

УДК 519.2

## О новой парадигме прикладной математической статистики

А.И. Орлов

Московский государственный технический университет им. Н.Э. Баумана, Институт высоких статистических технологий и эконометрики; Московский физико-технический институт; Москва, Россия;  
prof-orlov@mail.ru; <http://orlovs.pp.ru>; +7(916)8305117; 123104, Москва, Сытинский пер., д.7/14, кв.14.

**Аннотация.** Новая парадигма математической статистики основана на переходе от параметрических статистических методов к непараметрическим, от числовых данных - к нечисловым, на интенсивном использовании информационных технологий. Ее отличительные черты выявлены в сравнении со старой парадигмой математической статистики середины XX в.

**Ключевые слова:** *о математическая статистика, непараметрические методы, нечисловая статистика, информационные технологии, новая парадигма.*

Математическая статистика – это наука о том, как обрабатывать данные. Так начинается учебник "Прикладная статистика" [1]. Противоречия в терминах нет. В середине XX в. наука об обработке данных называлась математической статистикой. Затем – прикладной статистикой. Настало время вернуться к прежнему названию.

За сменой названий стоит смена парадигм. Под парадигмой понимаем систему идей, взглядов и понятий, различных моделей решения проблем, а также методов исследования. Цель работы - описать и сравнить две парадигмы в области методов статистического анализа данных – старую и новую.

### 1. Послевоенное развитие отечественной статистики

По нашей экспертной оценке, основные достижения математики в XX в. – аксиоматизация теории вероятностей, создание математической статистики и линейного программирования.

К середине XX в. в СССР, как и за рубежом, сформировалась научно-практическая дисциплина, которую называем классической *математической статистикой*. Специалисты-статистики учились теории по книге Г.Крамера [2], написанной в военные годы и впервые изданной на русском языке в 1948 г. Из прикладных руководств назовем учебник [3] и таблицы с комментариями [4].

Затем внимание многих специалистов сосредоточилось на изучении математических конструкций, используемых в статистике. Примером таких работ является монография [5]. В ней получены продвинутое математические результаты, но из них трудно выделить рекомендации для статистика, анализирующего конкретные данные.

Что же послужило причиной такого сдвига интересов? Некоторые особенности исторического развития нашей страны.

Большой вред развитию статистической науки и практики нанесло Всесоюзное совещание статистиков 1954 г. На нем было принято решение, что статистика – это одна из экономических наук, фактически – ведомственная наука ЦСУ–Госкомстата–Росстата РФ (Федеральной службы государственной статистики). При этом организаторы совещания не посмели покуситься на само существование математической статистики, но отнесли ее исключительно внутрь математики, в которой была выделена специальность "теория вероятностей и математическая статистика". Все остальные области применения статистических методов перестали замечаться официальными структурами, т.е. стали полуполевыми. Конечно, специалисты нашли способы противодействия. Например, статистические методы в химии относились к "химической кибернетике", статистические методы в медицине – к "математическому моделированию в медицине".

В результате решений Всесоюзного совещания статистиков 1954 г. работы по математической статистике стали рассматриваться исключительно с позиций математики. Стали цениться изолированные теоремы (типа полученных в монографии [5]), никак не связанные с анализом реальных данных. В то же время вопросы практики применения статистических методов стали отодвигаться на задний план и даже подвергаться гонениям. Типичным примером является провал при защите на мехмате МГУ им. М.В.Ломоносова в 1971 г. докторской диссертации В.В.Федорова, в которой были получены базовые результаты в области планирования эксперимента – одного из важнейших направлений прикладных статистических методов.

Как реакция на уход в математику выделилась новая научная дисциплина – *прикладная статистика*. В учебнике [1] в качестве рубежа, когда это стало очевидным, указан 1981 г. – дата выхода массовым тиражом (33 940 экз.) сборника [6], в названии которого использован термин "прикладная статистика" (полное название: "Современные проблемы кибернетики (прикладная статистика)"). С этого времени линии развития математической статистики и прикладной статистики окончательно разошлись. Первая из них полностью ушла в математику, перестав интересоваться практическими делами. Вторая позиционировала себя в качестве науки об обра-

ботке данных – результатов наблюдений, измерений, испытаний, анализов, опытов.

