Математические методы исследований

УДК 519.28

ТРИДЦАТЬ ЛЕТ СТАТИСТИКИ ОБЪЕКТОВ НЕЧИСЛОВОЙ ПРИРОДЫ (обзор)

© А. И. Орлов[1]

Статья поступила 29 декабря 2008 г.

Проведен анализ основных идей в области статистики нечисловой природы. Тридцать лет назад она была выделена как самостоятельная область математических методов исследования. Рассмотрены теоретические основы и методология, базовые подходы, концепции и научные результаты в этой области, дана обширная библиография (150 названий).

Термин «статистика объектов нечисловой природы» впервые появился в 1979 г. в монографии [1]. В том же году в статье [2] была предложена программа построения этой новой области статистических методов. В журнале «Заводская лаборатория» сразу же появилась обобщающая статья на эту тему [3]. Обсудим содержание, развитие и основные идеи статистики объектов нечисловой природы, особое внимание уделив исследованиям, опубликованным в нашем журнале.

Послевоенное развитие отечественной статистики

К 60-м годам XX в. в нашей стране сформировалась научно-практическая дисциплина, которую называем «классической математической статистикой». Статистики учились теории по книге Г. Крамера [4], написанной в военные годы и впервые изданной у нас в 1948 г. Из прикладных руководств назовем учебник [5] и таблицы с комментариями [6].

Затем внимание многих специалистов сосредоточилось на изучении математических конструкций, используемых в статистике. Этому посвящена, например, монография [7]. В ней представлены полученные математические результаты, но трудно выделить рекомендации для статистика, анализирующего конкретные данные.

Как реакция на уход в математику выделилась новая научная дисциплина — прикладная статистика [8]. Это стало очевидным в 1981 г., когда массовым тиражом (33 940 экз.) вышел сборник [9], 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 позиционировала как науку об обработке данных — результатов наблюдений, измерений, испытаний, анализов, опытов.

Естественно, что в прикладной статистике стали развиваться математические методы и модели, поскольку этого требуют конкретные прикладные исследования. Это математизированное ядро прикладной статистики хочется назвать 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, биологии — биометрика, химии — хемометрия, в технических исследованиях — технометри-

[1] Институт высоких статистических технологий и эконометрики Московского государственного технического университета им. Н. Э. Баумана, Москва, Россия.

ка, а также в геологии, демографии, социологии, медицине, истории и т.д.

Часто позиции 2 и 3 вместе называют прикладной статистикой, а позицию 1 — 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 уровню представлений середины XX века.

Примечание. Мы уточнили схему внутреннего деления статистической теории, предложенную в работе [10]. Естественный смысл приобрели термины «теоретическая статистика» и «прикладная статистика» (в узком смысле). Однако необходимо иметь в виду, что в учебнике [8] 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 позапрошлого века. Возможно, следовало бы от этого ведомства полностью отмежеваться и сменить название дисциплины, например, на «Анализ данных». В настоящее время компромиссным названием нашей научно-практической дисциплины является термин «статистические методы».

Во второй половине 80-х годов развернулось общественное движение, имеющее целью создание 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 (ВСА) оказалась парализованной в результате «развала» СССР. Некоторую активность проявили созданные на базе ВСА Российская ассоциация статистических методов, Российская академия статистических методов, Белорусская статистическая ассоциация.

В ходе создания ВСА были проанализированы состояние и перспективы развития теоретической и прикладной статистики. Обсудим их.

Новые идеи последних десятилетий: точки роста

В работе [11] выделено пять актуальных направлений, в которых развивается современная прикладная статистика, т.е. пять «точек роста» статистической науки: непараметрика, робастность, бутстреп, интервальная статистика, статистика объектов нечисловой природы. Кратко обсудим эти актуальные направления.

Непараметрика, или непараметрическая статистика [12], 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 [13]. Научная школа метролога П. В. Новицкого многочисленными экспериментами подтвердила отсутствие нормальности погрешностей измерений [14]. Опубликованная в «Заводской лаборатории» сводка [15] включена в учебники [8, 16]. В работе [17] установлено, что по выборкам объемов 6 – 50, как правило, не удается отличить нормальное распределение от других видов распределений.

Некритическое использование гипотезы нормальности часто приводит к значительным ошибкам, например, при отбраковке резко выделяющихся результатов наблюдений (выбросов), а также статистическом контроле качества и в других случаях [8]. Поэтому целесообразно использовать непараметрические методы, в которых на функции распределения результатов наблюдений наложены лишь 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 (речь идет, в частности, об оценивании характеристик распределения [18] и проверке гипотезы однородности для независимых [19 – 21] и связанных [22] выборок). Однако эта информация еще не вошла в массовое сознание. До сих пор тупиковой тематике параметрической статистики посвящены обширные разделы учебников и программных продуктов.

Основная идея работ по *робастности*, или *устойчивости*, состоит в том, что выводы, полученные на основе математических методов исследования, должны мало меняться при небольших изменениях исходных данных и малых отклонениях от предпосылок модели [1]. Здесь имеются два круга задач [23 – 26]: изучение устойчивости распространенных алгоритмов анализа данных и поиск робастных алгоритмов для решения тех или иных задач. Отметим, что термин «робастность» не имеет точно определенного смысла, всегда необходимо указывать конкретную вероятностно-статистическую модель. При этом модель «засорения» Тьюки – Хубера – Хампеля обычно не является практически полезной. Дело в том, что она ориентирована на «утяжеление хвостов», а в реальных ситуациях «хвосты» обрезаются априорными ограничениями на результаты наблюдений, связанными, например, с ограниченностью шкал используемых средств измерения.

