

СТАТИСТИЧЕСКИЕ МЕТОДЫ И АНАЛИЗ ДАННЫХ

СТАТИСТИКА ОБЪЕКТОВ НЕЧИСЛОВОЙ ПРИРОДЫ И АНАЛИЗ ДАННЫХ О НАУЧНОМ ПОТЕНЦИАЛЕ

Е.Г.Нечаева, А.И.Орлов, А.В.Соколов

(Москва)

В работе рассматриваются перспективы применения современных статистических методов, прежде всего методов статистики объектов нечисловой природы, в задачах изучения научного потенциала выборочными методами, а также на основе официальных статистических данных. Дан обзор основных постановок и результатов статистики объектов нечисловой природы применительно к методам анализа социологических данных и данных о российских организациях, выполняющих научно-исследовательские работы.

Ключевые слова: научный потенциал, статистика объектов нечисловой природы, показатели науки, разнотипные данные, выборочные методы, экспертные оценки, методы классификации, непараметрические оценки плотности, метрическое пространство.

1. Введение

Математические методы выборочных исследований - классическая область прикладной математической статистики. Наибольшее применение они находят в социологии и медицине. Начиная с 70-х годов в нашей стране развитие современных выборочных методов, в частности, статистики объектов нечисловой природы, стимулировалось запросами социологических и экспертных исследований [1]. Были разработаны новые подходы, сформулированы постановки, предложены алгоритмы анализа разнотипных данных, включающих значения количественных и качественных признаков, получены теоремы о свойствах этих алгоритмов, о состоятельности оценок и т.д. В частности, сводка теоретических результатов была издана в виде сборника [2], подготовленного подкомиссией "Статистика объектов нечисловой природы" Научного Совета АН СССР по комплексной проблеме "Кибернетика" и Институтом социологических исследований АН СССР. Однако до методик проведения конкретных выборочных исследований и анализа полученных данных с помощью соответствующих программных продуктов дело не дошло.

Переход к рыночной экономике в России, сопровождающийся резким спадом производства, высоким уровнем инфляции, дефицитом государственного бюджета, уменьшение спроса на исследования и разработки со стороны промышленности самым отрицательным образом сказались на состоянии российской науки. Процессы, происходящие в отечественной науке, требуют новых подходов как к методам сбора данных о состоянии науки (инфраструктура, кадры, финансирование, материально-техническая база, результативность и т.д.), так и к методам анализа полученных данных.

Такая работа проводится Центром исследований и статистики науки (ЦИСН) Миннауки РФ и РАН. За последние годы ЦИСН разработан ряд новых форм статистической отчетности, утвержденных Госкомстатом РФ, проведено несколько специальных

обследований (бюджетное финансирование науки, уровень жизни ученых, утечка умов - совместно с МВД РФ, деятельность аспирантуры и др.).

В качестве иллюстрации процессов, происходящих в последние годы в науке России, можно привести несколько цифр из статистического сборника [3], изданного ЦИСН: с 1990 по 1993 г. численность работников основной деятельности научных организаций сократилась на 40 %, расходы на НИР (в ценах 1989 г.) - на 75%, при этом, если в 1990 г. расходы на НИР составляли 2.03% ВВП (валового внутреннего продукта), то в 1993 г. - только 0.81%.

Сохранению потенциала российской науки уделяется большое внимание международной научной общественностью, международными организациями, в частности, Организацией экономического сотрудничества и развития (см. [4]), и научной общественностью отдельных стран. По некоторым оценкам зарубежная помощь российским ученым ("Фонд Сороса" и т.п.) в 1992 г. составила около 8% от объема государственных расходов на науку в России.

Кризис российской науки оживленно обсуждается и на страницах отечественных газет и журналов. Так, в No.1 за 1993 г. международной газеты "Наука и технология в России" была проведена дискуссия об эффективности научной деятельности и положении научных работников в современной России [5,6].

При обсуждении разнообразных задач изучения и управления наукой весьма важны исходные методологические принципы. Показательна упомянутая дискуссия, в которой один из авторов в качестве основного показателя использовал производительность труда научного работника [5], а другой - фондоемкость научной продукции [6], что явно затрудняло взаимопонимание. Характерно, что в обеих статьях широко использовались как статистические, так и экспертные данные. Поэтому бесспорным представляется вывод о том, что статистические данные о научном потенциале - база для теоретического и прикладного науковедения.

