

Математические методы исследования

Mathematical methods of investigation

DOI: <https://doi.org/10.26896/1028-6861-2019-85-11-69-84>

СТАТИСТИКА НЕЧИСЛОВЫХ ДАННЫХ ЗА Сорок ЛЕТ (ОБЗОР)

© Александр Иванович Орлов

Московский государственный технический университет им. Н. Э. Баумана, Россия, 105005, Москва, Бауманская 2-я, д. 5;
e-mail: prof-orlov@mail.ru

*Статья поступила 11 марта 2019 г. Поступила после доработки 11 марта 2019 г.
Принята к публикации 25 апреля 2019 г.*

Сорок лет назад статистика нечисловых данных была выделена как самостоятельная область математических методов исследования. Первоначально использовался термин «статистика объектов нечисловой природы». Статистика нечисловых данных — одна из четырех основных областей прикладной статистики (наряду со статистикой чисел, многомерным статистическим анализом, статистикой временных рядов и случайных процессов). Статистика нечисловых данных делится на статистику в пространствах общей природы и разделы, посвященные конкретным типам нечисловых данных (статистика интервальных данных, статистика нечетких множеств, статистика бинарных отношений и др.). В настоящее время статистика в пространствах общей природы — центральная часть прикладной статистики, а включающая ее статистика нечисловых данных — основная область прикладной статистики. Это утверждение подтверждается анализом публикаций в разделе «Математические методы исследования» нашего журнала. Данная статья посвящена анализу основных идей статистики нечисловых данных на фоне развития прикладной статистики. Основой является новая парадигма математических методов исследования. Рассмотрены различные виды нечисловых данных. Рассказано о развитии статистики нечисловых данных. Разобраны основные идеи статистики в пространствах общей природы: средние величины, законы больших чисел, экстремальные статистические задачи, непараметрические оценки плотности распределения вероятностей, методы классификации (диагностики и кластер-анализа), статистики интегрального типа. Кратко рассмотрены некоторые статистические методы анализа данных, лежащих в конкретных пространствах нечисловой природы: непараметрическая статистика (в подавляющем большинстве случаев реальные распределения существенно отличаются от нормальных), статистика нечетких множеств, теория экспертных оценок (медиана Кемени — выборочное среднее экспертных упорядочений) и др. Обсуждаются некоторые нерешенные задачи статистики нечисловых данных.

Ключевые слова: математические методы исследования; прикладная статистика; нечисловые данные; статистика в пространствах общей природы; задачи оптимизации; средние величины; законы больших чисел; непараметрические оценки плотности; статистики интегрального типа; непараметрическая статистика; экспертные оценки; нерешенные задачи.

STATISTICS OF NONNUMERIC DATA IN FORTY YEARS (REVIEW)

© Alexander I. Orlov

Bauman Moscow State Technical University, 5, 2-ya Baumanskaya ul., Moscow, 105005, Russia; e-mail: prof-orlov@mail.ru

Received March 11, 2019. Revised March 11, 2019. Accepted April 25, 2019.

Forty years ago, the statistics of nonnumeric data was singled out into an independent area of mathematical research methods. First the term “statistics of objects of nonnumeric nature” was used. The textbook on nonnumeric statistics is entitled “Nonnumeric Statistics.” Statistics of nonnumeric data is one of the four main fields of applied statistics (along with the statistics of numbers, multivariate statistical analysis, statistics of time series and random processes). Statistics of nonnumeric data is divided into statistics in spaces of general nature and sections devoted to specific types of nonnumeric data (statistics of interval data, statistics of fuzzy sets, statistics of binary relations, etc.). Currently, statistics in spaces of general nature is the central part of applied statistics, and the statistics of nonnumeric data that includes it is the

main area of applied statistics. This statement is confirmed by analysis of publications in the section “Mathematical Methods of Research” of our journal. This article is devoted to analysis of the main ideas of statistics of nonnumeric data against the background of the development of applied statistics on the basis of the new paradigm of the methods of mathematical research. Different types of nonnumeric data are considered along with the history of development of the statistics of nonnumeric data and statistical science. The basic ideas of statistics in spaces of general nature are analyzed: mean values, laws of large numbers, extreme statistical problems, nonparametric estimators of the probability distribution density, classification methods (diagnostics and cluster analysis), and statistics of the integral type. Some statistical methods for data analysis in specific nonnumeric spaces are briefly discussed: nonparametric statistics (in most cases, real distributions significantly differ from normal), statistics of fuzzy sets, expert estimation theory (Kemeny median as a sample average of expert orderings), etc. Some unsolved problems in statistics on nonnumeric data are also discussed.

Keywords: mathematical research methods; applied statistics; nonnumeric data; statistics in spaces of general nature; optimization problems; averages; laws of large numbers; nonparametric density estimates; integral type statistics; nonparametric statistics; expert estimation; unsolved problems.

Введение

В настоящее время статистика нечисловых данных — одна из четырех основных областей прикладной статистики (наряду со статистикой чисел, многомерным статистическим анализом, статистикой временных рядов и случайных процессов). Статистика нечисловых данных делится на статистику в пространствах общей природы и разделы, посвященные конкретным типам нечисловых данных (статистика интервальных данных, статистика нечетких множеств, статистика бинарных отношений и др.). Научные результаты, полученные в рамках статистики в пространствах общей природы, могут быть использованы для конкретных видов данных (например, непараметрические оценки плотности). Следовательно, статистика в пространствах общей природы — центральная часть прикладной статистики, а включающая ее статистика нечисловых данных — основная область прикладной статистики. Это утверждение подтверждается анализом публикаций в разделе «Математические методы исследования» нашего журнала — на первое место вышла именно статистика нечисловых данных. Так, за десять лет (2006 – 2015 гг.) ей посвящены 27,6 % всех публикаций, т.е. 63,0 % статей по прикладной статистике [1].

Первоначально использовался термин «статистика объектов нечисловой природы». Впервые он появился в 1979 г. в монографии [2] для обозначения совокупности некоторых полученных научных результатов. В том же году в статье [3] была развернута программа построения этой новой области статистических методов, приведены первоначальные формулировки ряда основных теорем. Через год в журнале «Заводская лаборатория» появилась обобщающая статья [4] авторов, занимавшихся различными аспектами статистики нечисловых данных. Итоги первых десяти лет развития новой области прикладной статистики были подведены в обстоятельном обзоре [5] (120 литературных ссылок). Дальнейшее

развитие было не менее плодотворным. Обзор [6] за тридцать лет содержал 150 литературных ссылок. К тридцатилетию вышел и первый учебник по нечисловой статистике [7]. Термин «нечисловая статистика» представляется слишком кратким, в то время как исходный термин «статистика объектов нечисловой природы» — слишком тяжеловесным. В данной статье будем называть рассматриваемую область прикладной статистики «статистикой нечисловых данных». Все три термина (статистика объектов нечисловой природы, статистика нечисловых данных, нечисловая статистика) — синонимы.

Обсудим содержание, развитие и основные идеи статистики нечисловых данных, особое внимание уделив исследованиям, опубликованным в нашем журнале (см. также сводку [1]).

Новая парадигма математических методов исследования

Появление и развитие статистики нечисловых данных знаменует переход к новой парадигме математических методов исследования.

Парадигма (от греч. *paradeigma* — пример, образец) — совокупность научных достижений, признаваемых всем научным сообществом в тот или иной период времени и служащих основой и образцом новых научных исследований. Понятие «парадигма» получило широкое распространение после выхода в свет книги [8] американского историка науки Т. Куна «Структура научных революций» (1962 г.).

Математические методы исследования используются для решения практических задач с давних времен. В Ветхом Завете рассказано о весьма квалифицированно проведенной переписи военнообязанных (Четвертая книга Моисеева «Числа»). В первой половине XX в. была разработана классическая парадигма методов обработки данных, полученных в результате измерений (наблюдений, испытаний, анализов, опытов). Математические методы исследования, со-

ответствующие классической парадигме, широко используются. Со стороны может показаться, что в этой области основное давно известно, современные работы направлены на мелкие усовершенствования. Однако это совсем не так. Новая парадигма математических методов исследования принципиально меняет прежние представления. Она зародилась в 1980-х годах, но была развита в серии монографий и учебников уже в XXI в. (см. [9 – 11] и др.).

Типовые исходные данные в новой парадигме — объекты нечисловой природы (элементы нелинейных пространств, которые нельзя складывать и умножать на число, например, множества, бинарные отношения), а в старой — числа, конечномерные векторы, функции [12]. Ранее (в классической старой парадигме) для расчетов использовали разнообразные суммы, однако объекты нечисловой природы нельзя складывать, поэтому в новой парадигме применяется другой математический аппарат, основанный на расстояниях между объектами нечисловой природы и решении задач оптимизации.

Изменились постановки задач анализа данных. Старая парадигма исходит из идей начала XX в., когда К. Пирсон предложил четырехпараметрическое семейство распределений для описания распределений реальных данных. В это семейство, в частности, входят подсемейства нормальных, экспоненциальных, Вейбулла – Гнеденко, гамма-распределений. Сразу было ясно, что распределения реальных данных, как правило, не входят в семейство распределений Пирсона (об этом говорил, например, академик С. Н. Бернштейн в 1927 г. в докладе на Всероссийском съезде математиков [13]). Однако математическая теория параметрических семейств распределений (методы оценивания параметров и проверки гипотез) оказалась достаточно интересной, и именно на ней до сих пор основано преподавание во многих вузах. Итак, в старой парадигме основной подход к описанию данных — распределения из параметрических семейств, а оцениваемые величины — их параметры, в новой парадигме рассматривают произвольные распределения, а оценивают характеристики и плотности распределений, зависимости, правила диагностики и др. Центральная часть теории — уже не статистика числовых случайных величин, а статистика в пространствах произвольной природы.

В старой парадигме источники постановок новых задач — традиции, сформировавшиеся к середине XX века, а в новой — современные потребности математического моделирования и анализа данных (XXI век), т.е. запросы практики. Конкретизируем это общее различие. В старой парадигме типовые результаты — предельные теоремы, в новой — рекомендации для кон-

кретных значений параметров, в частности, объемов выборок. Изменилась роль информационных технологий — ранее они использовались в основном для расчета таблиц (в частности, информатика находилась вне математической статистики), теперь же они — инструменты получения выводов (имитационное моделирование, датчики псевдослучайных чисел, методы размножения выборок, в том числе бутстреп, и др.). Вид постановок задач приблизился к потребностям практики — при анализе данных от отдельных задач оценивания и проверки гипотез перешли к статистическим технологиям (технологическим процессам анализа данных). Выявилась важность проблемы «стыковки алгоритмов» — влияния выполнения предыдущих алгоритмов в технологической цепочке на условия применимости последующих алгоритмов. В старой парадигме эта проблема не рассматривалась, для новой — весьма важна.

Если в старой парадигме вопросы методологии моделирования практически не обсуждались, достаточными признавались схемы начала XX в., то в новой парадигме роль методологии (учения об организации деятельности) [14] является основополагающей. Резко повысилась роль моделирования — от отдельных систем аксиом произошел переход к системам моделей. Сама возможность применения вероятностного подхода теперь — не «наличие повторяющегося комплекса условий» (реликт физического определения вероятности, использовавшегося до аксиоматизации теории вероятностей А. Н. Колмогоровым в 1930-х годах), а наличие обоснованной вероятностно-статистической модели. Если раньше данные считались полностью известными, то для новой парадигмы характерен учет свойств данных, в частности, интервальных и нечетких. Изменилось отношение к вопросам устойчивости выводов — в старой парадигме практически отсутствовал интерес к этой тематике, в новой разработана развитая теория устойчивости (робастности) выводов по отношению к допустимым отклонениям исходных данных и предпосылок моделей.