Бутстреп — направление непараметрической статистики, опирающееся на интенсивное использование информационных технологий [27]. Основная идея состоит в «размножении выборок», т.е. в нахождении набора из многих выборок, напоминающих полученную в эксперименте. По такому набору можно оценить свойства различных статистических процедур, не прибегая к излишне обременительным семействам вероятностно-статистических моделей. Простейший способ «размножения выборки» состоит в исключении из нее одного результата наблюдения. Исключаем первое наблюдение и получаем выборку, похожую на исходную, но с объемом, уменьшенным на единицу. Затем возвращаем результат первого наблюдения, но исключаем второе наблюдение. Получаем вторую выборку, похожую на исходную. Затем возвращаем результат второго наблюдения, и т.д. Есть и иные способы «размножения выборок». Например, по исходной выборке можно построить ту или иную оценку функции распределения, а затем методом статистических испытаний смоделировать ряд выборок из элементов, функция распределения которых совпадает с этой оценкой. Обобщая, можно сказать, что к настоящему времени в дополнение к классическим инструментам прикладной статистики — предельным теоремам теории вероятностей — добавились новые, основанные на интенсивном использовании компьютеров. Бутстреп — лишь один из таких инструментов. Метод статистических испытаний (Монте-Карло) — партнер и конкурент асимптотическим методам математической статистики [28].

Интервальная статистика — это совокупность методов анализа интервальных статистических данных. Вполне очевидно, что все средства измерения имеют погрешности. Однако до недавнего времени это очевидное обстоятельство никак не учитывалось в статистических процедурах. Только недавно начала развиваться теория интервальной статистики, в которой предполагается, что исходные данные — это не числа, а интервалы. Интервальную статистику можно рассматривать как часть интервальной математики. Выводы в ней часто принципиально отличны от классических.

Различным подходам к анализу интервальных данных посвящена дискуссия в «Заводской лаборатории» [29]. В научной школе А. П. Вощинина изучены проблемы регрессионного анализа, планирования эксперимента, сравнения альтернатив и принятия решений в условиях интервальной неопределенности [30 – 35]. В асимптотической статистике интервальных данных на значения случайных величин наложены малые интервальные неопределенности [36]. Основные результаты этого направления подробно изложены в учебниках [8, 37, 38]. Разрабатывались и иные подходы [39].

Статистика объектов нечисловой природы

Перейдем к статистике объектов нечисловой природы (она же — статистика нечисловых данных, или нечисловая статистика). 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 нечисловой статистике — это объекты нечисловой природы, которые нельзя складывать и умножать на числа. Другими словами, объекты нечисловой природы лежат в пространствах, не имеющих векторной структуры.

Примерами объектов нечисловой природы являются:

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

упорядочения (ранжировки) экспертами образцов продукции (при оценке ее технического уровня, качества и конкурентоспособности) или заявок на проведение научных работ (при проведении конкурсов на выделение грантов);

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

толерантности, т.е. бинарные отношения, описывающие сходство объектов между собой, например, сходства тематики научных работ, оцениваемого экспертами, с целью 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;

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.

Рассмотренные выше интервальные данные тоже можно рассматривать как пример объектов нечисловой природы, а именно, 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 [1], а также в учебниках [8, 16, 37, 38].

С 70-х годов в основном на основе запросов теории экспертных оценок (а также технических исследований, экономики, социологии и медицины) развивались различные направления статистики объектов нечисловой природы. Были установлены основные связи между конкретными видами таких объектов, разработаны для них базовые вероятностные модели [1, 40].

Следующий этап (80-е годы) — выделение статистики объектов нечисловой природы в качестве самостоятельной дисциплины в рамках математических методов исследования, ядром которого являются методы статистического анализа данных произвольной природы. Для работ этого периода характерна сосредоточенность на внутренних проблемах нечисловой статистики. Проводились всесоюзные конференции [41, 42], выпускались монографии [43 – 48], сборники трудов [49 – 51], защищались диссертации [52 – 58]. Наиболее представительным является сборник [59], подготовленный комиссией «Статистика объектов нечисловой природы» Научного Совета АН СССР по комплексной проблеме «Кибернетика» совместно с Институтом социологических исследований АН СССР.

К 90-м годам статистика объектов нечисловой природы 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. Поэтому наступило время перехода от теоретико-статистических исследований к применению полученных результатов на практике и включению их в учебный процесс, что и было сделано (см., например, работы [8, 16, 37, 38]). В 90-е годы в «Заводской лаборатории» опубликованы обзоры [60 – 62] по статистике объектов нечисловой природы и результаты многочисленных конкретных исследований.

Основные идеи и направления статистики объектов нечисловой природы

В чем принципиальная новизна нечисловой статистики? 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 нацелены на изучение сумм. В нечисловой же статистике нельзя использовать операцию сложения, поскольку элементы выборки лежат в пространствах, где нет операции сложения. Методы обработки нечисловых данных основаны на принципиально ином математическом аппарате — на применении различных расстояний в пространствах объектов нечисловой природы.

Следует отметить, что в статистике объектов нечисловой природы 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.

