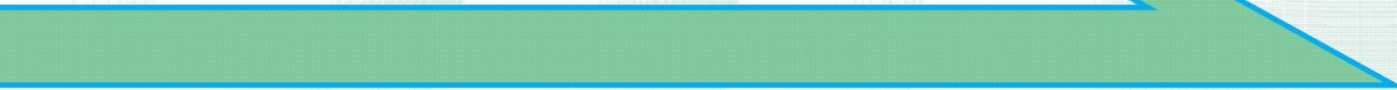
International Journal

INFORMATION

**CONTENT
&
PROCESSING**



2018 **Volume 5** **Number 1**



**International Journal
INFORMATION CONTENT & PROCESSING
Volume 5 / 2018, Number 1**

EDITORIAL BOARD

Editor in chief: **Krassimir Markov** (Bulgaria)

Abdel-Badeeh M. Salem (Egypt)	Gurgen Khachatryan (Armenia)	Oleksandr Stryzhak (Ukraine)
Abdelmgeid Amin Ali (Egypt)	Hasmik Sahakyan (Armenia)	Olena Chebanyuk (Ukraine)
Albert Voronin (Ukraine)	Irina Artemieva (Russia)	Oleksandr Trofymchuk (Ukraine)
Alexander Eremeev (Russia)	Yurii Krak (Ukraine)	Orly Yadid-Pecht (Israel)
Alexander Palagin (Ukraine)	Elguja Meqvabishvili (Georgia)	Pedro Marijuan (Spain)
Alexey Petrovskiy (Russia)	Juan Castellanos (Spain)	Rafael Yusupov (Russia)
Alexey Voloshin (Ukraine)	Koen Vanhoof (Belgium)	Sergey Krivii (Ukraine)
Alfredo Milani (Italy)	Krassimira B. Ivanova (Bulgaria)	Tatyana Gavrilova (Russia)
Anatoliy Gupal (Ukraine)	Levon Aslanyan (Armenia)	Vadim Vagin (Russia)
Anatoliy Krissilov (Ukraine)	Luis Fernando de Mingo (Spain)	Valeria Gribova (Russia)
Arnold Sterenharz (Germany)	Liudmila Cheremisnina (Belarus)	Vasil Sgurev (Bulgaria)
Benoa Depaire (Belgium)	Lyudmila Lyadova (Russia)	Vitalii Velychko (Ukraine)
Diana Bogdanova (Russia)	Mark Burgin (USA)	Vitaliy Snituk (Ukraine)
Dmitro Buy (Ukraine)	Martin P. Mintchev (Canada)	Vladimir Jotsov (Bulgaria)
Elena Zamyatina (Russia)	Mikhail Alexandrov (Russia)	Vladimir Ryazanov (Russia)
Ekaterina Solovyova (Ukraine)	Nadiia Volkovych (Ukraine)	Vladimir Shirokov (Ukraine)
Evgeniy Bodyansky (Ukraine)	Nataliia Kussul (Ukraine)	Xenia Naidenova (Russia)
Galyna Gayvoronska (Ukraine)	Natalia Ivanova (Russia)	Yuriy Zaichenko (Ukraine)
Galina Setlac (Poland)	Natalia Pankratova (Ukraine)	Yurii Zhuravlev (Russia)
Gordana Dodig Crnkovic (Sweden)	Olga Nevzorova (Russia)	Zurab Munjishvili (Georgia)

**IJ ICP is official publisher of the scientific papers of the members of the ITHEA®
International Scientific Society**

IJ ICP rules for preparing the manuscripts are compulsory.

The **rules for the papers** for ITHEA International Journals as well as the **subscription fees** are given on www.ithea.org.

The papers should be submitted by ITHEA® Submission system <http://ij.ithea.org>.

Responsibility for papers published in IJ IMA belongs to authors.

International Journal "INFORMATION CONTENT AND PROCESSING" Volume 5, Number 1, 2018

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria,
in collaboration with: University of Telecommunications and Posts, Bulgaria,
V.M.Glushkov Institute of Cybernetics of NAS, Ukraine, Universidad Politécnic de Madrid, Spain,
Hasselt University, Belgium, University of Perugia, Italy,
Institute for Informatics and Automation Problems, NAS of the Republic of Armenia
St. Petersburg Institute of Informatics, RAS, Russia,

Publisher: **ITHEA®** Sofia, 1000, P.O.B. 775, Bulgaria. www.ithea.org, e-mail: office@ithea.org

Technical editor: **Ina Markova**

Printed in Bulgaria

Copyright © 2018 All rights reserved for the publisher and all authors.

© 2014 – 2018 "Information Content and Processing" is a trademark of ITHEA®

© ITHEA is a registered trade mark of FOI-Commerce Co.

ISSN 2367-5128 (printed)

ISSN 2367-5152 (online)

ФРАКТАЛЬНЫЙ АНАЛИЗ САМОПОДОБНЫХ И МУЛЬТИФРАКТАЛЬНЫХ ВРЕМЕННЫХ РЯДОВ ЧАСТЬ 1

Людмила Кириченко, Тамара Радивилова

ВВЕДЕНИЕ

В настоящее время стало общепризнанным, что многие информационные, биологические, физические, технологические процессы обладают сложной фрактальной структурой. Сейчас фрактальный анализ используется для моделирования, анализа и контроля сложных систем в различных областях науки и техники: в телекоммуникациях – для предупреждения перегрузок компьютерных сетей; в геологии – для прогнозирования сейсмической активности и цунами; в биологии и медицине – для диагностики заболеваний и физиологического состояния по записям ЭКГ и ЭЭГ; в экономике – для прогнозирования кризисных ситуаций и оценивания риска по финансовым рядам; в физике – для исследования турбулентности и термодинамических процессов и т.д. Несмотря на множество публикаций, связанных с практическим применением фрактального анализа, в настоящее время не существует универсального подхода к оцениванию фрактальных характеристик, особенно для нестационарных временных рядов небольшой длины.

В серии из пяти статей будет предложена совокупность взаимосвязанных методов комплексного оценивания параметров фрактальных стохастических процессов для анализа временных рядов по выборочным данным небольших объемов.

В первой статье, *“Основные понятия, характеристики и модели самоподобных и мультифрактальных стохастических процессов”*, рассмотрены основные понятия, свойства и характеристики самоподобных

и мультифрактальных стохастических процессов и временных рядов; представлены основные модели фрактальных процессов.

Во второй статье, "*Оценивание параметра самоподобия для стационарных стохастических процессов*", рассмотрены основные методы определения показателя Херста; предложен подход к оцениванию параметра самоподобия по стационарным временным рядам, который обеспечивает несмещенное интервальное оценивание параметра самоподобия.

В третьей статье, "*Оценивание параметра самоподобия для нестационарных стохастических процессов*", предложены и исследованы методы оценивания показателя Херста для временных рядов со значительными трендовыми и циклическими составляющими; разработан комплексный метод оценивания параметра самоподобия для стационарных и нестационарных коротких временных рядов, позволяющий выбрать наиболее перспективную процедуру оценивания.

В четвертой статье, "*Оценивание мультифрактальных характеристик стохастических процессов*", проведено сравнительное исследование точности оценивания мультифрактальных характеристик для двух самых популярных методов мультифрактального анализа и разработан комплексный метод оценивания мультифрактальных характеристик для временных рядов малой длины.

В пятой статье, "*Фрактальный анализ реальных данных*", приведены результаты практического применения предложенных методов фрактального анализа в разных областях исследований, таких как инфокоммуникационные системы, финансовые рынки, биомедицинские сигналы, социальные сети и некоторых природных явления.

В этом номере журнала опубликованы первые три из них.

1. Основные понятия, характеристики и модели самоподобных и мультифрактальных стохастических процессов

2. Оценивание параметра самоподобия для стационарных стохастических процессов
3. Оценивание параметра самоподобия для нестационарных стохастических процессов

Остальные две будут опубликованы в следующем номере журнала.

4. Оценивание мультифрактальных характеристик стохастических процессов
5. Фрактальный анализ реальных данных

Благодарности

Работа опубликована при поддержке общества ITHEA ISS (www.ithea.org) и лично Крассимира Маркова, которому авторы выражают глубокую признательность за помощь и понимание.

Информация об авторах



Людмила Кириченко – д.т.н., профессор Харьковского национального университета радиозлектроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: lyudmyla.kirichenko@nure.ua.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, фрактальный анализ, вейвлет-анализ, детерминированные хаотические системы.



Тамара Радивилова – к.т.н., доцент Харьковского национального университета радиозлектроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: tamara.radivilova@gmail.com.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, телекоммуникационные системы, управление трафиком, информационная безопасность.

ОСНОВНЫЕ ПОНЯТИЯ, ХАРАКТЕРИСТИКИ И МОДЕЛИ САМОПОДОБНЫХ И МУЛЬТИФРАКТАЛЬНЫХ СТОХАСТИЧЕСКИХ ПРОЦЕССОВ

Людмила Кириченко, Тамара Радивилова

ВВЕДЕНИЕ

В настоящее время стало общепризнанным, что многие информационные, биологические, физические, технологические процессы обладают сложной фрактальной структурой. Фрактальный анализ используется для моделирования, анализа и контроля сложных систем в различных областях науки и техники. В частности, в геологии – для прогнозирования сейсмической активности и цунами, определения возраста геологических пород; в биологии и медицине – для диагностики заболеваний и физиологического состояния по записям ЭКГ и ЭЭГ, при исследовании мутаций и изменений на генетическом уровне; в экономике – для прогнозирования кризисных ситуаций и оценивания риска по финансовым рядам; в физике – для исследования турбулентности и термодинамических процессов. Этот список далеко не полон.

Процессы, обладающие фрактальными свойствами, можно разделить на две группы: самоподобные (монофрактальные) и мультифрактальные. Мультифрактальные объекты являются статистически неоднородными самоподобными объектами. Для мультифрактальных временных рядов статистическая неоднородность объекта выражается в неоднородности распределения данных ряда, т.е. наличии тяжелых хвостов плотности распределения вероятностей временного ряда. В сравнении с самоподобными, мультифрактальные процессы проявляют более сложное скейлинговое поведение.

1.1 Основные понятия и свойства самоподобных случайных процессов

Основные понятия и определения самоподобных случайных процессов были приведены в работах [Leland, 1993; Taqqu, 1995; Taqqu, 1997; Willinger, 1997; Mandelbrot, 1983; Федер, 1991].

1.1.1 Самоподобный стохастический процесс с непрерывным временем

Стохастический процесс $X(t)$, $t \in \mathbb{R}$ с непрерывной переменной времени называется самоподобным в узком смысле с параметром H , $0 < H < 1$, если для любого вещественного значения $a > 0$ конечномерные распределения для $X(at)$, $t \in \mathbb{R}$ идентичны конечномерным распределениям $a^{-H}X(at)$, $t \in \mathbb{R}$, т.е. если для любых $k \geq 1$, $t_1, t_2, \dots, t_k \in \mathbb{R}$ и любых $a > 0$:

$$\text{Law}\{X(t_2), \dots, X(t_k)\} = \text{Law}\{a^{-H}(X(at_1), a^{-H}X(at_2), \dots, a^{-H}X(at_k))\} \quad (1.1)$$

Кратко уравнение (1.1) можно записать в виде:

$$\text{Law}\{X(t), t \in \mathbb{R}\} = \text{Law}\{a^{-H}X(at), t \in \mathbb{R}\}. \quad (1.2)$$

Обозначение $\text{Law}\{\cdot\}$ означает конечномерные законы распределения случайного процесса. Параметр H , называемый параметром Херста, представляет собой меру самоподобия стохастического процесса.

Стохастический процесс $X(t)$ является статистически самоподобным или самоподобным в широком смысле, если процесс $a^{-H}X(at)$ обладает теми же статистическими характеристиками второго порядка (математическим ожиданием $M[X(t)]$, дисперсией $D[X(t)]$ и автокорреляционной функцией $r_x(t,s)$), что и $X(t)$:

$$M[X(t)] = \frac{M[X(at)]}{a^H}$$

$$D[X(t)] = \frac{D[X(at)]}{a^{2H}}$$

$$r_x(t,s) = \frac{r_x(at,as)}{a^{2H}}.$$

Легко показать, положив $a=1/t$, что для самоподобного процесса выполняется следующее равенство:

$$\text{Law}\{X(t)\} = \text{Law}\left\{\left(\frac{1}{t}\right)^{-H} X(1)\right\} = \text{Law}\{t^H X(1)\} \quad (1.3)$$

Учитывая (1.2), начальные моменты самоподобного случайного процесса можно выразить как

$$M\left[|X(t)|^q\right] = M\left[|t^H X(1)|^q\right] = t^{qH} M\left[|X(1)|^q\right] = C(q) \cdot t^{qH} \quad (1.4)$$

где величина $C(q) = M\left[|X(1)|^q\right]$.

Важным классом случайных процессов, часто встречающихся в различных областях естествознания и техники, являются случайные процессы, которые имеют стационарные приращения. Случайный процесс $\{X(t), t \in \mathbb{R}\}$, имеющий стационарные приращения, является самоподобным, если

$$\text{Law}\{X(t + \Delta t) - X(t)\} = \text{Law}\{X(t) - X(0)\}, \forall t, \Delta t \in \mathbb{R} \quad (1.5)$$

Такие процессы имеют среднее значение, равное 0, дисперсия определяется как $M\left[X^2(t)\right] = \sigma^2 |t|^{2H}$, а корреляционная функция зависит только от моментов времени t и s :

$$M\{X(t), X(s)\} = \frac{\sigma^2}{2} \left\{ |t|^{2H} + |s|^{2H} - |t - s|^{2H} \right\}. \quad (1.6)$$

1.1.2 Самоподобный стохастический процесс с дискретным временем

Рассмотрим понятие самоподобия для процессов с дискретным временем. Пусть $X = (X_1, X_2, \dots)$ – отрезок стационарного в широком смысле случайного процесса с дискретным временем $t \in N = \{1, 2, \dots\}$. Допустим, процесс X имеет автокорреляционную функцию следующего вида:

$$r(k) \sim k^{-\beta} L_1(k), \quad k \rightarrow \infty, \quad (1.7)$$

где $0 < \beta < 1$ и L_1 – медленно меняющаяся на бесконечности функция, то есть $\lim_{t \rightarrow \infty} \frac{L_1(tx)}{L_1(t)} = 1$ для всех $x > 0$.

Обозначим через $X^{(m)} = \{X_1^{(m)}, X_2^{(m)}, \dots\}$ – усредненный по блокам длины m процесс X , компоненты которого определяются равенством

$$X_t^{(m)} = \frac{1}{m} (X_{tm-m+1} + \dots + X_{tm}), \quad \dots \quad (1.8)$$

Такой ряд называется агрегированным. Обозначим через $r_m(k)$ и Var_m коэффициент корреляции и дисперсию процесса $X^{(m)}$ соответственно. Процесс X называется строго самоподобным в широком смысле с параметром $H = 1 - (\beta / 2)$, $0 < \beta < 1$, если

$$r_m(k) = r(k), \quad k \in Z_+, \quad m \in \{2, 3, \dots\},$$

то есть, процесс не меняет свой коэффициент корреляции после усреднения по блокам длины m . Другими словами, X является самоподобным в широком смысле, если агрегированный процесс $X^{(m)}$ неотличим от исходного процесса X в отношении статистических характеристик второго порядка.

Процесс X называется асимптотически самоподобным в широком смысле с параметром $H = 1 - (\beta / 2)$, $0 < \beta < 1$ если

$$\lim_{m \rightarrow \infty} r_m(k) = r(k), \quad k \in N.$$

Процесс X называется самоподобным в узком смысле с параметром $H = 1 - (\beta / 2)$, $0 < \beta < 1$, если справедливо выражение

$$\text{Law}\{m^{1-H} X^{(m)}\} = \text{Law}\{X\}, \quad m \in N. \quad (1.9)$$

1.1.3 Долгосрочная зависимость и тяжелые хвосты распределений

Понятие медленно убывающей зависимости имеет ключевое значение в теории самоподобных процессов и описывает долгосрочную память процесса: будущее процесса определяется его прошлым, причем с убывающей степенью влияния по мере того, как прошлое удалено от настоящего. Таким образом, процесс с продолжительной памятью как бы "забывает" свое относительно давнее прошлое по мере продвижения времени в будущее.

Процесс X обладает медленно убывающей зависимостью, если выполняется отношение (1.7). Таким образом, процессы с медленно убывающей зависимостью характеризуются автокорреляционной функцией, которая убывает гиперболически (по степенному закону) при увеличении временной задержки. Можно показать, что из (1.7) следует несуммируемость автокорреляционной функции, то есть $\sum_k r(k) = \infty$. В

отличие от процессов с медленно убывающей зависимостью, процессы с быстро убывающей зависимостью обладают экспоненциально спадающей автокорреляционной функцией вида

$$r(k) \sim \rho^k, \quad k \rightarrow \infty, \quad 0 < \rho < 1, .$$

и, как следствие, суммируемостью автокорреляционной функции:

$$0 < \sum_k r(k) < \infty.$$

Дискретные самоподобные процессы обладают медленно убывающей автокорреляционной функцией агрегированного процесса $X^{(m)}$ при $m \rightarrow \infty$ в отличие от распространенных стохастических моделей для которых выполняется

$$r_m(k) \rightarrow 0, \quad m \rightarrow \infty, \quad k \in N.$$

Задание автокорреляционной функции в виде (1.7) равносильно определению процесса со следующей дисперсией:

$$\sigma^2 \sim am^{-\beta}, \quad m \rightarrow \infty,$$

где a – конечная положительная константа, не зависящая от m ; $0 < \beta < 1$.

Т.е. долгосрочная зависимость подразумевает свойство самоподобия в широком смысле со значением показателя $H > 0.5$ и наоборот.

Большинство процессов, обладающих долгосрочной зависимостью, обладают тяжелыми хвостами одномерной функции распределения вероятностей. Случайная величина X имеет распределение с тяжелым хвостом, если

$$P[X > x] \sim c \cdot x^{-\alpha}, \quad x \rightarrow \infty, \quad (1.10)$$

где величина α , $0 < \alpha < 2$, называется параметром формы; c – некоторая положительная константа.

В отличие от распределений с легкими хвостами, такими как экспоненциальное или гауссовское, которые имеют экспоненциальное убывание хвоста, случайные величины с тяжелым хвостом обладают хвостами, спадающими по степенному (гиперболическому) закону. При $0 < \alpha < 2$ случайные величины имеют бесконечную дисперсию, а при $0 < \alpha \leq 1$ еще и обладают бесконечным средним. Основное свойство случайной величины, распределенной с тяжелым хвостом, состоит в том, что она проявляет высокую изменчивость. Иными словами, ее выборка представляет собой в основном относительно небольшие значения, однако также содержит и достаточное количество очень больших значений.

Параметр Херста H , называемый параметром самоподобия, находится в диапазоне $0 < H < 1$ и представляет собой ключевую меру самоподобия и меру длительности долгосрочной зависимости стохастического процесса.

В случае $0.5 < H < 1$ говорят о персистентном (поддерживаемом) поведении процесса, либо о том, что процесс обладает длительной памятью. Другими словами, если в течение некоторого времени в прошлом наблюдались положительные приращения процесса, то есть происходило

увеличение, то и впредь в среднем будет происходить увеличение. Иначе говоря, вероятность того, что процесс на $i+1$ шаге отклоняется от среднего в том же направлении, что и на i шаге настолько велика, насколько параметр H близок к 1. Таким образом, персистентные стохастические процессы обнаруживают четко выраженные тенденции изменения при относительно малом шуме.

В случае $0 < H < 0.5$ говорят об антиперсистентном процессе. Здесь высокие значения процесса следуют за низкими, и наоборот. Другими словами, вероятность того, что на $i+1$ шаге процесс отклоняется от среднего в противоположном направлении (по отношению к отклонению на i шаге) настолько велика, насколько параметр H близок к 0.

При $H = 0.5$ отклонения процесса от среднего являются действительно случайными и не зависят от предыдущих значений, что соответствует случаю обычного броуновского движения.

1.2 Мультифрактальные процессы

Мультифрактальные меры были введены Б.Мандельбротом [Mandelbrot, 1974] для описания распределения турбулентной диссипации. С тех пор мультифрактальные структуры были обнаружены во множестве природных и технических процессов и используются для моделирования и исследования турбулентных процессов, телекоммуникационных потоков, финансовых рынков, медицинских данных, геофизических процессов.

Мультифрактальность – это концепция, которая, с некоторыми незначительными изменениями, может быть в равной степени хорошо применена к функциям и мерам, детерминированным или стохастическим. В описании основных понятий и свойств мультифрактальных процессов существует несколько подходов: глобальный, при котором основные определения и свойства выводятся через свойства моментов случайных процессов [Calvet, 1997; Riedi, 2002; Kantelhardt, 2008; Abry, 2009];

локальный, описывающий локальные масштабные и сингулярные свойства фрактальных реализаций и функций и основанный на свойствах локальных экспонент Гельдера [Feder, 1988; Riedi, 2002; Kantelhardt, 2002; Abry, 2009]; и подход, основанный на свойствах фрактальных и мультифрактальных мер (множеств) [Mandelbrot, 2001; Riedi, 2002; Abry, 2009].

1.2.1. Законы распределения и моментные характеристики мультифрактальных процессов

В отличие от самоподобных процессов (1.2) мультифрактальные процессы обладают более разнообразным скейлинговым поведением [Riedi, 2002]:

$$\text{Law}\{X(at)\} = \text{Law}\{M(a) \cdot X(t)\}, \dots \quad (1.11)$$

где $M(a)$ - независимая от $X(t)$ случайная функция.

В случае самоподобного процесса $M(a) = a^H$. Для мультифрактальных процессов обобщенный показатель Херста $H(a) = \log_a M(a)$ является случайной функцией аргумента a . Соотношение (1.11) можно переформулировать следующим образом:

$$\text{Law}\{X(at)\} = \text{Law}\{a^{H(a)} \cdot X(t)\}, \quad (1.12)$$

Свойство мультифрактальности (1.12) устанавливает достаточно жесткие ограничения на законы распределения случайных функций $M(a)$ и $X(t)$. Для случайной скейлинговой функции $M(a)$ должно выполняться следующее условие:

$$\text{Law}\{M(ab)\} = \text{Law}\{M_1(a) \cdot M_2(b)\}, \quad (1.13)$$

где $M_1(a)$ и $M_2(a)$ независимы и имеют тот же закон распределения, что и $M(a)$.

