



СибГУТИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ«СИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ТЕЛЕКОММУНИКАЦИЙ И ИНФОРМАТИКИ»
(СибГУТИ)

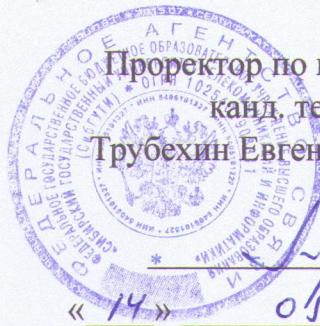
Кирова ул., д. 86, г. Новосибирск, 630102

т/ф: (383) 269-82-02, 269-82-03

http://www.sibsutis.ru, e-mail: rectorat@sibsutis.ru

ИНН 5405101327, КПП 540501001, ОКПО 01180010

«УТВЕРЖДАЮ»

Проректор по науке СибГУТИ,
канд. техн. наук, доцент
Трубехин Евгений Рудольфович

« 14 » 05 2019 г.

от « 14 » « 05 » 2019 г. № 13 - 1188

на № _____ от « _____ » « _____ » 20 _____ г.

Отзыв ведущей организации
на диссертацию Браништи Владислава Владимировича
на тему «Методы и алгоритмы настройки проекционной оценки плотности
вероятности случайного вектора в условиях малых выборок»,
представленную на соискание учёной степени
кандидата физико-математических наук по специальности
05.13.17 – Теоретические основы информатики

Актуальность темы исследования. Диссертация Браништи Владислава Владимировича «Методы и алгоритмы настройки проекционной оценки плотности вероятности случайного вектора в условиях малых выборок» посвящена непараметрическим методам восстановления закона распределения непрерывного случайного вектора – одной из главных задач математической статистики. Данной задаче посвящено большое количество работ, среди которых можно выделить таких авторов как А. Н. Колмогоров, Н. Н. Ченцов, М. Розенблатт, Э. Парзен, В. А. Епанечников, Ф. П. Тарасенко, А. В. Лапко, А. В. Медведев, Г. Вахба, А. А. Новосёлов. В большинстве работ по непараметрическим оценкам предполагается, что объём выборки исследуемого случайного вектора неограничен, что позволяет при анализе использовать центральную предельную теорему и другие асимптотические соотношения. Однако в представленной диссертации рассматривается случай малых выборок, что существенно усложняет задачу.

Проекционная оценка плотности вероятности представляет собой частичную сумму обобщённого ряда Фурье от истинной функции плотности вероятности. Это позволяет, как и указывается в диссертации, в отличие от других видов

непараметрических оценок, получать для оценки плотности компактное математическое выражение, а также ускорить работу основанных на ней алгоритмов классификации и восстановления зависимостей.

Вышеизложенное показывает актуальность проводимых в диссертации исследований.

Содержание работы и основные результаты. Первая глава диссертации «Обзор методов оценивания функции плотности вероятности» имеет вводный характер. В ней даются необходимые определения и обозначения функционального анализа, теории вероятностей и математической статистики, приводится общая постановка задачи настройки непараметрической оценки функции плотности вероятности в общем виде, а также указаны известные оценки, среди которых оценки проекционного типа, ядерные оценки, гистограмма и оценка k ближайших соседей, и методы их настройки. Приведённые результаты используются в следующих главах при сравнении качества восстановления плотности вероятности.

Во второй главе «Оптимизация проекционной оценки плотности вероятности» указывается на недостаток проекционной оценки – она неприменима при восстановлении плотности вероятности с несуммируемым квадратом, например, плотности вероятности квадрата случайной величины, подчинённой стандартному нормальному закону распределения. Указанный недостаток устраняется построением проекционной оценки в весовом гильбертовом пространстве. Указаны условия на весовую функцию, при которых данное пространство является гильбертовым.

Одним из основных результатов данной главы является предложение 2.4, утверждающее, что для любой функции плотности вероятности $f(\mathbf{x}) = f(x_1, \dots, x_k)$ случайного вектора $\xi = (\xi_1, \dots, \xi_k)$ и любого измеримого множества $\Omega \subseteq \mathbb{R}^k$, удовлетворяющего условию

$$\int_{\Omega} f(\mathbf{x}) d\mu(\mathbf{x}) = 1, \quad (1)$$

существует такая весовая функция $w(\mathbf{x})$, что $f \in L_{2,w}(\Omega)$. Следовательно, выбрав в этом пространстве подходящий базис, можно построить проекционную оценку функции $f(\mathbf{x})$, даже если она имеет несуммируемый квадрат.

Рассматривается вопрос о выборе весовой функции в условиях неизвестной функции плотности вероятности $f(\mathbf{x})$. Доказана теорема 2.4 о том, что весовое расширение $L_{2,w}(\Omega)$ гильбертова пространства $L_2(\Omega)$ гарантированно содержит более широкое множество функций плотности вероятности. Доказательство проводится конструктивно, что позволяет строить конкретные примеры таких плотностей. Кроме того, рассматривается случай построения весовой функции, когда вид функции плотности вероятности неизвестен, но известны точки, в окрестности которых функция имеет несуммируемый квадрат.