Рассмотрим кратко несколько идей, развиваемых в статистике объектов нечисловой природы для данных, 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 удается указать вид средних величин, соответствующих тем или иным шкалам измерения. Теория измерений [1, 63, 64], рассматривавшаяся в середине XX в. как часть математического обеспечения психологии, к настоящему времени признана общенаучной дисциплиной. Современные достижения представлены в статьях [65 – 68].

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 находится 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 математическое ожидание 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 [69]). При этом как эмпирическое, так и теоретическое средние как решения экстремальных задач могут быть не единственными элементами рассматриваемого пространства, а являться некоторыми множествами таких элементов, которые 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 [8, 16].

Оказалось, что методы доказательства законов больших чисел допускают существенно более широкую область применения, чем та, для которой они были разработаны. А именно, удалось изучить асимптотику решений экстремальных статистических задач, к которым, как известно, сводится большинство постановок прикладной статистики. В частности, кроме законов больших чисел установлена и состоятельность оценок минимального контраста, в том числе максимального правдоподобия и робастных. К настоящему времени подобные оценки изучены также и в интервальной статистике. Полученные результаты относительно асимптотики решений экстремальных статистических задач применены в работах [70 – 72].

В статистике в пространствах произвольной природы большую роль играют непараметрические оценки плотности, используемые, в частности, в различных алгоритмах регрессионного, дискриминантного, кластерного анализов. В нечисловой статистике 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].

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

Для анализа нечисловых данных, в частности экспертных, весьма важны методы классификации [76 – 82]. 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 распознавания образов как с учителем (т.е. дискриминантного анализа), так и без учителя (т.е. кластерного анализа). Аналогичным образом задачи многомерного шкалирования, т.е. визуализации данных [46, 47, 83], также естественно отнести к статистике объектов нечисловой природы.

Для проверки гипотез в пространствах нечисловой природы могут быть использованы статистики интегрального типа, в частности типа омега-квадрат [2, 21, 22]. Любопытно, что предельная теория таких статистик, построенная первоначально в классической постановке, приобрела естественный (завершенный, изящный) вид именно для пространств произвольного вида [84], 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.

Представляют практический интерес результаты, связанные с конкретными областями статистики объектов нечисловой природы, в частности, со статистикой нечетких [85] и случайных множеств (теория нечетких множеств в определенном смысле сводится к теории случайных множеств), с непараметрической теорией парных сравнений и люсианов (бернуллиевских бинарных векторов), аксиоматическим введением метрик в конкретных пространствах объектов нечисловой природы, а также с рядом других конкретных постановок. Отметим возрастающий интерес со стороны прикладников к математическому аппарату теории нечеткости [86 – 90].

Результаты контроля штучной продукции по альтернативному признаку представляют собой последовательности из 0 и 1 — объекты нечисловой природы, а потому теорию статистического контроля относят к нечисловой статистике [20 – 22]. В «Заводской лаборатории» постоянно публикуются работы по этой тематике, предназначенные для специалистов по статистическим методам управления качеством продукции [81, 91 – 95].

При статистическом анализе нечисловых данных возникает необходимость оценивать параметры модели. Вместо метода максимального правдоподобия целесообразно применять метод одношаговых оценок [96, 97].

Разрабатываются новые методы анализа конкретных видов нечисловых данных. Так, С. А. Смоляк рассматривает проблему восстановления функции многих переменных по ее точным или приближенным значениям в отдельных точках. Для функций числовых переменных — это обычная задача интерполяции, однако он решает задачу восстановления функции от номинальных или порядковых переменных и предлагает эвристические методы, основанные на формализации дискретного аналога понятия «гладкости» функции [98, 99]. А. Н. Горбач и Н. А. Цейтлин на основе практических потребностей обосновывают необходимость построения статистической теории спонтанных последовательностей, вводят расстояния между ними [100] и разрабатывают методы анализа этого нового вида объектов нечисловой природы [101]. Бурно развивается раздел нечисловой статистики, посвященный организационным структурам [102 – 108].

Статистика объектов нечисловой природы порождена потребностями практики, прежде всего в области экспертных оценок. Вполне естественно, что названия сборников трудов научного коллектива, развивающего нечисловую статистику, начинались со слов «Экспертные оценки» [109 – 112]. Различным вопросам теории и практики экспертных оценок посвящен ряд монографий этого коллектива [113 – 120]. Научные результаты последних лет постоянно публикуются в журналах «Заводская лаборатория» [121 – 129] и «Автоматика и телемеханика» [130 – 132].

Экспертные методы, как и статистические, активно используются при прогнозировании. Этой тематике наш журнал уделяет значительное место [133 – 135].

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. Это означает, что существенную часть непараметрической статистики [6, 136, 137] можно включить в нечисловую статистику. Тем более это касается статистики интервальных данных, изучающей методы анализа нечисловых данных конкретного вида — интервалов. В работе [38] она включена в нечисловую статистику. Однако в данной работе мы предпочли рассмотреть непараметрику — статистику интервальных данных и нечисловую статистику по отдельности. В частности, потому, что статистика в пространствах произвольной природы является центральной областью только для последнего из трех рассмотренных здесь направлений прикладной статистики.

Вопросы внедрения математических методов исследования всегда были в центре внимания журнала «Заводская лаборатория» [91, 138, 139]. Подчеркивалось большое теоретическое и прикладное значение статистики объектов нечисловой природы [140], необходимость перехода от отдельных методов анализа данных к разработке высоких статистических технологий [141] и использования современных систем внедрения математических методов, таких как «Шесть сигм» и ее аналоги [142]. Обсуждались проблемы программного обеспечения [143 – 145]. Однако приходится констатировать, что создание линейки современных программных продуктов по нечисловой статистике — пока дело будущего.

