

Отзыв
официального оппонента на диссертационную работу
Ковалевского Артема Павловича
«Статистические критерии апостериорного обнаружения
разладки временных рядов и их применения»,
представленной на соискание ученой степени
доктора физико-математических наук по специальности
05.13.17 – Теоретические основы информатики

Актуальность темы. Актуальность данного исследования обусловлена недостаточной изученностью статистических критериев апостериорного обнаружения разладки как в теоретическом плане, так и в прикладных аспектах. Для относительно простых вероятностных моделей существует множество статистических критериев, и проблема в том, как выбрать из них один. Если вероятностная модель сложно устроена, то зачастую возникает вопрос о том, как вообще построить состоятельный статистический критерий. Существующие на сегодняшний день алгоритмы обнаружения разладки в основном ориентированы на наискорейшее обнаружение в реальном времени, а не на анализ полного набора данных с принятием решения о том, произошла ли разладка на интервале наблюдения. Именно апостериорное обнаружение разладки является предметом интереса, например, при исследовании однородности текста и при анализе наличия линейных стохастических связей между переменными в многомерных массивах данных. Для успешного решения такого рода задач необходимы алгоритмы обнаружения изменения свойств случайной последовательности, основанные на соответствующих статистических критериях.

Целью диссертационной работы автор определяет разработку статистических критериев проверки гипотез для класса вероятностных моделей временных рядов. Этот класс вероятностных моделей включает модели с циклическим трендом и с фрактальным гауссовским шумом. Диссертант доказывает ряд теорем, которые служат для обоснования построенных в диссертации алгоритмов. Кроме того, А.П. Ковалевский расширяет теоретические основы методов обнаружения разладки, исследуя алгоритмы для моделей, основанных на фрактальном гауссовском шуме.

Научная значимость и новизна результатов, выносимых на защиту. Диссертация состоит из введения, шести глав, заключения и списка использованной литературы.

Во введении обоснована актуальность темы диссертации, сформулированы цели и задачи, перечислены методы исследования, указаны область, объект и предмет исследования, сформулирована научная новизна, выделены девять положений, которые выносятся на защиту, отмечаются практическая и теоретическая ценность полученных результатов, приводится информация об апробации (на всероссийских и международных конференциях) данной работы.

В первой главе автор строит и исследует класс статистических критериев для ситуации, когда в качестве основной гипотезы предлагается модель выборки, то есть совокупности одинаково распределенных наблюдений, а в качестве альтернативной – модель разладки, т. е. совокупности наблюдений, в которой произошло изменение математического ожидания в некоторый неизвестный исследователю момент времени. Для сравнения статистических критериев используется приближенный т.н. бахадуровский наклон, характеризующий мощность критерия при различении близких гипотез. Статистики рассматриваемых критериев строятся единым образом как функционалы от эмпирического моста – центрированной и нормированной ломаной, построенной по суммам значений временного ряда. Результатом первой главы является нахождение статистического критерия, имеющего наибольший бахадуровский наклон в широком классе изученных критериев.

Во второй главе автор строит новые оценки параметров фрактального броуновского движения и процесса его приращений – фрактального гауссовского шума. Построена оценка элементарным знаковым методом и ее модификации. Доказана асимптотическая нормальность предложенных оценок. Наилучшая (имеющая наименьший коэффициент асимптотической нормальности) из предложенных оценок – оценка бинарным знаковым методом.

В третьей главе автор строит алгоритмы проверки гипотез, связанных с моделью фрактального гауссовского шума. В параграфе 3.2 предложен статистический критерий проверки нормальности выборки, связанный с модификацией знакового метода. В параграфе 3.3 на основании бинарного знакового метода автор предлагает статистический критерий наличия фрактальности в гауссовском шуме. В параграфе 3.4 для целей следующего параграфа разработаны процедуры моделирования

фрактального гауссовского шума. В параграфе 3.5 автор строит критерий обнаружения разладки фрактального гауссовского шума с помощью результатов моделирования. В параграфе 3.6 строятся статистические критерии проверки однородности двух фрактальных броуновских движений. В параграфе 3.7 статистические критерии распространяются на случай, когда случайные величины получены из фрактального гауссовского шума приведением к негауссовскому устойчивому закону.

В четвертой главе предложены вероятностные модели текста и статистические критерии для их анализа. Вероятностные модели включают урновую модель для процесса числа разных слов с установлением функциональной центральной предельной теоремы для этого процесса, а также модели появления служебных слов в тексте. Построение статистических критериев использует результаты глав 1 и 3.