Другим результатом главы 2 является обоснование того, что предлагаемое автором обобщение метода моментов для настройки коэффициентов проекционной оценки приводит к улучшению качества восстановления плотности вероятности по сравнению с традиционным подходом (методом Ченцова). Кроме того, доказана представляющая теоретический интерес теорема 2.5 о том, что при увеличении одного из параметров оценки, предложенной соискателем, она сходится к оценке Ченцова. Таким образом, вскрывается связь между двумя на первый взгляд совершенно разными методами: методом Ченцова и методом моментов.

Задача настройки длины проекционной оценки также находит решение в главе 2 как для одномерного, так и для многомерного случая (формулы (2.25) и (2.37) соответственно). Работоспособность предложенных методов была проверена в ходе численных экспериментов на выбранных тестовых распределениях.

В третьей главе «Применение оценок плотности вероятности» рассматриваются некоторые приложения полученных в главе 2 результатов. Как известно, некоторые задачи анализа данных, теории управления и других областей науки сводятся к оцениванию плотности вероятности случайного вектора. К этим задачам относятся рассматриваемые в главе 3 задача оценивания функции регрессии, классификация и оценивание количества информации. Для решения каждой из этих задач были реализованы соответствующие алгоритмы, в которых в качестве оценки функции

плотности вероятности бралась предложенная автором оценка, а также некоторые другие известные оценки, включая оценку Розенблатта–Парзена. Таким образом, проводилось сравнение оценок при решении указанных задач.

В целом, изложения материала диссертации соответствует требованиям, предъявляемым на соискание учёной степени кандидата физико-математических наук: сформулированные утверждения снабжены полными доказательствами, используемые в тексте результаты других авторов имеют ссылки на цитируемые источники.

Научная и практическая ценность диссертации. На наш взгляд, научную и практическую ценность имеют следующие результаты диссертационной работы.

1. Доказательство сходимости проекционной оценки в пространстве $L_{2,w}(\Omega)$ позволяет использовать известные методы построения проекционных оценок для любых плотностей вероятности, в том числе, имеющих несуммируемый квадрат.
2. Предложенное обобщение метода моментов для настройки коэффициентов, а также предложенный метод настройки длины ряда позволяют повысить эффективность проекционной оценки.
3. Разработанные в главе 3 алгоритмы позволяют эффективно решать задачи восстановления функции регрессии, классификации и оценивания количества информации.

Замечания по работе. Несмотря на общее положительное впечатление о работе, она не лишена недостатков. По работе имеются следующие замечания.

1. В предложении 2.4 на множество Ω накладывается условие (1), тогда как в действительности оно является лишним. Таким образом, предложение 2.4 допускает более сильную формулировку.
2. Для применимости метода моментов необходимым условием является существование начальных моментов исследуемой случайной величины до требуемого порядка, что явно не указано в тексте.
3. Не проверена работоспособность предлагаемого метода на распределениях с «тяжёлыми хвостами», таких как распределение Коши.

4. В тексте диссертации недостаточно чётко сформулированы преимущества предлагаемых методов в случае малых выборок.
5. Не указаны затраты машинного времени, сделанные для получения численных результатов при расчёте функционалов качества.

Указанные замечания являются устранимыми, не влияют на правильность полученных результатов и не умаляют достижений соискателя.

Диссертация имеет теоретический характер. Результаты работы являются новыми, полностью доказаны и опубликованы в рецензируемых журналах из Перечня ВАК. Автореферат диссертации правильно отражает её содержание. Результаты диссертации докладывались автором на международных конференциях и научных семинарах.

Результаты работы могут быть использованы при разработке алгоритмов анализа данных, классификации, а также при решении других задач науки и техники, в которых требуется знание закона распределения исследуемых случайных величин или векторов. Используемый метод сравнения качества восстановления плотности вероятности может быть использован для сравнения любых непараметрических оценок плотностей.

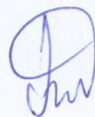
Вывод. Диссертация В. В. Браништи «Методы и алгоритмы настройки проекционной оценки плотности вероятности случайного вектора в условиях малых выборок» является научно-квалификационной работой, в которой содержится решение научной задачи, имеющей значение для развития методов анализа данных, удовлетворяет п. 9 «Положения о присуждении учёных степеней».

Диссертация удовлетворяет всем требованиям, предъявляемым к диссертациям на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 – Теоретические основы информатики, а ее автор, Браништи Владислав Владимирович, заслуживает присуждения учёной степени кандидата физико-математических наук.

Отзыв на диссертацию и автореферат диссертации рассмотрен и одобрен на заседании семинара кафедры высшей математики факультета информатики и вычислительной техники СибГУТИ 6 мая 2019 г., протокол № 3.

Отзыв составили:

Трофимов Виктор Куприянович,
декан факультета ИВТ,
заведующий кафедрой высшей математики,
Федеральное государственное бюджетное
образовательное учреждение
высшего образования
«Сибирский государственный университет
телекоммуникаций и информатики» (СибГУТИ)
ул. Кирова 86, г. Новосибирск, 630102, Россия
доктор технических наук, профессор
раб. телефон +7(383)2693935
e-mail: trofimov@sibsutis.ru



Лыткина Дарья Викторовна,
профессор кафедры высшей математики,
Федеральное государственное бюджетное
образовательное учреждение
высшего образования
«Сибирский государственный университет
телекоммуникаций и информатики» (СибГУТИ)
ул. Кирова 86, г. Новосибирск, 630102, Россия
доктор физико-математических наук, профессор
раб. телефон +7(383)2693936
e-mail: daria.lytkin@gmail.com