Для мультифрактального процесса $X(t)$ должно следовать, что если выполняется условие $\frac{a_2}{a_1} = \frac{a_3}{a_2}$, то отношение (1.12) влечет за собой следующее:

$$\text{Law}\left\{\frac{X(a_2t)}{X(a_1t)}\right\} = \text{Law}\left\{\frac{X(a_3t)}{X(a_2t)}\right\}.$$

Из этих свойств вытекает определяющее свойство мультифрактальных процессов [Calvet, 1997; Mandelbrot, 2001; Riedi, 2002]: процесс $X(t)$ является мультифрактальным, если выполняется следующее отношение:

$$\mathbb{M}\left[|X(t)|^q\right] = c(q) \cdot t^{\tau(q)+1}, \quad \forall t \in T, \quad \forall q \in \mathbb{Q}, \quad (1.14)$$

где $c(q)$ и $\tau(q)$ – некоторые детерминированные функции. Величины T и \mathbb{Q} являются интервалами ненулевой длины на действительной оси, и $0 \in T$, $[0,1] \subset \mathbb{Q}$.

Мультифрактальные процессы, таким образом, обладают глобальным скейлингом, в том смысле, что его моменты удовлетворяют скейлинговому

соотношению (1.14). Функция $\tau(q)$ называется скейлинговой экспонентой мультифрактального процесса. Из (1.14) вытекают некоторые свойства функции $\tau(q)$: $\tau(0) = -1$; $\tau(1) = 0$, $\tau(q)$ – выпуклая функция.

Обобщенный показатель Херста $h(q)$ связан с функцией соотношением [Kantelhardt, 2002]:

$$\tau(q) = qh(q) - 1. \quad (1.15)$$

Самоподобные процессы являются подклассом мультифрактальных. Для самоподобных процессов скейлинговая экспонента линейна: $\tau(q) = Hq - 1$

и значение $\frac{\tau+1}{q}$ при $q = 2$ совпадает со значением степени самоподобия

H .

1.2.2 Характеристики мультифрактальных множеств

Самоподобие фрактальных объектов заключается в сохранении структуры объекта при изменении масштаба. Рассмотрим основные характеристики мультифрактального множества [Божокин, 2001]. Пусть в общем случае мультифрактальный объект занимает некоторую ограниченную область в d -мерном Евклидовом пространстве и определяет собой множество из $N \rightarrow \infty$ точек. Разобьем всю область на кубические ячейки с ребром ε и объемом ε^d . Рассмотрим обобщенную статистическую сумму $Z(q, \varepsilon)$, характеризующую показателем степени q , который может принимать любые значения в интервале $-\infty < q < +\infty$:

$$Z(q, \varepsilon) = \sum_{i=1}^{N(\varepsilon)} p_i^q(\varepsilon), \quad (1.16)$$

где $p_i(\varepsilon) = \lim_{N \rightarrow \infty} \frac{n_i(\varepsilon)}{N}$; $n_i(\varepsilon)$ – количество точек, попавшее в ячейку с номером i ; $N(\varepsilon)$ – суммарное количество занятых ячеек, которое зависит от размера ячейки ε .

Вероятности p_i характеризуют относительную заселенность ячеек. В общем случае мультифрактальное множество характеризуется некоторой нелинейной функцией $\tau(q)$, определяющей поведение статистической суммы $Z(q, \varepsilon)$ при $\varepsilon \rightarrow 0$:

$$Z(q, \varepsilon) \propto \varepsilon^{\tau(q)}. \quad (1.17)$$

Функция $\tau(q)$ обычно называется скейлинговой экспонентой и определяется как

$$\tau(q) = \lim_{\varepsilon \rightarrow 0} \frac{\ln Z(q, \varepsilon)}{\ln \varepsilon}. \quad (1.18)$$

В случае однородного фрактального множества с фрактальной размерностью D во всех занятых ячейках содержится одинаковое

количество точек, т.е., $p_i(\varepsilon) = p(\varepsilon) = 1 / N(\varepsilon)$, и обобщенная статистическая сумма принимает вид:

$$Z(q, \varepsilon) = N^{1-q}(\varepsilon) = \varepsilon^{-D(1-q)},$$

а функция $\tau(q) = (q-1)D$ является линейной. Если распределение точек по ячейкам не одинаково, то фрактальное множество является неоднородным, т.е. мультифрактальным, и $\tau(q)$ является нелинейной функцией. При $q \rightarrow +\infty$ основной вклад в обобщенную статистическую сумму вносят ячейки, содержащие наибольшее число частиц n_i и, следовательно, характеризующиеся наибольшей вероятностью заполнения p_i . Наоборот, при $q \rightarrow -\infty$ основной вклад в сумму дают самые разреженные ячейки с малыми значениями p_i . Таким образом, функция $\tau(q)$ показывает, насколько неоднородным является исследуемое множество точек.

Функция мультифрактального спектра (спектра сингулярностей) $f(\alpha)$ в данном случае вводится следующим образом. Зависимость вероятности от размера ячейки $p_i(\varepsilon)$ имеет степенной характер

$$p_i(\varepsilon) \propto \varepsilon^{\alpha_i}, \quad (1.19)$$

где α_i представляет собой некоторый показатель степени, вообще говоря разный, для разных ячеек (показатель сингулярности). Для однородного

фрактала все показатели степени α_i одинаковы и равны фрактальной размерности D .

Функция мультифрактального спектра $f(\alpha)$ характеризует распределение вероятностей различных значений α_i . Если величина $n(\alpha)d\alpha$ является вероятностью того, что α_i находится в интервале $(\alpha, \alpha+d\alpha)$, т.е. представляет собой число ячеек i , обладающих одинаковой мерой $p_i(\varepsilon)$ с $\alpha_i \in (\alpha, \alpha+d\alpha)$, тогда

$$n(\alpha) \approx \varepsilon^{-f(\alpha)}. \quad (1.20)$$

Таким образом, функция $f(\alpha)$ представляет собой фрактальную размерность некоего однородного фрактального подмножества ξ_α из исходного множества ξ , характеризуемого одинаковыми вероятностями заполнения ячеек $p_i(\varepsilon) \approx \varepsilon^\alpha$.

Учитывая выражения (1.16) и (1.20), обобщенную статистическую сумму $Z(q, \varepsilon)$ можно выразить через функцию мультифрактального спектра $f(\alpha)$ следующим образом:

$$Z(q, \varepsilon) = \sum_{i=1}^{N(\varepsilon)} p_i^q(\varepsilon) \approx \int d\alpha n(\alpha) \varepsilon^{q\alpha} \approx \int d\alpha \varepsilon^{q\alpha - f(\alpha)}.$$

Формально переход от переменных $\{q, \tau(q)\}$ к переменным $\{\alpha, f(\alpha)\}$ может быть осуществлен при помощи следующих преобразований Лежандра:

$$\left\{ \begin{array}{l} \alpha = \frac{d\tau}{dq}, \\ f(\alpha) = q \frac{d\tau}{dq} - \tau \end{array} \right. \quad \text{и} \quad \left\{ \begin{array}{l} q = \frac{df}{d\alpha}, \\ \tau(q) = \alpha \frac{df}{d\alpha} - f. \end{array} \right. \quad (1.21)$$

1.3 Некоторые модели самоподобных и мультифрактальных процессов

Для анализа и сравнения методов оценивания, а также разработки новых, необходимо моделировать временные реализации самоподобных и мультифрактальных процессов.

1.3.1 Модели фрактального броуновского движения.

Фрактальное (обобщенное) броуновское движение (ФБД) впервые предложено и исследовано Б. Мандельбротом [Mandelbrot, 1968]. В дальнейшем класс моделей, основанных на ФБД, получил дальнейшее развитие во многих работах [Feder, 1988].

При обычном случайном блуждании частицы, для любой пары моментов времени t и t_0 , приращение ΔX координаты частицы задается выражением

$$X(t) - X(t_0) \sim \xi |t - t_0|^H,$$

где $H = 1/2$; ξ – случайная величина со стандартным нормальным распределением.

Обобщенное (фрактальное) броуновское движение $B_H(t)$ вводится путем замены показателя $H = 1/2$ на действительное число из интервала $[0, 1]$. Фрактальное броуновское движение также может быть получено из обычного броуновского движения взятием интеграла:

$$B_H(t) = \frac{1}{\Gamma(H + 0.5)} \int_{-\infty}^t (t - u)^{H-0.5} dB(u),$$

где $\Gamma(x)$ – гамма функция. Таким образом, значение $B_H(t)$ в момент времени t зависит от всех предыдущих ($u < t$) приращений $dB(u)$ простого броуновского движения. Приращения этого процесса (фрактальный гауссовский шум) имеют нулевое математическое ожидание и дисперсию, зависящую от времени

$$\text{Var}(B_H(t) - B_H(t_0)) = \text{Var}[\xi] \cdot |t - t_0|^{2H}.$$

Корреляционная функция фрактального гауссовского шума имеет вид

$$r(k) = \frac{1}{2} [(k+1)^{2H} - 2k^{2H} + (k-1)^{2H}]. \quad \text{Можно показать, что}$$

$$r(k) \sim H(2H-1)|k|^{-\beta}, \quad k \rightarrow \infty, \quad 0 < H < 1, \quad \beta = 2 - 2H, \quad \text{а также, что}$$

усредненный процесс $X^{(m)}$ имеет такое же распределение, как и процесс

X для всех $0 < H < 1$. Таким образом ФГШ является строго самоподобным процессом.

Методы построения моделей самоподобных временных рядов на основе ФБД для случая дискретного времени рассмотрены, в частности, в работах [Feder, 1988; Кроновер, 2000]. Модели, получаемые с использованием этих методов, обладают определенными недостатками, одним из которых является занижение (завышение) степени самоподобия процесса при малых или больших теоретических значениях параметра Херста и небольшой длине модельной реализации [Feder, 1988; Jeongy, 1998; Шелухин, 2008].

Широко используемым на практике методом построения моделей ФБД является метод последовательного случайного сложения Фосса. Опишем пошаговый алгоритм его реализации [Feder, 1988]. Исходные значения координат $X(t_i)$ в моменты времени $t_i = 0, 1/2, 1$ равны нулю. На первом шаге к значениям координат $X(t_1)$, $X(t_2)$, $X(t_3)$ прибавляются случайные числа, выбранные из нормального распределения с нулевым средним и начальной дисперсией σ_1^2 . Средние значения времени на каждом интервале рассматриваются как дополнительные узлы на оси времени; значения координаты в них оцениваются интерполяцией. На следующем шаге ко всем координатам $X(t_i)$ ($t_i = 0, 1/4, 1/2, 3/4, 1$) прибавляются случайные числа с нулевым средним значением и уменьшенной дисперсией $\sigma_2^2 = \frac{\sigma_1^2}{2^{2H}}$. На n -ом шаге алгоритма мы получаем значения реализации ФБД для $1+2^n$ значений времени t_i .

Дисперсия слагаемых n -го поколения равна $\sigma_n^2 = \frac{\sigma_{n-1}^2}{2^{2H}} = \frac{\sigma_0^2}{2^{2Hn}}$. Процесс, предложенный Фоссом, приводит к обобщенному броуновскому движению при любом разрешении.

Другим распространенным методом построения моделей ФБД для дискретного времени является метод Селлана-Мейера, использующий

биортогональные вейвлеты [Meyer, 1995; Sellan, 1995]. В этом случае построение реализации ФБД осуществляется с помощью дискретного вейвлет-преобразования, где детализирующие вейвлет-коэффициенты на каждом уровне являются независимыми нормально распределенными случайными величинами, а аппроксимирующие вейвлет-коэффициенты получены с помощью фрактального процесса авторегрессии и скользящего среднего FARIMA:

$$B_H(t) = \sum_{k=-\infty}^{\infty} \Phi_H(t-k) S_k^{(H)} + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} 2^{-jH} \Psi_H(2^j t - k) \varepsilon_{j,k} - b_0,$$

где Ψ_H – биортогональная базовая вейвлет-функция; Φ_H – соответствующая Ψ_H скейлинг-функция; $S_k^{(H)}$ – стационарный гауссовский процесс FARIMA с параметром дробного дифференцирования $d = H - 0.5$; $\varepsilon_{j,k}$ – независимые стандартные нормальные случайные величины; b_0 – константа, такая, что $B_H(0) = 0$.

В работе [Кириченко, 2009] было проведено исследование фрактальных свойств реализаций моделей ФБД, полученных как по методу последовательного случайного сложения (метод Фосса), так и по методу вейвлет-преобразования (метод Селлана-Мейера). Приведенные результаты свидетельствуют, что модель ФБД, реализованная с помощью вейвлет-преобразования, более корректно отображает степень самоподобия для выборок с диапазоном длины 200-10000 значений. Однако следует отметить, что время генерации данной модели в среднем на порядок больше времени генерации ФБД по методу Фосса. В дальнейшем будут приведены результаты численных экспериментов, где реализации фрактального процесса получены с помощью метода

Селлана-Мейера. На рис.1.1 представлены реализации ФБД и ФГШ, построенные при значениях показателя Херста $H = 0.3, 0.5, 0.8$.

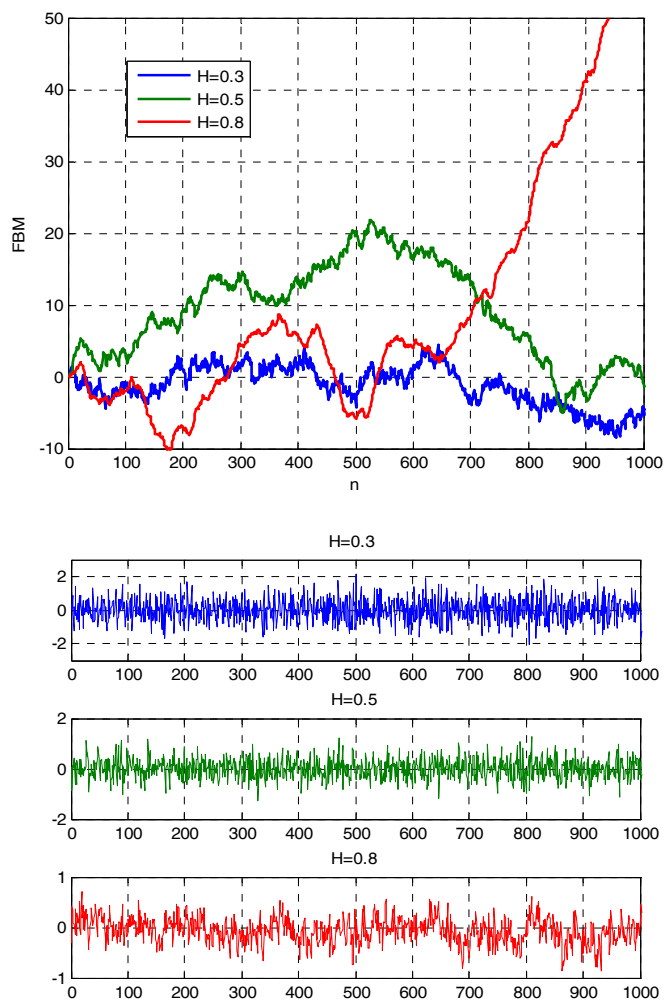


Рисунок 1.1 – Реализации ФБД (вверху) и ФГШ (внизу) для разных значений H

1.3.2 Модели, основанные на α -устойчивых распределениях

Случайная величина X называется α -устойчивой, если для любых двух положительных чисел a, b найдутся положительные числа c, d такие, что

$$\text{Law}(aX_1 + bX_2) = \text{Law}(cX + d), \quad (1.22)$$

где $\text{Law}(\cdot)$ обозначает закон распределения, X_1 и X_2 – независимые случайные величины; $\text{Law}(X_1) = \text{Law}(X_2) = \text{Law}(X)$.

Если выполняется (1.22), то существует такое $\alpha \in (0, 2]$, не зависящее от a и b , что для величины c в (1.22) выполняется равенство $c^\alpha = a^\alpha + b^\alpha$.

В общем случае, для α -устойчивых распределений в явном виде может быть записана только характеристическая функция $\varphi(t) = M[e^{itX}]$. Характеристическая функция устойчивой случайной величины $\ln \varphi(t)$ с параметрами $0 < \alpha \leq 2$, $\sigma \geq 0$, $-1 \leq \beta \leq 1$ и $\mu \in R$ имеет вид:

$$\ln \varphi(t) = \begin{cases} -\sigma^\alpha |t|^\alpha (1 - i\beta \text{sign}(t) \tan \frac{\pi\alpha}{2}) + i\mu t, & \alpha \neq 1, \\ -\sigma |t| (1 + i\beta \text{sign}(t) \frac{2}{\pi} \ln |t|) + i\mu t, & \alpha = 1. \end{cases}$$

Параметр α называется индексом устойчивости или характеристическим показателем и определяет, насколько выражен тяжелый хвост распределения. При $0 < \alpha < 2$ случайные величины имеют бесконечную дисперсию, а при $0 < \alpha \leq 1$ еще и обладают бесконечным средним. Параметр смещения β задает степень асимметричности распределения, при $\beta = 0$ распределение симметрично. Величина σ (масштабный параметр) выражает степень разброса значений относительно среднего значения, а μ при $\alpha > 1$ равно математическому ожиданию X . [Ворак, 2005]

Стохастический процесс $X(t)$ является α -устойчивым ординарным движением Леви, если он обладает стационарными независимыми α -

устойчивыми приращениями. Такой процесс является самоподобным с параметром $H = 1/\alpha$.

$$\text{Law}\{X(at)\} = a^{\frac{1}{\alpha}} \text{Law}\{X(t)\}. \quad (1.23)$$

Он используется в моделировании и исследовании стохастических процессов по двум причинам. Первой причиной является то, что движение Леви является обобщением ординарного броуновского движения. Математическое обоснование такого обобщения выходит в использовании главных свойств устойчивых законов вероятности. С точки зрения предельной теоремы, устойчивые распределения является естественным обобщением широко известных гауссовских распределений. Вторая причина важности движения Леви кроется в его свойствах масштабной инвариантности, то есть самоподобия. Более того приросты не только самоподобные, а еще имеют распределения с большими хвостами.

В работе [Соболев, 2012] рассмотрено моделирование реальных временных рядов на основе движения Леви. В основу моделирования положен алгоритм, предложенный в работе [Chamber, 1976]. Более поздние модификации алгоритма приведены в [Mittnik, 1999].

Алгоритм стартует от моделирования двух выборок случайных величин:

одной (γ_i) - равномерно распределенной на интервале $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ и второй

(W_i) – независимой от нее простейшей экспоненциально распределенной.

Для случая $\alpha \neq 1$ рассчитывается величина:

$$X_i = S_{\alpha,\beta} \frac{\sin(\alpha(\gamma_i + B_{\alpha,\beta}))}{(\cos \gamma_i)^{\frac{1}{\alpha}}} \left(\frac{\cos(\gamma_i - \alpha(\gamma_i + B_{\alpha,\beta}))}{W_i} \right)^{\frac{1-\alpha}{\alpha}}, \quad (1.24)$$

$$\text{где } S_{\alpha,\beta} = \left(1 + \beta^2 \tan^2\left(\frac{\pi\alpha}{2}\right)\right)^{\frac{1}{2\alpha}}, \quad B_{\alpha,\beta} = \frac{\arctan(\beta \tan \frac{\pi\alpha}{2})}{\alpha}.$$

Для случая $\alpha = 1$ формула (1.24) сводится к выражению

$$X_i = \frac{2}{\pi} \left(\left(\frac{\pi}{2} + \beta\gamma_i \right) \tan \gamma_i - \beta \ln \left(\frac{\frac{\pi}{2} W_i \cos \gamma_i}{\frac{\pi}{2} + \beta\gamma_i} \right) \right). \quad (1.25)$$

Формулы (1.24-1.25) моделируют значения стандартных устойчивых случайных величин $S_\alpha(1,0,0)$, однако легко перейти к общему случаю $S_\alpha(\sigma,\beta,\mu)$, используя соотношение

$$Y_i = \begin{cases} \sigma X_i + \mu, & \alpha \neq 1 \\ \sigma X_i + \frac{2}{\pi} \beta \sigma \ln \sigma + \mu, & \alpha = 1 \end{cases}.$$

На рис.1.2 приведены модельные временные реализации и соответствующие плотности распределений симметричных α -устойчивых случайных величин $S_\alpha(1,0,0)$ для значений параметра $\alpha=0.8,1.2,1.8$.

1.3.3 Каскадные модели мультифрактальных процессов

Основными математическими моделями мультифрактальных процессов являются каскадные процессы [Feder, 1988; Riedy, 2002]. Мультипликативные каскады первоначально были введены А.Н. Колмогоровым для описания нерегулярного распределения скорости диссипации энергии турбулентного потока. Дальнейшее интенсивное изучение и развитие каскадных моделей было вызвано работами Б. Мандельброта [Mandelbrot, 1974; Calvet, 1997].