Статистика нечисловых данных развивается в соответствии с новой парадигмой математических методов исследования.

Различные виды нечисловых данных

Типичный исходный объект в прикладной статистике — это выборка, т.е. совокупность независимых одинаково распределенных случайных элементов. Какова природа этих элементов? В классической математической статистике элементы выборки — это числа. В многомерном статистическом анализе — векторы. А в статистике

нечисловых данных элементы выборки — это объекты нечисловой природы, которые нельзя складывать и умножать на числа. Другими словами, объекты нечисловой природы лежат в пространствах, не имеющих векторной (линейной) структуры.

Примерами объектов нечисловой природы являются:

- значения качественных признаков, в том числе результаты кодировки объектов с помощью заданного перечня категорий (градаций);

- упорядочения (ранжировки) экспертами объектов экспертизы — образцов продукции (при оценке её технического уровня, качества и конкурентоспособности, ее характеристик), заявок на проведение научных работ (при проведении конкурсов на выделение грантов) и т.п.;

- классификации, т.е. разбиения объектов, сходных между собой, на группы (кластеры);

- толерантности, т.е. бинарные отношения, описывающие сходство объектов между собой, например, сходство тематик научных работ, оцениваемое экспертами в целях рационального формирования экспертных советов внутри определенной области науки;

- результаты парных сравнений или контроля качества продукции по альтернативному признаку («годен» – «брак»), т.е. последовательности из 0 и 1;

- множества (обычные или нечеткие), например, зоны, пораженные коррозией, или перечни возможных причин аварии, составленные экспертами независимо друг от друга;

- слова, предложения, тексты;

- графы;

- векторы, координаты которых — совокупность значений разнотипных признаков, например, результат составления статистического отчета о научно-технической деятельности организации или анкета эксперта, в которой ответы на часть вопросов носят качественный характер, а на часть — количественный;

- ответы на вопросы экспертной, медицинской, маркетинговой или социологической анкеты, часть из которых носит количественный характер (возможно, интервальный), часть сводится к выбору одной из нескольких подсказок, а часть представляет собой тексты, и т.д.

Все средства измерения имеют погрешности. Однако до недавнего времени это очевидное обстоятельство никак не учитывалось в статистических процедурах. Только с конца 1970-х годов начала развиваться статистика интервальных данных, в которой предполагается, что исходные данные — это не числа, а интервалы. Статистику интервальных данных можно рассматривать как часть интервальной математики. Выводы в ней часто принципиально отличны от классических.

Различным подходам к статистическому анализу интервальных данных посвящена дискуссия [15]. В научной школе А. П. Воицинина изучены проблемы регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности [16 – 22]. В разработанной нами асимптотической статистике интервальных данных на значения случайных величин наложены малые интервальные неопределенности (см., например, [23]). Основные результаты этого направления подробно изложены в главах учебников [7, 24, 25] и монографии [26]. Кроме двух основных направлений (полученные в них недавние результаты отражены в статьях [27 – 29]), разрабатывались и иные подходы (см., например, [30]).

Интервальные данные можно рассматривать как частный случай нечетких множеств. Если характеристическая функция нечеткого множества равна единице на некотором интервале и равна нулю вне этого интервала, то задание такого нечеткого множества эквивалентно заданию интервала. С методологической точки зрения важно, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств. Цикл соответствующих теорем приведен в монографии [2], а также в учебниках [7, 24, 25, 31], монографии [26], недавней статье [32]. Казалось бы, публикаций много. Но приходится констатировать, что отнюдь не все специалисты знакомы с теоремами о сведении теории нечетких множеств к теории вероятностей.

Исторический путь статистической науки

Развитие статистических методов в нашей стране показано в главе 2 монографии [33, с. 13 – 21]. Приведем краткую сводку, позволяющую выявить роль статистики нечисловых данных.

К 60-м годам XX в. в нашей стране сформировалась научно-практическая дисциплина, которую называем классической математической статистикой. Новое поколение училось теории по фундаментальной монографии шведского математика Г. Крамера [34], написанной в военные годы и впервые изданной на русском языке в 1948 г. Из прикладных руководств назовем учебник [35] и таблицы с комментариями [36].

Затем внимание многих специалистов сосредоточилось на изучении математических конструкций, используемых в статистике. Пример таких работ — монография [37], в которой представлены продвинутое математические результаты, но трудно выделить рекомендации для статистика, анализирующего конкретные данные.

Как реакция на уход в математику выделилась новая научная дисциплина — прикладная статистика. В базовом учебнике по прикладной статистике [24] в качестве рубежа, когда это стало очевидным, указан 1981 г. — дата выхода массовым тиражом (33 940 экз.) сборника [38], в названии которого использован термин «прикладная статистика». С этого времени линии развития математической статистики и прикладной статистики разошлись. Первая из этих дисциплин полностью ушла в математику, потеряв интерес к практическим делам. Вторая [24] позиционировала себя в качестве науки об обработке данных — результатов наблюдений, измерений, испытаний, анализов, опытов, обследований.

Вполне естественно, что в прикладной статистике стали развиваться математические методы и модели. Необходимость их развития вытекает из потребностей конкретных прикладных исследований. Это математизированное ядро прикладной статистики назовем теоретической статистикой. Тогда под собственно прикладной статистикой следует понимать обширную промежуточную область между теоретической статистикой и применением статистических методов в конкретных областях. В нее входят, в частности, вопросы формирования вероятностно-статистических моделей и выбора конкретных методов анализа данных (т.е. методология прикладной статистики и других статистических методов), проблемы разработки и применения информационных статистических технологий, организации сбора и анализа данных, т.е. разработки статистических технологий.

Таким образом, общая схема современной статистической науки выглядит следующим образом (от абстрактного к конкретному).

1. Математическая статистика — часть математики, изучающая статистические структуры. Сама по себе не дает рецептов анализа статистических данных, однако разрабатывает методы, полезные для использования в теоретической статистике.

2. Теоретическая статистика — наука, посвященная моделям и методам анализа конкретных статистических данных.

3. Прикладная статистика (в узком смысле) посвящена статистическим технологиям сбора и обработки данных. Она включает в себя методологию статистических методов, вопросы организации выборочных исследований, разработки статистических технологий, создания и использования статистических программных продуктов.

4. Применение статистических методов в конкретных областях (в экономике и менеджменте — эконометрика, в биологии — биометрика, в химии — хеометрия, в технических исследованиях — технометрика и т.д.).

Часто позиции 2 и 3 вместе называют прикладной статистикой. Иногда позицию 1 именуют теоретической статистикой. Эти терминологические расхождения связаны с тем, что описанное выше развитие рассматриваемой научно-прикладной области не сразу, не полностью и не всегда адекватно отражается в сознании специалистов. Так, до сих пор выпускают учебники, соответствующие старой парадигме — уровню представлений середины XX века.

Примечание. Мы уточнили схему внутреннего деления статистической теории, предложенную в [39]. Естественный смысл приобрели термины «теоретическая статистика» и «прикладная статистика» (в узком смысле). Однако необходимо иметь в виду, что в базовом учебнике [24] прикладная статистика понимается в широком смысле, т.е. как объединение позиций 2 и 3. К сожалению, в настоящее время невозможно отождествить теоретическую статистику с математической, поскольку последняя (как часть математики — научной специальности «теория вероятностей и математическая статистика») заметно оторвалась от задач практики.

Отметим, что математическая статистика, как и теоретическая с прикладной, заметно отличается от ведомственной науки органов официальной государственной статистики. ЦСУ, Госкомстат, Росстат применяли и применяют лишь проверенные временем приемы XIX века. Возможно, следовало бы от этого ведомства полностью отмежеваться и сменить название научной дисциплины, например, на «Анализ данных». В настоящее время компромиссным является термин «статистические методы».

Во второй половине 80-х годов развернулось общественное движение в целях создания профессионального объединения статистиков. Аналогами являются британское Королевское статистическое общество (основано в 1834 г.) и Американская статистическая ассоциация (создана в 1839 г.). К сожалению, деятельность учрежденной в 1990 г. Всесоюзной статистической ассоциации оказалась парализованной в результате развала СССР. Среди стран СНГ наибольшую активность в настоящее время проявляют узбекские исследователи, регулярно проводящие представительные конференции по статистике и ее применениям.

О развитии статистики нечисловых данных

С 70-х годов XX в. в основном на основе запросов теории экспертных оценок (а также технических исследований, экономики, социологии и медицины) развивались различные направления статистики нечисловых данных. Были уста-

новлены основные связи между конкретными видами таких объектов, разработаны для них базовые вероятностные модели [40]. Сводка полученных результатов дана в монографии [2], препринте [41]. Это — предыстория статистики нечисловых данных. А история начинается с осмысления созданного, констатации [2, 3] в 1979 г. появления новой области прикладной статистики.

Следующий этап (80-е годы) — развитие статистики нечисловых данных в качестве самостоятельного научного направления в рамках математических методов исследования, ядром которого являются методы статистического анализа данных произвольной природы. Для работ этого периода характерна сосредоточенность на внутренних (внутриматематических) проблемах статистики нечисловых данных. Проводили всесоюзные конференции, выпускали монографии, сборники трудов, защищали диссертации (А. И. Орлов, К. А. Пярна, Г. В. Рыданова, Г. А. Сатаров, В. А. Трофимов, А. П. Шер, Д. С. Шмерлинг и др.). Наиболее представительным является сборник [42], подготовленный комиссией «Статистика объектов нечисловой природы» Научного Совета АН СССР по комплексной проблеме «Кибернетика» совместно с Институтом социологических исследований АН СССР. Конкретная информация по работам 80-х годов имеется в обзорах [5, 6].

В настоящее время в связи с активным использованием наукометрических показателей разнообразными администраторами научной деятельности распространилась преувеличенная оценка роли журналов в развитии науки. Опыт статистики нечисловых данных показывает, что естественная цепочка развития научного результата такова: тезисы доклада — тематический сборник — монография — учебник — широкое использование [43, 44]. Для развития нового направления публикации в научных журналах, вообще говоря, не обязательны. Ясно, что издание собственных журналов или завоевание позиций в уже существующих возможно лишь на этапе зрелости нового направления, но не на этапе его создания.

К 90-м годам статистика нечисловых данных с теоретической точки зрения была достаточно хорошо развита, основные идеи, подходы и методы были разработаны и изучены математически, в частности, доказано достаточно много теорем. Однако она оставалась недостаточно апробированной на практике. И в 90-е годы наступило время перехода от теоретико-статистических исследований к применению полученных результатов на практике, а также включения их в учебный процесс, что и было сделано. В 90-е годы в нашем журнале опубликованы обзоры [5, 12, 40]

по статистике объектов нечисловой природы и многочисленные конкретные исследования, рассмотренные в [6].

2000-е годы отмечены появлением развернутых публикаций основных результатов статистики интервальных данных в учебниках по прикладной статистике, теории принятия решений, эконометрике (см., например, [24, 25, 31]). Выпущен первый учебник по статистике интервальных данных [7].

В 2010-е годы представленная научной общественности новая парадигма математических методов исследования закрепила положение статистики нечисловых данных как центральной быстро растущей части современной прикладной статистики [1].