Основой для проведения анализа научного потенциала России является информационная база данных по российским организациям, выполняющим научно-исследовательские работы, сформированная в Центре исследований и статистики науки Миннауки РФ и РАН.

Данная информационная база основана на использовании совокупности независимых баз данных как справочного, так и информационного характера и включает около 250 показателей научно-технической деятельности предприятия (организации). Благодаря подробной справочной информации, все объекты могут быть классифицированы по регионам, типам организации, ведомственной принадлежности, областям наук, формам собственности, секторам деятельности. Единицей учета является отдельный документ "Отчет предприятия (организации) о выполнении научно-технических работ" - статистическая форма "1-Наука".

В числе показателей формы присутствуют как количественные, так и качественные. К качественным показателям, имеющим нечисловую природу, относятся:

- область науки;
- сектор науки;
- отрасль народного хозяйства;
- территория;
- министерская (ведомственная) принадлежность;
- организационно-правовая форма;
- форма собственности;
- тип организации (предприятия);
- сектор деятельности.

Среди количественных показателей можно выделить данные об объемах выполненных исследований (фундаментальных, прикладных разработок), затратах на НИР, объемах основных средств, количестве персонала различных категорий и др.

База для применения методов статистики объектов нечисловой природы в наукометрии значительно расширится с началом проведения регулярных статистических обследований по форме "2-наука", которая, в отличие от формы "1-наука", предусматривает возможность группировок данных по социально-экономическим целям исследований и разработок (в соответствии с международными стандартами), что расширяет возможности как для построения принципиально новых моделей, так и для проведения более качественных межстрановых сопоставлений.

Вводится также краткая ежеквартальная форма статистической отчетности, позволяющая накапливать ряды данных и строить на их основе прогнозные модели.

Впервые в стране в 1995 г. начинается сбор данных по инновационной деятельности, что позволит увязывать данные о развитии науки с разработкой и внедрением технологий и более точно оценивать результативность науки.

2. Основные идеи статистики объектов нечисловой природы и перспективы их алгоритмизации и применения для анализа статистических данных о научных организациях России.

С начала 70-х годов в России активно развивается статистика объектов нечисловой природы, известная также как статистика нечисловых данных или нечисловая статистика. В развитии этого сравнительно нового направления прикладной математической статистики приоритет принадлежит российским ученым.

К настоящему времени статистика объектов нечисловой природы с теоретической точки зрения достаточно хорошо развита, основные идеи, подходы и методы описаны и изучены в математическом плане, в частности, доказано достаточно много теорем. Однако эта теория пока недостаточно апробирована.

Это связано как со сравнительной молодостью статистики объектов нечисловой природы, так и с общеизвестными особенностями организации науки в 80-е годы, когда не было достаточных стимулов к тому, чтобы теоретики занимались широким внедрением своих результатов. Однако за 15 лет развития статистики объектов нечисловой природы накопилось немало фактов, подтверждающих уверенность в ее большой практической значимости.

Показатели науки могут быть использованы для применения рассматриваемых теоретических достижений на практике. Более того, представляется, что при анализе столь важной с практической точки зрения информации, как данные о научном потенциале, именно методы статистики объектов нечисловой природы окажутся наиболее полезными, поскольку существенная часть данных носит нечисловой (в частности, качественный) характер.

Цель рассматриваемого научного направления - разработать новые статистические методы анализа нечисловых данных прежде всего для нужд социально-экономических исследований. На основе подходов и результатов статистики объектов нечисловой природы разрабатывается методическое, математическое и программное обеспечение для социологических, маркетинговых, экспертных, прогнозных и других исследований.