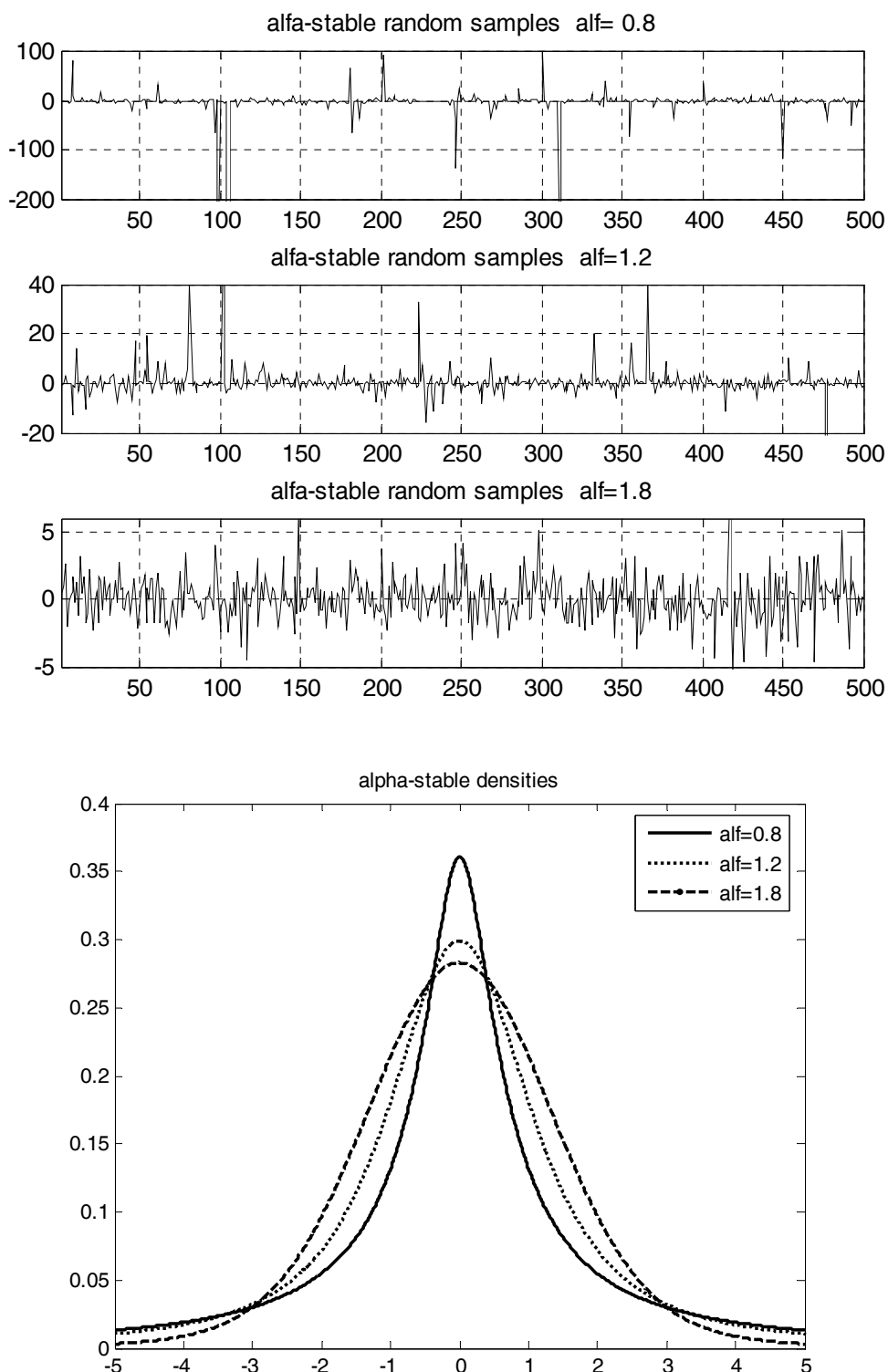


Рисунок 1.2 – Модельные реализации и плотности распределений для $S_{\alpha}(1,0,0)$

Детерминированные каскады

Простейшей моделью мультифрактального процесса с заданными свойствами является детерминированный биномиальный мультипликативный каскад [Feder, 1988; Riedy, 2002; Kantelhardt, 2008]. При его построении первоначальный единичный отрезок $[0,1]$ с мерой $\mu_0 = 1$ делится на два равных интервала $[0, 1/2]$ и $[1/2, 1]$, которым приписываются весовые коэффициенты p_1 и $p_2 = 1 - p_1$ соответственно, т.е. на 1-м шаге мы получаем следующие меры подинтервалов: $\mu_1[0, 1/2] = p_1$ и $\mu_1[1/2, 1] = 1 - p_1$. Затем с каждым из полученных подинтервалов прделывается аналогичная процедура. В результате на втором шаге имеется 4 подинтервала с соответствующими мерами $\mu_2[0, 1/4] = p_1 p_1$, $\mu_2[1/4, 1/2] = p_1 p_2$, $\mu_2[1/2, 3/4] = p_2 p_1$, $\mu_2[3/4, 1] = p_2 p_2$.

На n -м шаге подинтервал $[t_j, t_j + 2^{-n}]$, где $t_j = 0, 2^{-n}, \dots, 1 - 2^{-n}$ имеет меру $\mu_n[t_j, t_j + 2^{-n}] = p_1^{n \cdot \varphi_0} p_2^{n \cdot \varphi_1}$, где φ_0 - доля 0 в двоичном представлении t , φ_1 - доля единиц в двоичном представлении t . При числе шагов $n \rightarrow \infty$ и $p_1 \neq p_2$ мы приходим к предельной мере, являющейся неоднородным фрактальным множеством. На каждом n -м шаге мера всего первоначального отрезка остается неизменной

$$\sum_j \mu_n[t_j, t_j + 2^{-n}] = 1. \text{ Такие каскады называются консервативными.}$$

На рисунке 1.3 слева представлены теоретические скейлинговые экспоненты $\tau(q)$ для значений $p_1 = \{0.6, 0.7, 0.8, 0.9\}$, а справа показаны реализации биномиального каскада при значениях $p_1 = 0.6$ (вверху) и $p_1 = 0.8$ (внизу). Число итераций $n = 10$, т.е. длина реализации равна 2^{10} значений. Очевидно, что с увеличением первоначального весового коэффициента p_1 увеличивается неоднородность временного ряда.

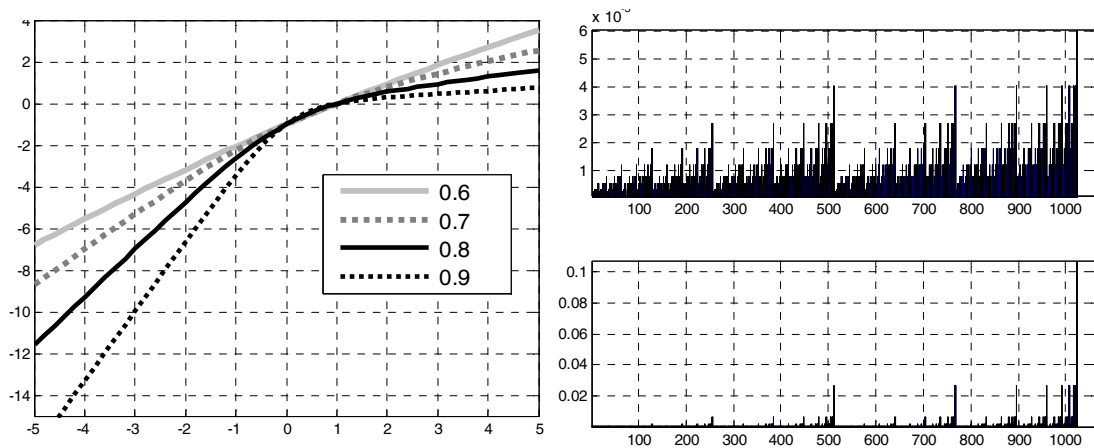


Рисунок 1.3 – Скейлинговые экспоненты и реализации детерминированного биномиального каскада

Стохастические каскады

Свойства детерминированного каскада полностью определяются величиной p_1 . При построении стохастических каскадов весовыми коэффициентами являются независимые значения некоторой заданной случайной величины W [Riedy, 2002]. В этом случае на n -м шаге подинтервал длины 2^{-n} можно обозначить как $I(k_1 \dots k_n)$, где $k_1 \dots k_n$ – последовательность 0 и 1 будет обладать мерой

$$\mu_n(I(k_1 \dots k_n)) = 2^{-n} \prod_{i=1}^n W(k_1 \dots k_i) = 2^{-n} \text{Exp} \left\{ \sum_{i=1}^n w(k_1 \dots k_i) \right\}, \quad (1.26)$$

где все $W(k_1 \dots k_i) = \text{Exp}\{w(k_1 \dots k_i)\}$ являются независимыми одинаково распределенными случайными величинами, такими что $M[W] = 1$. Доказано, что такая процедура сходится почти наверное к стохастической мультифрактальной мере при условии

$$M[W \cdot \ln W] < 1. \quad (1.27)$$

Важным свойством мультипликативных каскадов является их свойство самоподобия:

$$\mu(I(k_1 \dots k_n)) \doteq 2^{-1} W \mu(I(k_1 \dots k_{n-1})). \quad (1.28)$$

Из (1.26) и (1.28) следует скейлинговое соотношение

$$M\left[\mu(0, 2^{-n})^q\right] = 2^{-nq} M[W^q]^n M\left[\mu(0, 1)^q\right].$$

которое позволяет определить скейлинговую экспоненту $\tau(q)$.

Чтобы вычислить скейлинговую экспоненту $\tau(q)$, делаются следующие упрощающие предположения [Molchan, 1996; Riedi, 2002]:

- вместо предельной меры μ_∞ рассматривается мера μ_n на n -ом шаге;
- при вычислениях предполагается, что процесс обладает эргодическими свойствами и пространственное среднее величин μ_n^q заменяется средним по реализациям.

При сделанных допущениях

$$\tau(q) = -\log_2 M[W^q] + q - 1. \quad (1.29)$$

Для существования нетривиальной предельной меры μ_∞ требуется выполнение условия $\tau'(1) > 0$, что совпадает с (1.27).

При построении стохастических каскадов весовыми коэффициентами являются независимые значения некоторой заданной случайной величины W . Случайная величина выбирается таким образом, чтобы математическое ожидание суммы весовых коэффициентов на каждой итерации равнялось единице. Если выбрать случайную величину, определенную на интервале $[0,1]$, то сумма коэффициентов на каждой итерации будет равной единице.

В этом случае первым двум интервалам будут приписаны весовые коэффициенты w_1 и $1 - w_1$ соответственно. На втором шаге добавляются два новых независимых случайных значения w_2 и w_3 . Получится 4 интервала с весовыми коэффициентами $w_1 w_2$, $w_1(1 - w_2)$, $(1 - w_1)w_3$ и $(1 - w_1)(1 - w_3)$. При $n \rightarrow \infty$ мы приходим к предельной мере, являющейся неоднородным фрактальным множеством.

В работе [Kirichenko, 2012] в качестве случайной величины, порождающей весовые коэффициенты рассмотрена случайная величина, имеющую бета-распределение. Бета-распределением с параметрами $a > 0$, $b > 0$, называется распределение с плотностью вероятностей

$$p(x) = \begin{cases} \frac{1}{B(a,b)} (1-x)^{b-1}, & x \in (0,1), \\ 0, & x \notin (0,1) \end{cases},$$

где $B(a,b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$ – бета-функция.

Для бета-распределения с одинаковыми значениями параметров $a = b$, у которого функция плотности распределения симметрична, можно аналитически определить скейлинговую экспоненту $\tau(q)$ [Riedy, 2002]:

$$\tau(q) = -\log_2 \frac{\text{Beta}(\alpha+q, \alpha)}{\text{Beta}(\alpha, \alpha)} - 1. \quad (1.30)$$

На рис. 1.4 (а) приведены различные виды графиков плотности распределения вероятностей, для симметричного бета-распределения при значениях $a = \{0.5, 1, 1.5, 3\}$. При значениях параметров $a = b = 1$ мы получаем случайную величину, имеющую равномерное распределение на интервале $[0, 1]$. На рис. 1.4 (б) представлены графики скейлинговых экспонент для соответствующих значений параметра a симметричного бета-распределения.

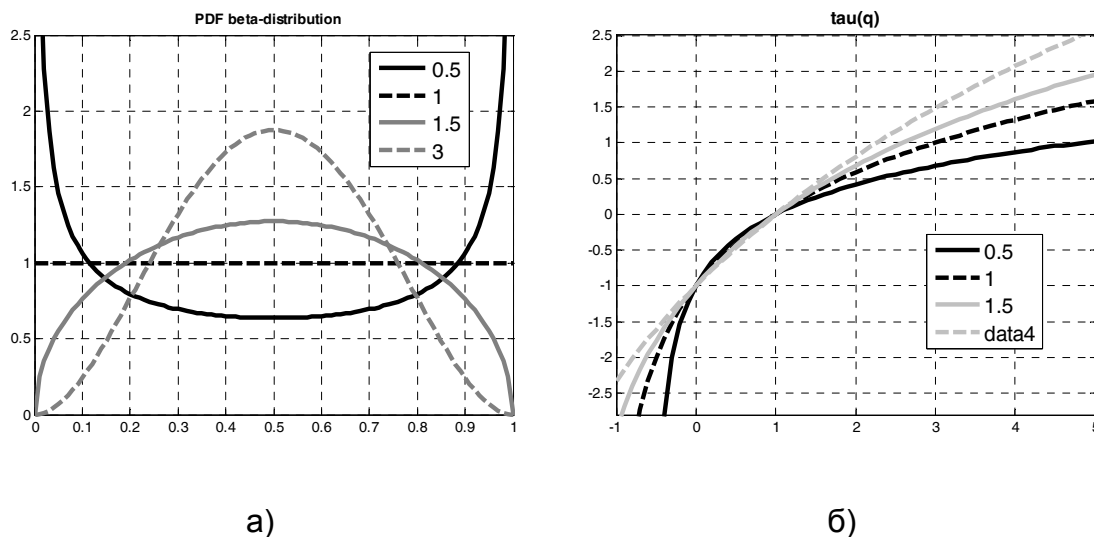


Рисунок 1.4 – Плотности распределения (а) и скейлинговые экспоненты $\tau(q)$ для разных значений параметра a симметричного бета-распределения

Очевидно, что с увеличением значения параметра a происходит ослабление мультифрактальных свойств временного ряда. На рис. 1.5 показаны соответствующие реализации биномиальных каскадов.

В случае симметричного бета-распределения мультифрактальные свойства каскада полностью определяются параметром a . Показатель Херста H , учитывая формулу (1.15), в этом случае равен

$$H = \frac{\tau(2) + 1}{2} = -\log_2 \frac{\text{Beta}(\alpha+q, \alpha)}{2 \text{Beta}(\alpha, \alpha)} .$$

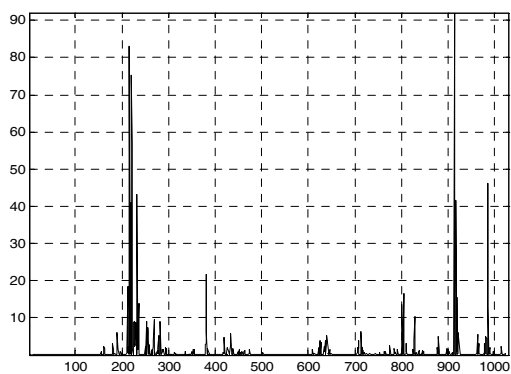
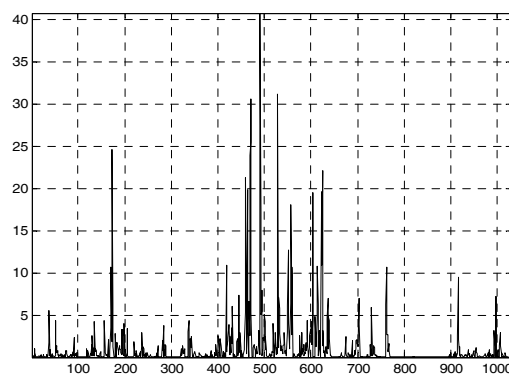
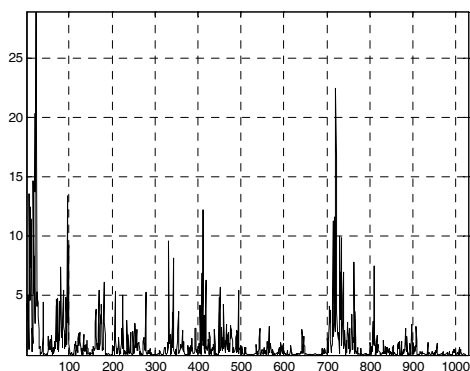
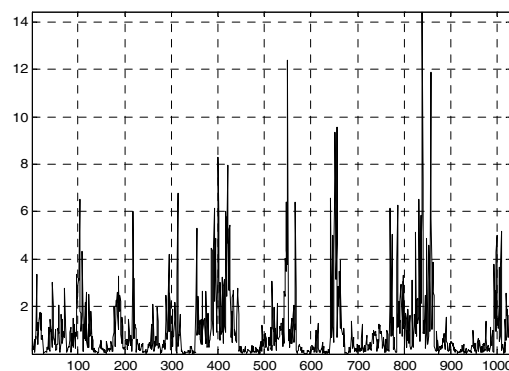
а) $a = 0.5$ б) $a = 1$ в) $a = 1.5$ г) $a = 3$

Рисунок 1.5 – Реализации биномиального каскада для разных значений a

Выводы

1. Многие информационные, биологические, физические, технологические процессы обладают сложной фрактальной структурой. Процессы, обладающие фрактальными свойствами, можно разделить на две группы: самоподобные (монофрактальные) и мультифрактальные. Монофрактальные процессы обладают одним показателем скейлинга – параметром Херста. Мультифрактальные процессы характеризуются спектром скейлинговых показателей, который обычно представлен скейлинговой экспонентой или функцией мультифрактального спектра, связанных между собой преобразованиями Лежандра.

2. Методы оценивания фрактальных характеристик находят все большее практическое применение при анализе случайных процессов, разработке математических моделей и проведении имитационного моделирования в самых разнообразных областях науки, техники, медицины. При оценивании показателя Херста наиболее популярными являются методы нормированного размаха, изменения дисперсии, флуктуационного анализа. При оценивании мультифрактальных характеристик наиболее востребованными являются методы мультифрактального детрендрованного флуктуационного анализа и максимумов модулей непрерывного вейвлет-преобразования.

3. Наиболее адекватным математическим аппаратом для исследования динамики и структуры стохастических фрактальных процессов является вейвлет-анализ. Модификация и дальнейшее развитие вейвлет-фрактальных методов являются важнейшими задачами обработки реализации самоподобных и мультифрактальных процессов.

4. Популярными и не сложными в реализации математическими моделями самоподобных и мультифрактальных процессов являются фрактальное

броуновское движение, ординарное движение Леви и стохастические биномиальные мультипликативные каскады.

Литература

- [Abry, 2009] Abry, P., Goncalves, P., Vehe J.L. Scaling, Fractals and Wavelets. London : John Wiley & Sons, 2009. – 464 p.
- [Borak, 2005] Borak, S., Hardle, W., Weron, R.. Stable distributions. In Statistical tools for finance and insurance. Berlin: Springer, 2005. pp. 21-44.
- [Calvet, 1997] Calvet, L., Fisher, A., Mandelbrot, B. Large deviation and the distribution of price changes. Cowles Foundation Discussion Paper, Yale: University, №1165, 1997. pp. 1–28.
- [Chamber, 1976] Chambers, J.M., Mallows, C.L., Stuck, B.W. A method for simulating stable random variables, Journal of the American Statistical Association, Vol. 71, 1976. pp.340-344.
- [Feder, 1988] Feder, J. Fractals. Plenum, New York, 1988.
- [Jeongy, 1998] Jeongy, H.-D. J., McNickle, D., Pawlikowski, K. A Comparative Study of Generators of Synthetic Self-Similar Teletraffic. Department of Computer Science and Management, University of Canterbury, 1998.
- [Kantelhardt, 2002] Kantelhardt, J.W., Zschiegner, S.A., Bunde, A., Havlin, S., Koscielny-Bunde, E., Stanley, H.E. Multifractal detrended fluctuation analysis of non-stationary time series. Physica A., vol. 316, 2002. pp. 87–114.
- [Kantelhardt, 2008] Kantelhardt J. W. Fractal and Multifractal Time Series. [Электронный ресурс]: 2008.– Режим доступа: <http://arxiv.org/abs/0804.0747>.
- [Kirichenko, 2012]. Kirichenko, L., Radivilova, T., Kayali, E. Modeling telecommunications traffic using the stochastic multifractal cascade process Problems of Computer Intellectualization. Kiev–Sofia: ITHEA, 2012. pp. 55–63.

- [Leland, 1993] Leland, W.E., Taqqu, M. S., Willinger, W., Wilson, D. V. Statistical analysis of high time-resolution Ethernet LAN traffic measurements. *Computing Science and Statistics*, vol. 25, 1993. pp. 146–155.
- [Mandelbrot, 1968] Mandelbrot, B., Van Ness J.W. Fractional Brownian motions, fractional noises and applications. *SIAM Review*, Vol. 10, №4, 1968. pp. 422–437.
- [Mandelbrot, 1974] Mandelbrot, B. Intermittent turbulence in self-similar cascades: divergence of high moments and dimension of the carrier. *Mathematical Sciences Department, IBM Thomas J. Watson Research Center, Yorktown Heights, 1974.*
- [Mandelbrot, 1983] Mandelbrot, B., Freeman, W. H. *The Fractal Geometry of Nature*. San Francisco, 1983.
- [Mandelbrot, 2001] Mandelbrot, B. Scaling in financial prices: II. Multifractals and the star equation. *Quantitative Finance*, Vol. 1, 2001. pp. 124–130.
- [Meyer, 1995] Meyer, Y., Sellan, F., Taqqu, M. S. Wavelets, generalized white noise and fractional integration: the synthesis of fractional Brownian motion. *The Journal of Fourier Analysis and Applications*, Vol. 5, №5, 1995. pp. 465–494.
- [Mittnik, 1999] Mittnik, S., Doganoglu, T., and Chenyao, D. Computing the probability density function of the stable Paretian distribution, *Mathematical and Computer Modelling*, Vol. 29, 1999. pp.235-240
- [Molchan, 1996] Molchan, G. M. Scaling Exponents and Multifractal Dimensions for Independent Random Cascades. *Commun. Math. Phys.*, Vol.179, 1996. pp.681–702.
- [Riedi, 2002] Riedi, Rudolf H. Multifractal processes, in Doukhan P., Oppenheim G., Taqqu M.S. (Eds.), *Long Range Dependence: Theory and Applications*, 2002. pp. 625–715.

- [Sellan, 1995] Sellan, F. Synthèse de mouvements browniens fractionnaires à l'aide de la transformation par ondelettes. Comptes Rendus de l'Académie des Sciences de Paris, Série I, Vol. 321, 1995. pp.351–358.
- [Taqqu, 1995] Taqqu, M.S., Teverovsky, V., Willinger, W. Estimators for long-range dependence: an empirical study. Fractals, № 3(4), 1995. pp. 785–798.
- [Taqqu, 1997] Taqqu, M., Willinger, M., Sherman, R. Proof of a Fundamental Result in Self-Similar Traffic Modeling. IEEE/ACM Transactions on Networking, № 5(2), 1997.
- [Willinger, 1997] Willinger W., Taqqu M. S., Sherman, R., and Wilson, D.V. Self-Similarity through high-variability: statistical analysis of ethernet LAN traffic at the source level. IEEE/ACM Transactions on Networking, № 5(1), 1997. pp. 43–62.
- [Божокин, 2001] Божокин, С.В., Паршин, Д.А. Фракталы и мультифракталы. Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001.
- [Кириченко, 2009] Кириченко, Л.О., Дейнеко, Ж.В. Оценивание самоподобия стохастического временного ряда методом вейвлет-анализа. Радіоелектронні і комп'ютерні системи, № 4 (38), 2009. сс. 99–105.
- [Кроновер, 2000] Кроновер, Р.М. Фракталы и хаос в динамических системах. Основы теории, М.: Постмаркет, 2000. 352 с.
- [Соболев, 2012] Соболев, В.І., Кіріченко, Л.О., Мережкін, Д.Г. Моделювання часових рядів за допомогою α -стійкого процесу Леві. Відбір і обробка інформації, №37 (113), 2012. сс.19-25.
- [Шелухин, 2008] Шелухин, О.И., Осин, А.В., Смольский, С.М. Самоподобие и фракталы. Телекоммуникационные приложения, М. : Физматлит, 2008. 368 с.

Информация об авторах



Людмила Кириченко – д.т.н., профессор Харьковского национального университета радиоэлектроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: lyudmyla.kirichenko@nure.ua.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, фрактальный анализ, вейвлет-анализ, детерминированные хаотические системы.



Тамара Радивилова – к.т.н., доцент Харьковского национального университета радиоэлектроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: tamara.radivilova@gmail.com.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, телекоммуникационные системы, управление трафиком, информационная безопасность.