Вполне естественно, что в прикладной статистике стали развиваться свои математические методы и модели. Необходимость их появления вытекает из потребностей конкретных прикладных исследований. Это математизированное ядро прикладной статистики вполне естественно назвать "теоретической статистикой" (или прикладной математической 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. Основным местом публикации работ по прикладной статистике с 1960-х гг. стал раздел "Математические методы исследования" журнала "Заводская лаборатория".

С рассматриваемой точки зрения общая схема современной статистической науки выглядит, по нашему мнению, 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 – хемометрия, в технических исследованиях – технометрика, в геологии, демографии, социологии, медицине, истории, 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 рассмат-

риваемой 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.

**Замечание.** Здесь мы уточнили схему внутреннего деления статистической теории, предложенную нами ранее в [7]. Естественный смысл приобрели термины "теоретическая статистика" и "прикладная статистика" (в узком смысле). Однако необходимо иметь в виду, что в недавнем учебнике [1] прикладная статистика понимается в широком смысле, т.е. как объединение позиций 2 и 3.

Отметим, что математическая статистика, как и теоретическая с прикладной, заметно отличается от ведомственной науки органов официальной государственной статистики. ЦСУ, Госкомстат, Росстат РФ применяли и применяют лишь проверенные временем приемы позапрошлого (19-го) века. Приходится с сожалением констатировать, что большинство учебных курсов в экономических ВУЗах и учебников с названиями "Статистика" или "Общая теория статистики" понимают под последней статистику в смысле Росстата. Возможно, следовало бы от этого ведомства полностью отмежеваться и сменить название дисциплины, например, на "Анализ данных". В настоящее время компромиссным самоназванием рассматриваемой научно-практической дисциплины является термин "*статистические методы*".

По нашей экспертной оценке, в последнее время стало ясно, что абстрактно - математическое направление в статистике в значительной мере снизило темпы развития. В то же время в теоретической и прикладной статистике - масса новых результатов. Настало время последовать словам члена-корреспондента АН СССР Л.Н. Большева (06.03.1922 - 29.08.1978) "Вся математическая статистика является прикладной" ([6]). Мы считаем необходимым ***объединить выделенные выше ветви статистики – математическую, теоретическую и прикладную – в единую математическую статистику.*** Первую из ветвей предлагаем называть "математическими методами в статистике как ее и называл основоположник Г. Крамер [2].

Во второй половине 80-х годов в СССР 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. Все-союзной статистической ассоциации (ВСА) [8] оказалась парализованной в результате развала СССР.

В ходе создания ВСА проанализировано состояние и перспективы раз-

вития статистических методов. Коллективными усилиями создана *новая парадигма математической статистики*, основанная, в частности, на переходе от параметрической статистики, основанной на использовании параметрических семейств распределений, к непараметрической и нечисловой статистике. Выявлена необходимость создания нового поколения учебной литературы, которая должна сменить издания на основе идей середины XX в.

Какова ситуация в других странах, в мире в целом? От англо-саксонских стран мы отстали в деле создания профессионального объединения статистиков. За 180 лет на Западе был пройден долгий путь организационного строительства, найдены способы эффективного сочетания теории и практики, объединения статистиков различных направлений – от теоретиков-вероятностников до работников официальной государственной статистики. Центральным ядром была и остается математическая статистика. Есть и развитие: основной статистический журнал "The Annals of Mathematical Statistics", созданный в 1930 г., в 1973 г. был разделен на два - "The Annals of Statistics" и "The Annals of Probability". Общее число научных журналов, публикующих работы по теории вероятностей, математической и прикладной статистике, Д.С. Шмерлинг [9] оценивает как 1000-1500.