Исходный объект в математической статистике - это выборка. В вероятностной теории статистики выборка - это 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 - 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: качественные признаки, например, пол человека или тип научной организации, вообще результат отнесения объекта в одну из заданных категорий (градаций); множество, например, совокупность научных организаций, занимающихся определенной тематикой; слово, предложение, текст, которые в памяти компьютера кодируются, как известно, с помощью цифр 0 и 1, но числами от этого не становятся; вектор, координаты которого - совокупность значений разнотипных признаков, например, результат составления статистического отчета о научно-технической деятельности (например, форма "1-наука"), в котором часть признаков носит качественный характер, а часть - количественный; ответы на вопросы социологической анкеты, часть из которых носит количественный характер (возможно, интервальный, т.е. респондент свое мнение выражает не числом, а интервалом), часть сводится к выбору одной из нескольких подсказок, а часть представляет собой тексты; упорядочение экспертом заявок на проведение научных работ при проведении конкурсов на выделение грантов; результаты контроля выполнения заданий по научно-техническим программам, вообще планов научных работ, по альтернативному признаку (выполнена или не выполнена конкретная позиция плана); разбиения объектов на группы сходных между собой (кластеры); ранжировки, например, упорядочения экспертами научных проектов по степени предпочтения (на одной из стадий процесса распределения грантов); толерантности, т.е. бинарные отношения, описывающие сходство объектов между собой, например, сходство тематики научных работ, оцениваемого экспертами с целью рационального формирования экспертных советов внутри определенной области науки; результаты парных сравнений и т.д. Интервальные данные тоже можно рассматривать как пример объектов нечисловой природы.

В чем принципиальная новизна статистики объектов нечисловой природы? В классической математической статистике постоянно используется операция сложения. При расчете

выборочных характеристик распределения (выборочное среднее арифметическое, выборочная дисперсия и др.), в регрессионном анализе и других областях этой научной дисциплины постоянно используются суммы. Математический аппарат - законы больших чисел, Центральная предельная теорема и другие теоремы - нацелен на изучение сумм. В нечисловой статистике нельзя использовать операцию сложения, поскольку элементы выборки лежат в пространствах, где нет этой операции. Методы обработки нечисловых данных основаны на принципиально ином математическом аппарате - на применении различных расстояний в пространствах объектов нечисловой природы.

Под влиянием запросов прикладных исследований, прежде всего в социально-экономических науках, мы развиваем статистику объектов нечисловой природы. Так, более чем двадцатилетний (с 1970 г.) опыт проведения социологических и маркетинговых исследований, а в последнее время - анализа и прогнозирования потребительских цен и индекса инфляции, исследований рынка товаров народного потребления, образовательных услуг, программного обеспечения - привел к постановкам ряда нерешенных задач в области эконометрических методов анализа и прогнозирования на основе нечисловых данных.

Теоретические исследования в области социологических и экспертных методов социально-экономических исследований публиковались А.И.Орловым с 1974 г. [7]. Они проводились в тесном контакте с сотрудниками Института социологии и ЦЭМИ РАН и отражены в монографии [8], в которой сведены результаты первого этапа изучения статистики в конкретных пространствах нечисловой природы, прежде всего результаты, связанные с теорией измерений и средними величинами, случайными и нечеткими множествами, парными сравнениями.

В 1979 г. была осознана необходимость выделения самостоятельного направления - статистики объектов нечисловой природы. Программа развития этого направления была сформулирована

в статье [9] и к настоящему времени в теоретическом плане в основном реализована.

За прошедшие более чем 15 лет различными авторами получено много интересных теоретически и полезных практически результатов в статистике объектов нечисловой природы. Работы координировала подкомиссия "Статистика объектов нечисловой природы" Научного Совета АН СССР по комплексной проблеме "Кибернетика", созданная в 1979 г. Различные направления статистики объектов нечисловой природы наиболее подробно охватывает уже упомянутый сборник [2]. Из недавних обзоров назовем [1,10-13].

Однако приходится с сожалением констатировать, что ряд теоретических разработок пока не доведен до уровня методик, математического и программного обеспечения. Этим мы занимаемся, в частности, в рамках работы по алгоритмизации и применению основных идей, подходов и методов рассматриваемого научного направления для анализа статистических данных о научных организациях России. На основе современных представлений о проведении социологических исследований и о методах анализа собранных данных будут разработаны соответствующие математико-статистические методы, создано математическое и программное обеспечение. Поскольку нечисловые данные составляют до 90% данных в социологии и большую часть - в экономике, то теоретические исследования в статистике нечисловых данных позволят получить новые результаты в той центральной области эконометрики, в которой отечественные работы имеют приоритет на мировом уровне.

3. Основные идеи статистики объектов нечисловой природы.

Кратко рассмотрим несколько идей, развиваемых в статистике объектов нечисловой природы для данных, лежащих в пространствах произвольного вида.

3.1. Теоретические и эмпирические средние

Первой обсудим проблему определения средних величин. В рамках репрезентативной теории измерений удастся указать вид средних величин, соответствующих тем или иным шкалам измерения [2,14]. В пространствах иной природы приходится действовать по-другому.