ОЦЕНИВАНИЕ ПАРАМЕТРА САМОПОДОБИЯ ДЛЯ СТАЦИОНАРНЫХ СТОХАСТИЧЕСКИХ ПРОЦЕССОВ

Людмила Кириченко, Тамара Радивилова

На практике оценивание показателя Херста является достаточно сложной задачей, поскольку в реальных условиях для такого оценивания необходимо использовать конечные наборы данных, механизм генерации которых зачастую неизвестен. В данной статье приведены результаты сравнительного анализа статистических характеристик оценок, полученных разными методами по стационарным модельным данным малой длины.

В общем случае оценку показателя Херста можно представить следующим образом:

$$\hat{H}_{meth} = f(method, parameters, N),$$

где *method* – метод оценивания показателя Херста; *parameters* – параметры данного метода; *N* – длина временного ряда.

2.1 Метод нормированного размаха

Данный метод был предложен Г.Херстом для решения задач, связанных с накоплением водных ресурсов [Hurst, 1951; Hurst, 1955] и получил дальнейшее развитие в работах Б.Мандельброта [Mandelbrot, 1969; Mandelbrot, 1972; Mandelbrot, 1979], В.Феллера [Feller, 1951], Е.Федера [Feder, 1988], Э.Петерса [Петерс, 2004]. В настоящее время метод нормированного размаха является самым известным и популярным методом фрактального анализа [Clegg, 2006; Willinger 1996]. Он широко используется в телекоммуникационных технологиях, при исследовании самоподобных свойств информационных трафиков [Шелухин, 2008;

Шелухин, 2011], в изучении динамики финансовых рынков [Петерс, 2004; Ширяев, 1998], исследованиях геофизических и радио сигналов [Feder, 1988; Доля, 2006], и т.д..

Оценивание параметра Херста методом нормированного размаха основано на том, что для случайного процесса, обладающего долгосрочной зависимостью (1.7) при больших значениях τ выполняется следующее соотношение:

$$M \left[\frac{R(\tau)}{S(\tau)} \right] = (c \cdot \tau)^H, \quad (2.1)$$

где c – некоторая постоянная величина;

$R(\tau)$ – размах кумулятивного ряда $x^{cum}(t, \tau)$;

$S(\tau)$ – среднее квадратичное отклонение исходного ряда.

Построим пошаговый алгоритм нахождения H по временному ряду.

1. Для расчета необходимо преобразовать исходный ряд $X(t)$, $t = 1, \dots, N$, путем нахождения первой разности: $x(t) = X(t+1) - X(t)$. Если исходный ряд уже представляет собой ряд приращений, данный шаг пропускается.

2. Полученный ряд делится на A смежных подпериодов длины τ , так что $A \cdot \tau = 2^N$. Обозначим каждый подпериод I_a с учетом того, что $a = 1, 2, 3, \dots, A$. Для каждого I_a длины τ определяется среднее значение:

$$\bar{x}(\tau) = \frac{1}{\tau} \sum_{t=1}^{\tau} x(t).$$

3. Временной ряд накопленных отклонений от среднего значения $x^{cum}(t, \tau)$ для каждого подпериода I_a определяется как:

$$x^{cum}(t, \tau) = \sum_{i=1}^t x(i) - \bar{x}(\tau), \quad t = \overline{1, \tau}.$$

4. Для каждого подпериода I_a определяется размах кумулятивного ряда

$$R(\tau) = \max(x^{cum}(t, \tau)) - \min(x^{cum}(t, \tau)), \quad t = \overline{1, \tau},$$

и выборочное стандартное отклонение ряда приращений $S(\tau) = \sqrt{\frac{1}{\tau-1} \sum_{t=1}^{\tau} (x(t) - \bar{x})^2}$.

5. Нормированный размах в течение каждого I_a подпериода равен $\frac{R(\tau)}{S(\tau)}$.

Среднее значение R/S для длины τ определяется как

$$\frac{R}{S}(\tau) = (1/A) \sum_{a=1}^A \frac{R_a(\tau)}{S_a(\tau)}.$$

6. Длина τ увеличивается до следующего более высокого значения. Используются значения τ , включающие начальные и конечные точки временного ряда, и шаги 1-6 повторяются до $\tau = N$.

7. Строится линейная регрессия методом наименьших квадратов на $\log(\tau)$, как независимой переменной, и $\log(R/S)$, как зависимой переменной (рис. 2.1). Тангенс угла наклона прямой является оценкой показателя Херста H .

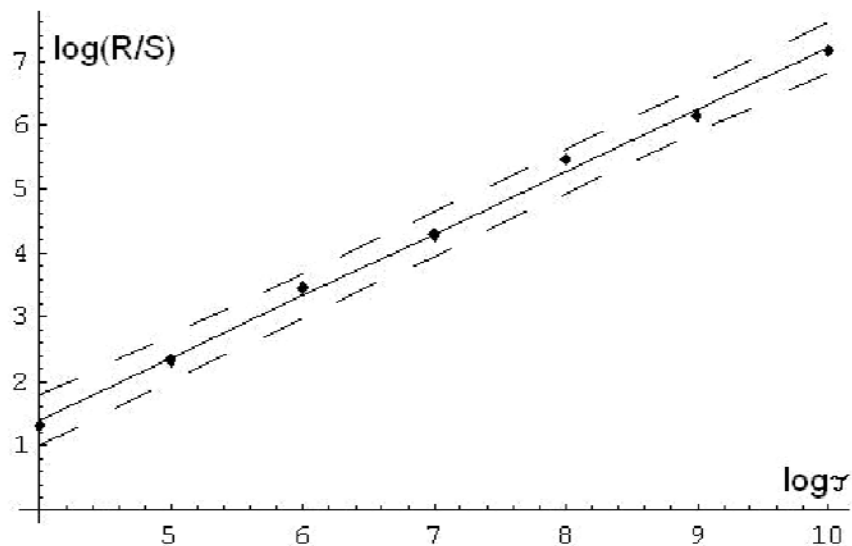


Рисунок 2.1 – Зависимость $\log(R/S)$ от $\log(\tau)$, полученная методом нормированного размаха

2.2 Метод изменения дисперсии

Метод изменения дисперсии агрегированного ряда непосредственно вытекает из определения и свойств самоподобного случайного процесса (1.4) и определения агрегированного временного ряда (1.8). Наибольшее распространение данный метод получил при анализе самоподобных свойств телекоммуникационного трафика [Шелухин, 2008; Шелухин, 2011; Willinger 1996].

Под агрегированием ряда $x(t)$ длиной τ по шкале времени с параметром

m понимается переход к такому процессу $x^{(m)}$, что $x_k^{(m)} = \frac{1}{m} \sum_{t=km-m+1}^{km} x(t)$, $k = \overline{1, \tau/m}$. Оценивание параметра Херста методом изменения дисперсии

основано на том, что для самоподобного процесса $X(t)$ дисперсия агрегированных временных серий x^m при больших значениях m подчиняется следующему соотношению:

$$Var[x^{(m)}] = \frac{Var[x]}{(c \cdot m)^\beta} \quad (2.2)$$

Параметр самоподобия $H = 1 - \beta / 2$.

Построим пошаговый алгоритм нахождения H по временному ряду.

1. Для расчета необходимо преобразовать исходный ряд $X(t)$, $t = 1, \dots, N$, путем нахождения первой разности: $x(t) = X(t+1) - X(t)$. Если исходный ряд уже представляет собой ряд приращений, данный шаг пропускается.

2. Полученный ряд агрегируется по минимальному уровню агрегации ΔT по формуле (1.8). Получается агрегированная эквидистантная реализация, содержащая m элементов, $x^{(m)} = \{x(\Delta T), x(2\Delta T), \dots, x(m\Delta T)\}$.

3. Для полученной выборки вычисляется выборочная дисперсия

$$\text{Var}[x^{(m)}] = \frac{1}{m-1} \sum_{t=1}^m (x(t) - \bar{x})^2 \quad (2.3)$$

4. Уровень агрегации увеличивается вдвое $\Delta T = 2\Delta T$. Для полученной выборки вычисляется выборочная дисперсия, и шаги 1–3 повторяются.

5. Строится линейная регрессия методом наименьших квадратов на $\log(m)$, как независимой переменной, и $\log[\text{Var}(x)]$, как зависимой переменной (рисунок 2.2). Тангенс угла наклона уравнения является оценкой величины β , параметр Херста определяется как $H = 1 - \frac{\beta}{2}$

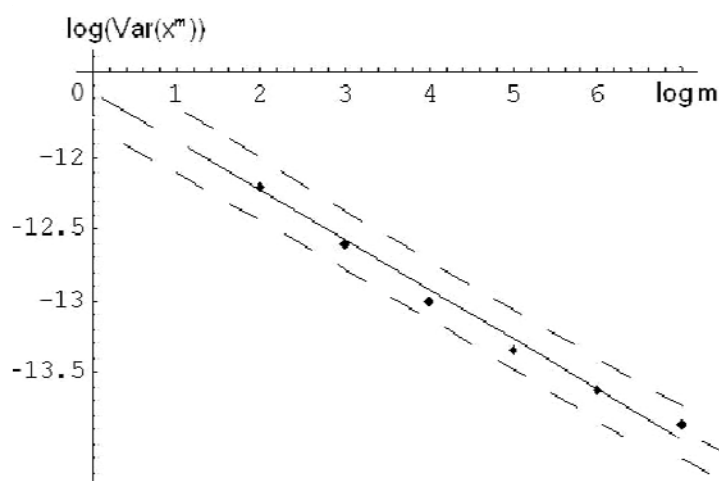


Рисунок 2.2 – Зависимость $\text{Var}[x^{(m)}]$ от $\log(m)$, полученная методом изменения дисперсии

2.3 Метод детрендрованного флуктуационного анализа (ДФА)

Данный метод изначально был предложен для анализа длительных корреляций в структуре сердечного ритма [Peng, 1994; Penzel, 2007]. В настоящее время является одним из наиболее широко используемых методов исследования длительных корреляций в нестационарных

процессах [Kantelhardt, 2001; 2002; 2008]. В частности, в исследованиях динамики биологических систем по экспериментальным данным [Hoerger, 1997; Gu, 2006], в последние годы ДФА активно используется во фрактальном анализе финансовых рядов для прогнозирования кризисных ситуаций [Cont, 2004; Ganchuk, 2006; Ramires, 2008].

Метод ДФА основан на скейлинговом соотношении флуктуационной функции $F(\tau)$ случайного процесса, обладающего долгосрочной зависимостью, и длины сегмента ряда τ :

$$F(\tau) = (c \cdot \tau)^H \quad (2.4)$$

Построим пошаговый алгоритм нахождения H по временному ряду.

1. Для расчета необходимо преобразовать исходный ряд $X(t)$, $t = 1, \dots, N$, путем нахождения первой разности: $x(t) = X(t+1) - X(t)$. Если исходный ряд уже представляет собой ряд приращений, данный шаг пропускается. Ряд $X(t)$ желательно привести к нулевому среднему: $X^*(t) = X(t) - \bar{X}(t)$.

2. Строится кумулятивный ряд $Y^{cum}(k) = \sum_{t=1}^k X^*(t)$.

3. Ряд $Y^{cum}(t)$ разбивается на N неперекрывающихся сегментов длиной τ и для каждого сегмента $Y^{cum}(t)$ вычисляется флуктуационная функция

$$F^2(\tau) = \frac{1}{\tau} \sum_{t=1}^{\tau} (Y^{cum}(t) - Y_m(t))^2, \quad (2.5)$$

где $Y_m(t)$ – локальный m -полиномиальный тренд в пределах данного сегмента.

4. Функция $F(\tau)$ усредняется по всему ряду $Y^{cum}(t)$:

$$F(\tau) = \sqrt{\frac{1}{N} \sum_{i=1}^N F_i^2(\tau)}. \quad (2.6)$$

6. Значение τ увеличивается и шаги 3–4 повторяются. Следует отметить, что максимальная длина сегмента не должна превышать значения $N/4$, иначе построение будет содержать большие погрешности [Kantelhardt, 2001; Павлов, 2007; Павлов, 2008].

7. Строится линейная регрессия методом наименьших квадратов на $\log(\tau)$, как независимой переменной, и $\log F(\tau)$, как зависимой переменной (рисунок 2.3). Тангенс угла наклона прямой является оценкой показателя Херста.

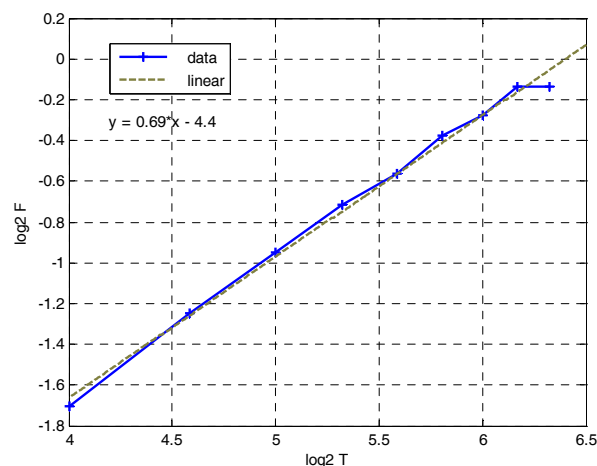


Рисунок 2.3 – Зависимость $\log F(\tau)$ от $\log(\tau)$, полученная методом ДФА

2.4 Метод с использованием дискретного вейвлет-преобразования

2.4.1 Дискретное вейвлет-преобразование одномерного сигнала

Вейвлет-преобразование одномерного сигнала – это его представление в виде обобщенного ряда или интеграла по системе базисных функций

$\psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$, полученных из материнского вейвлета $\psi(t)$,

обладающего определенными свойствами за счет операций сдвига во времени b и изменения временного масштаба a . Дискретное вейвлет-преобразование (ДВП) строится с помощью кратномасштабного анализа, основная идея которого состоит в представлении сигнала в виде совокупности его последовательных приближений [Смоленцев, 2005; Малла, 2005].

Кратномасштабный анализ заключается в разбиении исследуемого сигнала $X(t)$ на две составляющие: аппроксимирующую и детализирующую, с последующим аналогичным дроблением аппроксимирующей до заданного уровня декомпозиции сигнала N . В результате декомпозиции сигнал $X(t)$ представляется в виде суммы аппроксимирующей компоненты $\text{approx}_N(t)$ и детализирующих компонент $\text{detail}_j(t)$:

$$X(t) = \text{approx}_N(t) + \sum_{j=1}^N \text{detail}_j(t) = \sum_{k=1}^{N_a} \text{apr}(N, k) \phi_{Nk}(t) + \sum_{j=1}^N \sum_{k=1}^{N_j} \text{det}(j, k) \psi_{jk}(t),$$

где ψ – материнский вейвлет, ϕ – соответствующая ему скейлинг-функция, N – выбранный максимальный уровень разложения,

$\text{apr}(N, k) = \int_{-\infty}^{\infty} X(t) \phi_{Nk}(t) dt$ – аппроксимирующие вейвлет-коэффициенты

уровня N , $\text{det}(j, k) = \int_{-\infty}^{\infty} X(t) \psi_{jk}(t) dt$ – детализирующие вейвлет-

коэффициенты уровня j , N_j – количество детализирующих коэффициентов на уровне j , N_a – количество аппроксимирующих коэффициентов на уровне N .

Величина вейвлет-энергии на заданном уровне вейвлет-разложения j

определяется как $E_j = \frac{1}{N_j} \sum_{k=1}^{N_j} \det^2(j, k)$. Набор величин E_j для каждого

уровня разложения составляет спектр вейвлет-энергии ряда. Полная вейвлет-энергия спектра представляет собой сумму энергий каждого

уровня $E_{tot} = \sum_{j=1}^N E_j$. Относительная вейвлет-энергия показывает

распределение энергии по уровням разложения: $p_j = \frac{E_j}{E_{tot}}$.

Одной из важных характеристик вейвлет-функции $\psi(t)$ является наличие

нулевых моментов высокого порядка: $\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0$, $k = \overline{0, p-1}$. Число

нулевых моментов p характеризует способность вейвлет-преобразования анализировать мелкомасштабные флуктуации и особенности высокого порядка, игнорируя регулярные полиномиальные составляющие степени $m < p$.

2.4.2. Свойства вейвлет-коэффициентов самоподобных процессов

Главным инструментом для исследования изучаемых процессов является обработка вейвлет-коэффициентов, полученных на разных масштабах. Полученный в результате ДВП ряд коэффициентов обладает определенными свойствами, позволяющими исследовать поведение стохастических процессов, обладающих свойствами самоподобия. Анализируя свойства вейвлет-коэффициентов для самоподобных процессов со стационарными приращениями, можно показать применение свойств вейвлет-коэффициентов для исследований и других процессов с долгосрочной зависимостью [Abry, 2009].

Свойство 1. Если случайный процесс $X(t)$ является самоподобным процессом со стационарными приращениями (1.5), то детализирующие

коэффициенты $d_X(j,k)$, $k = 1 \div N_j$ на каждом уровне разложения j самоподобны, что означает равенство законов распределения для ряда вейвлет-коэффициентов на каждом уровне разложения с некоторым масштабom:

$$\begin{aligned} & \text{Law}\{(d_X(j,0), d_X(j,1), \dots, d_X(j, N_j - 1))\} = \\ & = 2^{j\left(\frac{H+1}{2}\right)} \text{Law}\{(d_X(0,0), d_X(0,1), \dots, d_X(0, N_j - 1))\} \end{aligned} \quad (2.7)$$

где $d_X(j,k)$ – k -й детализирующий коэффициент уровня разложения j ; N_j – количество вейвлет-коэффициентов на соответствующем уровне разложения; H – показатель Херста.

Это свойство детализирующих коэффициентов вытекает из самоподобности процесса, определяемого свойством масштабирования (сжатия/растяжения) самих материнских вейвлетов. Если некоторый процесс является самоподобным, т.е. $X(2^j u) \doteq 2^{jH} X(u)$, то:

$$\begin{aligned} d_X(j,k) &= \int X(u) \psi(2^j u - k) 2^{j/2} du = \int 2^{-j/2} X(2^{-j} u) \psi(u - k) du \doteq \\ &\doteq 2^{-j\left(\frac{H+1}{2}\right)} \int X(u) \psi(u - k) du = 2^{-j\left(\frac{H+1}{2}\right)} d_X(0,k). \end{aligned}$$

Свойство 2. Вейвлет-коэффициенты, полученные в результате декомпозиции процесса со стационарными приращениями, являются стационарными на каждом масштабе 2^j . Это свойство вытекает из свойства вейвлет-функций $\int \psi(t) dt = 0$, которое гарантирует стационарность коэффициентов для процессов со стационарными приращениями.

Свойство 3. Если существуют моменты p -го порядка, то для вейвлет-коэффициентов, полученных в результате декомпозиции процесса $X(t)$, выполняется следующее равенство:

$$M|d_X(j,k)|^p = M|d_X(0,k)|^p 2^{-jp\left(H+\frac{1}{2}\right)} \quad (2.8)$$

Как следствие, для процессов с конечной дисперсией, что чаще всего встречается на практике, выражение (2.8) преобразуется в следующую форму:

$$M|d_X(j,k)|^2 = M|d_X(0,k)|^2 2^{-j(2H+1)} \quad (2.9)$$

Свойство 4. В случае, если $X(t)$ является самоподобным процессом со стационарными приращениями, корреляционная функция вейвлет-коэффициентов уровня j убывает в соответствии с соотношением:

$$M(d_X(j,k)d_X(j,k+n)) \sim n^{2(H-m)}, \quad n \rightarrow \infty,$$

где m – число нулевых моментов материнского вейвлета ψ .

Свойство 5. Для разных уровней разложения $j_1 \neq j_2$ корреляция детализирующих коэффициентов этих уровней $M(d_X(j_1,k)d_X(j_2,k+n)) = 0$ для всех n .

Свойство 6. Детализирующие коэффициенты ДВП на каждом уровне разложения j имеют нормальное распределение с нулевым математическим ожиданием.

Все перечисленные свойства вейвлет-разложения широко используются для анализа реальных временных рядов на практике.

2.4.3 Метод оценивания показателя Херста на основе дискретного вейвлет-преобразования

Метод оценивания показателя Херста на основе ДВП является относительно новым и не приобрел еще массового применения, поскольку для его корректного использования необходимо владеть математическим

аппаратом вейвлет-анализа. Основное развитие метод получил в работах французских исследователей P. Abry и P. Flandrin [Abry, 1998; Abry, 2000; Abry, 2009; Flandrin, 2002]. В настоящее время применяется при анализе самоподобного трафика [Шелухин, 2008; Шелухин, 2011].

Метод оценивания степени самоподобия H базируется на свойствах 1–6 детализирующих вейвлет-коэффициентов, полученных в результате декомпозиции временного ряда. Построим пошаговый алгоритм нахождения H по временному ряду.

1. Для исходного ряда $X(t)$, $t = 1, 2, \dots, n$ с помощью заданного набора вейвлет-функций выполняем вейвлет-декомпозицию с максимальным уровнем разложения $N_{\max} = \lceil \log_2 n \rceil$; где операция $\lceil \cdot \rceil$ обозначает взятие целой части.

2. На каждом уровне разложения j вычисляем сумму квадратов вейвлет-коэффициентов. Согласно формуле (2.17) показатель Херста находится из скейлингового соотношения

$$\frac{1}{N_j} \sum_{k=1}^{N_j} \det^2(j, k) \propto 2^{(2H+1)j} \quad (2.10)$$

где $\det(j, k)$ – детализирующие вейвлет-коэффициенты уровня j ; N_j – количество детализирующих коэффициентов на уровне j .

3. Каждому номеру уровня разложения j ставится в соответствие

логарифм величины $\frac{1}{N_j} \sum_{k=1}^{N_j} \det^2(j, k)$. Аппроксимируя вычисленные

значения вейвлет-коэффициентов методом наименьших квадратов, получаем линейную зависимость, тангенс угла наклона соответствующей прямой определяет значение оценки показателя Херста.

На рисунке 2.4 представлен типичный спектр вейвлет-энергии самоподобного процесса (вверху) и зависимость $\log_2(E_j)$ от номера уровня разложения j (внизу).