О некоторых нерешенных проблемах нечисловой статистики

За каждым новым научным результатом открывается многообразие неизвестного. Рассмотрим несколько конкретных постановок.

В статистике в пространствах общей природы получены аналоги классического закона больших чисел. Но нет аналога центральной предельной теоремы. Какова скорость сходимости эмпирических средних к теоретическим? Как сравнить различные способы усреднения? В частности, что лучше применять для усреднения упорядочений — медиану Кемени или среднее по Кемени (среднее отличается от медианы тем, что в качестве показателя различия берется не расстояние Кемени, а его квадрат)? Какие конкретные представители различных классов непараметрических оценок плотности достойны рекомендации для использования в нацеленных на практическое применение алгоритмах и программных продуктах анализа нечисловых данных?

До сих пор не проведена полная классификация классических статистических методов с точки зрения теории измерений. Законченные результаты получены только для теории средних [8, 16, 37]. Установлено, что для измерений в порядковой шкале в качестве средних можно использовать только порядковые статистики, например медиану (при нечетном объеме выборки). Среднее арифметическое применять нельзя. Однако многочисленные эксперименты показывают, что упорядочения объектов по средним арифметическим рангов и по медианам рангов в подавляющем большинстве случаев совпадают или близки. Нужна теория, объясняющая этот экспериментальный факт.

Все более широкое распространение получает теория нечеткости. Давно установлено, что она в определенном смысле сводится к теории случайных множеств [8, 16, 37]. На основе этого сведения требуется проанализировать различные теоретические и прикладные постановки теории нечеткости и рассмотреть их в рамках вероятностно-статистического моделирования.

Перейдем к классическим областям статистики. 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 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 [8, 16].

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 обрывочные сведения. Приведем три примера.

Пример 1. Построение доверительного интервала для математического ожидания обычно проводят с использованием распределения Стьюдента (при справедливости гипотезы нормальности). Как следует из Центральной предельной теоремы (ЦПТ) теории вероятностей, в асимптотике (при большом объеме выборки) такие расчетные методы дают правильные результаты (из ЦПТ вытекает использование квантилей нормального распределения, а из классической теории — квантилей распределения Стьюдента, но при росте объема выборки квантили распределения Стьюдента стремятся к соответствующим квантилям нормального распределения).

Пример 2. Для проверки однородности двух независимых выборок (на самом деле — для проверки равенства математических ожиданий) обычно рекомендуют использовать двухвыборочный критерий Стьюдента. Предпосылки его использования — это нормальность распределений, соответствующих выборкам, и равенство их дисперсий. Что будет при отклонении от нормальности распределений, из которых взяты выборки? Если объемы выборок равны или если дисперсии совпадают, то в асимптотике (когда объемы выборок безгранично возрастают) классический метод является корректным. Если же объемы выборок существенно отличаются или дисперсии различны, то критерий Стьюдента проверки гипотезы однородности применять нельзя, поскольку распределение двухвыборочной статистики Стьюдента будет существенно отличаться от классического. Поскольку проверка равенства дисперсий — более сложная задача, чем проверка равенства математических ожиданий, то для выборок разного объема использовать двухвыборочную статистику Стьюдента не следует, целесообразно применять критерий Крамера – Уэлча [8, 16, 21].

Пример 3. В задаче отбраковки (исключения) резко выделяющихся наблюдений (выбросов) расчетные методы, основанные на нормальности, являются крайне неустойчивыми по отношению к отклонениям от нормальности, что полностью лишает эти методы научной обоснованности [8, 16, 146].

Примеры 1 – 3 показывают весь спектр возможных свойств классических расчетных методов в случае отклонения от нормальности. Методы примера 1 оказываются вполне пригодными при таких отклонениях, примера 2 — пригодными в некоторых случаях, примера 3 — полностью непригодными.

Итак, имеется *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.

Почему необходимо изучение классических алгоритмов, а не построение новых, специально предназначенных для работы в условиях отклонения от классических предпосылок? Во-первых, потому что классические алгоритмы в настоящее время наиболее распространены (благодаря сложившейся системе образования прикладников). Например, для проверки однородности двух независимых выборок традиционно используют критерий Стьюдента, при этом условия его применимости не проверяют. Насколько обоснованными являются выводы? Как следует из примера 2, во многих случаях подвергать их сомнению нет оснований, хотя они получены с помощью некорректной процедуры. Во-вторых, более новые подходы зачастую методологически уязвимы. Так, известная робастная модель засорения Тьюки – Хубера нацелена на борьбу с большими выбросами, которые зачастую физически невозможны из-за ограниченности интервала значений измеряемой характеристики, в котором работает конкретное средство измерения. Следовательно, модель Тьюки – Хубера – Хампеля [25, 26] имеет скорее теоретическое значение, чем практическое. Сказанное, конечно, не означает, что следует прекратить разработку, изучение и внедрение непараметрических и устойчивых методов, выделенных выше как «точки роста» современной прикладной статистики.

Нерешенным проблемам статистики посвящены статьи [147, 148]. 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.

Другая проблема — обоснование выбора одного из многих критериев для проверки конкретной гипотезы. Например, для проверки однородности двух независимых выборок можно предложить 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. [19, 21]. Какой выбрать?