Статистические методы основаны на теории вероятностей. В СССР была создана самая мощная в мире научная школа по теории вероятностей. К сожалению, это стратегическое преимущество не удалось реализовать в виде адекватного развития статистических методов. Число специалистов (на миллион граждан страны) в СНГ на порядок меньше, чем в других развитых странах. Если бы удалось ввести в средней школе полноценный курс вероятности и статистики – а такой курс есть в Японии и США, Швейцарии, Кении и Ботсване, почти во всех странах, кроме стран бывшего социалистического лагеря (см. подготовленный ЮНЕСКО сборник докладов [10]) – то ситуация могла бы быть резко улучшена. Такой курс должен использовать высокие статистические технологии [11], а не на устаревшие. Другими словами, он должен отражать современные достижения, а не концепции пятидесятилетней давности.

## **2. Основные черты новой парадигмы математической статистики**

Демонстрацией необходимости новой парадигмы является само появление новой научной области – прикладной статистики. Сравнение парадигм удобно провести с помощью табл.1, в которой выделены 17 основных характеристик систем идей, взглядов и понятий.

Таблица 1. Сравнение новой и старой парадигм

№	Характеристика	Старая парадигма	Новая парадигма
1	Типовые исходные данные	Числа, вектора, функции	Объекты нечисловой природы
2	Основной подход к описанию данных	Распределения из параметрических семейств	Произвольные(непрерывные) функции распределения
3	Математический аппарат	Суммы	Расстояния и алгоритмы оптимизации
4	Источник постановок задач	Математические традиции, сформировавшиеся к середине XX века	Современные потребности анализа данных (XXI век)
5	Отношение к вопросам устойчивости выводов	Практически отсутствует интерес к устойчивости выводов	Развитая теория устойчивости (робастности)
6	Оцениваемые величины	Параметры распределений	Характеристики и плотности распределений, зависимости и др.
7	Возможность применения	Наличие повторяющегося комплекса условий	Наличие обоснованной вероятностно-статистической модели
8	Центральная часть теории	Статистика числовых случайных величин	Статистика в пространствах произвольной природы
9	Роль информационных технологий	Только для расчета таблиц	Инструмент получения выводов (датчики псевдослучайных чисел, размножение выборок, в т.ч. бутстреп, и др.)
10	Учет соображений информатики	Информатика находится вне статистики	Учет возможностей проведения расчетов
11	Точность данных	Данные полностью известны	Учет свойств данных, в частности, интервальных
12	Типовые результаты	Предельные теоремы	Рекомендации для конкретных объемов выборок
13	Вид постановок задач	Отдельные задачи	Статистические технологии
14	Стыковка алгоритмов	Не рассматривается	Весьма важна
15	Роль моделирования	Отдельные системы аксиом	Системы моделей
16	Анализ экспертных оценок	Отдельные алгоритмы	Прикладное "зеркало" общей теории
17	Роль методологии	Практически отсутствует	Основополагающая

Основные составляющие новой парадигмы подробно разбирались на страницах журнала "Заводская лаборатория". Развитие нечисловой статистики за 30 лет проанализировано в [12]. Пять актуальных направлений, в которых развивается современная прикладная статистика, т.е. пять "точек роста" – непараметрика, робастность, бутстреп, статистика интервальных данных, статистика объектов нечисловой природы – разобраны в [13]. Проблемы устойчивости (робастности) рассмотрены в [14]. Большую роль стала играть методология [15, 16] – при постановках задач, построении систем моделей.

Стало ясно, что надо разделять четыре уровня работ - методологический, теоретический, методический, прикладной (см. табл.2).