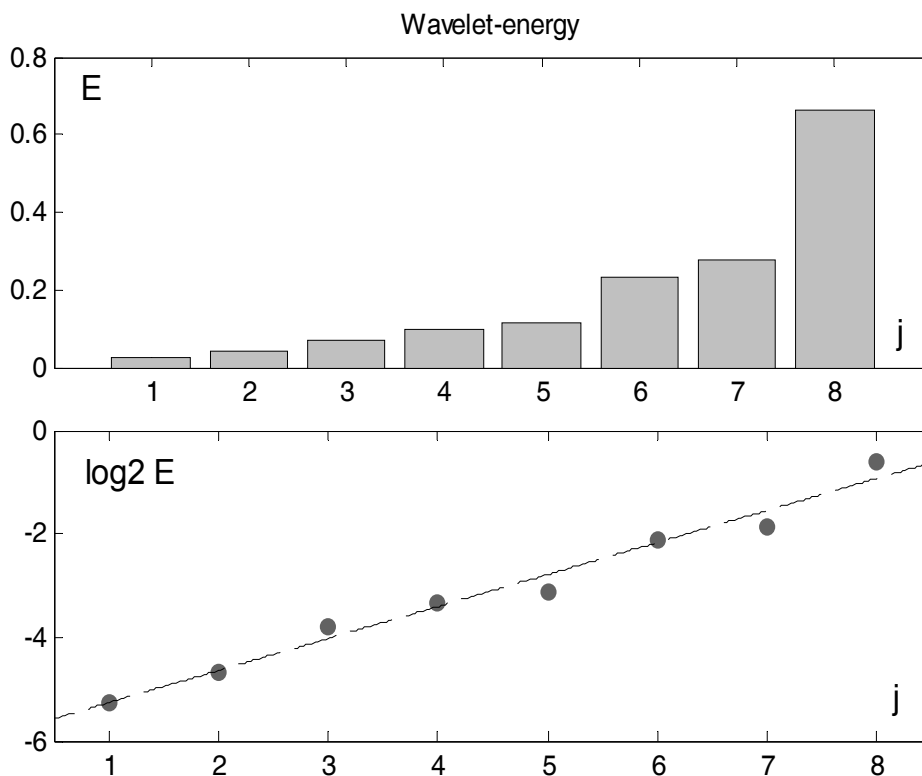


Рисунок 2.4 – Спектр вейвлет-энергии самоподобного процесса (вверху) и зависимость $\log_2(E_j)$ (внизу)

В виду нелинейности, полученной на этапе вычисления линии регрессии, которая образуется из-за функции логарифмирования, происходит смещение математического ожидания оценки показателя Херста. Свойства 5–6 детализирующих коэффициентов вейвлет-разложения, заключающиеся в некоррелируемости и нормальности, позволяют определить величину смещения аналитически и ввести поправочные коэффициенты для каждого уровня детализации j [Abry, 2009]:

$$w_j = (n \log^2 2) / 2^{j+1} \quad (2.11)$$

где w_j – поправочный коэффициент j -го уровня детализации;

n – число значений исследуемого временного ряда.

Тогда, следуя методу наименьших квадратов при линейной аппроксимации, получаем следующее:

$$\sum (w_j y_j - (b_0 + b_1 \cdot j))^2 \rightarrow \min,$$

где j – значение уровня детализации; y_j – логарифм величины

$\frac{1}{N_j} \sum_{k=1}^{N_j} \det^2(j, k)$ j -го уровня; b_0, b_1 – соответствующие коэффициенты в

уравнении аппроксимирующей прямой $b_0 + b_1 x_j$; w_j – поправочный коэффициент j -го уровня детализации. Несмещенная оценка показателя Херста определяется по следующей зависимости:

$$\hat{H}(j_1, j_2) \equiv \frac{1}{2} \left[\frac{\sum_{j=j_1}^{j_2} w_j j E_j - \sum_{j=j_1}^{j_2} w_j j \sum_{j=j_1}^{j_2} w_j E_j}{\sum_{j=j_1}^{j_2} w_j \sum_{j=j_1}^{j_2} w_j j^2 - \left(\sum_{j=j_1}^{j_2} w_j j \right)^2} + 1 \right] \quad (2.12)$$

где j_1, j_2 – уровни разложения, для которых рассчитывается линейная аппроксимация; w_j – поправочные коэффициенты; E_j – средние значения квадратов модулей вейвлет-коэффициентов; j – уровень вейвлет-разложения.

2.5 Статистические характеристики оценок показателя Херста

В данном подразделе приводятся результаты сравнительного анализа статистических характеристик оценок показателя Херста, полученных вышеописанными методами, по модельным фрактальным временным рядам небольшой длины [Kirichenko, 2010].

2.5.1 Описание численного эксперимента

В эксперименте моделировалось фрактальное броуновское движение с заданным показателем H . При этом значения H изменялись во всем возможном диапазоне $0 < H < 1$. Длина реализаций была выбрана равной от 256 до 4096 отсчетов.

Для каждого теоретического значения показателя Херста H были получены массивы оценок и исследованы их статистические характеристики, полученные по множеству различных реализаций. Для каждого сгенерированного временного ряда рассчитывалась оценки методом R/S -анализа ($\hat{H}_{R/S}$), изменения дисперсии агрегированного ряда (\hat{H}_d), детерендированного флуктуационного анализа (\hat{H}_{fa}) и вейвлет-оцениванием с помощью ДВП (\hat{H}_w). Показателями, по которым проводилось сравнение, были:

- среднее значение оценок показателя Херста $\bar{H}_{R/S}$, \bar{H}_d , \bar{H}_{fa} и \bar{H}_w ;
- среднее квадратическое отклонение оценок $S_{R/S}$, S_d , S_{fa} и S_w ;
- коэффициент корреляции между оценками, полученными разными методами;
- выборочный закон распределения оценок показателя Херста.

2.5.2 Сравнительный анализ средних значений и среднеквадратических отклонений оценок параметра Херста

На рисунке 2.5 показана зависимость средних значений оценок параметра Херста, полученных всеми исследуемыми методами, от его теоретического значения. По оси абсцисс отложены задаваемые теоретические значения показателя Херста, а по оси ординат – средние значения соответствующих оценок. Модельные реализации содержали 1024 отсчетов. Сплошная жирная прямая линия соответствует теоретическим значениям H . Очевидно, что средние значения оценок имеют смещение, которое зависит от теоретического значения показателя Херста. Оценки показателя Херста смещены, как в области персистентности, так и для антиперсистентного диапазона. Поскольку подавляющее большинство фрактальных процессов имеют долгосрочную зависимость, дальнейшие результаты, рассмотренные нами, будут приведены только для диапазона $0,5 < H < 1$. Из графика, представленного на рисунке 2.5, видно, что более смещенными являются оценки, полученные методами R/S -анализа и агрегированной дисперсии.

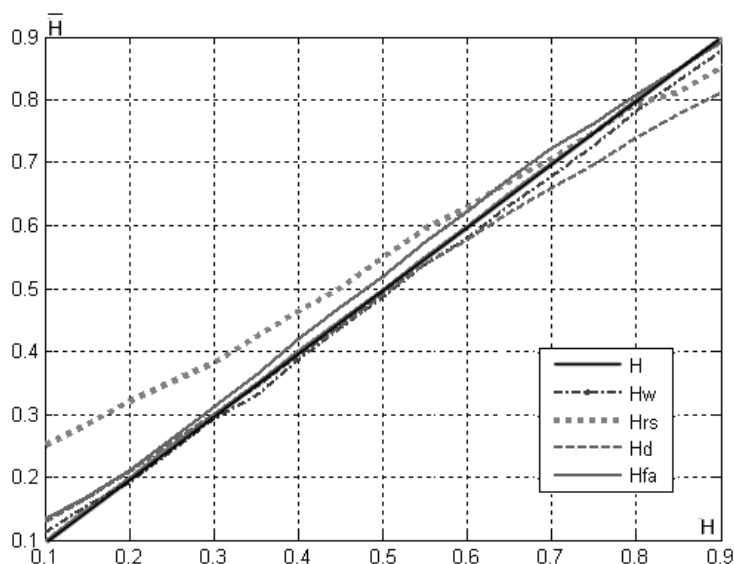


Рисунок 2.5 – Зависимости средних значений оценок показателя Херста, полученных разными методами, от теоретических значений H

Рассмотрим результаты оценивания параметра H методом R/S - анализа. На рисунке 2.6 (а) представлены зависимости средних значений оценок \hat{H}_{rs} , от теоретических значений показателя H для модельных рядов разной длины. Очевидно, что средние значения оценок достаточно хорошо можно аппроксимировать прямыми линиями $\hat{H}_N = k_N H + b_N$, где коэффициенты k_N и b_N зависят от длины реализации N , по которой проводится оценивание. Эти прямые пересекают линию теоретических значений параметра H в районе $H \approx 0,75$; и до этого значения оценки являются завышенными, а после – заниженными. Результаты проведенных исследований подтверждают данные, полученные при анализе оценивания R/S -методом по фрактальным реализациям других моделей, представленных в [Федер, 1991; Шелухин, 2005; Шелухин, 2011]. С возрастанием длины реализации N угол наклона аппроксимированной прямой k_N медленно увеличивается и стремится к теоретическому углу наклона $\pi/4$.

На рисунке 2.6 (б) представлены зависимости средних значений оценок \hat{H}_d , полученных методом изменения дисперсии. Эти зависимости также можно аппроксимировать прямыми $\hat{H}_N = k_N H + b_N$, где коэффициенты k_N и b_N зависят от длины реализации N . В данном случае аппроксимирующие прямые пересекают линию теоретических значений в районе $H \approx 0,5$ и, фактически, на всем диапазоне персистентности значения оценки показателя Херста являются заниженными. Смещение увеличивается с ростом показателя Херста, особенно для значений $H > 0,9$. Можно отметить, что смещения оценок \hat{H}_d больше, чем соответствующие смещения \hat{H}_{rs} . С возрастанием длины реализации N смещение очень медленно уменьшается.

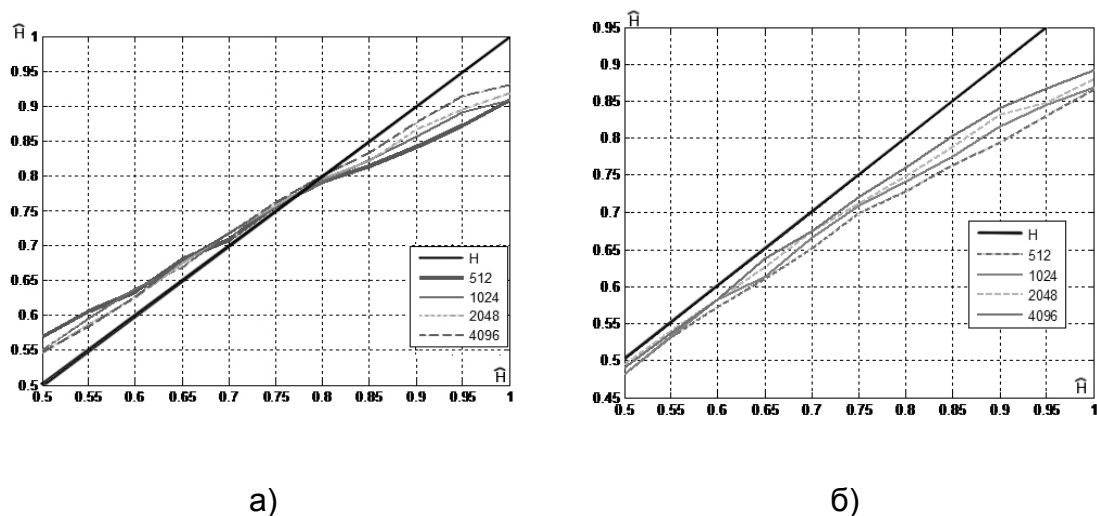


Рисунок 2.6 – Зависимости средних значений оценок от теоретического значения H :

а) метод нормированного размаха; б) метод изменения дисперсии

Метод изменения дисперсии агрегированного ряда в связи с простотой и прозрачной интерпретацией наиболее часто используется при оценивании самоподобия трафика в информационных сетях. Однако для процессов с долгосрочной зависимостью этот метод дает заниженные оценки [Шелухин, 2005; Шелухин, 2008; Шелухин, 2011]. Это может оказаться фатальным, например, при оценивании нагрузки СЕТИ при прохождении самоподобного трафика [Столинс, 2003].

Метод ДФА широко используется при анализе биоэлектрических сигналов. Для оценок \hat{H}_{fa} , полученных методом ДФА, характерно очень небольшое смещение (см. рисунок 2.7(а)) даже для реализаций небольшой длины. Это смещение меняет знак и увеличивается для значений $H > 0.9$. Стоит отметить, что большинство природных и информационных фрактальных процессов имеют степень самоподобия, меньшую 0.9.

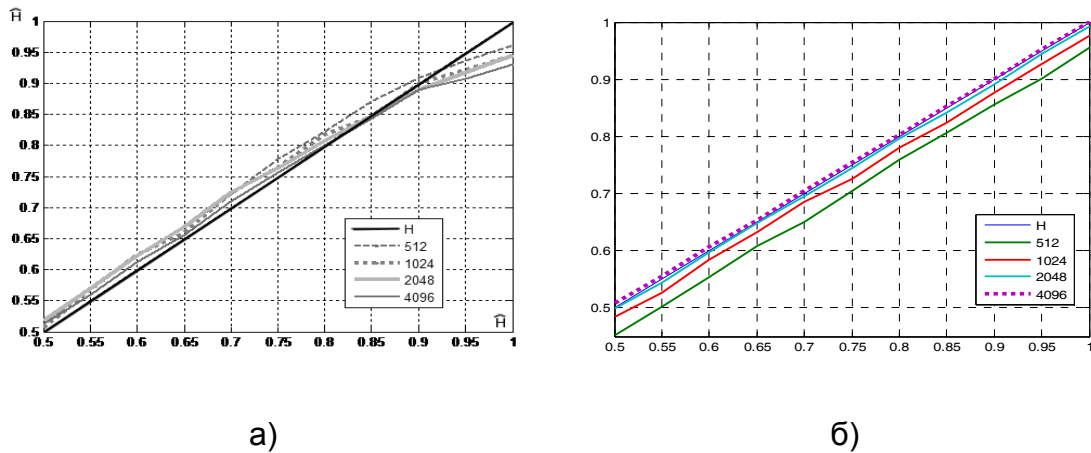


Рисунок 2.7 – Зависимости средних значений оценок от теоретического значения H :

а) метод флуктуационного анализа; б) метод вейвлет-оценивания

Вейвлет-оценивание проводилось на основе материнского вейвлета db4. Как показано в работе [Кириченко, 2009], использование этой функции дает наименьшее смещение и среднеквадратичное отклонение при оценивании показателя Херста. При сравнении основных методов оценивания параметра Херста с методом, основанным на ДВП, очевидны его преимущества по сравнению с другими методами оценки. При правильном выборе базовых вейвлетов оценки \hat{H}_w являются асимптотически несмещенными, что демонстрирует рисунок 2.7 (б). При увеличении длины ряда N смещение уменьшается, и уже при $N \approx 4000$ отсчетов оно достаточно близко к 0.

Для количественного оценивания смещения оценки \hat{H}_{meth} , полученной с помощью определенного метода, от теоретического значения показателя H на всем диапазоне значений, была рассчитана величина смещения Δ^{meth} , которая рассчитывалась по формуле:

$$\Delta^{meth} = \frac{1}{k} \sum_{i=1}^k (\bar{H}_i^{meth} - H_i), \quad (2.13)$$

где \bar{H}_i^{meth} – среднее значение оценки \hat{H}_{meth} , полученной для реализаций с теоретическим показателем H_i с помощью заданного метода расчета показателя Херста.

Для каждого значения H были получены массивы оценок показателя Херста $H_{R/S}$, H_d , H_{fa} и H_w , и рассчитаны средние значения оценок $\bar{H}_{R/S}$, \bar{H}_d , \bar{H}_{fa} и \bar{H}_w , полученные по множеству различных реализаций.

На рисунке 2.8 представлена столбцовая диаграмма распределения смещения оценок показателя Херста для разных методов в зависимости от длины реализации. Очевидно, что наименьшую погрешность при различной длине временной реализации имеет метод, основанный на дискретном вейвлет-преобразовании.

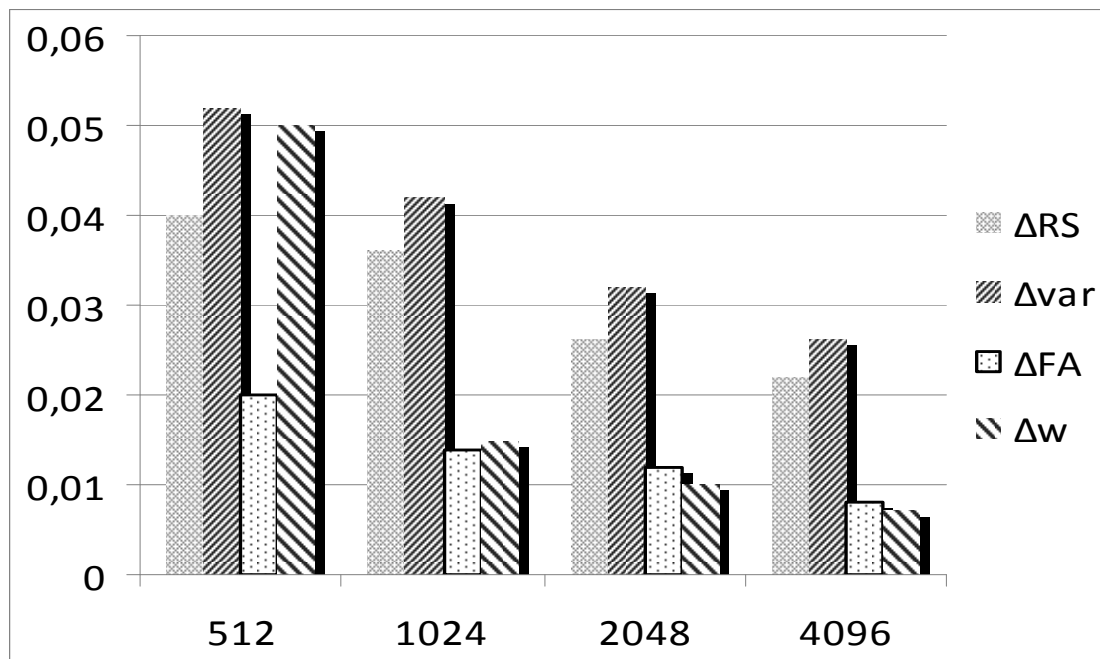


Рисунок 2.8 – Распределение смещения оценок показателя Херста для разных методов

В работах [Кириченко, 2009; Kirichenko, 2010] исследована зависимость средних квадратических отклонений оценок показателя Херста от заданных значений N и длины модельного фрактального ряда для каждого метода. Рисунок 2.9 иллюстрирует, как уменьшаются средние квадратические отклонения, полученные при расчете \hat{H} , при увеличении длины ряда. В данном случае теоретический показатель Херста $H = 0.8$.

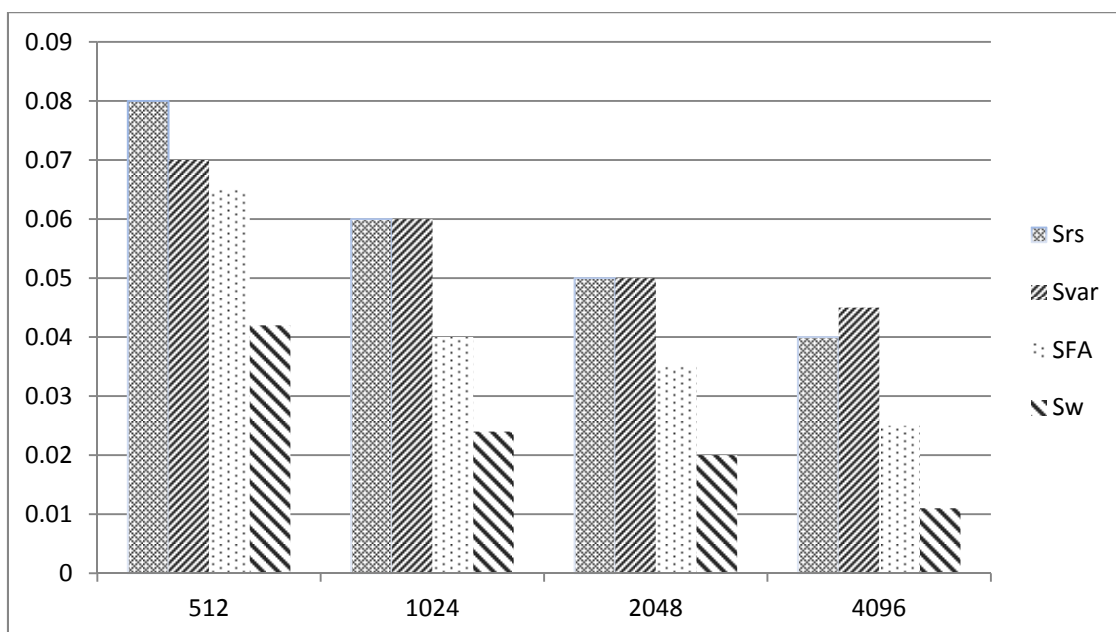


Рисунок 2.9 – Средние квадратические отклонения для разных методов

Исходя из данных столбцовой диаграммы очевидно, что оценки, полученные вейвлет-методом, имеют значительно меньший разброс, чем оценки, полученные другими основными методами оценивания показателя Херста.

2.5.3 Сравнительный анализ выборочных законов распределения и корреляции оценок параметра Херста

Вопрос о законе распределения оценок. был рассмотрен в ряде работ. где было показано, численно или аналитически, что оценки являются

нормальными для конкретного метода или конкретных значений показателя Херста. Аналитические и численные исследования, проведенные в работах [Feller, 1951; Mandelbrot, 1969; Mandelbrot, 1972; Mandelbrot, 1979], показали, что для случая $H=0,5$ оценка показателя Херста, полученная методом нормированного размаха, представляет собой случайную величину с нормальным распределением. В работах [Петерс, 2004] численно показано, что выборочный закон распределения $\hat{H}_{R/S}$ является нормальным. Вопрос о законе распределения оценок \hat{H}_w , полученных методом, основанным на ДВП, рассмотрен в ряде работ [Абгу, 1998; Абгу, 2009], где показано аналитически, что оценки являются нормально распределенными. Нормальное распределение оценок \hat{H}_d , полученных методом изменения дисперсии, численно продемонстрировано в работе [Кириченко, 2005]. Свойства оценок, в том числе выборочное нормальное распределение для некоторых значений показателя Херста, рассмотрены в работах [Kantelhardt, 2001; Kantelhardt, 2002; Kantelhardt, 2008].

В работах [Kirichenko, 2010; Кириченко, 2009] были исследованы законы распределения оценок \hat{H} для каждого метода и разных значений параметра Херста. Для всех рассмотренных методов была выдвинута гипотеза о нормальном распределении выборочных значений оценок с параметрами $N(\bar{H}, S_{\hat{H}})$. Практически для всех выборочных данных гипотеза была принята с уровнем значимости $\alpha=0.05$ по нескольким критериям согласия.

Поскольку выборочное распределение оценок показателя Херста является нормальным, то оценка показателя Херста H для исследуемого временного ряда длины N может быть представлена интервалом значений, внутри которого с заданной вероятностью находится истинное значение оцениваемого показателя H . Таким образом, истинное значение показателя Херста попадает в следующий доверительный интервал:

$$\hat{H} + \Delta - t_{\alpha}S < H < \hat{H} + \Delta + t_{\alpha}S, \quad (2.14)$$

где N – длина исследуемого временного ряда; $method$ – выбранный метод оценивания; $\hat{H} = \hat{H}(N, method)$ – полученное значение оценки показателя Херста по реализации длины N ; $\Delta = \Delta(N, method)$ – величина смещения оценки; $S = S(N, method)$ – среднеквадратическое отклонение, рассчитанное по модельным реализациям длины N ; α – требуемый уровень значимости; t_{α} – квантиль стандартного нормального распределения.