Критерии однородности проанализированы в работе [149]. Естественных подходов к сравнению критериев несколько — на основе асимптотической относительной эффективности по Бахадуру, Ходжесу – Леману, Питмену. И каждый критерий 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 [149] не позволяет дать рекомендации для выбора критерия проверки однородности при анализе реальных данных.

Проблемы разработки высоких статистических технологий поставлены в статье [141] (см. также одноименный сайт http://orlovs.pp.ru). Используемые при обработке реальных данных статистические технологии состоят из последовательности операций, каждая из которых, как правило, хорошо изучена, поскольку сводится к оцениванию (параметров, характеристик, распределений) или проверке той или иной гипотезы. Однако статистические свойства результатов обработки, полученных путем последовательного применения таких операций, мало изучены. Необходима теория, позволяющая оценивать 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 [8, 16, 21, 38]. При дальнейшем развитии исследований важно опираться на современную методологию [150].

ЛИТЕРАТУРА

1. *Орлов А. И.* Устойчивость в социально-экономических моделях. — М.: Наука, 1979. — 296 с.
2. *Орлов А. И.* Экспертные оценки. Вопросы кибернетики. Вып. 58. — М.: Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1979. С. 17 – 33.
3. *Тюрин Ю. Н., Литвак Б. Г., Орлов А. И., Сатаров Г. А., Шмерлинг Д. С.* / Заводская лаборатория. 1980. Т. 46. № 10. С. 931 – 935.
4. *Крамер Г.* Математические методы статистики. — М.: Мир, 1975. — 648 с.
5. *Смирнов Н. В., Дунин-Барковский И. В.* Курс теории вероятностей и математической статистики для технических приложений. Изд. 3-е, стереотипное. — М.: Наука, 1969. — 512 с.
6. *Большев Л. Н., Смирнов Н. В.* Таблицы математической статистики. 3-е изд. — М.: Наука, 1983. — 416 с.
7. *Каган А. М., Линник Ю. В., Рао С. Р.* Характеризационные задачи математической статистики. — М.: Наука, 1972. — 656 с.
8. *Орлов А. И.* Прикладная статистика. — М.: Экзамен, 2006. — 672 с.
9. Современные проблемы кибернетики (прикладная статистика). — М.: Знание, 1981. — 64 с.
10. *Орлов А. И.* / Вестник статистики. 1990. № 1. С. 65 – 71.
11. *Орлов А. И.* / Заводская лаборатория. Диагностика материалов. 1998. Т. 64. № 3. С. 52 – 60.
12. *Тюрин Ю. Н.* Непараметрические методы статистики. — М.: Знание, 1978. — 64 с.
13. *Налимов В. В.* Применение математической статистики при анализе вещества. — М.: Физматгиз, 1960. — 430 с.
14. *Новицкий П. В., Зограф И. А.* Оценка погрешностей результатов измерений. — Л.: Энергоатомиздат, 1985. — 248 с.
15. *Орлов А. И.* / Заводская лаборатория. 1991. Т. 57. № 7. С. 64 – 66.
16. *Орлов А. И.* Эконометрика. — М.: Экзамен, 2002. — 576 с.
17. *Селезнев В. Д., Денисов К. С.* / Заводская лаборатория. Диагностика материалов. 2005. Т. 71. № 1. С. 68 – 73.
18. *Орлов А. И.* / Там же. 2004. Т. 70. № 5. С. 65 – 70.
19. *Камень Ю. Э., Камень Я. Э., Орлов А. И.* / Там же. 1986. Т. 52. № 12. С. 55 – 57.
20. *Орлов А. И.* / Там же. 1999. Т. 65. № 1. С. 51 – 55.
21. *Орлов А. И.* / Там же. 2003. Т. 69. № 1. С. 55 – 60.
22. *Орлов А. И.* / Там же. 2004. Т. 70. № 7. С. 57 – 61.
23. *Смоляк С. А., Титаренко Б. П.* Устойчивые методы оценивания: Статистическая обработка неоднородных совокупностей. — М.: Финансы и статистика, 1980. — 208 с.
24. Устойчивые статистические методы оценки данных. — М.: Машиностроение, 1984. — 230 с.
25. *Хьюбер П.* Робастность в статистике. — М.: Мир. 1984. — 304 с.
26. *Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В.* Робастность в статистике. Подход на основе функций влияния. — М.: Мир, 1989. — 512 с.
27. *Орлов А. И.* / Заводская лаборатория. 1987. Т. 53. № 10. С. 82 – 85.
28. *Орлов А. И.* / Заводская лаборатория. Диагностика материалов. 1998. Т. 64. № 5. С. 64 – 67.
29. Дискуссия по анализу интервальных данных / Заводская лаборатория. 1990. Т. 56. № 7. С. 75 – 95.
30. *Вощинин А. П.* Метод оптимизации объектов по интервальным моделям целевой функции. — М.: МЭИ, 1987. — 109 с.
31. *Вощинин А. П., Акматбеков Р. А.* Оптимизация по регрессионным моделям и планирование эксперимента. — Бишкек: Изд-во «Илим», 1992. — 164 с.
32. *Вощинин А. П.* / Заводская лаборатория. Диагностика материалов. 2000. Т. 66. № 3. С. 51 – 64.
33. *Вощинин А. П.* / Там же. 2002. Т. 68. № 1. С. 118 – 126.
34. *Вощинин А. П., Бронз П. В.* / Там же. 2007. № 1. С. 101 – 109.
35. *Вощинин А. П., Скибицкий Н. В.* / Там же. 2007. Т. 73. № 11. С. 66 – 71.
36. *Гуськова Е. А., Орлов А. И.* / Там же. 2005. Т. 71. № 3. С. 57 – 63.
37. *Орлов А. И.* Теория принятия решений. — М.: Экзамен, 2006. — 576 с.
38. *Орлов А. И.* Организационно-экономическое моделирование. Ч. 1. Нечисловая статистика. — М.: Изд-во МГТУ им. Н. Э. Баумана, 2009. — 544 с.
39. *Таранцев А. А.* / Заводская лаборатория. Диагностика материалов. 2004. Т. 70. № 3. С. 60 – 65.
40. *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.
41. Первое Всесоюзное совещание по статистическому и дискретному анализу нечисловой информации, экспертным оценкам и дискретной оптимизации. / Тезисы докладов. — М. — Алма-Ата: ВИНИТИ, 1981. — 439 с.
42. Вторая Всесоюзная конференция по анализу нечисловой информации. / Тезисы докладов. — М.: Таллин: ВИНИТИ, 1984. — 348 с.
43. *Лбов Г. С.* Методы обработки разнотипных экспериментальных данных. — Новосибирск: Наука, 1981. — 160 с.
44. *Маамяги А. В.* Некоторые задачи статистического анализа классификаций. — Таллин: АН ЭССР, 1982. — 24 с.
45. *Миркин Б. Г.* Анализ качественных признаков и структур. — М.: Статистика, 1980. — 319 с.

