

DOI: <https://doi.org/10.26896/1028-6861-2024-90-7-5-7>

## РЕВОЛЮЦИЯ В МАТЕМАТИЧЕСКИХ МЕТОДАХ ИССЛЕДОВАНИЯ

© Александр Иванович Орлов

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

## THE PARADIGM SHIFT IN MATHEMATICAL METHODS OF RESEARCH

© Alexander I. Orlov

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

Математические методы исследования — интеллектуальные инструменты для применения в различных областях. На математические методы можно взглянуть с точек зрения как прикладников, применяющих такие методы, так и теоретиков, их разрабатывающих.

Прикладники обычно считают, что совокупность нужных им математических методов давно разработана, все необходимое для практического применения изложено в учебниках и справочниках, для проведения расчетов достаточно распространённых программных продуктов, а теоретики занимаются отдельными мелкими улучшениями и вникать в их работы нет необходимости, нецелесообразно, поскольку времени всегда не хватает.

Теоретики знают, что за последние десятилетия в области математических методов исследования произошла принципиально важная научная революция, в ходе которой создана новая методология, разработаны резко отличающиеся от прежних модели и методы. Усилиями этой категории исследователей научная революция осуществлена и продолжает развиваться. Весьма важно добиться широкого использования новых методов.

В настоящее время между воззрениями прикладников и теоретиков в области математических методов исследований наблюдаем значительное различие. Для его уменьшения необходимо разъяснить научному сообществу существо обсуждаемой научной революции.

В хорошо знакомых прикладникам учебниках и справочниках середины XX в. в качестве статистических данных рассматривались числовые величины, т.е. действительные числа, конечномерные вектора (т.е. конечные последовательности чисел), функции с числовыми значениями (временные ряды, случайные процессы). Термин «числовые» означает, что элементы выборки

можно складывать и умножать на число, т.е. эти элементы лежат в некотором линейном пространстве.

В результате научной революции конца XX — начала XXI вв. произошел отказ от предположения линейности. В качестве выборочных данных стали рассматривать элементы пространств произвольной природы. Центром математических методов исследования стала статистика нечисловых данных. Внутри нее надо выделить статистику в пространствах произвольной природы, которая включает научные результаты в наиболее общих формулировках, и области, посвященные анализу нечисловых данных конкретных видов. При таком анализе используются как результаты статистики в пространствах произвольной природы, так и специфические методы, предназначенные для данных конкретных видов нечисловых данных.

Другая принципиально важная черта научной революции — обобщение классических типов чисел путем явного учета размытости (нечеткости, расплывчатости) реальных статистических данных. Для всех видов измерений их результаты имеют погрешности, однако классические статистические методы не учитывают наличие погрешностей. Для преодоления этого недостатка разработана статистика интервальных данных, в которых элементы выборки — не числа, а интервалы. Учет погрешностей измерений может быть проведен и путем перехода к анализу нечетких данных, например, треугольных нечетких чисел, как это продемонстрировано в [1] на примере аддитивно-мультипликативной модели оценки рисков.

Научная революция в области математических методов исследования началась в 1980-е годы прошлого столетия. В ходе создания Всесоюзной статистической ассоциации [2] профессиональным сообществом был проведен тщатель-

ный анализ состояния и перспектив развития рассматриваемой научной области. Эта работа проводилась в рамках отечественной научной школы в области теории вероятностей и математической статистики, созданной акад. А. Н. Колмогоровым. История ее дальнейшего развития отражена в [3].

Основные идеи научной революции сформулированы в новой парадигме математических методов исследования [4]. Их реализации посвящен новый раздел теоретической и прикладной математики — системная нечеткая интервальная математика [5, 6], которую мы рассматриваем как основу математики XXI в.

В соответствии с новой парадигмой выпущено довольно много книг и статей. Полные тексты большинства из них можно найти, например, с помощью Российского индекса научного цитирования. Значительная часть таких статей опубликована в разделе «Математические методы исследования» нашего журнала.

Однако наблюдается значительное отставание научного уровня основной массы работ по статистическому анализу конкретных данных от достигнутого на переднем крае фронта современных научных исследований. Причины и последствия такого отставания целесообразно обсудить.

В настоящее время активно используются три парадигмы математических методов исследования — примитивная, устаревшая и современная (новая). Примитивная основана на рецептах XIX в. Разнообразные таблицы Росстата выполнены в рамках этой парадигмы. Необходимо признать, что во многих случаях непосредственный анализ табличных статистических данных позволяет получить нужные исследователю научные и практические выводы.