**Таблица 2. Четыре уровня работ по математической статистике**

<b>Методологический уровень.</b> Как ставить задачу (как организовать моделирование)? Основные идеи метода
<b>Теоретический уровень.</b> Проработка основных идей, доказательство теорем
<b>Методический уровень.</b> Разработка методик, алгоритмов, программных продуктов, рекомендаций по практическому использованию
<b>Прикладной уровень.</b> Практическое использование: адаптация и применение разработанного метода при решении конкретных практических задач

Особенно важным представляется разделение методологического и теоретического уровней, а именно, выявление актуальных для практики и подлежащих решению постановок задач и цепочек теорем, посвященных все более продвинутому математическим результатам решения поставленной на методологическом уровне задачи (в качестве общеизвестного примера можно указать на двухсотлетнюю историю Центральной предельной теоремы).

Выявилась роль статистики в пространствах произвольной природы [12, 17], позволяющей единообразным образом анализировать как результаты измерений, наблюдений, испытаний, анализов, опытов, так и экспертные оценки разнообразных видов [18, 19]. В частности, оказалось, что задачи классификации [20] наиболее естественно ставить и решать в рамках статистики в пространствах произвольной природы и тем самым относить их к нечисловой статистике [17], а не к многомерному статистическому анализу.

Принципиально важным является понятие "высокие статистические технологии" [11], вводящее в прикладную статистику производственный

процесс анализа данных, состоящий из последовательностей операций, каждая из которых - оценивание, проверка гипотезы или иная операция, используемая и в математической статистике, но по отдельности, без объединения в процесс. Важной оказалась проблема "стыковки алгоритмов", т.е. проверки условий применимости последующей статистической операции (алгоритма) после окончания выполнения предыдущей. Например, перед проведением регрессионного анализа иногда рекомендуют разбить совокупность данных на однородные части, т.е. провести классификацию. Однако после классификации выделенные подсовкупности сосредоточены в отдельных частях исходного пространства, следовательно, моделирующие их плотности распределения равны 0 вне соответствующих частей, а потому не могут иметь нормальное распределение. Следовательно, к полученным подсовкупностям результатов измерений (наблюдений, испытаний, анализов, опытов) принципиально нельзя применять методы регрессионного анализа, предполагающие нормальность погрешностей (ошибок, невязок). Следует использовать непараметрические методы регрессии, не опирающиеся на заведомо не выполненную в рассматриваемом случае гипотезу нормальности погрешностей.

За рубежом по каждому из перечисленных узких научных направлений ведутся научные исследования. Однако в единое целое – в новую научную парадигму – они интегрированы именно в нашей стране. Важно эффективно использовать это наше конкурентное преимущество - общее достояние российского научного сообщества.

### **3. Последствия перехода на новую парадигму**

В качестве примера рассмотрим подготовку специалистов, поскольку именно она определяет уровень выполняемых ими в дальнейшем научных и прикладных работ. При переходе на преподавание согласно новой парадигме математической статистики необходимо существенно изменить содержание традиционного для университетов и технических вузов учебного курса "Теория вероятностей и математическая статистика". В частности, необходимо изучать случайные величины (если угодно, случайные элементы) со значениями в произвольных пространствах (в пространстве бинарных отношений, конечных множеств, других нелинейных пространствах); центральные предельные теоремы в полном объеме – для разнораспределенных слагаемых, в многомерном пространстве; средние величины в произвольных пространствах как решения оптимизационных задач; законы больших чисел в произвольных пространствах; непараметрический подход ко всем основным задачам прикладной математической статистики.



В то же время почти отпадает необходимость в изучении таких традиционных тем, как геометрические вероятности; параметрические семейства распределений (за исключением нормального распределения, появляющегося в центральных предельных теоремах); параметрические постановки в математической статистике, достаточные статистики, неравенство Рао-Крамера, метод максимального правдоподобия и метод одношаговых оценок; проверка параметрических гипотез с использованием распределений Стьюдента и Фишера.