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 операции сложения. Теоретические и эмпирические средние приходится вводить как решения экстремальных задач. Теоретическое среднее - это решение задачи минимизации математического ожидания (в классическом смысле) расстояния от случайного элемента со значениями в рассматриваемом пространстве до фиксированной точки этого пространства. Для 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. При этом как эмпирическое, так и теоретическое средние как решения экстремальных задач могут быть не единственными элементами пространства, а состоять из множества таких элементов, которое может оказаться и пустым. Тем не менее удалось сформулировать и доказать законы больших чисел для средних

величин, определенных указанным образом, т.е. сходимость эмпирических средних к теоретическим при росте объема выборки [2,8,15].

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 [15]. В частности, кроме законов больших чисел установлена и состоятельность оценок минимального контраста, в том числе оценок максимального правдоподобия и робастных оценок. К настоящему времени подобные оценки изучены также и в интервальной статистике.

3.2. Непараметрические оценки плотности

В статистике в пространствах произвольной природы большую роль играют непараметрические оценки плотности, используемые, в частности, в различных алгоритмах регрессионного, дискриминантного, кластерного анализа. В [9] было предложено использовать непараметрические оценки плотности в пространствах произвольной природы. В [16] были введены и изучены еще несколько видов таких оценок, в частности, установлена их состоятельность. Затем была оценена скорость сходимости и установлен примечательный факт совпадения наилучшей скорости сходимости в произвольном случае с той, что имеет место в классической постановке для числовых случайных величин [17].

Дискриминантный, кластерный, регрессионный анализ в пространствах произвольной природы основаны либо на параметрической теории [2,15] - и тогда применяется подход, связанный с асимптотикой решения экстремальных статистических задач, либо на непараметрической теории [9,16,17] - и

тогда используются алгоритмы на основе непараметрических оценок плотности.

Для проверки гипотез могут быть использованы статистики интегрального типа. Любопытно, что предельная теория таких статистик, построенная первоначально в классической постановке [8], приобрела естественный (завершенный, изящный) вид именно для пространств произвольного вида [18], поскольку при этом удалось провести рассуждения, опираясь на базовые математические соотношения, а не на те случайные (с общей точки зрения), что были связаны с конечномерностью пространства.

3.3. Конкретные виды объектов нечисловой природы

Представляют интерес результаты, связанные с конкретными областями статистики объектов нечисловой природы, в частности, со статистикой нечетких множеств, со случайными множествами (следует отметить, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств [8,19]), с непараметрической теорией парных сравнений, с аксиоматическим введением метрик в конкретных пространствах объектов нечисловой природы.

Для анализа данных о научных организациях России, об их научном потенциале весьма важны методы классификации, в том числе типологии. Проблемами теории и практики классификации в нашей стране занимались многие научные работники. С 1984 г. в рамках Союза научных и инженерных обществ действует "Комиссия по классификации", объединившая усилия нескольких десятков специалистов различных научных областей. Итоги десяти лет работы нашей комиссии показали, в частности, что наиболее естественно ставить и решать задачи классификации в рамках статистики объектов нечисловой природы. Это касается как распознавания образов с учителем (дискриминантного анализа), так и распознавания образов без учителя

(кластерного анализа). Современное состояние дискриминантного и кластерного анализа отражено с точки зрения статистики объектов нечисловой природы в работах [17,20,21].

Отметим, что некоторые из высказанных нами [22] десять лет назад предложений по применению статистики объектов нечисловой природы в задачах изучения научного потенциала и управления наукой продолжают сохранять интерес.

4. Примеры возможных применений современных статистических методов анализа выборочных данных в задачах изучения и управления научным потенциалом.

В силу целого ряда причин (высокая динамика изменений, потребность в оперативной информации для принятия решений и т.д.) возникает острая необходимость регулярного проведения оперативных обследований научных организаций и социологических опросов ученых. Эта задача может быть успешно решена только путем проведения выборочных обследований. Обсудим несколько примеров возможных применений современных статистических методов анализа выборочных данных в задачах изучения и управления научным потенциалом.

4.1. Изучение показателей научной деятельности, в частности, изучение предпочтений научных работников, руководителей научных организаций и профессионалов-управленцев, экспертно-статистический подход к построению интегральных показателей (рейтингов).

Статистический и экспертный анализ показателей науки, выявление их рейтинга (какие показатели являются наиболее значимыми с учетом общепринятых международных стандартов в статистике науки [23]).