Такое интервальное оценивание особенно необходимо при малой длине временной реализации, когда точечная оценка малонадежна.

В работах [Kirichenko, 2010; Кириченко, 2009] был исследован вопрос о корреляционной связи между оценками, полученными разными методами. Численный анализ корреляционной зависимости между оценками параметра Херста показал, что для всех рассмотренных случаев, кроме корреляции с оценками, полученными с помощью вейвлет-преобразования, выборочные коэффициенты корреляции оказались значимыми и лежащими в диапазоне абсолютных значений в основном меньше 0.5 (таблица 2.1).

Таким образом, исходя из результатов корреляционного, оценки показателя Херста, полученные с помощью разных методов, обладают слабой корреляционной зависимостью, коэффициенты корреляции находятся в диапазоне $[0, 0.5]$.

Таблица 2.1 – Выборочные коэффициенты корреляции для разных методов оценки показателя Херста

Методы, которыми получены оценки показателя Херста	Выборочный коэффициент корреляции r
R/S -анализ и изменение дисперсии	$0.3 \leq r \leq 0.6$
ДФА и R/S -анализ	$0.1 \leq r \leq 0.4$
ДФА и изменение дисперсии	$0.1 \leq r \leq 0.3$
Вейвлет-анализ и остальные методы	Является незначимым

Для увеличения точности оценки показателя Херста можно использовать среднее арифметическое оценок, полученных с помощью нескольких методов оценивания.

2.5.4 ВРЕМЕННЫЕ ЗАТРАТЫ ВЫПОЛНЕНИЯ МЕТОДОВ ОЦЕНИВАНИЯ

Сравнительная оценка эффективности рассмотренных в этой статье методов производилась экспериментальным путем на основании оценок времени решения контрольных задач. В ходе исследования моделировалось фрактальное броуновское движение с заданным показателем Херста H .

Для каждого сгенерированного ряда соответствующим методом определялось время решения задачи в цикле. Временные затраты исследуемых методов для модельных реализаций длиной 4096 отсчетов для цикла числом 100 представлены в таблице 2.2.

Таблица 2.2 – Временные затраты методов оценивания показателя Херста

Метод	RS-анализ	Изменение дисперсии	ДФА	ДВП
Время	9.7969	10.6081	138.6693	3.6036

Данные показывают, что метод, основанный на применении ДВП, является самым быстродействующим. Это связано с тем, что в нем использован алгоритм быстрого вейвлет-преобразования. Наибольшее время счета алгоритма ДФА во многом связано с тем, что размер сегмента изменяется не как степень двойки, что используется в методе нормированного размаха и изменения дисперсии, а выбирается некоторое произвольное фиксированное значение. Увеличив шаг изменения длины сегмента, можно добиться существенного снижения времени счета, однако это влечет за собой снижение точности оценивания.

С увеличением длины модельной реализации время оценивания показателя Херста разными методами увеличивается. На рисунке 2.9 показано изменение времени расчета для разных методов в зависимости от длины фрактального временного ряда. Как видно из представленного графика для метода оценивания показателя Херста с помощью ДФА-метода время счета с увеличением длины реализации резко возрастает. Это связано с вышеуказанными причинами.

Используемый алгоритм быстрого вейвлет-преобразования позволяет с минимальными затратами по времени и объему информации применять метод фрактального анализа в различных технических и информационных системах в режиме реального времени.

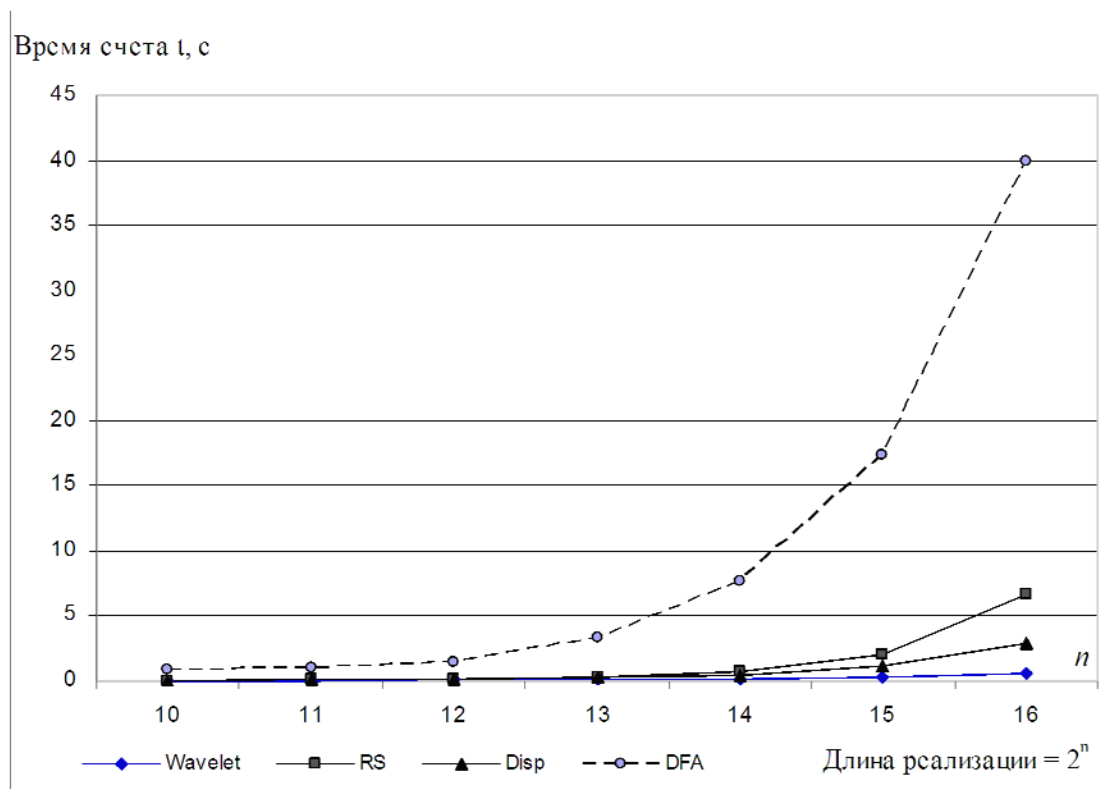


Рисунок 2.9 – Изменение временных затрат для методов оценки показателя Херста при увеличении длины выборки

Выводы

1. Приведены пошаговые алгоритмы основных методов оценивания показателя Херста: метод нормированного размаха, изменения дисперсии агрегированного ряда, детрендированного флуктуационного анализа и вейвлет-оценивания.
2. Проведен сравнительный анализ статистических характеристик оценок показателя Херста, полученных разными методами оценивания. Результаты исследования показали, что оценки являются смещенными. Для каждого метода смещение зависит от истинного значения степени самоподобия процесса и длины временного ряда. Минимальное смещение имеют оценки, полученные методом флуктуационного анализа и с помощью вейвлет-преобразования. Средние квадратические отклонения

оценок зависят от метода оценивания и уменьшаются с ростом длины ряда. Минимальные средние квадратические отклонения имеют вейвлет-оценки.

3. Результаты численного исследования показали, что оценки показателя Херста, полученные рассмотренными методами оценивания, имеют нормальное распределение и могут быть представлены доверительным интервалом значений, характеристики которого учитывают длину исследуемого временного ряда и метод оценивания.

4. Численный анализ корреляционной зависимости между вейвлет-оценками параметра Херста показал, что оценки показателя Херста, полученные с помощью разных методов, обладают слабой корреляционной зависимостью, а вейвлет-оценки не коррелируют с другими. Таким образом, для увеличения точности оценки показателя Херста можно использовать среднее арифметическое оценок, полученных с помощью нескольких методов оценивания.

5. Сравнительное оценивание быстродействия рассмотренных в этой статье методов показало, что метод, основанный на применении вейвлет-преобразования, является самым быстродействующим.

Литература

[Abry, 1998] Abry, P., Veitch, D. Wavelet analysis of long-range dependent traffic. IEEE/ACM Transactions Information Theory, Vol. № 1 (44), 1998. pp. 2–15.

[Abry, 2000] Abry, P., Flandrin, P., Taqqu, M. S., Veitch, D. Self-Similar Network Traffic and Performance Evaluation. New-York: John Wiley & Sons, 2000. pp. 39–88.

[Abry, 2009] Abry, P., Goncalves P., Veitch, J. L. Scaling, Fractals and Wavelets. London: John Wiley & Sons, 2009. – 464 p.

- [Clegg, 2006] Clegg, R. G. A practical guide to measuring the hurst parameter. International Journal of Simulation. Systems, Science & Technology, 2006, Vol. 7, № 2. pp. 3–14.
- [Cont, 2004] Cont, R., Tankov, P. Financial modelling with jump processes. Chapman & Hall / CRC Press, 2004. - 527 p.
- [Feller, 1951] Feller, W. The asymptotic distribution of the range of sums of independence variables. Annals of mathematics and statistics, France, University of Nancy, 1951. – 306 p.
- [Feder, 1988] Feder, J. Fractals. Plenum, New York, 1988.
- [Flandrin, 2002] Flandrin P. Wavelet Tools for Scaling Processes. Turbulence: measurements and signals: Lecture given at the Summer School, Lyon, 2002. pp. 123–128.
- [Ganchuk, 2006] Ganchuk, A., Derbentsev, V., Soloviev, V. Multifractal Properties of the Ukraine Stock Market. [Электронный ресурс]: 2006. – Режим доступа: <http://arxiv.org/pdf/physics/0608009.pdf>.
- [Gu, 2006] Gu, G. F., Zhou, W. X. Detrended fluctuation analysis for fractals and multifractals in higher dimensions. Phys. Rev. E 74, 2006. pp.61-104.
- [Hoyer, 1997] Hoyer, D., Schmidt, K., Bauer, R., Zwiener, U., Kohler, M., Luthke, B., Eiselt, M. Nonlinear analysis of heart rate and respiratory dynamics. IEEE Engineering in Medicine and Biology Magazine, 1997, № 16. pp. 31–39.
- [Hurst, 1951] Hurst, H. E. The Long-Term Storage Capacity of Reservoirs. Transactions of the American Society of Civil Engineers, 1951, Vol. 116. pp. 770–799.
- [Hurst, 1955] Hurst, H. E. Methods of using long-term storage in reservoirs. Proceedings of the Institution of Civil Engineers, London, 1955, Part 1. pp. 519–577.
- [Kantelhardt, 2001] Kantelhardt, J. W., Koscielny-Bunde, E., Havlin, S. and et al. Detecting long-range correlations with detrended fluctuation analysis.

- Physica A., Statistical Mechanics and its Applications, 2001, Vol. 295, № 3–4. pp. 441–454.
- [Kantelhardt, 2002] Kantelhardt, J.W., Zschiegner, S.A., Bunde, A., Havlin, S., Koscielny-Bunde, E., Stanley, H.E. Multifractal detrended fluctuation analysis of non-stationary time series. Physica A., 2002, № 316. pp. 87–114.
- [Kantelhardt, 2008] Kantelhardt, J.W. Fractal and Multifractal Time Series. [Электронный ресурс]: 2008.– Режим доступа: <http://arxiv.org/abs/0804.0747>.
- [Kirichenko, 2010] Kirichenko, L., Radivilova, T. Comparative analysis of statistical properties of the Hurst exponent estimates obtained by different methods. Information Models of Knowledge / ed. K. Markov, V. Velychko, O. Voloshin. – Kiev–Sofia : ITHEA, 2010. pp. 451–459
- [Mandelbrot, 1969] Mandelbrot, B., Wallis, J. R. Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long-Run Statistical Dependence. Water Resources Research, 1969, Vol. 5. pp. 967–988.
- [Mandelbrot, 1972] Mandelbrot, B. Statistical methodology for non-periodic cycles: from the covariance to R/S analysis. Annals of economic social measurement, № 1, 1972. pp. 2–11.
- [Mandelbrot, 1979] Mandelbrot, B., Taqqu, M. S. Robust R/S analysis of long-run serial correlation. Bulletin of the International Statistical Institute. Proceedings of the 42nd Session of the International Statistical Institute, Manila. Book 2, Vol.48, 1979. pp. 69–104.
- [Peng, 1996] Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., Goldberger, A.L. Mosaic organization of DNA nucleotides. Phys. Rev. E., Vol. 49, 1994. pp. 1685-1689.
- [Penzel, 2007] Penzel, T., Kantelhardt, J.W., Becker, H.F., Peter, J.H., Bunde, A. Detrended Fluctuation Analysis and Spectral Analysis of Heart Rate Variability for Sleep Stage and Sleep Apnea Identification. Computers in Cardiology, 2007. pp. 307–310.
- [Ramires, 2008] Ramires, D., Ceballos, D., Cortez, K., Rodrigues, M. Financial crisis: Theory and Practice. Based on Workshop IAFI at Financial Crisis, 2008.

- [Willinger, 1996] Willinger, W., Taqqu, M. S., Erramilli, A.A. Bibliographical guide to self-similar traffic and performance modeling for modern high-speed network in «Stochastic networks: theory and applications». Oxford: Claredon Press, 1996. – 286 p.
- [Доля, 2006] Доля, Н. Г., Иванов, В. К., Пащенко, Р. Э., Стадник, А. М. и др. Фрактальный анализ процессов, структур и сигналов: Коллективная монография. под ред. Р.Э. Пащенко. Х. : ХООО “НЭО “ЭкоПерспектива”, 2006. – 348 с.
- [Кириченко, 2005] Кириченко, Л. О., Радивилова, Т. А., Синельникова, М. И. Сравнительный анализ методов оценки параметра Херста самоподобных процессов. Системи обробки інформації, Вип. 8(48), 2005. сс.36-40.
- [Кириченко, 2009] Кириченко, Л. О., Дейнеко, Ж. В. Оценивание самоподобия стохастического временного ряда методом вейвлет-анализа. Радіоелектронні і комп'ютерні системи, № 4 (38), 2009. сс. 99–105.
- [Малла, 2005] Малла, С. Вэйвлеты в обработке сигналов. М.: Мир, 2005. – 671 с.
- [Петерс, 2004] Петерс, Э. Фрактальный анализ финансовых рынков. Применение теории хаоса в инвестициях и экономике : пер. с англ. М : Интернет-трейдинг, 2004. – 304 с.
- [Смоленцев, 2005] Смоленцев, Н.К. Основы теории фракталов. Вейвлеты в MATLAB. М.: ДМК Пресс, 2005. – 304 с.
- [Столингс, 2003] Столингс В. Современный компьютерные сети. СПб.: Питер, 2003. – 783 с.
- [Шелухин, 2005] Шелухин, О. И., Тенякшев, А. В., Осин, А. В. Моделирование информационных систем. М. : Сайнс-Пресс, 2005. – 368 с.

[Шелухин, 2008] Шелухин, О. И., Осин, А. В., Смольский, С. М. Самоподобие и фракталы. Телекоммуникационные приложения. М. : Физматлит, 2008. – 368 с.

[Шелухин, 2011] Шелухин, О. И. Мультифракталы. Инфокоммуникационные приложения. М. : Горячая линия – Телеком, 2011. – 576 с.

[Ширяев, 1998] Ширяев, А. Н. Основы стохастической финансовой математики. М. : Фазис, Т. 1: Факты. Модели, 1998. –512 с.

Информация об авторах



Людмила Кириченко – д.т.н., профессор Харьковского национального университета радиоэлектроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: lyudmyla.kirichenko@nure.ua.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, фрактальный анализ, вейвлет-анализ, детерминированные хаотические системы.



Тамара Радивилова – к.т.н., доцент Харьковского национального университета радиоэлектроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: tamara.radivilova@gmail.com.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, телекоммуникационные системы, управление трафиком, информационная безопасность.

ОЦЕНИВАНИЕ ПАРАМЕТРА САМОПОДОБИЯ ДЛЯ НЕСТАЦИОНАРНЫХ СТОХАСТИЧЕСКИХ ПРОЦЕССОВ

Людмила Кириченко, Тамара Радивилова

В общем случае, временной ряд можно рассматривать в виде суммы разнотипных составляющих: системной составляющей функции тренда – средних значений по большим интервалам усреднения; циклических компонент с определенным периодом повторения; локальных особенностей; аномалий разного порядка; флуктуаций более высокого порядка (шумов) вокруг всех перечисленных составляющих. Наличие трендовых и циклических компонент требует применения соответствующих подходов для определения самоподобия исследуемого процесса. Эти подходы рассмотрены в статье.

3.1 Предварительное исследование структуры самоподобного временного ряда на основе R/S-анализа

Согласно общепринятым методам расчета показателя Херста по исследуемому временному ряду получают точечную оценку этого параметра – число, которое характеризует степень долгосрочной зависимости: чем ближе параметр H к 1, тем медленнее при увеличении временной задержки k убывают корреляции между временными отсчетами. Однако возможно показатель Херста рассматривать как функцию числа временных отсчетов $H(n)$. Такой подход возможен при применении метода нормированного размаха, когда интервалы времени τ изменяются малыми приращениями.

3.1.1 Проверка гипотезы о наличии самоподобия

Важнейшим положением, с которого необходимо начинать исследование фрактальной структуры случайного процесса, является проверка гипотезы о наличии самоподобия (долгосрочной зависимости). В качестве нулевой гипотезы обычно постулируется, что приращения случайного процесса имеют независимый характер. Зная результаты для случайного процесса с независимыми приращениями, мы можем сравнить исследуемый процесс с нулевой гипотезой и измерить значимость степени самоподобия.

Г. Херст основывал нулевую гипотезу на биномиальном распределении и подбрасывании монет [Hurst, 1951; Hurst, 1955]. Его результат для случайных блужданий:

$$M \left[\frac{R}{S}(n) \right] = (n \cdot \pi/2)^{0.5}, \quad (3.1)$$

где n – количество наблюдений.

На основе (3.1) и дальнейших модификаций данного результата в работе [Петерс, 2004] было предложено следующее уравнение для математического ожидания нормированного размаха:

$$M \left[\frac{R}{S}(n) \right] = ((n - 0.5)/n)(n \cdot \pi/2)^{-0.5} \sum_{r=1}^{n-1} \sqrt{(n-1)/r}. \quad (3.2)$$

Исследования, проведенные в работах [Feller, 1951; Mandelbrot, 1969; Mandelbrot, 1979; Петерс, 2004], показали, что для случая $H = 0,5$ оценка показателя Херста представляет собой случайную величину с нормальным распределением. Математическое ожидание этой величины определяется формулой (3.2). Дисперсия в общем случае зависит от вида распределения данных, однако численные исследования [Taqqi, 1995; Taqqi, 1996; Федер, 1991; Петерс, 2004] позволяют грубо считать среднее

квадратичное отклонение оценки s для случая $H = 0,5$ равным $s = \sqrt{\frac{1}{N}}$, где N – число значений ряда.

Логарифм функции (3.2) показывает теоретические значения $\log(R/S)$ для случая независимых случайных данных (при $H = 0,5$). При наличии долгосрочной зависимости (персистентности процесса) значения $\log(R/S)$ будут находиться выше, а в случае антиперсистентности – ниже линии, описанной уравнением (3.2).

На рис 3.1 представлены типичные зависимости $\log(R/S)$ от логарифма длины ряда для модельных временных рядов с различными типами памяти для длиной $n = 2048$. Пунктирная линия показывает теоретические значения $\log(R/S)$ для случая независимых случайных данных. Показаны зависимости $\log(R/S)$ для классического броуновского движения (линия 2), фрактального броуновского движения с параметром Херста $H = 0,8$ (линия 1) и ФБД с параметром Херста $H = 0,2$ (линия 3),.

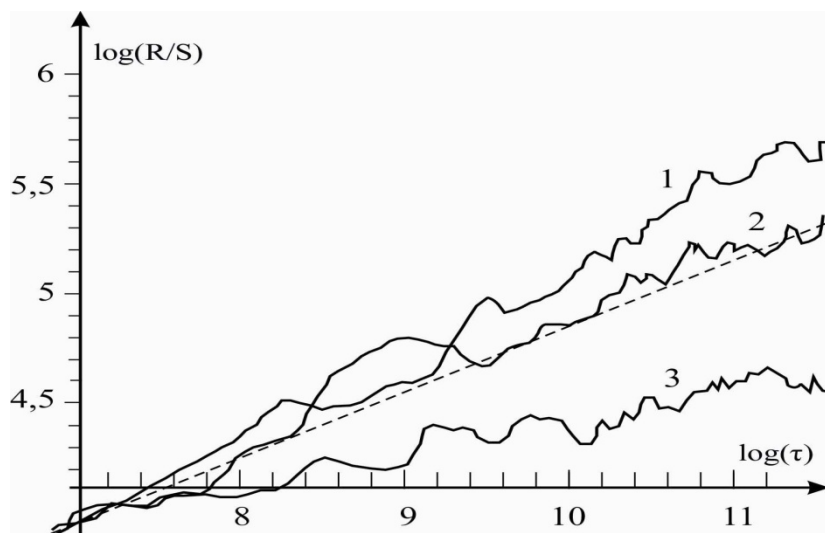


Рисунок 3.1 – Зависимости $\log(R/S)$ для процессов с разными типами памяти

Одним из методов проверки значимости показателя Херста является тест на перемешивание данных [Петер, 2000; Петерс, 2004]. В этом случае порядок наблюдений становится полностью отличным от исходных. Вычисляется показатель Херста для перемешанных данных. Если ряд действительно являлся независимым, то значение H не изменится, поскольку корреляции между наблюдениями не существует. Если же имел место эффект долговременной памяти, то порядок между наблюдениями весьма важен. Перемешивая данные, мы разрушаем структуру системы. Оценка H при этом окажется значительно ниже и будет близка к 0,5. На рисунке 3.2 показаны зависимости $\log(R/S)$ для исходного и перемешанного временного ряда. В первом случае оценка показателя Херста равна 0,57, во втором – 0,48. Отсюда можно сделать вывод о наличии долгосрочной зависимости для исследуемого ряда.