46. *Перекрест В. Т.* Нелинейный типологический анализ социально-экономической информации: Математические и вычислительные методы. — Л.: Наука, 1983. — 176 с.
47. *Терехина А. Ю.* Анализ данных методами многомерного шкалирования. — М.: Наука, 1986. — 168 с.
48. *Хованов Н. В.* Математические основы теории шкал измерения качества. — Л.: Изд-во ЛГУ, 1982. — 185 с.
49. Анализ нечисловых данных в системных исследованиях / Сб. трудов. Вып. 10. — М.: ВНИИСИ, 1982. — 155 с.
50. Методы анализа данных, оценивания и выбора / Сб. трудов. Вып. 11. — М.: ВНИИСИ, 1984. — 92 с.
51. Методы анализа данных, оценивания и выбора в системных исследованиях. / Сб. трудов. Вып. 14. — М.: ВНИИСИ, 1986. — 108 с.
52. *Орлов А. И.* Разработка и исследование статистических методов моделирования и анализа объектов нечисловой природы. Дисс. в форме научного доклада докт. техн. наук. — М.: МЭИ, 1992. — 40 с.
53. *Пярна К. А.* Оптимальное разбиение метрического вероятностного пространства. Автореф. дисс. канд. физ.-мат. наук. Вильнюс, 1987. 20 с.
54. *Рыданова Г. В.* Некоторые вопросы статистического анализа случайных бинарных векторов. Автореф. дисс. канд. физ.-мат. наук. М., 1988. 16 с.
55. *Сатаров Г. А.* Многомерное шкалирования при анализе дихотомических данных о социально-экономических системах. Автореф. канд. техн. наук. М., 1985. 20 с.
56. *Трофимов В. А.* Модели и методы качественного факторного анализа матриц связи. Автореф. канд. техн. наук. Новосибирск, 1982. 18 с.
57. *Шер А. П.* Исследование тестовых методов диагностики и разработка на их основе алгоритмов обработки океанологической информации для задач рыбопромыслового прогнозирования. Автореф. дисс. канд. техн. наук. Владивосток, 1984. 19 с.
58. *Шмерлинг Д. С.* Разработка и исследование ранговых методов анализа информации для задач упорядочения элементов сложных систем. Автореф. дисс. канд. физ.-мат. наук. М., 1982. 21 с.
59. *Андреенков В. Г.*, *Орлов А. И.*, *Толстова Ю. Н.* (ответственные редакторы). Анализ нечисловой информации в социологических исследованиях. — М.: Наука, 1985. — 220 с.
60. *Орлов А. И.* / Заводская лаборатория. 1990. Т. 56. № 3. С. 76 – 83.
61. *Орлов А. И.* / Там же. 1995. Т. 61. № 3. С. 43 – 52.
62. *Орлов А. И.* / Там же. 1995. Т. 61. № 5. С. 43 – 51.
63. Психологические измерения: Сб. статей. — М.: Мир, 1967. — 195 с.
64. *Пфанцагль И.* Теория измерений. — М.: Мир, 1976. — 166 с.
65. *Толстова Ю. Н.* / Заводская лаборатория. Диагностика материалов. 1999. Т. 65. № 3. С. 49 – 56.
66. *Орлов А. И.* / Там же. 1999. Т. 65. № 3. С. 57 – 62.
67. *Барский Б. В.*, *Соколов М. В.* / Там же. 2006. Т. 72. № 1. С. 59 – 66.
68. *Орлов А. И.* / Там же. 2006. Т. 72. № 1. С. 67 – 70.
69. *Кемени Дж.*, *Снелл Дж.* Кибернетическое моделирование: Некоторые приложения. — М.: Советское радио, 1972. — 192 с.
70. *Орлов А. И.* / Заводская лаборатория. Диагностика материалов. 1995. Т. 61. № 1. С. 56 – 58.
71. *Орлов А. И.* / Там же. 1996. Т. 62. № 10. С. 45 – 46.
72. *Тырсин А. Н.* / Там же. 2005. Т. 71. № 11. С. 53 – 58.
73. *Богданов Ю. И.* / Там же. 1998. Т. 64. № 7. С. 56 – 61.
74. *Орлов А. И.* / Там же. 2003. Т. 69. № 3. С. 53 – 64.
75. *Богданов Ю. И.* / Там же. 2004. Т. 70. № 3. С. 51 – 59.
76. *Абусев Р. А.* Групповая классификация. Решающие правила и их характеристики. — Пермь: Изд-во Пермского университета, 1992. — 219 с.
77. *Апраушева Н. Н.* Новый подход к обнаружению кластеров. — М.: Вычислительный центр РАН, 1993. — 48 с.