Как оценивать эффективность науки? В какой мере общепринятые показатели результативности науки (публикации, патенты, лицензии, индексы цитируемости и т.д.) отражают реальную эффективность науки?

4.2. Кластеризация научных организаций, выделение типов. При обсуждении вопросов финансирования, организационных преобразований, перспективного развития необходимо к различным типам научных организаций подходить дифференцированно. Насколько принятая типология научных организаций (см. [3, стр. 228]) соответствует тем или иным аналитическим задачам? Сколько типов существует реально? Какова естественная типология научных организаций? Если такая типология излишне сложна, не удастся ли ее упростить, рассматривая научные организации с определенных практических позиций - с точки зрения научного уровня, отдачи, финансирования, перспективности и выживания и т.д.

Представляет также интерес структура научных советов, комиссий и др. форм деятельности ученых, особенно в ситуации, когда через такие структуры идет финансирование.

4.3. Рейтинги проектов, научных организаций и др. В рамках достаточно однородной совокупности проектов НИР, заявок на гранты, научных организаций и т.д. с помощью методов многомерного статистического анализа можно выявить основные направления вариации. Однако главный фактор, вопреки распространенному способу интерпретации результатов факторного анализа (или метода главных компонент), не всегда соответствует оси "эффективность - неэффективность". Тем не менее, идея рейтинга (интегрального показателя качества) заслуживает проработки. Можно указать несколько подходов. Для решения этой задачи можно использовать три варианта экспертно-статистического метода интегральной оценки перспективности научного направления или научной организации:

по интегральному показателю (линейная функция, параметры которой - показатели науки, а значения коэффициентов получаются от экспертов);

по обучающим выборкам, полученным от квалифицированных экспертов;

с помощью оценки параметров интегрального показателя по обучающим выборкам (собственно экспертно-статистический метод).

4.4. Расчет и краткосрочное прогнозирование показателей науки. Важнейшее значение для принятия управленческих решений в сфере науки имеет прогнозирование таких показателей как занятость, средняя зарплата в секторе "Наука и научное обслуживание" и др., особенно с учетом эффекта изменения макроэкономических показателей (в частности, индекса инфляции, курса доллара и других валют) и влияния тех или иных мер (экономических, законодательных, налоговых, таможенных и т.д.), предпринимаемых на государственном уровне.

4.5. Применение современных статистических методов анализа нечисловых и интервальных данных. Использование статистики объектов нечисловой природы в выборочных обследованиях даст возможность реализовать идеи, изложенные в упомянутом выше сборнике [2]. В частности, регрессионный анализ в пространствах разнотипных признаков (объектов нечисловой природы) даст возможность оценить эффективность финансирования, а статистика интервальных данных позволит учесть неизбежные неточности в имеющихся данных.

4.6. Анализ существующей системы приоритетов отечественной науки с точки зрения объемов базового финансирования отдельных институтов, структуры государственных научно-технических программ, системы бюджетных и внебюджетных фондов, проведения экспертных опросов ведущих ученых и специалистов. Выявление наиболее перспективных направлений развития российской науки, выделение приоритетных

технологий, сопоставление научно-технического развития России и развитых стран.

4.7. Оценка основных тенденций развития научной сферы за последние годы (изменение структуры сети научных организаций по ведомственному признаку, численности, эффективности функционирования, направлениям исследований и т.д.). Выявление закономерностей между основными количественными и качественными показателями развития науки и базовыми макроэкономическими показателями с последующим моделированием процессов, происходящих в сфере науки, в их взаимосвязи с инновационными процессами в экономике и социальной сфере.

4.8. Выделение "групп риска". Частный случай обсуждаемых выше задач - выделение (по формальным - отчетным - признакам) научных организаций, само существование которых оказывается под вопросом в ближайшем будущем.

Прогноз "выживаемости" НИИ может быть построен с помощью обучающих выборок на основе непараметрических оценок плотности в пространстве разнотипных признаков, часть координат которых - количественные признаки, а часть - качественные.

Существует много других интересных проблем [24,25]. Например, выявление цикличности развития научно-технического потенциала страны, изучение динамики реальной и формальной структуры науки [26], форм примитивизации научной деятельности и научной продукции в условиях резкого спада производства и сокращения финансирования научного труда, проблемы отражения в общественном сознании и в самосознании научного сообщества специфики научной деятельности (включая анализ распространенных догм) и др.