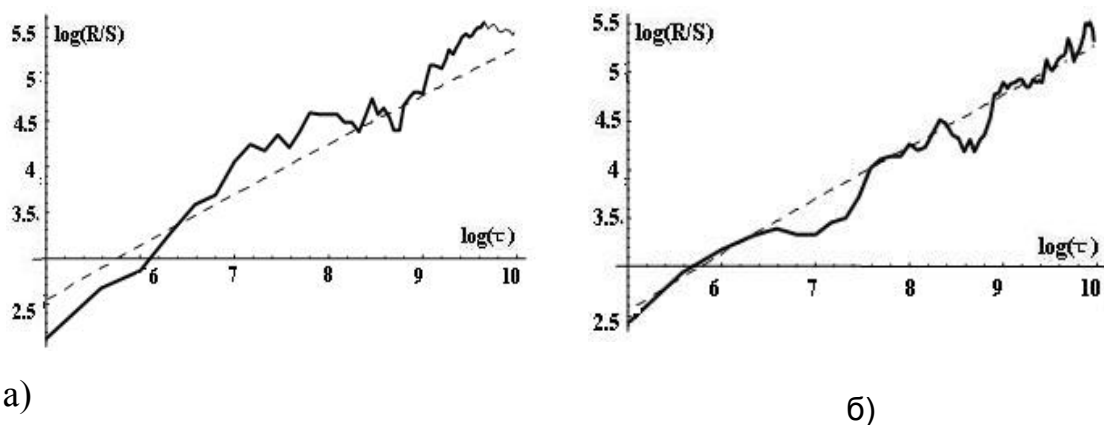


Рисунок 3.2 – Зависимости $\log(R/S)$ в тесте на перемешивание:

а) исходный ряд; б) перемешанный ряд

3.1.2 Определение интервалов долгосрочной зависимости и удаление краткосрочной зависимости

Анализ поведения показателя Херста относительно функции (3.2) позволяет определить значение временного интервала, начиная с которого ряд обладает долгосрочной памятью [Ширяев, 1998; Петерс, 2004]. На рис. 3.3 показаны два интервала зависимости $\log(R/S)$. На временных интервалах меньших момента τ_L значения ряда являются независимыми, а на больших интервалах ряд обладает долгосрочной зависимостью.

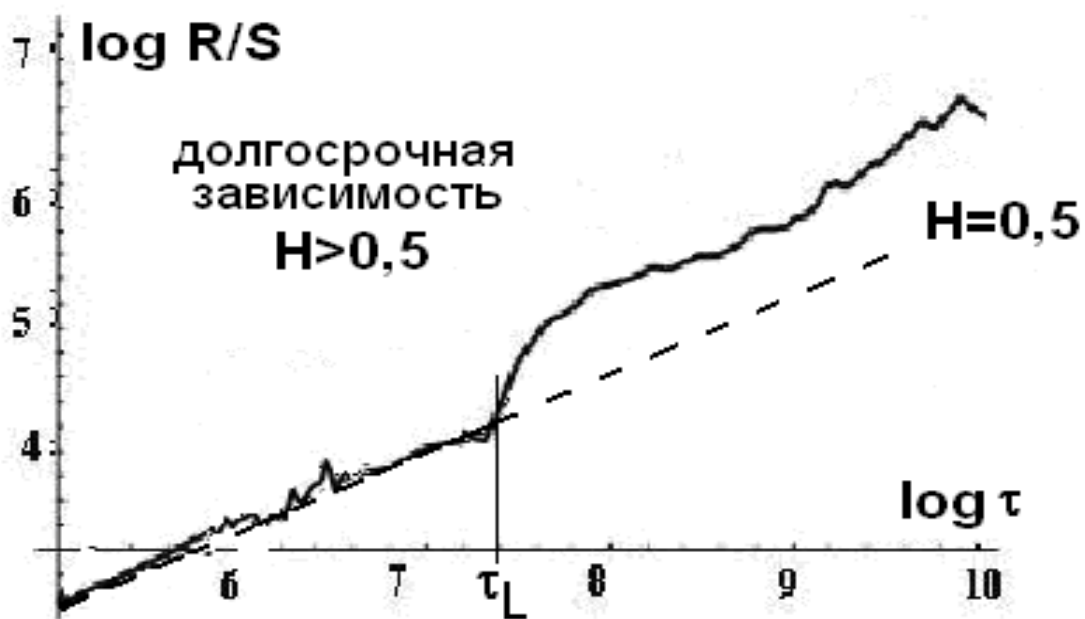


Рисунок 3.3 – Зависимость $\log(R/S)$ для двух интервалов памяти

R/S -анализ позволяет обнаружить и устранить краткосрочную зависимость, характерную для авторегрессионных процессов, в простейшем случае имеющих зависимость:

$$X(t+1) = \phi X(t) + \varepsilon(t),$$

где ϕ – параметр авторегрессии; t – дискретное время; $\varepsilon(t)$ – независимые от $X(t)$ значения случайной величины с нормальным распределением $N(0, \sigma)$.

В подобных случаях линейная зависимость смещает значения показателя Херста и демонстрирует ложную долговременную память, что показано на рисунке 3.4 слева (линия 1).

Поэтому, при выяснении фрактальной структуры временного ряда необходимо сначала выяснить наличие краткосрочной зависимости. Для этого надо значения временного ряда $X(t)$ регрессировать как зависимую переменную против $X(t-1)$ и найти линейную зависимость между ними:

$$X(t) = a + b \cdot X(t-1). \quad (3.3)$$

Значимость коэффициента b свидетельствует о наличии краткосрочной зависимости. Для ее устранения определяется остаток:

$$S(t) = X(t) - (a + b \cdot X(t-1)). \quad (3.4)$$

После этого проводится R/S -анализ остаточного ряда $S(t)$. Если исходный ряд $X(t)$ имел долговременную зависимость, то она сохраняется, в то время, как краткосрочная зависимость устраняется. На рисунке 3.4

показано устранение краткосрочной зависимости для процесса авторегрессии (слева) и ФБД с параметром Херста $H = 0,8$ (справа).

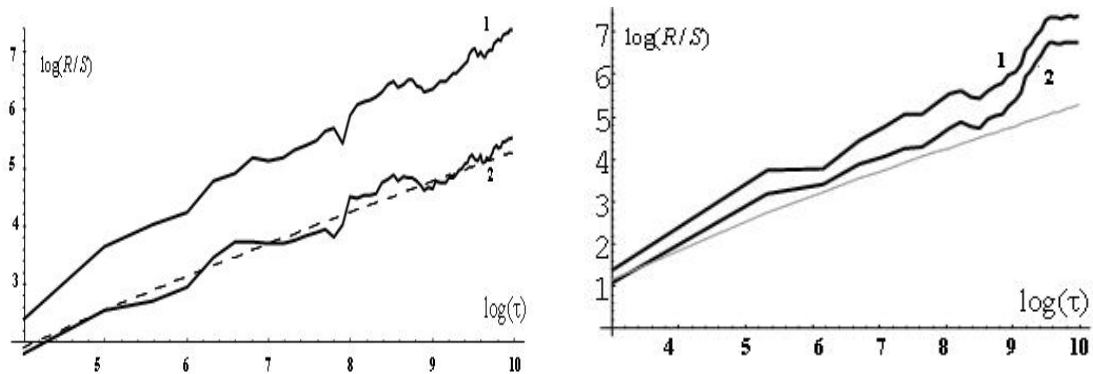


Рисунок 3.4 – Удаление краткосрочной зависимости для процесса авторегрессии (слева) и ФБД с параметром Херста $H = 0.8$ (справа)

Оценка показателя Херста в первом случае была равна 0,73. После удаления краткосрочной зависимости для процесса авторегрессии оценка показателя Херста снизилась до значения 0,52 (линия 2). Оценка показателя Херста для реализации ФБД изменилась незначительно: с 0,83 (линия 1) до 0,81 (линия 2).

3.2 Применение метода детрендированного флуктуационного анализа для временных рядов с трендовыми компонентами

Метод нормированного размаха и другие, не удаляющие тренда методы, хорошо работают в случае стационарных временных рядов. Но если во временном ряде присутствует тренд, эти методы могут дать неверные результаты. Очень часто причины возникновения основных трендов во

временных рядах не известны. ДФА является методом определения скейлингового поведения зашумленных данных при наличии трендов, не зная их происхождение и формы. Метод ДФА традиционно используется для анализа фрактальной структуры и оценивания степени самоподобия временных рядов с трендами (например, реализации энцефалограмм) или кумулятивных рядов с нестационарными приращениями (например, финансовые ряды).

Самым важным этапом выполнения ДФА является устранение локального полиномиального тренда $Y_m(t)$ в каждом сегменте кумулятивного ряда X^{cum} длиной τ (см. рисунок 3.5), содержащего N значений:

$$F_{DFA}^2(\tau) = \frac{1}{N} \sum_{t=1}^N \left(X^{cum}(t) - Y_m(t) \right)^2. \quad (3.5)$$

Флуктуационная функция зависит от степени полинома локального тренда $Y_m(t)$. Каждый раз, изменяя порядок полиномиального тренда m мы получаем другое значение флуктуационной функции $F_{DFA}^{(m)}(\tau)$. Надо отметить, что удаление локального тренда $Y_m(t)$ для кумулятивного ряда X^{cum} соответствует удалению тренда порядка $m-1$ для начального исследуемого ряда.

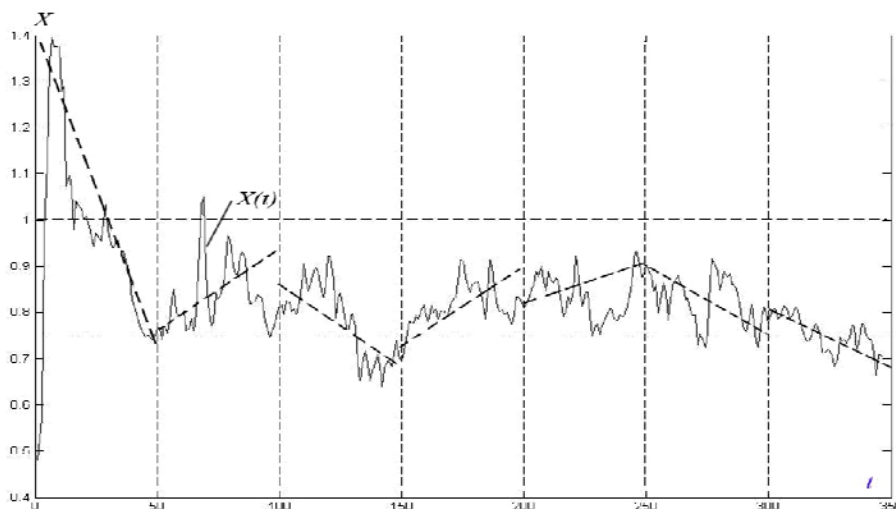


Рисунок 3.5 – Этап построения флуктуационной функции

В данном подразделе представлены результаты исследования влияния выбора степени полинома m локального тренда Y_m на оценку показателя Херста для временных рядов небольшой длины. Для таких рядов имеет существенное значение, что для минимальной длины сегмента τ должно выполняться условие [Kantelhardt 2001]: $\tau \geq m + 2$, в тоже время максимальная длина сегмента не должна превышать значения $N / 4$, где N - длина исследуемого временного ряда.

Для проведения численных исследований в работах использовались модельные временные реализации, представленные суммой фрактального броуновского движения с заданным показателем Херста, и различными трендовыми составляющими: полиномиальными, трансцендентными, синусоидальными. Модельный сигнал можно представить суммой:

$$Y(t) = aT(t) + \varepsilon_H(t), \quad (3.6)$$

где a – коэффициент, регулирующий отношение тренда к шуму; $T(t)$ – тренд; $\varepsilon_H(t)$ – фрактальный гауссовский шум с заданным показателем Херста H .

Проведенные численные исследования для фрактальных рядов с полиномиальными трендовыми составляющими порядка p показали, что адекватное оценивание параметра Херста достигается при использовании локального полиномиального тренда порядка $m > p$. На рисунке 3.6 слева представлены флуктуационные функции $F_{DFA}^{(m)}(\tau)$ порядка $m = 1, 2, 3, 4$ для фрактальных рядов с линейным трендом (теоретическое значение показателя Херста $H = 0.7$).

Исследования [Kirichenko 2011], проведенные для временных реализаций с рациональными полиномами и трансцендентными функциями показали корректное оценивание при использовании локального полиномиального тренда порядка $m \geq 2$. В правой части рис. 3.6 показано изменение

соответствующих флуктуационных функций $F_{DFA}^{(m)}(\tau)$ при изменении порядка локального полиномиального тренда $m=2,3,4,5$ для модельного ряда, который в соответствии с (3.6) имеет выражение $Y(t) = 0.5 * t^{\frac{1}{2}} + \text{fng}(t)$.

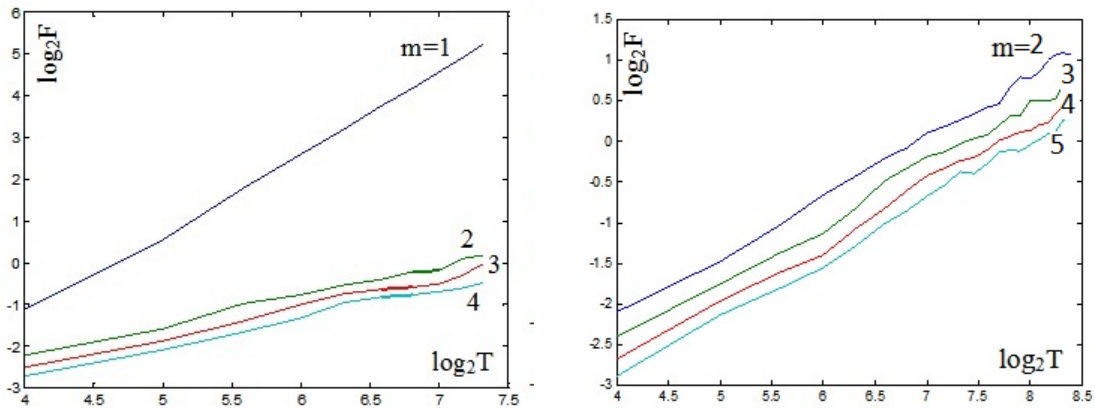


Рисунок 3.6 – Флуктуационные функции данных с линейным (слева) трендом и трендом вида квадратного корня (справа)

Значения оценок показателя Херста, которые получены с помощью флуктуационных функций, для некоторых полиномиальных и рациональных функций трендов, приведены в таблице 3.1.

Очевидно, что корректная оценка параметра H для фрактальных рядов с полиномиальным трендом порядка p достигается при вычитании из кумулятивного ряда локального тренда порядка $m > p$ и не изменяется (в пределах погрешностей), при увеличении порядка локального тренда m . Корректная оценка параметра H для трендов рациональных и трансцендентных функций достигается при вычитании локального тренда порядка $m \geq 2$ в зависимости от вида тренда.

Наибольшую сложность в оценивании степени самоподобия представляют ряды с периодическими и циклическими трендовыми компонентами. На

рисунке 3.7 представлены флуктуационные функции для трендов с синусоидальными составляющими разных периодов.

Таблица 3.1 – Оценка H для полиномиальных и рациональных трендов разного порядка (длина ряда 1024)

Тренд	Порядок локального тренда	Оценка H	Тренд	Порядок локального тренда	Оценка H
t	1	1.935	\sqrt{t}	2	0.784
	2	0.713		3	0.703
	3	0.712		4	0.718
	4	0.691		5	0.689
t^3	2	2.15	$\sqrt[5]{t}$	2	0.693
	3	0.793		3	0.709
	4	0.718		4	0.718
	5	0.723		5	0.688

Численный анализ показал, что чем больше периодов содержит реализация, тем большую степень локального полинома необходимо использовать (см. табл. 3.2).

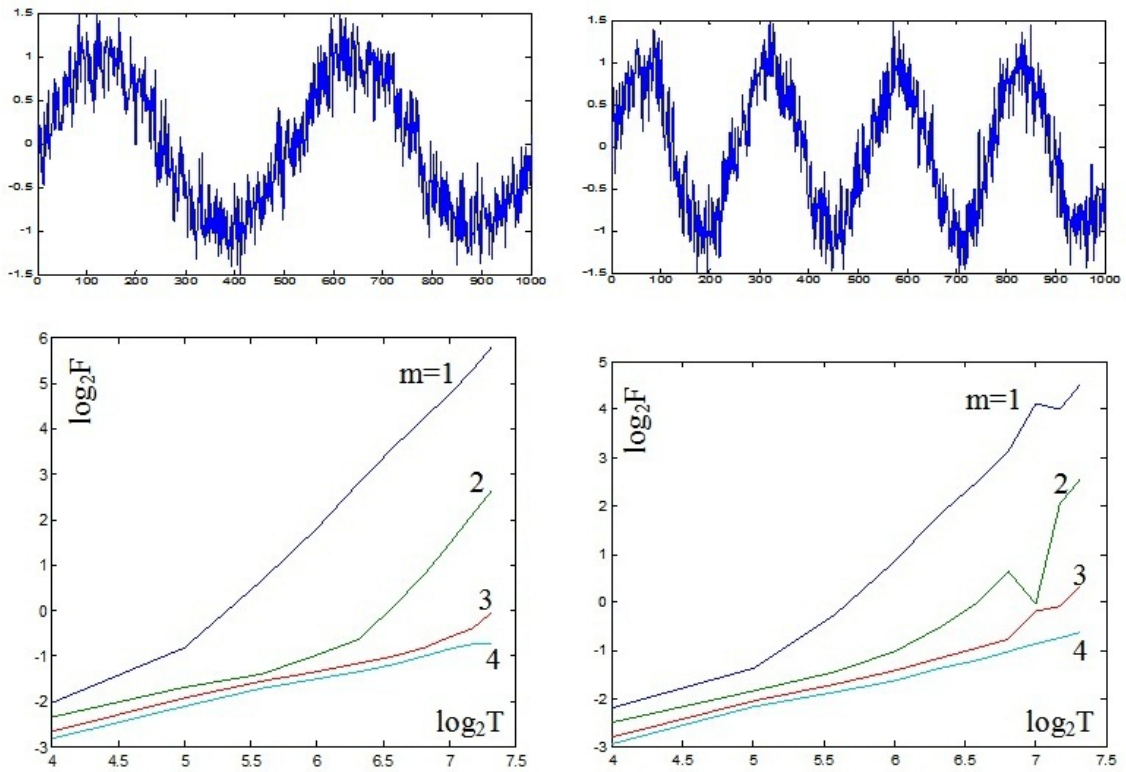


Рисунок 3.7 – Флуктуационные функции для трендов с синусоидальными составляющими

Таблица 3.2 – Оценка H для рядов с разными функциями тренда (длина ряда 1024)

Тренд	Порядок локального тренда	Оценка H	Тренд	Порядок локального тренда	Оценка H
1 период $\sin 2\pi t$	2	2.44	3 периода $\sin 2\pi t$	2	2.67
	3	0.733		3	1.23
	4	0.697		4	0.731
	5	0.701		5	0.7101

Обобщая результаты численного исследования, можно сделать вывод, что при оценивании показателя Херста методом ДФА, необходимо сначала провести прикидочное оценивание, используя локальные полиномы увеличивающейся степени и определить наименьшую степень полинома, начиная с которой оценка показателя Херста перестает изменяться. После этого для полного проведения исследования самоподобия временного ряда необходимо удалять локальный полиномиальный тренд найденной степени.

3.3 Вейвлет-оценивание параметра Херста для рядов с существенными трендовыми и циклическими компонентами

Вейвлет-разложение позволяет эффективно выделять тренды, содержащие сложные циклические составляющие. Для выделения трендовой составляющей анализируемый сигнал декомпозируется с помощью дискретного вейвлет-преобразования (см. раздел 2.4) на заданное число уровней.

3.3.1 Определение диапазона корректного оценивания показателя Херста при наличии трендовых и периодических компонент

Спектр вейвлет-энергии характеризует энергию сигнала на каждом из уровней разложения. Каждый уровень разложения соответствует определенному диапазону частот. Таким образом, коэффициенты этого уровня показывают, насколько мощно представлен этот диапазон частот в сигнале. С помощью вейвлет-разложения можно выделить трендовые и циклические составляющие в исследуемом временном ряде.

Однако удаление тренда из временного ряда в большинстве случаев не дает корректную оценку показателя Херста, поскольку при этом нарушаются корреляционные связи. Во многих случаях нельзя упрощенно разделять случайный процесс или исследуемый временной ряд на регулярную медленно меняющуюся трендовую составляющую и остаток,

который в свою очередь находится под влиянием множества составляющих ряда.

В качестве одного из основных критериев для выбора конкретного вейвлета можно использовать число нулевых моментов вейвлет-функции:

$$\int_{-\infty}^{+\infty} t^k \psi(t) dt = 0, \quad k = \overline{0, m-1}. \quad (3.7)$$

Число нулевых моментов m характеризует способность вейвлета выявлять локальные особенности анализируемого сигнала, поскольку следует, что вейвлет ψ ортогонален любому многочлену степени $m-1$. Обладающие большим числом нулевых моментов вейвлеты позволяют, игнорируя наиболее регулярные полиномиальные составляющие сигнала, анализировать мелкомасштабные флуктуации и особенности высокого порядка. Считается, что если вейвлет имеет нулевые моменты m -ого порядка, то влияние трендовых полиномиальных составляющих до m -ого порядка включительно не проявляется на анализируемом сигнале [Малла, 2005].

В данном подразделе представлены результаты исследований [Дейнеко, 2010; Kirichenko, 2011]. В таблице 3.3 приведены значения оценок показателя Херста, полученные для модельных рядов (3.6) с полиномиальным трендом порядка m . Оценки получены по самоподобным реализациям с $H = 0,7$ длиной 1000 значений с использованием вейвлетов семейства Добеши.

Таблица 3.3 – Оценка показателя H , полученная с помощью вейвлетов с разным числом нулевых моментов

Полиномиальный тренд	H	оценка H				
		db1	db2	db3	db4	db5
		$p = 1$	$p = 2$	$p = 3$	$p = 4$	$p = 5$
$m = 1$	0.7	0.7350	0.7105	0.7127	0.6899	0.7024
$m = 2$	0.7	0.7908	0.7353	0.7013	0.7031	0.7036
$m = 3$	0.7	1.2272	0.9572	0.7344	0.7888	0.7805

Однако корректное оценивание показателя H с применением вейвлетов, имеющих нулевые моменты высокого порядка, возможно только при небольшом значении отношения полинома к фрактальной составляющей. Величиной, характеризующей соотношение тренда и фрактального шума, является величина:

$$Ratio = \frac{S_{trend}}{S_{noise}}, \quad (3.8)$$

где S_{trend} – среднеквадратическое отклонение тренда; S_{noise} – среднеквадратическое отклонение фрактального процесса.

Результаты исследований модельных сигналов, содержащих полиномиальные тренды разных порядков, а также циклическую составляющую, с использованием материнских вейвлетов с разным числом нулевых моментов демонстрирует таблица 3.4.