78. Группировки и корреляции в экономико-статистических исследованиях: Сер. «Ученые записки по статистике». Т. 43. — М.: Наука, 1982. — 373 с.
79. *Миркин Б. Г.* Группировки в социально-экономических исследованиях: Методы построения и анализа. — М.: Финансы и статистика. 1985. — 223 с.
80. *Абусев Р. А.* / Заводская лаборатория. Диагностика материалов. 2003. Т. 69. № 3. С. 65 – 68.
81. *Штремель М. А.*, *Кудря А. В.*, *Иващенко А. В.* / Там же. 2006. Т. 72. № 5. С. 53 – 62.
82. *Борисова И. А.*, *Загоруйко Н. Г.*, *Кутненко О. А.* / Там же. 2008. Т. 74. № 1. С. 68 – 71.
83. *Лагутин М. Б.* / Там же. 2005. Т. 71. № 7. С. 53 – 57.
84. *Орлов А. И.* / Вероятностные процессы и их приложения. Межвуз. сб. науч. тр. — М.: МИЭМ, 1989. С. 118 – 123.
85. *Орлов А. И.* Задачи оптимизации и нечеткие переменные. — М.: Знание, 1980. — 64 с.
86. *Заде Д.* Понятие лингвистической переменной и его применение к принятию приближенных решений. — М.: Мир, 1976. — 168 с.
87. *Таранцев А. А.* / Заводская лаборатория. Диагностика материалов. 1999. Т. 65. № 1. С. 67 – 68.
88. *Хургин Я. И.* / Там же. 2000. Т. 66. № 1. С. 64 – 66.
89. *Гермашев И. В.*, *Дербишер В. Е.*, *Морозенко Т. Ф.*, *Орлова С. А.* / Там же. 2001. Т. 67. № 1. С. 65 – 68.
90. *Клементьева С. В.* / Там же. 2006. Т. 72. № 11. С. 65 – 67.
91. *Орлов А. И.* / Там же. 1997. Т. 63. № 3. С. 55 – 62.
92. *Орлов А. И.* / Там же. 1999. Т. 65. № 11. С. 51 – 55.
93. *Орлов А. И.* / Там же. 2000. Т. 66. № 1. С. 58 – 62.
94. *Корхин А. С.* / Там же. 2003. Т. 69. № 7. С. 52 – 58.
95. *Митрохин И. Н.*, *Орлов А. И.* / Там же. 2007. Т. 73. № 5. С. 74 – 78.
96. *Орлов А. И.* / Там же. 1986. Т. 52. № 5. С. 67 – 69.
97. *Струков Т. С.* / Там же. 2004. Т. 70. № 5. С. 60 – 64.
98. *Смоляк С. А.* / Там же. 2007. № 3. С. 69 – 77.
99. *Смоляк С. А.* / Там же. 2007. № 5. С. 67 – 73.
100. *Горбач А. Н.*, *Цейтлин Н. А.* / Там же. 2008. Т. 74. № 11. С. 62 – 68.
101. *Горбач А. Н.*, *Цейтлин Н. А.* / Там же. 2009. Т. 75. № 1. С. 66 – 69.
102. *Новиков Д. А.* Теория управления организационными системами. — М.: Московский психолого-социальный институт, 2005. — 584 с.
103. *Губко М. В.* Математические модели оптимизации иерархических структур. — М: ЛЕНАНД, 2006. — 264 с.
104. *Новиков Д. А.*, *Иващенко А. А.* Модели и методы организационного управления инновационным развитием фирмы. — М.: ЛЕНАНД, 2006. — 336 с.
105. *Анисимов С. Н.*, *Колобов А. А.*, *Омельченко И. Н.* и др. Проектирование интегрированных производственно-корпоративных структур: эффективность, организация, управление / Под ред. *А. А. Колобова*, *А. И. Орлова*. — М.: МГТУ им. Н. Э. Баумана, 2006. — 728 с.
106. *Новиков Д. А.* Управление проектами: организационные механизмы. — М.: ПМСОФТ, 2007. — 140 с.
107. *Новиков Д. А.* Математические модели формирования и функционирования команд. — М.: Физматлит, 2008. — 184 с.
108. *Шадрин А. П.* / Заводская лаборатория. Диагностика материалов. 2008. Т. 74. № 11. С. 68 – 72.
109. Статистические методы анализа экспертных оценок / Ученые записки по статистике. Т. 29. — М.: Наука, 1977. — 384 с.
110. Экспертные оценки / Вопросы кибернетики. Вып. 58. — М.: Научный Совет АН СССР по комплексной проблеме «Кибернетика», 1979. — 200 с.
111. Экспертные оценки в системных исследованиях / Сб. трудов. Вып. 4. — М.: ВНИИСИ, 1979. — 120 с.
112. Экспертные оценки в задачах управления / Сб. трудов. — М.: Институт проблем управления, 1982. — 106 с.
113. *Дэвид Г.* Метод парных сравнений. — М.: Статистика, 1978. — 144 с.