В начале XX в. возникла необходимость в более углубленном анализе статистических данных, в частности, в области биометрии. В ответ на запросы практики возникла классическая математическая статистика. Его основоположник К. Пирсон предложил для описания реальных данных использовать функцию распределения из четырехпараметрического семейства (в настоящее время обычно используют подсемейства с меньшим числом параметров — нормальные, экспоненциальные, Вейбулла — Гнеденко и др.). Это предложение имело как положительные, так и отрицательные последствия.

Использование параметрических семейств распределений вероятностей позволило к середине XX в. разработать развитую математическую теорию, предназначенную для оценивания параметров и проверки гипотез. Ее обычно называют параметрической статистикой (в соответствии с базовым предположением, лежащим в ее основе). Эту теорию обычно излагают в вузовских курсах

по теории вероятностей и математической статистике, и ее основы обычно знакомы исследователям в прикладных областях. Отметим, что развитие отдельных аспектов этой теории продолжается. Так, сравнительно недавно установлено, что вместо оценок максимального правдоподобия для оценивания параметров целесообразно использовать одношаговые оценки, а также получены новые результаты в области оценивания параметров гамма-распределений и бета-распределений.

Однако давно известно, что базовое предположение параметрической статистики обычно не выполняется. В качестве примера обсудим часто принимаемое без достаточных оснований предположение, что рассматриваемые статистические данные получены в соответствии с нормальным законом распределения. По крайней мере с середины XX в. (из работ основателя раздела «Математические методы исследования» нашего журнала проф. В. В. Налимова) известно, что распределения реальных данных, как правило, не являются нормальными (современные обоснования этого принципиального утверждения приведены в [7]). Следовательно, научные результаты параметрической статистики являются чисто математическими (относятся к математике как науке о формальных системах), они не позволяют получать обоснованные выводы для реальных явлений и процессов. В соответствии с новой парадигмой следует использовать методы непараметрической статистики, которые к настоящему времени позволяют решать тот же круг задач, что и методы параметрической статистики.

Переход от устаревшей парадигмы к современной требует усилий от исследователей, занимающихся конкретными прикладными задачами. Необходимо изучить непараметрические модели и методы, соответствующее им теоретическое обоснование, перейти на новое программное обеспечение или разработать его. Приходится менять алгоритмы расчетов, нормативно-техническую документацию, а преподавателям — содержание читаемых курсов и соответствующую учебно-методическую литературу. Естественно, возникает сопротивление, как всегда при переходе от старого к новому.

Обсудим типовые возражения приверженцев устаревшей парадигмы.

Предлагают проверить нормальность распределения элементов выборки с помощью того или иного статистического критерия — показателей асимметрии и эксцесса, Шапиро – Уилка, типа Колмогорова, типа омега-квадрат и др. Если гипотеза нормальности не отклоняется, то считают возможным использовать методы, основанные на нормальности.

Почему это рассуждение не является корректным? Дело в том, что для тех же данных можно

проверить их соответствие другим распределениям. И для типовых объемов выборки (десятки или сотни наблюдений) ответ будет положительным. Другими словами, столь же обоснованно можно принять не только нормальность, но и многие другие распределения. Известно, что для достоверного (на уровне значимости 0,05) обнаружения различия между нормальным и логистическим распределениями необходима выборка объемом не менее 2500 (см., например, [7]).

Отклонения от нормальности могут сильно влиять на свойства статистических процедур, разработанных в предположении нормальности. В результате полученные на их основе выводы могут не иметь ничего общего с действительностью. Примером являются процедуры отбраковки выбросов, нацеленных на обнаружение в выборке резко выделяющихся ошибочных значений. При отклонении от нормальности свойства методов отбраковки крайне сильно меняются [7].

В пользу параметрической статистики приводят, например, такое рассуждение. «В задачах менеджмента безопасности требуется подтверждение сверхмалых рисков порядка одной миллионной — десяти миллиардных единицы (вероятность 0,999999 – 0,99999999). Уже для подтверждения вероятности 0,999999 непараметрическим методом требуется более двух миллионов статистических испытаний, что много даже при моделировании средствами современной вычислительной техники. Применение параметрического подхода позволяет снизить этот объем на порядок». Это рассуждение принципиально неверно. Выводы на основе необоснованной вероятностно-статистической модели сами являются необоснованными [8]. В известной притче сказано так: «Человек потерял ключи в кустах. Где их искать? Под фонарем. Почему под фонарем? Потому что там светлее». Параметрическую статистику можно сравнить с поиском потерянных в кустах ключей под фонарем.