Статистика в пространствах произвольного вида

В чем принципиальная новизна статистики нечисловых данных? Для классической математической статистики характерна операция сложения. При расчете выборочных характеристик распределения (выборочного среднего арифметического, выборочной дисперсии и др.), в регрессионном анализе и других областях этой научной дисциплины постоянно используются суммы. Математический аппарат — законы больших чисел, Центральная предельная теорема и другие теоремы — нацелен на изучение сумм. Принципиально важно, что в статистике нечисловых данных нельзя использовать операцию сложения, поскольку элементы выборки лежат в пространствах, где нет этой операции. Методы обработки нечисловых данных основаны на принципиально ином математическом аппарате — на применении различных расстояний в пространствах объектов нечисловой природы.

Следует отметить, что в статистике нечисловых данных одна и та же математическая схема может с успехом применяться во многих прикладных областях, для анализа данных различных типов, а потому ее лучше всего формулировать и изучать в наиболее общем виде, для объектов произвольной природы.

Кратко рассмотрим несколько идей, развиваемых в статистике нечисловых данных для элементов выборок, лежащих в пространствах произвольного вида. Они нацелены на решение классических задач описания данных, оценивания, проверки гипотез, но — для неклассических данных, а потому неклассическими методами.

Первой обсудим проблему определения средних величин. В рамках теории измерений удастся указать вид средних величин, соответствующих тем или иным шкалам измерения [45]. Теория измерений [2, 46, 47], в середине XX в. рас-

смаатривавшаяся как часть математического обеспечения психологии, к настоящему времени признана общенаучной дисциплиной. Проблемы теории измерений постоянно рассматриваются в разделе журнала «Математические методы исследования» [48 – 51].

В классической математической статистике средние величины вводят с помощью операций сложения (выборочное среднее арифметическое, математическое ожидание) или упорядочения (выборочная и теоретическая медианы). В пространствах произвольной природы средние значения нельзя определить с помощью операций сложения или упорядочения. Теоретические и эмпирические средние приходится вводить как решения экстремальных задач. Теоретическое среднее определяется как решение задачи минимизации математического ожидания (в классическом смысле) расстояния от случайного элемента со значениями в рассматриваемом пространстве до фиксированной точки этого пространства (минимизируется указанная функция от этой точки). Для получения эмпирического среднего математического ожидания берется по эмпирическому распределению, т.е. берется сумма расстояний от некоторой точки до элементов выборки и затем минимизируется по этой точке (примером является медиана Кемени [52], методам нахождения которой посвящены недавние работы [53 – 55]). При этом и эмпирическое, и теоретическое средние как решения экстремальных задач могут быть не единственными элементами рассматриваемого пространства, а являться некоторыми множествами таких элементов. Они могут оказаться и пустыми. Тем не менее удалось сформулировать и доказать законы больших чисел для средних величин, определенных указанным образом, т.е. установить сходимость (в специально определенном смысле) эмпирических средних к теоретическим [7, 56, 57].

Оказалось, что методы доказательства законов больших чисел допускают более широкую область применения, чем та, для которой они были разработаны. А именно, удалось изучить [7, 58] асимптотику решений экстремальных статистических задач, к которым, как известно, сводится большинство постановок прикладной статистики. В частности, кроме законов больших чисел, установлена состоятельность оценок минимального контраста, в том числе оценок максимального правдоподобия и робастных оценок. К настоящему времени подобные оценки изучены также и в статистике интервальных данных. Полученные результаты относительно асимптотики решений экстремальных статистических задач применяются, например, в работах [59 – 61].

В статистике в пространствах произвольного вида большую роль играют непараметрические

оценки плотности распределения вероятностей, используемые, в частности, в различных алгоритмах регрессионного, дискриминантного, кластерного анализов. В статистике нечисловых данных предложен и изучен ряд типов непараметрических оценок плотности в пространствах произвольной природы, в том числе в дискретных пространствах [73 – 75]. В частности, доказана их состоятельность, изучена скорость сходимости и установлен (для ядерных оценок плотности) примечательный факт совпадения наилучшей скорости сходимости в произвольном пространстве с той, которая имеет место в классической теории для числовых случайных величин [74].

Введем обозначения. Пусть (Z, A) — измеримое пространство, p и q — сигма-конечные меры на (Z, A) , причем p абсолютно непрерывна относительно q , т.е. из $q(B) = 0$ следует $p(B) = 0$ для любого множества B из сигма-алгебры A . В этом случае на (Z, A) существует неотрицательная измеримая функция $f(x)$ такая, что

$$q(C) = \int_C f(x) p(dx) \quad (1)$$

для любого множества C из сигма-алгебры измеримых множеств A . Функция $f(x)$ называется производной Радона – Никодима меры q по мере p , а в случае, когда q — вероятностная мера, эта функция является также плотностью вероятности q по отношению к мере p [62, с. 460].

Пусть $X_1, X_2, \dots, X_n$ — независимые одинаково распределенные случайные элементы (величины), распределение которых задается вероятностной мерой q . В статьях [63, 64] введено несколько видов непараметрических оценок плотности вероятности q по выборке $X_1, X_2, \dots, X_n$. Подробнее изучены линейные оценки. В статьях [65, 66] рассмотрены их частные случаи — ядерные оценки плотности в пространствах произвольной природы. В статьях [67, 68] асимптотическая теория ядерных оценок плотности развита, прежде всего, для нужд статистики конкретных видов объектов нечисловой природы, в которой основной интерес представляют конечные пространства Z . Мера p при этом не непрерывная, а дискретная, например считающая. Таким образом, в рамках единого подхода удаётся рассмотреть оценки плотностей и оценки вероятностей.

В предположении непрерывности неизвестной плотности $f(x)$ представляется целесообразным «размазать» каждый атом эмпирической меры, т.е. рассмотреть линейные оценки, введен-

ные в первой работе автора по нечисловой статистике [3, с. 24],

$$f_n(x) = \frac{1}{n} \sum_{1 \leq i \leq n} g_n(x, X_i), \quad g_n: Z^2 \rightarrow R^1, \quad (2)$$

в которых действительнзначные функции g_n удовлетворяют некоторым условиям регулярности.

Пусть d — показатель различия (синоним — мера близости) на Z [4] (в наиболее важных случаях — метрика на Z). В [69] введены ядерные оценки плотности — оценки вида (2) с

$$g_n(x, X_i) = \frac{1}{b(h_n, x)} K\left(\frac{d(x, X_i)}{h_n}\right),$$

$$K: [0, +\infty) \rightarrow R^1, \quad (3)$$

где $K = K(u)$ — ядро (ядерная функция); h_n — последовательность положительных чисел (показателей размытости); $b(h_n, x)$ — нормировочный множитель. В [70] линейные оценки (2) с функциями g_n из (3) названы «обобщенными оценками типа Парзена – Розенблатта», так как в частном случае $Z = R^1$, $d(x, X_i) = |x - X_i|$, $b(h_n, x) = h_n$ они переходят в известные оценки, введенные Розенблаттом [71] и Парзеном [72].

Цель статей [73, 74] — завершение цикла работ, посвященного математическому изучению асимптотических свойств различных видов непараметрических оценок плотности распределения вероятности в пространствах общей природы. Изучен средний квадрат ошибки ядерной оценки плотности. С целью максимизации порядка его убывания обоснован выбор ядерной функции и последовательности показателей размытости. Основные понятия — круговая функция распределения и круговая плотность. Порядок сходимости в общем случае тот же, что и при оценивании плотности числовой случайной величины [75], но основные условия наложены не на плотность случайной величины, а на круговую плотность. Далее рассматриваем другие виды непараметрических оценок плотности — гистограммные оценки и оценки типа Фикс – Ходжеса. Затем изучаем непараметрические оценки регрессии и их применение для решения задач дискриминантного анализа в пространстве общей природы.

Дискриминантный, кластерный, регрессионный анализы в пространствах произвольной природы основаны либо на параметрической теории (и тогда применяется подход, связанный с асимптотикой решения экстремальных статистических задач), либо на непараметрической теории (и тогда используются алгоритмы на основе непараметрических оценок плотности).

Для анализа нечисловых, в частности, экспертных данных, весьма важны методы классификации [76 – 78]. Интересно движение мысли в обратном направлении — наиболее естественно ставить и решать задачи классификации, основанные на использовании расстояний или показателей различия, именно в рамках статистики объектов нечисловой природы (а не, скажем, многомерного статистического анализа). Это касается как распознавания образов с учителем (другими словами, дискриминантного анализа), так и распознавания образов без учителя (кластерного анализа). Аналогичным образом задачи многомерного шкалирования, т.е. визуализации данных [79 – 81], также естественно отнести к статистике объектов нечисловой природы. Важны методы оценки истинной размерности признакового пространства [82].

Отметим несколько конкретных научных результатов математической теории классификации. В задачах диагностики (дискриминантного анализа), как следует из леммы Неймана – Пирсона, целесообразно строить алгоритмы на основе отношения непараметрических оценок плотностей распределения вероятностей, соответствующих классам [83]. Установлено, что наилучшим показателем качества алгоритма диагностики является прогностическая сила [84]. Устойчивость классификации относительно выбора метода кластер-анализа обосновывает вывод о реальности кластеров [85], и т.д. (см. соответствующий раздел в обзоре [1]).

Для проверки гипотез в пространствах нечисловой природы могут быть использованы статистики интегрального типа [3, 86], в частности, типа омега-квадрат [87, 88]. Отметим, что предельная теория таких статистик, построенная первоначально в классической постановке, приобрела естественный (завершенный, изящный) вид именно для пространств произвольного вида [89, 90], поскольку при этом удалось провести рассуждения, опираясь на базовые математические соотношения, а не на те частные (с общей точки зрения), что были связаны с конечномерным пространством.

О некоторых областях статистики нечисловых данных

Кратко рассмотрим некоторые статистические методы анализа данных [91], лежащих в конкретных пространствах нечисловой природы.

Непараметрическая статистика — это прежде всего ранговая статистика, т.е. основанная на рангах — номерах элементов выборок в вариационных рядах. Ранги измерены в порядковых шкалах, а значения ранговых статистик инвариантны относительно любых строго возрастающих

преобразований — допустимых преобразований в таких шкалах. Непараметрическая статистика позволяет делать статистические выводы, оценивать характеристики и плотность распределения, проверять статистические гипотезы без слабо обоснованных предположений о том, что функция распределения элементов выборки входит в то или иное параметрическое семейство. Например, широко распространена вера в то, что статистические данные часто подчиняются нормальному распределению. Математики думают, что это — экспериментальный факт, установленный в прикладных исследованиях. Прикладники уверены, что математики доказали нормальность результатов наблюдений. Между тем анализ конкретных результатов наблюдений, в частности, погрешностей измерений, приводит всегда к одному и тому же выводу — в подавляющем большинстве случаев реальные распределения существенно отличаются от нормальных. На этот объективный факт обращал внимание В. В. Налимов в своей классической монографии [92]. Научная школа метролога П. В. Новицкого многочисленными экспериментами подтвердила отсутствие нормальности погрешностей измерений [93]. Опубликованная в журнале сводка [94] (см. также [95]) включена в учебники [24, 31]. В [96] установлено, что по выборкам объемов 6 – 50, как правило, не удастся отличить нормальное распределение от других видов распределений.