Ряд проблем связан с использованием распространенных программных продуктов при преподавании. Очевидно, что математические методы исследования, в том числе методы статистического анализа данных, требуют больших вычислений и зачастую невозможны без компьютеров. Применение новой парадигмы прикладной статистики, продвинутое применение высоких статистических технологий [11] предполагает использование соответствующих программных продуктов. Статистические пакеты - постоянно используемые интеллектуальные инструменты исследователей, инженеров, управленцев, занимающихся анализом больших массивов данных. Более 20 статистических пакетов, разработанных под нашим руководством Всесоюзным центром статистических методов и информатики, в том числе пакеты СПК, АТСТАТ-ПРП, СТАТКОН, АВРОРА-РС, ЭКСПЛАН, ПАСЭК, НАДИС, проанализированы в [21, 22]. Сравнительному анализу четырех диалоговых систем по статистическому контролю посвящена статья [23], и т.д. Однако наряду с очевидной пользой статистические пакеты могут приносить вред неискушенному пользователю. Например, в них зачастую пропагандируется применение двухвыборочного критерия Стьюдента, когда условия его применимости не проверены, а зачастую и не выполнены. Между тем хорошо известно, каковы отрицательные последствия использования критерия Стьюдента вне сферы его применимости, а также и то, что применять его нет необходимости, поскольку разработаны более адекватные критерии [24].

Другой пример. Малограмотность переводчиков в русифицированной версии MS Excel (по крайней мере в разделе "Анализ данных") шокирует специалиста по прикладной статистике: например, "объем выборки" именуется "счет". С сожалением приходится констатировать, что не соответствует современным требованиям и электронный учебник - обзор методов, реализованных в пакете STATISTICA-6. Анализ допущенных в документации к пакету недочетов занял бы не меньше места, чем сама документация. В [21] продемонстрировано, насколько трудоемким оказался критический анализ всего лишь нескольких десятков ГОСТов по статистическим мето-

дам управления качеством.

Это замечание касается, конечно, не только пакетов. Из одной публикации в другую "кочуют" одни и те же ошибки. Для анализа каждой нужна развернутая публикация. Например, распространенная ошибка при использовании критериев Колмогорова и омега-квадрат разобрана в статье [25], ошибочные утверждения о том, какие гипотезы можно проверять с помощью двухвыборочного критерия Вилкоксона, рассмотрены в [26].

Основное противоречие в области разработки статистических пакетов на настоящий момент таково. Те, кто программирует, не являются специалистами по прикладной статистике, поскольку это не входит в их профессиональные обязанности. С другой стороны, специалисты по статистическим методам не берутся реализовывать их в программных продуктах, поскольку такая работа, весьма трудоемкая и ответственная, обычно не соответствует их профессиональным устремлениям. Судя по опыту Всесоюзного центра статистических методов и информатики, стоимость разработки (на профессиональном уровне) пакета среднего уровня сложности - порядка 70 тыс. руб. (в ценах 1990 г.), что соответствует 10,5 млн. руб. РФ в ценах 2012 г. (индекс инфляции за 22 года примерно равен 150 при расчете по методике [27]). Это означает, что разработкой, распространением и сопровождением статистических пакетов должны заниматься специализированные на этом организации или подразделения, в том числе коммерческие организации - как во всем мире.

В нашей стране активная работа по созданию развернутой системы отечественных статистических пакетов развернулась в 80-х годах. Как уже отмечалось, только Всесоюзным центром статистических методов и информатики было разработано более 20 программных продуктов по прикладной статистике и другим статистическим методам. Эта работа проводилась в рамках нашего более широкого проекта, нацеленного на объединение усилий специалистов по статистическим методам с целью повышения эффективности теоретических и прикладных исследований. Важным промежуточным итогом было создание в 1990 г. Всесоюзной организации по статистическим методам и их применениям и Всесоюзной статистической ассоциации [8]. Планы тех лет отражены в статье [28]. Итогом выдилось создание (развертывание, организационное оформление) новой отрасли прикладной науки по образцу метрологии.