Не всегда исследователи осознают и признают сам факт научной революции. У некоторых из них возникает желание объявить порождающие ее новые научные результаты малозначительными, находящимися на периферии математических методов исследования, а потому не требующими осмысления, не заслуживающими внимания, изучения и применения. Одна из причин этого явления — включение психологической защиты против нового, требующего решительного пересмотра привычных старых подходов. Важно также углубление в отдельные узкие проблемы, связанные с частными постановками задач, отсутствие широкого кругозора, а также желания и возможности проанализировать динамику развития математических методов исследования. Такому анализу мешает в том числе и громадный

объем накопленных к настоящему времени научных публикаций в рассматриваемой области. По нашей оценке, математическим методам исследования посвящены миллионы статей и книг. Именно понимание рассматриваемых причин разрыва между прикладниками и теоретиками послужило причиной написания данной статьи.

Обращаю внимание читателей всех разделов нашего журнала на происшедшую революцию в математических методах исследований, на необходимость овладения ее результатами. В частности, ссылки на незнание, например, непараметрической статистики, не могут оправдать применение устаревших неадекватных методов, образно говоря, поиска потерянных в кустах ключей под фонарем, где светлее.

## ЛИТЕРАТУРА

1. Орлов А. И. Обобщенная аддитивно-мультипликативная модель оценки рисков на основе нечетких и интервальных исходных данных / Заводская лаборатория. Диагностика материалов. 2023. Т. 89. № 1. С. 74 – 84. DOI: 10.26896/1028-6861-2023-89-1-74-84
2. Орлов А. И. Создана единая статистическая ассоциация / Вестник Академии наук СССР. 1991. № 7. С. 152 – 153.
3. Орлов А. И. Отечественная научная школа в области организационно-экономического моделирования, эконометрики и статистики / Контроллинг. 2019. № 73. С. 28 – 35.
4. Орлов А. И. Новая парадигма математических методов исследований / Заводская лаборатория. Диагностика материалов. 2015. Т. 81. № 7. С. 5 – 5.
5. Орлов А. И., Луценко Е. В. Системная нечеткая интервальная математика. Монография (научное издание). — Краснодар, КубГАУ, 2014. — 600 с.
6. Орлов А. И., Луценко Е. В. Анализ данных, информации и знаний в системной нечеткой интервальной математике: научная монография. — Краснодар: КубГАУ, 2022. — 405 с.
7. Орлов А. И. Прикладной статистический анализ. — М.: Ай Пи Ар Медиа, 2022. — 812 с.
8. Орлов А. И. О требованиях к статистическим методам анализа данных (обобщающая статья) / Заводская лаборатория. Диагностика материалов. 2023. Т. 89. № 11. С. 98 – 106. DOI: 10.26896/1028-6861-2023-89-11-98-106

## REFERENCES

1. Orlov A. I. Generalized additive-multiplicative risk estimation model based on fuzzy and interval data / Industr. Lab. Mater. Diagn. 2023. Vol. 89. N 1. P. 74 – 84 [in Russian]. DOI: 10.26896/1028-6861-2023-89-1-74-84
2. Orlov A. I. The unified statistical association has been created / Vestn. AN SSSR. 1991. N 7. P. 152 – 153 [in Russian].
3. Orlov A. I. Native scientific school in the field of organizational and economic modeling, econometrics and statistics / Kontroling. 2019. N 73. P. 28 – 35 [in Russian].
4. Orlov A. I. New paradigm of mathematical research methods / Industr. Lab. Mater. Diagn. 2015. Vol. 81. N 7. P. 5 – 5 [in Russian].
5. Orlov A. I., Lutsenko E. V. System Fuzzy Interval Mathematics. — Krasnodar: KubGAU. 2014. — 600 p. [in Russian].
6. Orlov A. I., Lutsenko E. V. Analysis of data, information and knowledge in system fuzzy interval mathematics. — Krasnodar: KubGAU, 2022. — 405 p. [in Russian].
7. Orlov A. I. Applied Statistical Analysis. — Moscow: Ai Pi Ar Media, 2022. — 812 p. [in Russian].
8. Orlov A. I. On the requirements for statistical methods on data analysis (generalizing article) / Industr. Lab. Mater. Diagn. 2023. Vol. 89. N 11. P. 98 – 106 [in Russian]. DOI: 10.26896/1028-6861-2023-89-11-98-106