Некритическое использование гипотезы нормальности часто приводит к значительным ошибкам, например, при отбраковке резко выделяющихся результатов наблюдений (выбросов), при статистическом контроле качества и в других случаях [24]. Поэтому целесообразно использовать непараметрические методы, в которых на функции распределения результатов наблюдений наложены весьма слабые требования. Обычно предполагается лишь их непрерывность. К настоящему времени с помощью непараметрических методов можно решать практически тот же круг задач, что ранее решался параметрическими методами. Примеры — оценивание характеристик распределения [97] и проверка гипотезы однородности для независимых [87, 99, 100] и связанных [88] выборок. Однако эта информация еще не вошла в массовое сознание. До сих пор тупиковой тематике параметрической статистики посвящены обширные разделы учебников и программных продуктов.

Современное состояние непараметрической статистики рассмотрено в [101, 102]. Эта область исследований продолжает активно развиваться. Отметим научные результаты как внутри неё [103, 104], так и на стыке с другими областями [105, 106].

Представляют практический интерес результаты, связанные с конкретными областями статистики объектов нечисловой природы, в частности, со статистикой нечетких множеств [107] и со статистикой случайных множеств (напомним, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств [32]), с непараметрической теорией парных сравнений и люсианов (бернуллиевских бинарных векторов) [108], с аксиоматическим введением метрик в конкретных пространствах объектов нечисловой природы [109], а также с рядом других конкретных постановок.

Результаты контроля штучной продукции по альтернативному признаку представляют собой последовательности из 0 и 1 — объекты нечисловой природы (люсианы), а потому теорию статистического контроля относят к статистике нечисловых данных [5, 6]. Работы по этой тематике предназначены для специалистов по статистическим методам управления качеством продукции [110 – 112].

Статистика нечисловых данных порождена потребностями практики, прежде всего в области экспертных оценок. Можно констатировать, что анализ экспертных оценок [113] — это прикладное «зеркало» общей теории. Решения задач теории экспертных оценок обобщались в статистике нечисловых данных. При движении мысли в обратном направлении результаты статистики в пространствах общей природы интерпретировались для анализа экспертных оценок. Как и для статистики нечисловых данных в целом, публикации шли по траектории: тезисы доклада — тематический сборник — монография — учебник — широкое использование [43, 44]. Вполне естественно, что названия сборников трудов неформального научного коллектива, развивающего статистику нечисловых данных, начинались со слов «Экспертные оценки» [114 – 117]. Отметим, что публикации в журналах не сыграли значительной роли в развитии рассматриваемых научных направлений. Обзор развития экспертных технологий в нашей стране дан в статьях [118 – 120].

Вопросы внедрения математических методов исследования всегда были в центре внимания журнала «Заводская лаборатория. Диагностика материалов» [121 – 123]. Подчеркивались большое теоретическое и прикладное значение статистики нечисловых данных [124], необходимости перехода от отдельных методов анализа данных к разработке высоких статистических технологий [125, 126] и использования современных систем внедрения математических методов, таких как система «Шесть сигм» и ее аналоги [127]. Обсуждались проблемы программного обеспечения [128 – 131]. Однако приходится констатировать,

что создание линейки современных программных продуктов по статистике нечисловых данных — пока дело будущего.

Нерешенные задачи статистики нечисловых данных

Начнем с обсуждения влияния отклонений от традиционных предпосылок. В вероятностной теории статистических методов выборка обычно моделируется как конечная последовательность независимых одинаково распределенных случайных величин или векторов. В устаревшей парадигме середины XX в. часто предполагают, что эти величины (векторы) имеют нормальное распределение.

При внимательном взгляде совершенно ясна нереалистичность приведенных классических предпосылок. Независимость результатов измерений обычно принимается «из общих предположений», между тем во многих случаях очевидна их коррелированность. Одинаковая распределенность также вызывает сомнения из-за изменения во времени свойств измеряемых образцов, средств измерения и психофизического состояния специалистов, проводящих измерения (испытания, анализы, опыты). Даже обоснованность самого применения вероятностных моделей иногда вызывает сомнения, например, при моделировании уникальных измерений (согласно классическим воззрениям, теорию вероятностей обычно привлекают при изучении массовых явлений). И уж совсем редко распределения результатов измерений можно считать нормальными [24, 31].

Итак, методы классической математической статистики обычно используют вне сферы их обоснованной применимости. Каково влияние отклонений от традиционных предпосылок на статистические выводы? В настоящее время об этом имеются лишь отрывочные сведения. Так, три примера в статье [6] показывают весь спектр возможных свойств классических расчетных методов в случае отклонения от нормальности. Методы построения доверительного интервала для математического ожидания оказываются вполне пригодными при таких отклонениях. Методы проверки однородности двух независимых выборок с помощью двухвыборочного критерия Стьюдента пригодны в некоторых случаях. В задаче отбраковки (исключения) резко выделяющихся наблюдений (выбросов) расчетные методы, основанные на нормальности, оказались полностью непригодными.

Итак, отметим необходимость изучения свойств расчетных методов классической математической статистики, опирающихся на предположение нормальности, в ситуациях, когда это

предположение не выполнено. Аппаратом для такого изучения наряду с методом Монте-Карло могут послужить предельные теоремы теории вероятностей, прежде всего ЦПТ, поскольку интересующие нас расчетные методы обычно используют разнообразные суммы. Пока подобное изучение не проведено, остается неясной научная ценность, например, применения основанного на предположении многомерной нормальности факторного анализа к векторам из переменных, принимающих небольшое число градаций и к тому же измеренных в порядковой шкале.

Нерешенным проблемам статистики посвящены статьи [132, 133]. Одна из важных проблем — использование асимптотических результатов при конечных объемах выборок. Конечно, естественно изучить свойства алгоритма с помощью метода Монте-Карло. Однако из какого конкретного распределения брать выборки при моделировании? От выбора распределения зависит результат. Кроме того, датчики псевдослучайных чисел лишь имитируют случайность. До сих пор неизвестно, каким датчиком целесообразно пользоваться в случае возможного безграничного роста размерности пространства. В нашем журнале продолжается дискуссия, начатая статьями [134, 135].

Другая проблема — обоснование выбора одного из многих критериев для проверки конкретной гипотезы. Например, для проверки однородности двух независимых выборок можно использовать критерии Стьюдента, Крамера – Уэлча, Лорда, хи-квадрат, Вилкоксона (Манна – Уитни), Ван-дер-Вардена, Сэвиджа, Н. В. Смирнова, типа омега-квадрат (Лемана – Розенблатта), Реньи, Г. В. Мартынова и др. [87, 103]. Какой выбрать?

Критерии однородности проанализированы в [136]. Естественных подходов к сравнению критериев несколько — на основе асимптотической относительной эффективности по Бахадуру, Ходжесу – Леману, Питмену. И каждый из перечисленных критериев является оптимальным при соответствующей альтернативе или подходящем распределении на множестве альтернатив. При этом математические выкладки обычно используют альтернативу сдвига, сравнительно редко встречающуюся в практике анализа реальных статистических данных. Итог печален — блестящая математическая техника, продемонстрированная в [136], не позволяет дать рекомендации для выбора критерия проверки однородности при анализе реальных данных.

Проблемы разработки высоких статистических технологий поставлены в [125, 126] (см. также одноименный сайт <http://orlovs.pp.ru>). Используемые при обработке реальных данных статистические технологии состоят из последовательности операций, каждая из которых, как пра-

вило, хорошо изучена, поскольку сводится к оцениванию (параметров, характеристик, распределений) или проверке той или иной гипотезы. Однако статистические свойства результатов обработки, полученных в ходе последовательного применения таких операций, мало изучены [137]. Необходима теория, позволяющая изучать свойства статистических технологий и так их конструировать, чтобы обеспечить высокое качество обработки данных.

В заключение отметим, что развернутое описание статистики нечисловых данных дано в монографиях [7, 24, 31]. При дальнейшем развитии исследований важно опираться на современную методологию [14, 138].

ЛИТЕРАТУРА

- Орлов А. И. Развитие математических методов исследования (2006 – 2015 гг.) / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 1. Ч. 1. С. 78 – 86.
- Орлов А. И. Устойчивость в социально-экономических моделях. — М.: Наука, 1979. — 296 с.
- Орлов А. И. Статистика объектов нечисловой природы и экспертные оценки / Экспертные оценки. Вопросы кибернетики. Вып. 58. — М.: Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1979. С. 17 – 33.
- Тюрин Ю. Н., Литвак Б. Г., Орлов А. И., Сатаров Г. А., Шмерлинг Д. С. Анализ нечисловой информации / Заводская лаборатория. 1980. Т. 46. № 10. С. 931 – 935.
- Орлов А. И. Статистика объектов нечисловой природы (обзор) / Заводская лаборатория. 1990. Т. 56. № 3. С. 76 – 83.
- Орлов А. И. Тридцать лет статистики объектов нечисловой природы (обзор) / Заводская лаборатория. Диагностика материалов. 2009. Т. 75. № 5. С. 55 – 64.
- Орлов А. И. Организационно-экономическое моделирование. Ч. 1. Нечисловая статистика. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2009. — 544 с.
- Кун Т. Структура научных революций. — М.: АСТ, 2003. — 605 с.
- Орлов А. И. Новая парадигма прикладной статистики / Заводская лаборатория. Диагностика материалов. 2012. Т. 78. № 1. Ч. I. С. 87 – 93.
- Орлов А. И. Новая парадигма математических методов исследования / Заводская лаборатория. Диагностика материалов. 2015. Т. 81. № 7. С. 5.
- Орлов А. И. О новой парадигме математических методов исследования / Научный журнал КубГАУ. 2016. № 122. С. 807 – 832.
- Орлов А. И. Объекты нечисловой природы / Заводская лаборатория. Диагностика материалов. 1995. Т. 61. № 3. С. 43 – 52.
- Бернштейн С. Н. Современное состояние теории вероятностей и ее приложений / Труды Всероссийского съезда математиков в Москве 27 апреля – 4 мая 1927 г. — М. — Л.: ГИЗ, 1928. С. 50 – 63.
- Новиков А. М., Новиков Д. А. Методология. — М.: СИНТЕГ, 2007. — 668 с.
- Дискуссия по анализу интервальных данных / Заводская лаборатория. 1990. Т. 56. № 7. С. 75 – 95.
- Воцинин А. П. Метод оптимизации объектов по интервальным моделям целевой функции. — М.: МЭИ, 1987. — 109 с.
- Воцинин А. П., Сотиров Г. Р. Оптимизация в условиях неопределенности. — М.: МЭИ; — София: Техника, 1989. — 224 с.
- Воцинин А. П., Акматбеков Р. А. Оптимизация по регрессионным моделям и планирование эксперимента. — Бишкек: Илим, 1991. — 164 с.
- Воцинин А. П. Метод анализа данных с интервальными ошибками в задачах проверки гипотез и оценивания параметров неявных линейно параметризованных функций / Заводская лаборатория. Диагностика материалов. 2000. Т. 66. № 3. С. 51 – 64.
- Воцинин А. П. Интервальный анализ данных: развитие и перспективы / Заводская лаборатория. Диагностика материалов. 2002. Т. 68. № 1. С. 118 – 126.
- Воцинин А. П., Бронз П. В. Построение аналитических моделей по данным вычислительного эксперимента в задачах анализа чувствительности и оценки экономических рисков / Заводская лаборатория. Диагностика материалов. 2007. Т. 73. № 1. С. 101 – 109.
- Воцинин А. П., Скибицкий Н. В. Интервальный подход к выражению неопределенности измерений и калибровке цифровых измерительных систем / Заводская лаборатория. Диагностика материалов. 2007. Т. 73. № 11. С. 66 – 71.
- Гуськова Е. А., Орлов А. И. Интервальная линейная парная регрессия (обобщающая статья) / Заводская лаборатория. Диагностика материалов. 2005. Т. 71. № 3. С. 57 – 63.
- Орлов А. И. Прикладная статистика. — М.: Экзамен, 2006. — 671 с.
- Орлов А. И. Теория принятия решений. — М.: Экзамен, 2006. — 576 с.
- Орлов А. И., Луценко Е. В. Системная нечеткая интервальная математика. — Краснодар: КубГАУ, 2014. — 600 с.
- Орлов А. И. Статистика интервальных данных (обобщающая статья) / Заводская лаборатория. Диагностика материалов. 2015. Т. 81. № 3. С. 61 – 69.
- Скибицкий Н. В., Севальнев Н. В. Интервальные модели в задачах оптимального управления с дифференциальными связями / Заводская лаборатория. Диагностика материалов. 2015. Т. 81. № 11. С. 73 – 80.
- Скибицкий Н. В. Построение прямых и обратных статистических характеристик объектов по интервальным данным / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 1. Ч. I. С. 87 – 93.
- Таранцев А. А. О связи интервального анализа с теорией вероятностей / Заводская лаборатория. 2004. Т. 70. № 3. С. 60 – 65.
- Орлов А. И. Эконометрика. — М.: Экзамен, 2002. — 576 с.
- Орлов А. И. Теория нечетких множеств — часть теории вероятностей / Научный журнал КубГАУ. 2013. № 92. С. 51 – 60.
- Орлов А. И., Луценко Е. В., Лойко В. И. Организационно-экономическое, математическое и программное обеспечение контроллинга, инноваций и менеджмента: монография. — Краснодар: КубГАУ, 2016. — 600 с.
- Крамер Г. Математические методы статистики. — М.: Мир, 1975. — 648 с.
- Смирнов Н. В., Душин-Барковский И. В. Курс теории вероятностей и математической статистики для технических приложений. Изд. 3-е, стереотипное. — М.: Наука, 1969. — 512 с.
- Большев Л. Н., Смирнов Н. В. Таблицы математической статистики. Изд. 3-е. — М.: Наука, 1983. — 416 с.
- Каган А. М., Линник Ю. В., Рао С. Р. Характеризационные задачи математической статистики. — М.: Наука, 1972. — 656 с.
- Современные проблемы кибернетики (прикладная статистика). — М.: Знание, 1981. — 64 с.
- Орлов А. И. О перестройке статистической науки и её применений / Вестник статистики. 1990. № 1. С. 65 – 71.
- Орлов А. И. Вероятностные модели конкретных видов объектов нечисловой природы / Заводская лаборатория. Диагностика материалов. 1995. Т. 61. № 5. С. 43 – 51.
- Тюрин Ю. Н., Литвак Б. Г., Орлов А. И., Сатаров Г. А., Шмерлинг Д. С. Анализ нечисловой информации. — М.: Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1981. — 80 с.
- Андреев В. Г., Орлов А. И., Толстова Ю. Н. (ответственные редакторы). Анализ нечисловой информации в социологических исследованиях. — М.: Наука, 1985. — 220 с.