Развал СССР, либерализация цен и гиперинфляция начала 90-х положили конец рассматриваемому проекту. Из плана работ реализована только подготовка серии современных учебников [1, 17, 19, 29, 30, 33 и др.], составленных на основе наших научных статей (учебники выложе-

ны в свободном доступе на сайте "Высокие статистические технологии" <http://orlovs.pp.ru> и странице <http://ibm.bmstu.ru/nil/biblio.html> Лаборатории экономико-математических методов в контроллинге МГТУ им. Н.Э. Баумана). Предприятия и организации, лишившись оборотных средств из-за инфляции, перестали покупать статистические программные продукты, коллективы разработчиков распались, перестали поддерживать статистические пакеты в условиях быстрого обновления технических средств и базового программного обеспечения. В результате многообразие продуктов на отечественном рынке статистических пакетов резко сократилось, и монополистами оказались SPSS, STATISTICA, STATGRAPHICS (и немногие другие). При анализе программных продуктов целесообразно сказать несколько слов об аналитических надстройках над распространенными системами широкого назначения, такими, как, например, Oracle и т.п. Приходится констатировать, что в них примитивная статистика сочетается с хорошей визуализацией, облегчающей практическое использование подобных систем и при этом, увы, дающей импульс распространению устаревших статистических методов.

На опасность бездумного применения статистических пакетов В.В. Налимов обращал внимание еще около 40 лет назад [31]. Он имел в виду прежде всего склонность к проведению расчетов без знакомства с сутью применяемых методов. Необходимо обратить внимание также на научно-технический уровень самих пакетов и сопровождающей документации. Дополнительно к сказанному ранее приходится констатировать, что в популярных в настоящее время в России статистических пакетах нет примерно половины того, что разработано представителями отечественной вероятно-статистической научной школы и включено в современные учебники [1, 17,27], подготовленные в соответствии с рекомендациями Всесоюзной статистической ассоциации и позже – Российской ассоциации статистических методов. Сказанное легко проверить, сопоставив содержание указанных учебников и перечень методов, включенных в распространенные пакеты. Поэтому в научно-учебном комплексе "Инженерный бизнес и менеджмент" МГТУ им. Н.Э. Баумана мы сознательно избегаем использования в учебном процессе популярных в настоящее время пакетов, чтобы не приучать студентов к статистике 60-70-х годов прошлого века. Однако, поскольку нет современных пакетов, приходится для практических расчетов использовать устаревшие программные продукты.

Тиражи пакетов и учебников сопоставимы. Пакет STATGRAPHICS имеет более 40 тыс. зарегистрированных пользователей, учебник [1] выпущен тиражом 3 тыс. экземпляров, его электронную версию только с сайта

”Высокие статистические технологии” на 29 апреля 2013 г. скачали около 43 тыс. пользователей

Поэтому состав пакетов и качество документации имеют большое значение. Они во многом определяют качество прикладных научных работ и обоснованность хозяйственных решений.

Отметим, что по сравнению с 80-ми годами к настоящему времени наметился рост внимания к статистическим технологиям [11], а не только к их составляющим - конкретным методам обработки данных. В этом суть популярного ныне подхода Data Mining (на русском – ”добыча данных”, ”интеллектуальный анализ данных”). Термин Data Mining введен эмигрантом из СССР Г. Пятецким-Шапиро в 1989 г. Задачи, решаемые Data Mining, – классификация, кластеризация, регрессия, ассоциация (поиск повторяющихся паттернов, например, поиск устойчивых связей) – это типичные задачи прикладной статистики. Новизна состоит в разработке технологий добычи данных путем решения не одной, а ряда таких задач.

Итак, статистические пакеты – интеллектуальные инструменты, необходимые широким кругам научных работников, инженеров, менеджеров. Однако распространенные в настоящее время статистические программные продукты отстают от современного уровня научных исследований примерно на 30 лет. Весьма актуальна задача разработки статистических пакетов нового поколения, соответствующих современному научному уровню и одновременно обеспечивающих удобства пользователей, достигнутые в популярных ныне пакетах. Эта задача должна решаться одновременно с созданием систем обучения, сопровождения и внедрения пакетов нового поколения, в частности, в соответствии с технологиями типа ”Шесть сигм” [32].