Таблица 3.4 – Оценивание с помощью разных вейвлетов для сигналов с трендом

Тренд	<i>Ratio</i>	оценка <i>H</i>				
		db1	db2	db4	db8	db10
$0.01t^2$	0.7107	0.7030	0.7305	0.7127	0.6899	0.7024
$0.02t^2$	1.4389	0.7908	0.7353	0.7013	0.7031	0.7036
$0.5t^2$	36.4223	1.2272	0.9572	0.7344	0.7888	0.7805
$0.005t^3$	2.7621	0.8010	0.7604	0.7194	0.7058	0.7045
$0.01t^3$	5.3329	0.8791	0.8042	0.7128	0.7066	0.7137
$0.02t^3$	10.8966	0.9623	0.8748	0.7398	0.7193	0.7313
$1.5 \sin(0.5\pi t)$	3.8520	0.9768	0.9720	0.7493	0.7635	0.7692
$5.5 \sin(0.5\pi t)$	14.2869	1.1938	1.0142	0.7841	0.7851	0.7978

Из таблицы видно, что при небольшом отношении тренда к шуму, оценки показателя Херста, полученные материнскими вейвлетами Добеши db4, db8 и db10, имеющих соответственно 4, 8 и 10 нулевых моментов, дают корректные значения показателя *H*, в то время как вейвлет Хаара (haar) и Добеши 2 порядка (db2) показывают завышенный результат, показатель Херста довольно быстро растет при возрастании значения *Ratio* и становится выше единицы.

Оценка показателя Херста *H*, заданного при моделировании фрактального процесса, вычисляется корректно либо нет, в зависимости от коэффициента *Ratio*, определяющего уровень шума. При возрастании

значения $Ratio$, оценка H становится завышенной при любом материнском вейвлете и достаточно скоро становится больше единицы. Однако существует некоторый диапазон значений отношения тренда к шуму, при котором параметр Херста оценивается корректно.

Определить момент появления тренда, соответствующий значению $Ratio^*$, можно с помощью простых и известных статистических критериев случайности [Кендэл, 1981]. Критерии позволяют проверить гипотезу, являются ли отдельные результаты независимыми наблюдениями или временной ряд имеет тренд, а также определить момент его возникновения. Значение $Ratio^*$ определялось как максимальное значение $Ratio^*$, для которого принимается гипотеза об отсутствии тренда для обоих критериев одновременно

На рисунке 3.8 представлены значения оценок параметра H , полученные с помощью вейвлета db4, в зависимости от показателя $Ratio$, полученные для фрактальных временных рядов с линейным трендом. На представленной зависимости выделен диапазон, где оценка является корректной.

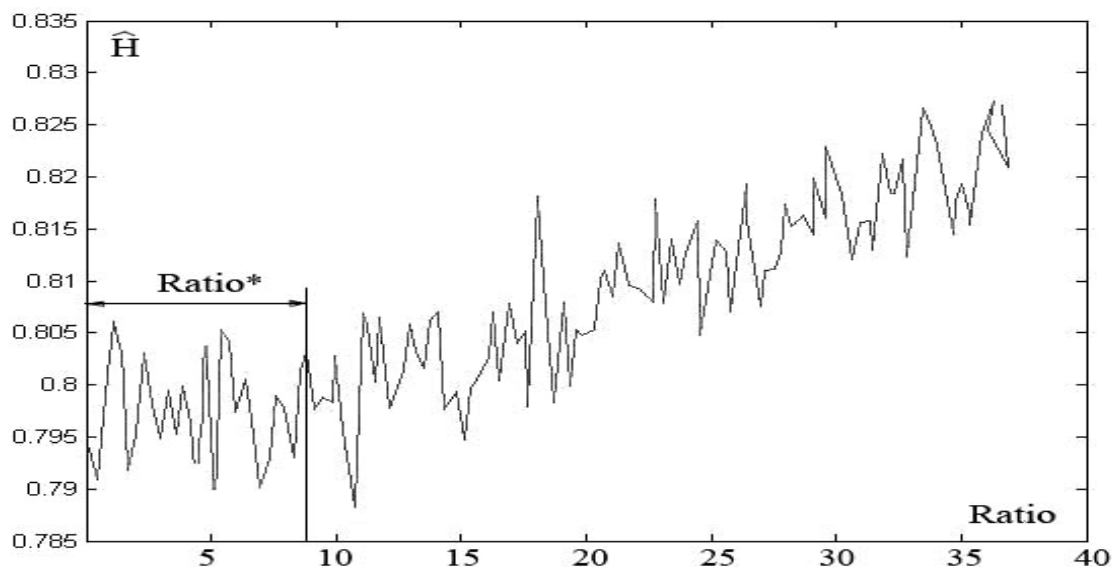


Рисунок 3.8 – Зависимость оценок показателя Херста от коэффициента $Ratio$

В таблице 3.5 представлены результаты эксперимента для модельных фрактальных временных рядов с различными трендовыми компонентами и соответствующее значение коэффициента *Ratio*^{*}.

Таблица 3.5 – Значения коэффициента *Ratio*^{*} для модельных временных рядов с трендовыми компонентами

Тренд	Значение <i>Ratio</i> [*]
$T(t) = t$	8.27
$T(t) = t^2$	7.25
$T(t) = t^3$	4.74
$T(t) = \sqrt{t}$	1.82
$T(t) = t^{3/4}$	0.45
$T(t) = \log(t)$	0.30
$T(t) = \sin(2\pi t)$, 1 период	0.18
$T(t) = \sin(2\pi t)$, 2 периода	0.08

Очевидно, что для полиномиального тренда невысокого порядка оценка показателя *H* является корректной даже при большом соотношении тренда и фрактального шума. При гиперболическом и логарифмическом тренде корректное оценивание возможно только при небольшом тренде.

3.3.2 Вейвлет-оценивание показателя Херста, на основе предварительного анализа спектра вейвлет-энергии

Для оценивания показателя H могут быть выбраны вейвлет-функции семейства Добеши (db2-db20), которые обладают необходимым числом нулевых моментов, (соответственно 2-20). При небольших значениях отношения тренда к ФГШ ($Ratio \leq 2$) метод вейвлет-оценивания позволяет получить адекватные оценки показателя Херста при использовании вейвлет-функций с большим числом нулевых моментов. Однако, с увеличением значения $Ratio$ или числа периодов в циклической компоненте, требуются вейвлет-функции с все большим числом нулевых моментов, что сложно реализуемо на практике. Кроме того, использование вейвлет-функций, обладающих большим числом нулевых моментов, приводит к менее точному оцениванию показателя H .

В том случае, когда сигнал содержит трендовую составляющую, спектр вейвлет-коэффициентов демонстрирует большой всплеск коэффициентов на низких частотах. В этом случае преобладающие в сигнале низкочастотные составляющие подавляют все остальные. Однако, при оценивании параметра Херста H можно не учитывать коэффициенты на высоких уровнях и оценивать угол наклона только по значениям вейвлет-энергии коэффициенты на более низких уровнях разложения.

На рис. 3.9 представлены аддитивные сигналы (3.6) различного типа (заданный показатель Херста $H = 0,8$, длина временного ряда – 4096 значений). В средней части рис. 3.9 показаны соответствующие спектры вейвлет-энергий (материнский вейвлет – db4), а в нижней - зависимость $\log_2(E_j)$ от уровня разложения j . Как видно из приведенных рисунков, можно исключить влияние низкочастотных компонент на оценивание показателя Херста, не учитывая при вычислениях коэффициенты на более высоких уровнях (низких частотах) и вновь оценивая показатель Херста. На рис. 3.9 показано определение показателя Херста без учета (сплошная линия 2) и с учетом (пунктирная линия 1) трендовых компонент. Оценивание H без учета уровней разложения, где проявляется влияние

трендовых и циклических компонент, дает корректные результаты вычислений.

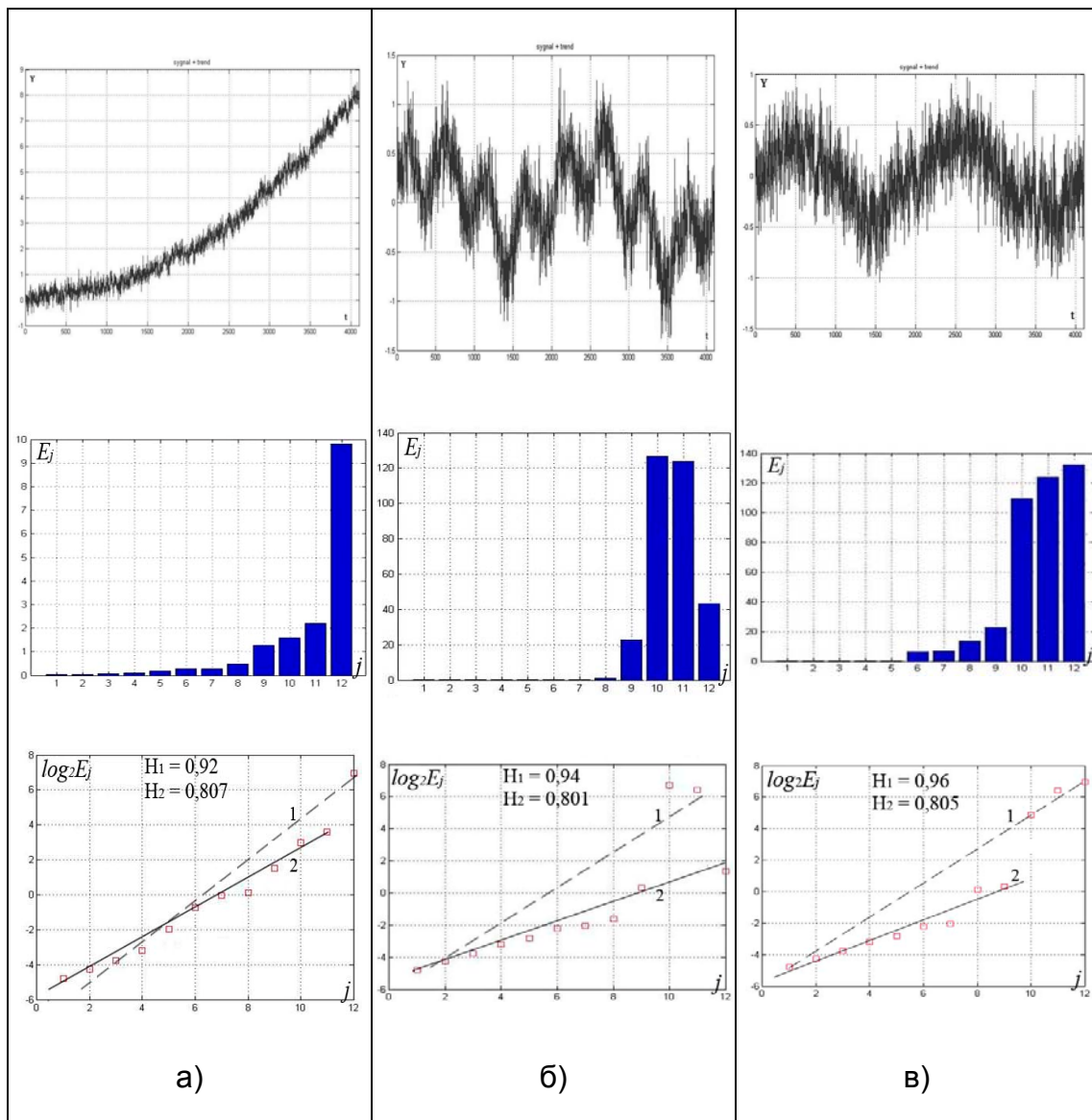


Рисунок 3.9 – Модельные сигналы, спектры вейвлет-энергии, логарифмические зависимости с трендовыми компонентами: а) кубический полином порядка; б) линейный полином и сумма синусов; в) синусоида

В работе [Дейнеко, 2010] предложен метод вейвлет-оценивания показателя Херста для временных рядов со значительными трендовыми и циклическими составляющими, состоящий из нескольких этапов:

Этап 1. Предварительное исследование временного ряда. Анализ временных рядов путем вейвлет-разложения на составные компоненты, выделение трендовых и циклических компонент.

Этап 2. Анализ трендовых и циклических компонент, оценивание порядка m полиномиального тренда с целью выбора эффективной материнской вейвлет-функции.

Этап 3. Определение значения параметра $Ratio$. Если значение лежит внутри диапазона корректного оценивания ($Ratio < Ratio^*$), то показатель Херста оценивается по временному ряду с помощью выбранного материнского вейвлета с числом нулевых моментов $p > m$.

Этап 4. Если значение $Ratio > Ratio^*$, то необходимо сделать преобразование спектра вейвлет-энергии: выявить уровни разложения, соответствующие трендовым компонентам и отбросить их при оценке наклона графика функции $\log_2(E_j)$ от j .

3.4 Комплексный подход к оцениванию фрактальных свойств самоподобных временных рядов

Подводя итоги исследований, представленных во 2-ом и 3-ем разделах, можно предложить следующую схему проведения фрактального анализа некоторого случайного фрактального процесса, представленного временным рядом длины N [Кириченко, 2014]. В основных этапах фрактального анализа задействованы методы нормированного размаха, ДФА и вейвлет-оценивания. Поскольку для применения аппарата вейвлет-преобразований необходимо соответствующее программное обеспечение и опыт работы, описание алгоритма построено таким образом, что

использование методов вейвлет-оценивания является желательным, но не обязательным элементом. Однако применение метода ДФА является необходимым по двум причинам: этот метод обладает достаточной точностью и предназначен для работы с нестационарными временными рядами. Рассмотрим поэтапную реализацию комплексного подхода к оцениванию фрактальных свойств самоподобных временных рядов.

Этап 1. Предварительное исследование структуры временного ряда.

1. Прежде чем приступать к фрактальному анализу временного ряда, необходимо выяснить из априорно известной информации, является ли ряд кумулятивным (например, курс валюты) или нет (например, информационный трафик). Если по своей природе ряд является кумулятивным, то все ниже указанные этапы фрактального анализа относятся к соответствующему ряду приращений.

2. Определение интервалов различных скейлингов.

Если самоподобный процесс обладает несколькими скейлингами, зависящими от временных интервалов (например дневные и часовые данные для валютных рядов), то на каждом таком интервале динамика временного ряда определяется соответствующим показателем Херста. Для определения таких интервалов надо построить функциональную зависимость $H(\tau) = f[\log \frac{R}{S}(\tau)]$ совместно с функцией (3.2).

Кроме метода нормированного размаха для выявления интервалов разных скейлингов можно использовать построение флуктуационной функции $F_{DFA}(\tau)$ методом ДФА (3.5). Поскольку данный этап является прикидочным, степень локального полиномиального тренда можно выбрать завышенной. Если существует несколько скейлингов, функция

$F_{DFA}(\tau)$ изменит угол наклона. Однако, в отличие от метода нормированного размаха, который исследует долгосрочную зависимость временного ряда в диапазоне всей его длины, флуктуационная функция может быть корректно построена только на интервале до значений $N / 4$.

3. Выявление и удаление краткосрочной авторегрессионной зависимости. Проверка гипотезы о наличии самоподобия.

Наличие авторегрессионной зависимости смещает значения показателя Херста и демонстрирует ложную долговременную память. При устранении авторегрессионной зависимости по формуле (3.5) определяется остаточный ряд. Если авторегрессионная зависимость является значимой, то все ниже указанные этапы фрактального анализа относятся к остаточному временному ряду.

Если значение степени самоподобия близко к 0.5, необходимо проверить гипотезу о наличии самоподобия. Качественной проверкой наличия свойств статистического самоподобия является построение агрегированных временных рядов согласно формуле и построение для них выборочных функций распределения. В случае самоподобия агрегированные ряды имеют одинаковое распределение, подтвержденное статистическими критериями.

Этап 2. Исследование стационарности временного ряда. Оценивание показателя Херста по стационарному ряду.

Для оценивания показателя Херста надо определить, является ли ряд стационарным или нет известными статистическими методами. Если ряд стационарен, то степень самоподобия и интервальную оценку показателя Херста можно определить методами, рассмотренными в преидущей статье.

Этап 3. Оценивание показателя Херста по нестационарному ряду.

1. Если исследуемый временной ряд нестационарен, необходимо исследовать его структуру с помощью построения корреляционной функции (спектральной плотности) и спектра вейвлет-энергии, которые позволяют выявить трендовые и циклические составляющие ряда.

2. При оценивании показателя Херста методом ДФА, необходимо сначала провести прикидочное оценивание, используя локальные полиномы увеличивающейся степени и определить наименьшую степень полинома, начиная с которой оценка показателя Херста перестает изменяться. После этого для оценивания самоподобия временного ряда необходимо удалять локальный полиномиальный тренд найденной степени. Интервальную оценку показателя Херста для временного ряда длиной N можно определить согласно формуле (2.22).

3. Вейвлет–оценивание нестационарного временного ряда проводится согласно методу, представленному в подразделе 3.3. Интервальная оценка показателя определяется по формуле (2.22).

Этап 4. Уточнение полученной оценки показателя Херста.

Если оценки были получены разными методами, то для увеличения точности оценки показателя Херста можно использовать среднее арифметическое этих оценок, полученных с помощью нескольких методов оценивания. Если проводилось вейвлет-оценивание с разными вейвлет-функциями, то более точной будет среднее арифметическое этих оценок.

Выводы

1. Рассмотрены методы исследований, предваряющие проведение фрактального анализа, которые включают в себя проверку гипотезы о наличии самоподобия, определение интервалов различных скейлингов, выявление и удаление краткосрочной зависимости.

2. Получил дальнейшее развитие метод оценивания показателя Херста ДФА для фрактальных рядов с трендовыми компонентами различного вида. Проведенные в статье численные исследования для фрактальных рядов с полиномиальными трендовыми составляющими порядка p показали, что адекватное оценивание параметра Херста достигается при использовании локального полиномиального тренда порядка $m > p$. В общем случае для рядов с трендовыми компонентами различного вида необходимо сначала провести прикидочное оценивание, используя локальные полиномы увеличивающейся степени, и определить наименьшую степень полинома, начиная с которой оценка показателя Херста перестает изменяться.

3. Исследовано воздействие трендовых и циклических составляющих на свойства оценок показателя Херста в зависимости от вида трендовой компоненты и параметров материнских вейвлет-функций. Определен диапазон отношения трендовой компоненты к фрактальному шуму, на котором корректно оценивается показатель Херста.

4. Предложен метод вейвлет-оценивания показателя Херста для временных рядов со значительными трендовыми и циклическими составляющими, основанный на анализе компонент спектра вейвлет-энергии и выделении диапазонов частот трендовой и циклической компонент ряда.

5. Предложен комплексный подход к оцениванию фрактальных свойств самоподобных временных рядов и предложены практические рекомендации по реализации данного подхода.

Литература

- [Feller, 1951] Feller W. The asymptotic distribution of the range of sums of independence variables. *Annals of mathematics and statistics*, France, University of Nancy, 1951. – 306 p.
- [Hurst, 1951] Hurst H. E. The Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, Vol. 116, 1951. pp. 770–799.
- [Hurst, 1955] Hurst H. E. Methods of using long-term storage in reservoirs. *Proceedings of the Institution of Civil Engineers*, Part 1, London, 1955. pp. 519–577.
- [Kantelhardt, 2001] Kantelhardt, J. W., Koscielny-Bunde, E., Havlin, S., et al. Detecting long-range correlations with detrended fluctuation analysis. *Physica A., Statistical Mechanics and its Applications*, Vol. 295, № 3–4, 2001. pp. 441–454.
- [Kirichenko, 2011] Kirichenko, L., Radivilova, T., Deineko, Zh. Comparative Analysis for Estimating of the Hurst Exponent for Stationary and Nonstationary Time Series. *Information Technologies & Knowledge*, Vol.5, № 4, 2011. pp. 371–388.
- [Mandelbrot, 1969] Mandelbrot, B., Wallis, J. R. Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long-Run Statistical Dependence. *Water Resources Research*, Vol. 5, 1969. pp. 967–988.
- [Mandelbrot, 1979] Mandelbrot, B., Taqqu, M. S. Robust R/S analysis of long-run serial correlation. *Bulletin of the International Statistical Institute*,

Proceedings of the 42nd Session of the International Statistical Institute, Manila. Book 2, Vol.48,1979. pp. 69–104.

[Taqqu, 1995] Taqqu, M. S., Teverovsky, V., Willinger, W. Estimators for long-range dependence: an empirical study. *Fractals*, Vol. 3, № 4, 1995. P. 785–798.

[Taqqu, 1996] Taqqu M. S., Teverovsky, V. Semi-parametric graphical estimation techniques for long-memory data. *Lecture Notes in Statistics*, Vol. 115– 1996. pp. 420–432.

[Дейнеко, 2010] Дейнеко, Ж. В., Кириченко, Л. О. Оценивание параметра Хёрста для временных рядов с трендом методом вейвлет-преобразования. *Системы управління, навігації та зв'язку*, Вип. 4 (16), 2010. сс. 85–89.

[Кендэл, 1981] Кендэл М. Временные ряды. М. :Финансы и статистика, 1981. 198с.

[Кириченко, 2014] Кириченко Л.О., Чалая Л.Э. Комплексный подход к исследованию фрактальных временных рядов. *Information Technologies & Knowledge, International Journal*, Vol. 8, № 1, 2014. pp. 22-28.

[Малла, 2005] Малла С. Вэйвлеты в обработке сигналов. М.: Мир, 2005. 671 с.

[Петерс, 2000]. Петерс Э. Хаос и порядок на рынках капитала. Новый аналитический взгляд на циклы, цены и изменчивость рынка : пер. с англ. М. : Мир, 2000. 333 с.

[Петерс, 2004] Фрактальный анализ финансовых рынков. Применение теории хаоса в инвестициях и экономике : пер. с англ. М : Интернет-трейдинг, 2004. 304 с.

[Федер, 1991] Федер Е. Фракталы. М.: Мир, 1991. 254 с.

[Ширяев, 1998] Ширяев А. Н. Основы стохастической финансовой математики. М. : Фазис, Т. 1: Факты. Модели, 1998. 512 с.

Информация об авторах



Людмила Кириченко – д.т.н., профессор Харьковского национального университета радиотехники и электроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: lyudmyla.kirichenko@nure.ua.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, фрактальный анализ, вейвлет-анализ, детерминированные хаотические системы.



Тамара Радивилова – к.т.н., доцент Харьковского национального университета радиотехники и электроники; пр. Науки 14, 61166, Харьков, Украина; e-mail: tamara.radivilova@gmail.com.

Основные области научных исследований: самоподобные и мультифрактальные временные ряды, телекоммуникационные системы, управление трафиком, информационная безопасность.

TABLE OF CONTENTS

<i>Фрактальный анализ самоподобных и мультифрактальных временных рядов Часть 1</i>	
Людмила Кириченко, Тамара Радивилова.....	3
<i>Основные понятия, характеристики и модели самоподобных и мультифрактальных стохастических процессов</i>	
Людмила Кириченко, Тамара Радивилова.....	6
<i>Оценивание параметра самоподобия для стационарных стохастических процессов</i>	
Людмила Кириченко, Тамара Радивилова.....	41
<i>Оценивание параметра самоподобия для нестационарных стохастических процессов</i>	
Людмила Кириченко, Тамара Радивилова.....	72
Table of contents.....	100