43. Орлов А. И. Критерии выбора показателей эффективности научной деятельности / Контроллинг. 2013. № 3(49). С. 72 – 78.
44. Лойко В. И., Луценко Е. В., Орлов А. И. Современные подходы в наукометрии. — Краснодар: КубГАУ, 2017. — 532 с.
45. Орлов А. И. Характеризация средних величин шкалами измерения / Научный журнал КубГАУ. 2017. № 134. С. 877 – 907.
46. Психологические измерения: Сб. статей. — М.: Мир, 1967. — 196 с.
47. Пфанцгаль И. Теория измерений. — М.: Мир, 1976. — 248 с.
48. Толстова Ю. Н. Краткая история развития репрезентативной теории измерений / Заводская лаборатория. Диагностика материалов. 1999. Т. 65. № 3. С. 49 – 56.
49. Орлов А. И. Репрезентативная теория измерений и ее применения / Заводская лаборатория. Диагностика материалов. 1999. Т. 65. № 3. С. 57 – 62.
50. Барский Б. В., Соколов М. В. Средние величины, инвариантные относительно допустимых преобразований шкалы измерения / Заводская лаборатория. Диагностика материалов. 2006. Т. 72. № 1. С. 59 – 66.
51. Орлов А. И. Математические методы исследования и теория измерений / Заводская лаборатория. Диагностика материалов. 2006. Т. 72. № 1. С. 67 – 70.
52. Кемени Дж., Снелл Дж. Кибернетическое моделирование: Некоторые приложения. — М.: Советское радио, 1972. — 192 с.
53. Жуков М. С., Орлов А. И. Использование экспертных ранжировок при расчетах кредитного риска в банке / Инновации в менеджменте. 2017. № 1. С. 18 – 25.
54. Жуков М. С., Орлов А. И. Задача исследования итогового ранжирования мнений группы экспертов с помощью медианы Кемени / Научный журнал КубГАУ. 2016. № 122. С. 785 – 806.
55. Жуков М. С., Орлов А. И., Фалько С. Г. Экспертные оценки в рисках / Контроллинг. 2017. № 4(66). С. 24 – 27.
56. Орлов А. И. Средние величины и законы больших чисел в пространствах произвольной природы / Научный журнал КубГАУ. 2013. № 89. С. 556 – 586.
57. Орлов А. И. О средних величинах / Управление большими системами. Выпуск 46. — М.: ИПУ РАН, 2013. С. 88 – 117.
58. Орлов А. И. Предельная теория решений экстремальных статистических задач / Научный журнал КубГАУ. 2017. № 133. С. 579 – 600.
59. Орлов А. И. Методы поиска наиболее информативных множеств признаков в регрессионном анализе / Заводская лаборатория. Диагностика материалов. 1995. Т. 61. № 1. С. 56 – 58.
60. Орлов А. И. Асимптотическое поведение решений экстремальных статистических задач / Заводская лаборатория. Диагностика материалов. 1996. Т. 62. № 10. С. 45 – 46.
61. Тырсин А. Н. Робастное построение линейных регрессионных моделей по экспериментальным данным / Заводская лаборатория. Диагностика материалов. 2005. Т. 71. № 11. С. 53 – 58.
62. Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю. В. Прохоров. — М.: Большая Российская Энциклопедия, 1999. — 910 с.
63. Орлов А. И. Оценки плотности в пространствах произвольной природы / Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. Вып. 25. — Пермь: Перм. гос. нац. иссл. ун-т, 2013. С. 21 – 33.
64. Орлов А. И. Оценки плотности распределения вероятностей в пространствах произвольной природы / Научный журнал КубГАУ. 2014. № 99. С. 15 – 32.
65. Орлов А. И. Ядерные оценки плотности в пространствах произвольной природы / Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. Вып. 26. — Пермь: Перм. гос. нац. иссл. ун-т, 2015. С. 43 – 57.
66. Орлов А. И. Предельные теоремы для ядерных оценок плотности в пространствах произвольной природы / Научный журнал КубГАУ. 2015. № 108. С. 316 – 333.
67. Орлов А. И. Непараметрические ядерные оценки плотности вероятности в дискретных пространствах / Научный журнал КубГАУ. 2016. № 122. С. 833 – 855.
68. Орлов А. И. Ядерные оценки плотности в конечных пространствах / Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. Вып. 27. — Пермь: Перм. гос. нац. иссл. ун-т, 2016. С. 24 – 37.
69. Орлов А. И. Статистика объектов нечисловой природы / Теория вероятностей и ее применения. 1980. Т. XXV. № 3. С. 655 – 656.
70. Орлов А. И. Непараметрические оценки плотности в топологических пространствах / Прикладная статистика. Ученые записки по статистике. Т. 45. — М.: Наука, 1983. С. 12 – 40.
71. Rosenblatt M. Remarks on some nonparametric estimates of a density function / Ann. Math. Statist. 1956. Vol. 27. N 5. P 832 – 837.
72. Parzen E. On estimation of a probability density function and mode / Ann. Math. Statist. 1962. Vol. 33. N 6. P 1065 – 1076.
73. Орлов А. И. Асимптотика оценок плотности распределения вероятностей / Научный журнал КубГАУ. 2017. № 131. С. 845 – 873.
74. Орлов А. И. Скорость сходимости ядерных оценок плотности в пространствах произвольной природы / Статистические методы оценивания и проверки гипотез: межвуз. сб. науч. тр. Вып. 28. — Пермь: Перм. гос. нац. иссл. ун-т, 2018. С. 35 – 45.
75. Ибрагимов И. А., Хасьминский Р. З. Асимптотическая теория оценивания. — М.: Наука, 1979. — 528 с.
76. Орлов А. И. О развитии математических методов теории классификации / Заводская лаборатория. Диагностика материалов. 2009. Т. 75. № 7. С. 51 – 63.
77. Орлов А. И., Толчеев В. О. Об использовании непараметрических статистических критериев для оценки точности методов классификации (обобщающая статья) / Заводская лаборатория. Диагностика материалов. 2011. Т. 77. № 3. С. 58 – 66.
78. Орлов А. И. Три основных результата математической теории классификации / Заводская лаборатория. Диагностика материалов. 2016. Т. 82. № 5. С. 63 – 70.
79. Перекрест В. Т. Нелинейный типологический анализ социально-экономической информации: Математические и вычислительные методы. — Л.: Наука, 1983. — 176 с.
80. Терехина А. Ю. Анализ данных методами многомерного шкалирования. — М.: Наука, 1986. — 168 с.
81. Лагутин М. Б. Визуальное представление тесноты связей / Заводская лаборатория. Диагностика материалов. 2005. Т. 71. № 7. С. 53 – 57.
82. Луценко Е. В., Орлов А. И. Методы снижения размерности пространства статистических данных / Научный журнал КубГАУ. 2016. № 119. С. 92 – 107.
83. Орлов А. И. Математические методы исследования и диагностика материалов (Обобщающая статья) / Заводская лаборатория. Диагностика материалов. 2003. Т. 69. № 3. С. 53 – 64.
84. Орлов А. И. Прогностическая сила — наилучший показатель качества алгоритма диагностики / Научный журнал КубГАУ. 2014. № 99. С. 33 – 49.
85. Орлов А. И. Устойчивость классификации относительно выбора метода кластер-анализа / Заводская лаборатория. Диагностика материалов. 2013. Т. 79. № 1. С. 68 – 71.
86. Орлов А. И. Асимптотическое поведение статистик интегрального типа / Доклады АН СССР. 1974. Т. 219. № 4. С. 808 – 811.
87. Орлов А. И. О проверке однородности двух независимых выборок / Заводская лаборатория. Диагностика материалов. 2003. Т. 69. № 1. С. 55 – 60.
88. Орлов А. И. Методы проверки однородности связанных выборок / Заводская лаборатория. Диагностика материалов. 2004. Т. 70. № 7. С. 57 – 61.