Подчеркнем, что математическая статистика является общенаучной дисциплиной. В США статистиков больше, чем математиков [31]. Этот перекос необходимо исправлять. У математической статистики должен быть такой же статус и такая же инфраструктура, как у (чистой) математики. В частности, система научно-исследовательских институтов, статистические факультеты в ведущих вузах, отделение в составе РАН, и т.д., и т.п.

Должна быть организована подготовка студентов всех специальностей по современным методам прикладной статистики и смежным дисциплинам, а также переподготовка специалистов. Важно от старой парадигмы 1950-х годов, распространенный при обучении теории вероятностей и математической статистики, перейти к новой парадигме, выраженной, например, в учебниках [1, 17, 19, 27, 29, 30, 33]. Повторим основную мысль: новая парадигма основана, в частности, на непараметрической и нечисловой ста-

тистике, а старая – на нереалистических предположениях о возможности применения семейств параметрических распределений данных (результатов измерений, наблюдений, испытаний, анализов, опытов) в инженерных, управленческих, экономических, медицинских и других практических задачах, встающих перед научными и практическими работниками, поэтому для модернизации народного хозяйства и эффективного внедрения инноваций необходимо сменить парадигму анализа данных – от представлений середины XX в. перейти к современным.

В настоящей статье использованы материалы доклада [34] на конференции "Статистика и ее применения – 2012" в Ташкенте.

### Библиографический список

1. Орлов А.И. Прикладная статистика. – М.: Экзамен, 2006. – 671 с.
2. Крамер Г. Математические методы статистики. – М.: Мир, 1975. – 648 с.
3. Смирнов Н.В., Дунин-Барковский И.В. Курс теории вероятностей и математической статистики для технических приложений. Изд. 3-е, стереотипн. – М.: Наука, 1969. – 512 с.
4. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики / 3-е изд. – М.: Наука, 1983. – 416 с.
5. Каган А.М., Линник Ю.В., Рао С.Р. Характеризационные задачи математической статистики. – М.: Наука, 1972. – 656 с.
6. Современные проблемы кибернетики (прикладная статистика). – М.: Знание, 1981. – 64 с.
7. Орлов А.И. О перестройке статистической науки и ее применений // Вестник статистики. – 1990. – №1. – С.65 – 71.
8. Орлов А.И. Создана единая статистическая ассоциация // Вестник Академии наук СССР. – 1991. – №7. – С.152–153.
9. Шмерлинг Д.С. Журналы по теории вероятностей, математической статистике и их применениям / Вероятность и математическая статистика: Энциклопедия / Гл. ред. Ю.В. Прохоров. – М.: Большая Российская энциклопедия, 1999. – С.893-910.
10. The teaching of statistics / Studies in mathematics education. Paris, UNESCO, – V.7. – 1989. – 258 pp.
11. Орлов А.И. Высокие статистические технологии // Заводская лаборатория. – 2003. – Т.69. – №11. – С.55-60.
12. Орлов А.И. Тридцать лет статистики объектов нечисловой природы (обзор) // Заводская лаборатория. Диагностика материалов. – 2009. – Т.75. – №5. – С.55-64.
13. Горский В.Г., Орлов А.И. Математические методы исследования: итоги и перспективы // Заводская лаборатория. – 2002. – Т.68. – №1. – С.108-112.