114. ГОСТ 23554.2–81. Экспертные методы оценки качества промышленной продукции. Обработка значений экспертных оценок качества продукции. — М.: Изд-во стандартов, 1981. — 69 с.
115. *Литвак Б. Г.* Экспертная информация: Методы получения и анализа. — М.: Радио и связь. 1982. — 184 с.
116. *Левин М. Ш.* Современные подходы к оценке эффективности плановых и проектных решений в машиностроении. — М.: ВНИИ информации и технико-экономических исследований по машиностроению и робототехнике, 1987. — 56 с.
117. *Сидельников Ю. В.* Теория и организация экспертного прогнозирования. — М.: Институт мировой экономики и международных отношений, 1990. — 196 с.
118. *Литвак Б. Г.* Экспертные технологии управления. 2-е изд. — М.: Дело, 2004. — 398 с.
119. *Сидельников Ю. В.* Технология экспертного прогнозирования: Учебное пособие. Изд. 2-е, исправл. — М.: Доброе слово, 2004. — 284 с.
120. *Сидельников Ю. В.* Системный анализ в экспертных технологиях. — М.: Доброе слово, 2008. — 354 с.
121. *Орлов А. И.* / Заводская лаборатория. Диагностика материалов. 1996. Т. 62. № 1. С. 54 – 60.
122. *Крушенко Г. Г.*, *Кокшаров И. И.*, *Торшилова С. И.*, *Крушенко С. Г.* / Там же. 2000. Т. 66. № 5. С. 64 – 67.
123. *Литвак Б. Г.* / Там же. 2000. Т. 66. № 7. С. 61 – 65.
124. *Шахнов И. Ф.* / Там же. 2005. Т. 71. № 5. С. 59 – 64.
125. *Файн В. Б.*, *Дель М. В.* / Там же. 2005. Т. 71. № 7. С. 58 – 59.
126. *Орлов А. И.* / Там же. 2005. Т. 71. № 7. С. 60 – 61.
127. *Стрижов В. В.* / Там же. 2006. Т. 72. № 11. С. 59 – 63.
128. *Глухов А. И.* / Там же. 2007. № 3. С. 77 – 79.
129. *Стрижов В. В.*, *Казакова Т. В.* / Там же. 2007. Т. 73. № 7. С. 72 – 75.
130. *Шахнов И. Ф.* / Автоматика и телемеханика. 2004. № 10. С. 67 – 84.
131. *Дорофеюк А. А.*, *Покровская И. В.*, *Чернявский А. Л.* / Автоматика и телемеханика. 2004. № 10. С. 172 – 188.
132. *Подиновский В. В.* / Там же. 2004. № 11. С. 141 – 159.
133. *Сидельников Ю. В.*, *Танасова А. С.* / Заводская лаборатория. Диагностика материалов. 2006. Т. 72. № 11. С. 59 – 64.
134. *Муравьева В. С.*, *Орлов А. И.* / Там же. 2008. Т. 74. № 1. С. 63 – 68.
135. *Муравьева В. С.* / Там же. 2008. Т. 74. № 3. С. 70 – 73.
135. *Крюкова Е. М.* / Там же. 2008. Т. 74. № 7. С. 67 – 72.
136. *Кендэл М.* Ранговые корреляции. — М.: Статистика, 1975. — 216 с.
137. *Холлендер М.*, *Вулф Д. А.* Непараметрические методы статистики. — М.: Финансы и статистика, 1983. — 518 с.
138. *Гнеденко Б. В.*, *Орлов А. И.* / Заводская лаборатория. 1988. Т. 54. № 1. С. 1 – 4.
139. *Орлов А. И.* / Там же. 1992. Т. 58. № 1. С. 67 – 74.
140. *Горский В. Г.*, *Орлов А. И.* / Там же. 2002. Т. 68. № 1. С. 108 – 112.
141. *Орлов А. И.* / Там же. 2003. Т. 69. № 11. С. 55 – 60.
142. *Орлов А. И.* / Там же. 2006. Т. 72. № 5. С. 50 – 53.
143. *Орлов А. И.* / Там же. 1996. Т. 62. № 7. С. 46 – 49.
144. *Смирнова О. С.* / Там же. 2008. Т. 74. № 5. С. 68 – 75.
145. *Орлов А. И.* / Там же. 2008. Т. 74. № 5. С. 76 – 78.
146. *Орлов А. И.* / Там же. 1992. Т. 58. № 7. С. 40 – 42.
147. *Загоруйко Н. Г.*, *Орлов А. И.* Современные проблемы кибернетики (прикладная статистика). — М.: Знание, 1981. С. 53 – 63.
148. *Орлов А. И.* / Заводская лаборатория. Диагностика материалов. 2002. Т. 68. № 3. С. 52 – 56.
149. *Никитин Я. Ю.* Асимптотическая эффективность непараметрических критериев. — М.: Наука, 1995. — 240 с.
150. *Новиков А. М.*, *Новиков Д. А.* Методология. — М.: СИНТЕГ, 2007. — 668 с.