89. Орлов А. И. Асимптотическое поведение статистик интегрального типа / Вероятностные процессы и их приложения. Межвузовский сборник научных трудов. — М.: МИЭМ, 1989. С. 118 – 123.
90. Орлов А. И. Предельная теория непараметрических статистик / Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета. 2014. № 100. С. 31 – 52.
91. Орлов А. И. Организационно-экономическое моделирование: учебник: в 3 ч. Ч. 3. Статистические методы анализа данных. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2012. — 624 с.
92. Налимов В. В. Применение математической статистики при анализе вещества. — М.: Физматгиз, 1960. — 430 с.
93. Новицкий П. В., Зограф И. А. Оценка погрешностей результатов измерений. — Л.: Энергоатомиздат, 1985. — 248 с.
94. Орлов А. И. Часто ли распределение результатов наблюдений является нормальным? / Заводская лаборатория. 1991. Т. 57. № 7. С. 64 – 66.
95. Орлов А. И. Распределения реальных статистических данных не являются нормальными / Научный журнал КубГАУ. 2016. № 117. С. 71 – 90.
96. Селезнев В. Д., Денисов К. С. Исследование свойств критериев согласия функции распределения данных с гауссовой методом Монте-Карло для малых выборок / Заводская лаборатория. Диагностика материалов. 2005. Т. 71. № 1. С. 68 – 73.
97. Орлов А. И. Непараметрическое точечное и интервальное оценивание характеристик распределения / Заводская лаборатория. Диагностика материалов. 2004. Т. 70. № 5. С. 65 – 70.
98. Орлов А. И. Проверка статистической гипотезы однородности математических ожиданий двух независимых выборок: критерий Крамера – Уэлча вместо критерия Стьюдента / Научный журнал КубГАУ. 2015. № 110. С. 197 – 218.
99. Орлов А. И. Двухвыборочный критерий Вилкоксона — анализ двух мифов / Научный журнал КубГАУ. 2014. № 104. С. 91 – 111.
100. Орлов А. И. Состоятельные критерии проверки абсолютной однородности независимых выборок / Заводская лаборатория. Диагностика материалов. 2012. Т. 78. № 11. С. 66 – 70.
101. Орлов А. И. Структура непараметрической статистики (обобщающая статья) / Заводская лаборатория. Диагностика материалов. 2015. Т. 81. № 7. С. 62 – 72.
102. Орлов А. И. Современное состояние непараметрической статистики / Научный журнал КубГАУ. 2015. № 106. С. 239 – 269.
103. Орлов А. И. Реальные и номинальные уровни значимости при проверке статистических гипотез / Научный журнал КубГАУ. 2015. № 114. С. 42 – 54.
104. Орлов А. И. Модель анализа совпадений при расчете непараметрических ранговых статистик / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 11. С. 66 – 72.
105. Муравьева В. С., Орлов А. И. Непараметрическое оценивание точки пересечения регрессионных прямых / Заводская лаборатория. Диагностика материалов. 2008. Т. 74. № 1. С. 63 – 68.
106. Орлов А. И. Непараметрический метод наименьших квадратов с периодической составляющей / Заводская лаборатория. Диагностика материалов. 2014. Т. 80. № 1. С. 65 – 75.
107. Орлов А. И. Статистика нечетких данных / Научный журнал КубГАУ. 2016. № 119. С. 75 – 91.
108. Орлов А. И. Теория люсианов / Научный журнал КубГАУ. 2014. № 101. С. 275 – 304.
109. Орлов А. И. Расстояния в пространствах статистических данных / Научный журнал КубГАУ. 2014. № 101. С. 227 – 252.
110. Орлов А. И. Статистический контроль по двум альтернативным признакам и метод проверки их независимости по совокупности малых выборок / Заводская лаборатория. Диагностика материалов. 2000. Т. 66. № 1. С. 58 – 62.
111. Орлов А. И. Основные проблемы контроллинга качества / Научный журнал КубГАУ. 2015. № 111. С. 20 – 52.
112. Орлов А. И. Предельные теоремы в статистическом контроле / Научный журнал КубГАУ. 2016. № 116. С. 462 – 483.
113. Орлов А. И. Организационно-экономическое моделирование: учебник. В 3 ч. Ч. 2. Экспертные оценки. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2011. — 486 с.
114. Статистические методы анализа экспертных оценок / Ученые записки по статистике. Т. 29. — М.: Наука, 1977. — 385 с.
115. Экспертные оценки / Вопросы кибернетики. Вып. 58. — М.: Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1979. — 200 с.
116. Экспертные оценки в системных исследованиях / Сб. тр. Вып. 4. — М.: ВНИИСИ, 1979. — 120 с.
117. Экспертные оценки в задачах управления / Сб. тр. — М.: Институт проблем управления, 1982. — 106 с.
118. Орлов А. И. Экспертные оценки / Заводская лаборатория. Диагностика материалов. 1996. Т. 62. № 1. С. 54 – 60.
119. Орлов А. И. О развитии экспертных технологий в нашей стране / Заводская лаборатория. Диагностика материалов. 2010. Т. 76. № 11. С. 64 – 70.
120. Орлов А. И. Теория экспертных оценок в нашей стране / Научный журнал КубГАУ. 2013. № 93. С. 1 – 11.
121. Гнеденко Б. В., Орлов А. И. Роль математических методов исследования в кардинальном ускорении научно-технического прогресса / Заводская лаборатория. 1988. Т. 54. № 1. С. 1 – 4.
122. Орлов А. И. О современных проблемах внедрения прикладной статистики и других статистических методов / Заводская лаборатория. 1992. Т. 58. № 1. С. 67 – 74.
123. Орлов А. И. Сертификация и статистические методы (обобщающая статья) / Заводская лаборатория. Диагностика материалов. 1997. Т. 63. № 3. С. 55 – 62.
124. Горский В. Г., Орлов А. И. Математические методы исследования: итоги и перспективы / Заводская лаборатория. Диагностика материалов. 2002. Т. 68. № 1. С. 108 – 112.
125. Орлов А. И. Высокие статистические технологии / Заводская лаборатория. Диагностика материалов. 2003. Т. 69. № 11. С. 55 – 60.
126. Орлов А. И. О высоких статистических технологиях / Научный журнал КубГАУ. 2015. № 105. С. 14 – 38.
127. Орлов А. И. «Шесть сигм» — новая система внедрения математических методов исследования / Заводская лаборатория. Диагностика материалов. 2006. Т. 72. № 5. С. 50 – 53.
128. Орлов А. И. Математическое обеспечение сертификации: сравнительный анализ диалоговых систем по статистическому контролю / Заводская лаборатория. Диагностика материалов. 1996. Т. 62. № 7. С. 46 – 49.
129. Смирнова О. С. Программное обеспечение для статистического анализа / Заводская лаборатория. Диагностика материалов. 2008. Т. 74. № 5. С. 68 – 75.
130. Орлов А. И. Статистические пакеты — инструменты исследователя / Заводская лаборатория. Диагностика материалов. 2008. Т. 74. № 5. С. 76 – 78.
131. Жуков М. С. Об алгоритмах расчета медианы Кемени / Заводская лаборатория. Диагностика материалов. 2017. Т. 83. № 7. С. 72 – 78.
132. Загоруйко Н. Г., Орлов А. И. Некоторые нерешенные математические задачи прикладной статистики / Современные проблемы кибернетики (прикладная статистика). — М.: Знание, 1981. С. 53 – 63.
133. Орлов А. И. Некоторые нерешенные вопросы в области математических методов исследования / Заводская лаборатория. Диагностика материалов. 2002. Т. 68. № 3. С. 52 – 56.
134. Орлов А. И. Предельные теоремы и метод Монте-Карло / Заводская лаборатория. Диагностика материалов. 2016. Т. 82. № 7. С. 67 – 72.
135. Григорьев Ю. Д. Метод Монте-Карло: вопросы точности асимптотических решений и качества генераторов псевдослучайных чисел / Заводская лаборатория. Диагностика материалов. 2016. Т. 82. № 7. С. 72 – 84.

136. Никитин Я. Ю. Асимптотическая эффективность непараметрических критериев. — М.: Наука, 1995. — 240 с.
137. Орлов А. И. Проблема множественных проверок статистических гипотез / Заводская лаборатория. Диагностика материалов. 1996. Т. 62. № 5. С. 51 – 54.
138. Орлов А. И. О методологии статистических методов / Научный журнал КубГАУ. 2014. № 104. С. 53 – 80.