14. Орлов А.И. Устойчивые математические методы и модели// Заводская лаборатория. Диагностика материалов. – 2010. – Т.76. – №3. – С.59-67.
15. Новиков А.М., Новиков Д.А. Методология. - М.: СИНТЕГ, 2007. - 668 с.
16. Орлов А.И. О развитии методологии статистических методов// Статистические методы оценивания и проверки гипотез. Межвузовский сборник научных трудов. - Пермь: Изд-во Пермского гос. университета, 2001. – С.118-131.
17. Орлов А.И. Организационно-экономическое моделирование: учебник : в 3 ч. Часть 1: Нечисловая статистика. - М.: Изд-во МГТУ им. Н.Э. Баумана. 2009. – 541 с.
18. Орлов А.И. О развитии экспертных технологий в нашей стране// Заводская лаборатория. Диагностика материалов. – 2010. – Т.76. – №11. – С.64-70.
19. 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.
20. Орлов А.И. О развитии математических методов теории классификации// Заводская лаборатория. Диагностика материалов. – 2009. – Т.75. – №7. – С.51-63.
21. Орлов А.И. Сертификация и статистические методы (обобщающая статья)// Заводская лаборатория. – 1997. – Т.63. – №3. – С.55-62.
22. Орлов А.И. Внедрение современных статистических методов с помощью персональных компьютеров// Качество и надежность изделий. - М.: Знание, 1992. – №5(21). – С.51-78.
23. Орлов А.И. Математическое обеспечение сертификации: сравнительный анализ диалоговых систем по статистическому контролю// Заводская лаборатория. – 1996. – Т.62. – №7. – С.46-49.
24. Орлов А.И. О проверке однородности двух независимых выборок// Заводская лаборатория. – 2003. – Т.69. – №1. – С.55-60.
25. Орлов А.И. Распространенная ошибка при использовании критериев Колмогорова и омега-квадрат// Заводская лаборатория. – 1985. – Т.51. – №1. – С.60-62.
26. Орлов А.И. Какие гипотезы можно проверять с помощью двухвыборочного критерия Вилкоксона?// Заводская лаборатория. – 1999. – Т.65. – №1. – С.51-55.
27. Орлов А.И. Эконометрика. Изд. 4-е. - Ростов-на-Дону: Феникс, – 2009. – 572 с.
28. Орлов А.И. О современных проблемах внедрения прикладной статистики и других статистических методов// Заводская лаборатория. – 1992. – Т.58. – №1. – С.67-74.
29. Орлов А.И. Теория принятия решений. - М.: Экзамен, 2006. – 576 с.
30. Орлов А.И. Организационно-экономическое моделирование: теория принятия решений. - М. : КноРус, 2011. – 568 с.

31. *Налимов В.В.* О преподавании математики экспериментаторам// О преподавании математической статистики экспериментаторам. Препринт Межфакультетской лаборатории статистических методов - М.: Изд-во МГУ им. М.В. Ломоносова, 1971. – №17. – С.5-39.
32. *Орлов А.И.* . "Шесть сигм"- новая система внедрения математических методов исследования// Заводская лаборатория. Диагностика материалов. – 2006. – Т.72. – №5. – С.50-53
33. *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.
34. *Орлов А.И.* Новая парадигма математической статистики// Материалы республиканской научно-практической конференции "Статистика и ее применения - 2012". Под редакцией профессора А.А. Абдушукурова. - Ташкент: НУУз, – 2012. – С.21-36.

## A new paradigm of Applied Mathematical Statistics

A.I. Orlov

Bauman Moscow State Technical University, Institute of high statistical technologies and econometrics; Moscow Physics-Technical Institute; Moscow, Russia;  
prof-orlov@mail.ru; <http://orlovs.pp.ru>; +7(916)8305117; 123104, Moscow, the Sytinsky lane, house 7/14, apartment 14.

**Abstract.** The new paradigm of mathematical statistics based on the transition from parametric to nonparametric statistical methods, the numerical data - to non-numeric, on the intensive use of information technology. Its distinctive features are revealed in comparison with the old paradigm of mathematical statistics in the mid-twentieth century.

**Key words:** mathematical statistics, nonparametric methods, non-numeric statistics, information technology, the new paradigm.