В пятой главе автор строит статистические критерии обнаружения разладки в регрессионных моделях. Для процесса последовательных сумм остатков регрессии на порядковые статистики предложена теорема о слабой сходимости траекторий процесса к центрированному гауссовскому процессу. Для процесса последовательных сумм остатков регрессии с циклическим трендом соответствующий результат был известен ранее. Автор сравнивает мощность разных статистических критериев на основании результатов моделирования и обосновывает превосходство критерия, использующего в качестве статистики модуль наибольшего отклонения эмпирического моста от оси абсцисс.

В шестой главе результаты пятой главы используются для анализа линейных зависимостей на примерах медицинских и экономических данных. Показано, как разработанные методы приводят к использованию более эффективных моделей.

В заключении автор приводит основные результаты работы, обосновывает достоверность результатов и формулирует положения, выносимые на защиту.

Достоверность полученных в работе результатов не вызывает сомнений. В диссертации используется строгий математический аппарат, предлагаемые методы и алгоритмы апробированы на практике.

Выполненные исследования показывают, что их автор, с одной стороны, является весьма глубоким специалистом в области теории вероятностей и математической статистики, что позволило ему с успехом предложить новые статистические критерии, исследовать их вероятностные свойства и сравнительные характеристики. С другой

стороны, им продемонстрировано глубокое понимание особенностей прикладных задач и умение строить адекватные математические модели для разнообразных приложений, разрабатывать алгоритмы и программное обеспечение. Здесь отдельно хочется отметить поразительные результаты статистического анализа однородности текстов, широко представленные в четвертой главе диссертации.

Публикации. По теме представленной диссертации автором опубликовано 48 работ, в том числе одиннадцать статей в авторитетных международных журналах. Автореферат отражает основные положения диссертационной работы, материалы диссертации достаточно полно освещены в публикациях автора.

Диссертация написана хорошим литературным языком и почти не содержит опечаток. Есть и некоторые замечания.

- 1) Фраза из "Введения" о том, что для скорейшего обнаружения разладки используется последовательный критерий отношения правдоподобия Вальда, нуждается в корректировке. На самом деле критерий Вальда предназначен для других целей. А метод обнаружения разладки, основанный на использовании отношения правдоподобия, был впервые предложен в работе Page E. S. (1954). Continuous inspection schemes. *Biometrika*, 1954, и получил название процедуры кумулятивных сумм (CUSUM procedure).
- 2) Утверждение об условиях состоятельности критерия в Замечании 1.1 на стр. 25 требует пояснения о каком критерии идет речь.
- 3) Не указан адрес ссылки перед теоремой 6.1 (стр. 229).

Отмеченные недостатки не снижают общую высокую оценку работы. Она выполнена на высоком научном уровне и хорошо оформлена.

Заключение. Диссертационная работа А. П. Ковалевского «Статистические критерии апостериорного обнаружения разладки временных рядов и их применения» представляет собой целостную научную работу, в которой автором разработаны теоретические основы, методы и алгоритмы, имеющие важное научное и практическое значение. Разработанные алгоритмы и написанные на их основе программы защищены авторскими свидетельствами и патентами. Тем самым диссертация А. П. Ковалевского являет собой существенное продвижение в данной области.

Считаю, что диссертационная работа А. П. Ковалевского «Статистические критерии апостериорного обнаружения разладки временных рядов и их применения» соответствует «Положению о порядке присуждения ученых степеней». Автор работы Ковалевский Артем Павлович достоин присуждения ученой степени доктора физико-математических наук по специальности 05.13.17 – Теоретические основы информатики.

Официальный оппонент
д.ф.-м.н., профессор,
заведующий лабораторией
теории вероятностей
и математической статистики
ИМ СО РАН



Лотов

В. И. Лотов

Дата *24.02.2019*

Подпись д.ф.-м.н., профессора В.И. Лотова заверяю

Ученый секретарь ИМ СО РАН к.ф.-м.н.

Светов

И. Е. Светов

Лотов Владимир Иванович, доктор физико-математических наук по специальности 01.01.05 – теория вероятностей и математическая статистика, профессор, заведующий лабораторией теории вероятностей и математической статистики Федерального государственного бюджетного учреждения науки Института математики им. С. Л. Соболева Сибирского отделения Российской академии наук (ИМ СО РАН), 630090, г. Новосибирск, пр. Академика Коптюга, 4, телефон: (8-383) 333-28-92 факс: (8-383) 333-25-98 адрес электронной почты: im@math.nsc.ru веб-сайт: <http://math.nsc.ru>

Контактные данные:
телефон: (8-383) 329-76-04
адрес электронной почты: lotov@math.nsc.ru