REFERENCES

1. Orlov A. I. Development of the Methods of Mathematical Research (2006 – 2015) / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 1. Part 1. P. 78 – 86 [in Russian].
2. Orlov A. I. Stability in socio-economic models. — Moscow: Nauka, 1979. — 296 p. [in Russian].
3. Orlov A. I. Statistics of Non-Numeric Objects and Expert Estimation / Expert estimates. Questions of cybernetics. Vol. 58. — Moscow: Nauch. Sovet AN SSSR po kompleksnoi probleme “Kibernetika”, 1979. P. 17 – 33 [in Russian].
4. Tyurin Yu. N., Litvak B. G., Orlov A. I., Satarov G. A., Shmerling D. S. Analysis of non-numeric information / Zavod. Lab. 1980. Vol. 46. N 10. P. 931 – 935 [in Russian].
5. Orlov A. I. Statistics of Non-Numeric Objects (review) / Zavod. Lab. 1990. Vol. 56. N 3. P. 76 – 83 [in Russian].
6. Orlov A. I. Thirty Years of the Statistics of Non-Numeric Objects (review) / Zavod. Lab. Diagn. Mater. 2009. Vol. 75. N 5. P. 55 – 64 [in Russian].
7. Orlov A. I. Organizational-economic modeling: Textbook. Part 1. Nonnumeric statistics. — Moscow: Izd. MGTU im. N. É. Bauman, 2009. — 542 p. [in Russian].
8. Kuhn T. The Structure of Scientific Revolutions. — Moscow: AST, 2003. — 605 p. [in Russian].
9. Orlov A. I. The New Paradigm of Applied Statistics / Zavod. Lab. Diagn. Mater. 2012. Vol. 78. N 1. Part I. P. 87 – 93 [in Russian].
10. Orlov A. I. The New Paradigm of Mathematical Methods of Research / Zavod. Lab. Diagn. Mater. 2015. Vol. 81. N 7. P. 5 [in Russian].
11. Orlov A. I. About the New Paradigm of Mathematical Methods of Research / Nauch. Zh. KubGAU. 2016. N 122. P. 807 – 832 [in Russian].
12. Orlov A. I. Objects of non-numeric nature / Zavod. Lab. Diagn. Mater. 1995. Vol. 61. N 3. P. 43 – 52 [in Russian].
13. Bernshtein S. N. The current state of probability theory and its applications / Proc. of the All-Russian Congress of Mathematicians in Moscow, April 27 – May 4, 1927. — Moscow – Leningrad: GIZ, 1928. P. 50 – 63 [in Russian].
14. Novikov A. M., Novikov D. A. Methodology. — Moscow: SINTEG, 2007. — 668 p. [in Russian].
15. Discussion on the analysis of interval data / Zavod. Lab. 1990. Vol. 56. N 7. P. 75 – 95 [in Russian].
16. Voshchinin A. P. The method of optimization of objects by interval models of the objective function. — Moscow: MÊI, 1987. — 109 p. [in Russian].
17. Voshchinin A. P., Sotirov G. R. Optimization under uncertainty. — Moscow: MÊI, Sofiya: Tekhnika, 1989. — 224 p. [in Russian].
18. Voshchinin A. P., Akmatbekov R. A. Optimization of regression models and design of experiments. — Bishkek: Ilim, 1991. — 164 p. [in Russian].
19. Voshchinin A. P. Method of data analysis with interval errors in problems of hypothesis testing and estimation of parameters of implicit linearly parameterized functions / Zavod. Lab. Diagn. Mater. 2000. Vol. 66. N 3. P. 51 – 64 [in Russian].
20. Voshchinin A. P. Interval analysis of data: development and prospects / Zavod. Lab. Diagn. Mater. 2002. Vol. 68. N 1. P. 118 – 126 [in Russian].
21. Voshchinin A. P., Bronz P. V. Build-up of analytical models according to simulating experiment in problems of sensitivity analysis and economic risk estimation / Zavod. Lab. Diagn. Mater. 2007. Vol. 73. N 1. P. 101 – 109 [in Russian].
22. Voshchinin A. P., Skibitskii N. V. Interval Approach to the Uncertainty of Measurements and Calibration of Digital Measurement Systems / Zavod. Lab. Diagn. Mater. 2007. Vol. 73. N 11. P. 66 – 71 [in Russian].
23. Gus'kova E. A., Orlov A. I. Interval Linear Pair Regression (generalizing paper) / Zavod. Lab. Diagn. Mater. 2005. Vol. 71. N 3. P. 57 – 63 [in Russian].
24. Orlov A. I. Applied statistics. — Moscow: Ékzamen, 2006. — 671 p. [in Russian].
25. Orlov A. I. Decision theory. — Moscow: Ékzamen, 2006. — 576 p. [in Russian].
26. Orlov A. I., Lutsenko E. V. System fuzzy interval mathematics. — Krasnodar: KubGAU, 2014. — 600 p. [in Russian].
27. Orlov A. I. Statistics of Interval Data (generalizing paper) / Zavod. Lab. Diagn. Mater. 2015. Vol. 81. N 3. P. 61 – 69 [in Russian].
28. Skibitskii N. V., Seval'nev N. V. Interval Model in Optimal Control Problems / Zavod. Lab. Diagn. Mater. 2015. Vol. 81. N 11. P. 73 – 80 [in Russian].
29. Skibitskii N. V. Construction of Direct and Inverse Static Characteristics of the Objects by Interval Data / Zavod. Lab. Diagn. Mater. 2017. Vol. 83. N 1. Part 1. P. 87 – 93 [in Russian].
30. Tarantsev A. A. On the relationship of interval analysis with probability theory / Zavod. Lab. Diagn. Mater. 2004. Vol. 70. N 3. P. 60 – 65 [in Russian].
31. Orlov A. I. Econometrics. — Moscow: Ékzamen, 2002. — 576 p. [in Russian].
32. Orlov A. I. Fuzzy Sets Theory as the Part of Probability Theory / Nauch. Zh. KubGAU. 2013. N 92. P. 51 – 60 [in Russian].
33. Orlov A. I., Lutsenko E. V., Loiko V. I. Organizational-economic methods, mathematical modeling and software for controlling, innovation and management: monograph. — Krasnodar: KubGAU, 2016. — 600 p. [in Russian].
34. Kramer G. Mathematical methods of statistics. — Moscow: Mir, 1975. — 648 p. [in Russian].
35. Smirnov N. V., Dunin-Barkovskii I. V. Course of probability theory and mathematical statistics for technical applications. 3rd Edition. — Moscow: Nauka, 1969. — 512 p. [in Russian].
36. Bol'shev L. N., Smirnov N. V. Tables of mathematical statistics. 3rd edition. — Moscow: Nauka, 1983. — 416 p. [in Russian].
37. Kagan A. M., Linnik Yu. V., Rao P. R. Characterization problems of mathematical statistics. — Moscow: Nauka, 1972. — 656 p. [in Russian].
38. Actual problems of cybernetics (applied statistics) — Moscow: Znanie, 1981. — 64 p. [in Russian].
39. Orlov A. I. On the restructuring of statistical science and its applications / Vestn. Statist. 1990. N 1. P. 65 – 71 [in Russian].
40. Orlov A. I. Probabilistic models of specific types of objects of non-numeric nature / Zavod. Lab. Diagn. Mater. 1995. Vol. 61. N 5. P. 43 – 51 [in Russian].
41. Tyurin Yu. N., Litvak B. G., Orlov A. I., Satarov G. A., Shmerling D. S. Analysis of non-numeric information. — Moscow: Nauch. Sovet AN SSSR po kompleksnoi probleme “Kibernetika”, 1981. — 80 p. [in Russian].
42. Andreenkov V. G., Orlov A. I., Tolstova Yu. N., Eds. Analysis of non-numerical information in sociological research. — Moscow: Nauka, 1985. — 220 p. [in Russian].
43. Orlov A. I. The criteria for selection of indicators of efficiency of scientific activity / Kontrolling. 2013. N 3(49). P. 72 – 78 [in Russian].
44. Loiko V. I., Lutsenko E. V., Orlov A. I. Actual approaches in scientometrics. — Krasnodar: KubGAU, 2017. — 532 p. [in Russian].
45. Orlov A. I. Characterization of average values by means of measurement scales / Nauch. Zh. KubGAU. 2017. N 134. P. 877 – 907 [in Russian].
46. Psychological measurements. Digest of article. — Moscow: Mir, 1967. — 196 p. [in Russian].
47. Pfanzagl I. Measurement Theory. — Moscow: Mir, 1976. — 248 p. [Russian translation].

48. **Tolstova Yu. N.** A brief history of the development of representative measurement theory / *Zavod. Lab. Diagn. Mater.* 1999. Vol. 65. N 3. P. 49 – 56 [in Russian].
49. **Orlov A. I.** Representative Theory of Measurement and Its Applications / *Zavod. Lab. Diagn. Mater.* 1999. Vol. 65. N 3. P. 57 – 62 [in Russian].
50. **Barskii B. V., Sokolov M. V.** Medial values, invariant relative to permissible transformations of measuring scale / *Zavod. Lab. Diagn. Mater.* 2006. Vol. 72. N 1. P. 59 – 66 [in Russian].
51. **Orlov A. I.** Mathematical testing methods and theory of measurings / *Zavod. Lab. Diagn. Mater.* 2006. Vol. 72. N 1. P. 67 – 70 [in Russian].
52. **Kemeni G., Snell G.** Cybernetic modeling: Some applications. — Moscow: Sovetskoe radio, 1972. — 192 p. [in Russian].
53. **Zhukov M. S., Orlov A. I.** Using expert rankings in calculating credit risk in a bank / *Innovats. Menedzhm.* 2017. N 1. P. 18 – 25 [in Russian].
54. **Zhukov M. S., Orlov A. I.** The problem of research of final ranking for group of experts by means of Kemeny median / *Nauch. Zh. KubGAU.* 2016. N 122. P. 785 – 806 [in Russian].
55. **Zhukov M. S., Orlov A. I., Fal'ko S. G.** Expert assessment in risks / *Kontrolling.* 2017. N 4(66). P. 24 – 27 [in Russian].
56. **Orlov A. I.** Average values and rules of large numbers in the spaces of arbitrary origin / *Nauch. Zh. KubGAU.* 2013. N 89. P. 556 – 586 [in Russian].
57. **Orlov A. I.** On mean values / *Large systems management.* Vol. 46. — Moscow: IPU RAN, 2013. P. 88 – 117 [in Russian].
58. **Orlov A. I.** The limit theory of the solutions of extremal statistical problems / *Nauch. Zh. KubGAU.* 2017. N 133. P. 579 – 600 [in Russian].
59. **Orlov A. I.** Methods for finding the most informative sets of characteristics in regression analysis / *Zavod. Lab. Diagn. Mater.* 1995. Vol. 61. N 1. P. 56 – 58 [in Russian].
60. **Orlov A. I.** Asymptotic behavior of solutions of extremal statistical problems / *Zavod. Lab. Diagn. Mater.* 1996. Vol. 62. N 10. P. 45 – 46 [in Russian].
61. **Tyrstin A. N.** Robust Construction of Linear Regressive Models by Experimental Data / *Zavod. Lab. Diagn. Mater.* 2005. Vol. 71. N 11. P. 53 – 58 [in Russian].
62. **Prokhorov Yu. V., Ed.** Probability and Mathematical Statistics: Encyclopedia. — Moscow: Bol'shaya Rossiiskaya Ėntsiklopediya, 1999. — 910 p. [in Russian].
63. **Orlov A. I.** Density estimates in spaces of arbitrary nature / *Statistical methods for evaluating and testing hypotheses: an inter-university collection of scientific papers.* Vol. 25. — Perm: Perm. Gos. Nats. Issl. Univ., 2013. P. 21 – 33 [in Russian].
64. **Orlov A. I.** Estimates of probability density function in spaces of arbitrary nature / *Nauch. Zh. KubGAU.* 2014. N 99. P. 15 – 32 [in Russian].
65. **Orlov A. I.** Nuclear density estimates in spaces of arbitrary nature / *Statistical methods for evaluating and testing hypotheses: an inter-university collection of scientific papers.* Vol. 26. — Perm: Perm. Gos. Nats. Issl. Univ., 2015. P. 43 – 57 [in Russian].
66. **Orlov A. I.** Limit theorems for kernel density estimators in spaces of arbitrary / *Nauch. Zh. KubGAU.* 2015. N 108. P. 316 – 333 [in Russian].
67. **Orlov A. I.** Nonparametric kernel estimators of probability density in the discrete spaces / *Nauch. Zh. KubGAU.* 2016. N 122. P. 833 – 855 [in Russian].
68. **Orlov A. I.** Nuclear density estimates in finite spaces / *Statistical methods for evaluating and testing hypotheses: an inter-university collection of scientific papers.* Vol. 27. — Perm: Perm. Gos. Nats. Issl. Univ., 2016. P. 24 – 37 [in Russian].
69. **Orlov A. I.** Statistics of Non-Numeric Objects / *Teor. Veroyatn. Primen.* 1980. Vol. XXV. N 3. P. 655 – 656 [in Russian].
70. **Orlov A. I.** Nonparametric density estimates in topological spaces / *Applied statistics. Scientists notes on statistics.* Vol. 45. — Moscow: Nauka, 1983. P. 12 – 40 [in Russian].
71. **Rosenblatt M.** Remarks on some nonparametric estimates of a density function / *Ann. Math. Statist.* 1956. Vol. 27. N 5. P. 832 – 837.
72. **Parzen E.** On estimation of a probability density function and mode / *Ann. Math. Statist.* 1962. Vol. 33. N 6. P. 1065 – 1076.
73. **Orlov A. I.** Asymptotics of estimates of probability density function / *Nauch. Zh. KubGAU.* 2017. N 131. P. 845 – 873 [in Russian].
74. **Orlov A. I.** The rate of convergence of kernel density estimates in spaces of arbitrary nature / *Statistical methods for evaluating and testing hypotheses: an inter-university collection of scientific papers.* Vol. 28. — Perm: Perm. Gos. Nats. Issl. Univ., 2018. P. 35 – 45 [in Russian].
75. **Ibragimov I. A., Khas'minskiĭ R. Z.** Asymptotic estimation theory. — Moscow: Nauka, 1979. — 528 p. [in Russian].
76. **Orlov A. I.** On the Development of Mathematical Methods in the Theory of Classification (review) / *Zavod. Lab. Diagn. Mater.* 2009. Vol. 75. N 7. P. 51 – 63 [in Russian].
77. **Orlov A. I., Tolcheev V. O.** On the Use of Nonparametric Statistical Tests to Assess the Accuracy of Classification Methods (Summarizing Article) / *Zavod. Lab. Diagn. Mater.* 2011. Vol. 77. N 3. P. 58 – 66 [in Russian].
78. **Orlov A. I.** Three Main Results of the Mathematical Theory of Classification / *Zavod. Lab. Diagn. Mater.* 2016. Vol. 82. N 5. P. 63 – 70 [in Russian].
79. **Perekrest V. T.** Nonlinear typological analysis of socioeconomic information: Mathematical and computational methods. — Leningrad: Nauka, 1983. — 176 p. [in Russian].
80. **Terekhina A. Yu.** Data analysis using multidimensional scaling methods. — Moscow: Nauka, 1986. — 168 p. [in Russian].
81. **Lagutin M. B.** Visual representation of tight connections / *Zavod. Lab. Diagn. Mater.* 2005. Vol. 71. N 7. P. 53 – 57 [in Russian].
82. **Lutsenko E. V., Orlov A. I.** Methods of reducing space dimension of statistical data / *Nauch. Zh. KubGAU.* 2016. N 119. P. 92 – 107 [in Russian].
83. **Orlov A. I.** Mathematical methods of research and diagnosis materials (generalizing paper) / *Zavod. Lab. Diagn. Mater.* 2003. Vol. 69. N 3. P. 53 – 64 [in Russian].
84. **Orlov A. I.** Predictive power — the best indicator of the quality of the diagnostic algorithm / *Nauch. Zh. KubGAU.* 2014. N 99. P. 33 – 49 [in Russian].
85. **Orlov A. I.** Stability of Classification Regarding the Choice of the Method of Cluster Analysis / *Zavod. Lab. Diagn. Mater.* 2013. Vol. 79. N 1. P. 68 – 71 [in Russian].
86. **Orlov A. I.** Asymptotic behavior of integral type statistics / *Dokl. AN SSSR.* 1974. Vol. 219. N 4. P. 808 – 811 [in Russian].
87. **Orlov A. I.** About the testing of homogeneity for two independent samples / *Zavod. Lab. Diagn. Mater.* 2003. Vol. 69. N 1. P. 55 – 60 [in Russian].
88. **Orlov A. I.** Methods for testing the homogeneity of the associated samples / *Zavod. Lab. Diagn. Mater.* 2004. Vol. 70. N 7. P. 57 – 61 [in Russian].
89. **Orlov A. I.** Asymptotic behavior of integral type statistics / *Probabilistic processes and their applications. Interuniversity collection of scientific papers.* — Moscow: MIEHM, 1989. P. 118 – 123 [in Russian].
90. **Orlov A. I.** Limit Theory of Nonparametric Statistics / *Nauch. Zh. KubGAU.* 2014. N 100. P. 31 – 52 [in Russian].
91. **Orlov A. I.** Organizational-economic modeling: Textbook. Part 3: Statistical data analysis methods. — Moscow: Izd. MGTO im. N. Ė. Bauman, 2012. — 624 p. [in Russian].
92. **Nalimov V. V.** The Application of Mathematical Statistics to Chemical Analysis. — Moscow: Fizmatgiz, 1960. — 430 p. [in Russian].
93. **Novitskii P. V., Zograf I. A.** Estimation of measurement errors. — Leningrad: Ėnergoatomizdat, 1985. — 248 p. [in Russian].
94. **Orlov A. I.** How often the distribution of the results of observations is normal? / *Zavod. Lab.* 1991. Vol. 57. N 7. P. 64 – 66 [in Russian].
95. **Orlov A. I.** Distributions of real statistical data are not normal / *Nauch. Zh. KubGAU.* 2016. N 117. P. 71 – 90 [in Russian].
96. **Seleznev V. D., Denisov K. P.** Study of the properties of consent criterion of data distribution function with the Gaussian by Monte-Carlo method for small samples / *Zavod. Lab. Diagn. Mater.* 2005. Vol. 71. N 1. P. 68 – 73 [in Russian].