По нашей оценке, применение современных статистических методов в выборочных исследованиях научных организаций позволит получать результаты, интересные с теоретической точки зрения и полезные для практики управляющих воздействий на развитие российской науки.

ЛИТЕРАТУРА

1. Орлов А.И. Нечисловая статистика/Наука и технология в России. 1994, No.3(5), с.7-8.
2. Анализ нечисловой информации в социологических исследованиях (под ред. В.Г.Андреевкова, А.И.Орлова, Ю.Н.Толстовой).- М.: Наука, 1985.
3. Наука России:1993. Статистический сборник.- М. ЦИСН, 1994.
4. Научно-техническая и инновационная политика. Российская федерация. Том 1. Оценочный доклад. - Организация экономического сотрудничества и развития, 1994.
5. Орлов А.И. Социологический прогноз развития Российской науки на 1993-1995 годы. - Международная газета "Наука и технология в России", 1993, N1.
6. Страхов В.Н. Нужны ли подобные прогнозы? - Международная газета "Наука и технология в России", 1993, No.1, с.30-31.
7. Орлов А.И. Допустимые средние в некоторых задачах экспертных оценок и агрегирования показателей качества. - В сб. "Многомерный статистический анализ в социально-экономических исследованиях", М.: Наука, 1974, с.388-393.
8. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979.
9. Орлов А.И. Статистика объектов нечисловой природы и экспертные оценки. - В сб. "Экспертные оценки. Вопросы кибернетики, вып. 58". - М.: Научный Совет АН СССР по комплексной проблеме "Кибернетика", 1979, с.17-33.
10. Орлов А.И. Статистика объектов нечисловой природы. Обзор/ Заводская лаборатория, 1990, т.56, No.3, с.76-83.
11. Orlov A.I. On the Development of the Statistics of Nonnumerical Objects. - In: DESIGN OF EXPERIMENTS AND DATA ANALYSIS: NEW TRENDS AND RESULTS. Ed. by prof.E.K.Letzky. Moscow: ANTAL, 1993. p.52-90.

12. Орлов А.И. Объекты нечисловой природы/ Заводская лаборатория. 1995. Т.61. No.3.
13. Орлов А.И. Вероятностные модели объектов нечисловой природы/ Заводская лаборатория. 1995. Т.61. No.5.
14. Орлов А.И. Связь между средними величинами и допустимыми преобразованиями шкалы/Математические заметки, 1981, т.30, No.4, с.361-368.
15. Орлов А.И. Асимптотика решений экстремальных статистических задач. - В сб."Анализ нечисловых данных в системных исследованиях. Труды ВНИИСИ, 1982, вып.10." - М.: ВНИИСИ, 1982, с.4-12.
16. Орлов А.И. Непараметрические оценки плотности в топологических пространствах. - В сб.: Прикладная статистика. - М.: Наука, 1983, с.12-40.
17. Орлов А.И. Классификация объектов нечисловой природы на основе непараметрических оценок плотности. - В сб. "Проблемы компьютерного анализа данных и моделирования: Сб. науч. ст." - Минск: Белорусский государственный университет, 1991, с.141-148.
18. Орлов А.И. Асимптотическое поведение статистик интегрального типа. - В сб. "Вероятностные процессы и их приложения". - М.: МИЭМ, 1989, с.118-123.
19. Орлов А.И. Задачи оптимизации и нечеткие переменные. - М.: Знание, 1980, 64 с.
20. Орлов А.И. Некоторые вероятностные вопросы теории классификации.- В сб.: Прикладная статистика. - М.: Наука, 1983, с.166-179.
21. Орлов А.И. Заметки по теории классификации/ Социология: методология, методы, математические модели. 1992. No.2. С.28-50.
22. Орлов А.И. Организационные методы управления наукой и статистика объектов нечисловой природы. - Тезисы докладов Всесоюзного симпозиума "Медицинское науковедение и автоматизация информационных процессов". - М., 1984, с.215-216.

23. Frascati Manual: 1993. The Measurement of Scientific and Technological Activities. Paris: OECD, 1994 - 261 с.
24. Развитие науки в России./ЦИСН. - М.: 1993.
25. *Налимов В.В., Мульченко А.Б.* Наукометрия. - М.: Наука, 1969.
26. *Орлов А.И.* Прикладная статистика - "Золушка" научно-технической революции. - Международная газета "Наука и технология в России". 1994, No.1(3), с.13-14.