97. **Orlov A. I.** Nonparametric point and interval estimation of the characteristics of distribution/ *Zavod. Lab. Diagn. Mater.* 2004. Vol. 70. N 5. P 65 – 70 [in Russian].
98. **Orlov A. I.** Statistical hypothesis testing of homogeneity of mathematical expectations of two independent samples: the Cramer – Welch instead of the Student's test / *Nauch. Zh. KubGAU.* 2015. N 110. P 197 – 218 [in Russian].
99. **Orlov A. I.** Two-sample Wilcoxon test — analysis of two myths / *Nauch. Zh. KubGAU.* 2014. N 104. P 91 – 111 [in Russian].
100. **Orlov A. I.** Consistent Tests of Absolute Homogeneity of Independent Samplings / *Zavod. Lab. Diagn. Mater.* 2012. Vol. 78. N 11. P 66 – 70 [in Russian].
101. **Orlov A. I.** Structure of Nonparametric Statistics (generalizing paper) / *Zavod. Lab. Diagn. Mater.* 2015. Vol. 81. N 7. P 62 – 72 [in Russian].
102. **Orlov A. I.** Current status of nonparametric statistics / *Nauch. Zh. KubGAU.* 2015. N 106. P 239 – 269 [in Russian].
103. **Orlov A. I.** Real and nominal significance levels in statistical hypothesis testing / *Nauch. Zh. KubGAU.* 2015. N 114. P 42 – 54 [in Russian].
104. **Orlov A. I.** The Model of Coincidence Analysis in the Calculation of Nonparametric Rank Statistics / *Zavod. Lab. Diagn. Mater.* 2017. Vol. 83. N 11. P 66 – 72 [in Russian].
105. **Murav'eva V. S., Orlov A. I.** Nonparametric estimation of the point of intersection of the regression lines / *Zavod. Lab. Diagn. Mater.* 2008. Vol. 74. N 1. P 63 – 68 [in Russian].
106. **Orlov A. I.** Nonparametric Method of Least Squares with the Periodic Component (generalizing article) / *Zavod. Lab. Diagn. Mater.* 2014. Vol. 80. N 1. P 65 – 75 [in Russian].
107. **Orlov A. I.** Statistics of fuzzy data / *Nauch. Zh. KubGAU.* 2016. N 119. P 75 – 91 [in Russian].
108. **Orlov A. I.** Theory of Lusians / *Nauch. Zh. KubGAU.* 2014. N 101. P 275 – 304 [in Russian].
109. **Orlov A. I.** Distances in the spaces of statistical data / *Nauch. Zh. KubGAU.* 2014. N 101. P 227 – 252 [in Russian].
110. **Orlov A. I.** Statistical control of two alternative variables and a method for verifying their independence from a set of small samples / *Zavod. Lab. Diagn. Mater.* 2000. Vol. 66. N 1. P 58 – 62 [in Russian].
111. **Orlov A. I.** Main problems of controlling of the quality / *Nauch. Zh. KubGAU.* 2015. N 111. P 20 – 52 [in Russian].
112. **Orlov A. I.** Limit theorems in statistical control / *Nauch. Zh. KubGAU.* 2016. N 116. P 462 – 483 [in Russian].
113. **Orlov A. I.** Organizational-economic modeling: Textbook. Part 2: Expert estimators. — Moscow: Izd. MGTU im. N. É. Bauman, 2009. — 486 p. [in Russian].
114. Statistical methods for analyzing expert estimators / *Scientists notes on statistics.* Vol. 29. — Moscow: Nauka, 1977. — 385 p. [in Russian].
115. Expert estimators / *Voprosy kibernetiki.* — Vyp. 58. — Moscow: Nauch. Sovet AN SSSR po kompleksnoi probleme “Kibernetika”, 1979. — 200 p. [in Russian].
116. Expert estimates in system studies. Collection of works. Vol. 4. — Moscow: VNIISI, 1979. — 120 p. [in Russian].
117. Expert estimates in management. Collection of works. — Moscow: Inst. Probl. Upravl., 1982. — 106 p. [in Russian].
118. **Orlov A. I.** Expert estimation / *Zavod. Lab. Diagn. Mater.* 1996. Vol. 62. N 1. P 54 – 60 [in Russian].
119. **Orlov A. I.** On the Development of Expert Technology in Our Country (review) / *Zavod. Lab. Diagn. Mater.* 2010. Vol. 76. N 11. P 64 – 70 [in Russian].
120. **Orlov A. I.** Theory of expert estimates in our country / *Nauch. Zh. KubGAU.* 2013. N 93. P 1 – 11 [in Russian].
121. **Gnedenko B. V., Orlov A. I.** The role of mathematical research methods in the fundamental acceleration of scientific and technological progress / *Zavod. Lab.* 1988. Vol. 54. N 1. P 1 – 4 [in Russian].
122. **Orlov A. I.** On modern problems of implementing applied statistics and other statistical methods / *Zavod. Lab.* 1992. Vol. 58. N 1. P 67 – 74 [in Russian].
123. **Orlov A. I.** Certification and statistical methods (generalizing paper) / *Zavod. Lab. Diagn. Mater.* 1997. Vol. 63. N 3. P 55 – 62 [in Russian].
124. **Gorskii V. G., Orlov A. I.** Mathematical methods of research: results and prospects / *Zavod. Lab. Diagn. Mater.* 2002. Vol. 68. N 1. P 108 – 112 [in Russian].
125. **Orlov A. I.** High statistical technologies / *Zavod. Lab. Diagn. Mater.* 2003. Vol. 69. N 11. P 55 – 60 [in Russian].
126. **Orlov A. I.** About high statistical technologies / *Nauch. Zh. KubGAU.* 2015. N 105. P 14 – 38 [in Russian].
127. **Orlov A. I.** “Six sigmas” — new system of introduction of mathematical research methods / *Zavod. Lab. Diagn. Mater.* 2006. Vol. 72. N 5. P 50 – 53 [in Russian].
128. **Orlov A. I.** Mathematical software certification: a comparative analysis of dialogue systems for statistical control / *Zavod. Lab. Diagn. Mater.* 1996. Vol. 62. N 7. P 46 – 49 [in Russian].
129. **Smirnova O. S.** Software for Statistic Analysis / *Zavod. Lab. Diagn. Mater.* 2008. Vol. 74. N 5. P 68 – 75 [in Russian].
130. **Orlov A. I.** Statistic Packages: Tools of the Researcher / *Zavod. Lab. Diagn. Mater.* 2008. Vol. 74. N 5. P 76 – 78 [in Russian].
131. **Zhukov M. S.** On the Algorithms for Kemeny Median Calculation / *Zavod. Lab. Diagn. Mater.* 2017. Vol. 83. N 7. P 72 – 78 [in Russian].
132. **Zagoruiko N. G., Orlov A. I.** Some unsolved mathematical problems of applied statistics / *Modern problems of cybernetics (applied statistics).* — Moscow: Znanie, 1981. P 53 – 63 [in Russian].
133. **Orlov A. I.** Some outstanding problems in the field of mathematical research methods / *Zavod. Lab. Diagn. Mater.* 2002. Vol. 68. N 3. P 52 – 56 [in Russian].
134. **Orlov A. I.** Limit Theorems and Monte Carlo Method / *Zavod. Lab. Diagn. Mater.* 2016. Vol. 82. N 7. P 67 – 72 [in Russian].
135. **Grigor'ev Yu. D.** Monte Carlo Method: Accuracy of Asymptotic Solutions and Quality of Pseudorandom Number Generators / *Zavod. Lab. Diagn. Mater.* 2016. Vol. 82. N 7. P 72 – 84 [in Russian].
136. **Nikitin Ya. Yu.** Asymptotic efficiency of nonparametric criteria. — Moscow: Nauka, 1995. — 240 p. [in Russian].
137. **Orlov A. I.** The problem of multiple tests of statistical hypotheses / *Zavod. Lab. Diagn. Mater.* 1996. Vol. 62. N 5. P 51 – 54 [in Russian].
138. **Orlov A. I.** About the methodology of statistical methods / *Nauch. Zh. KubGAU.* 2014. N 104. P 53 – 80 [in Russian].