ЗНАНИЕ

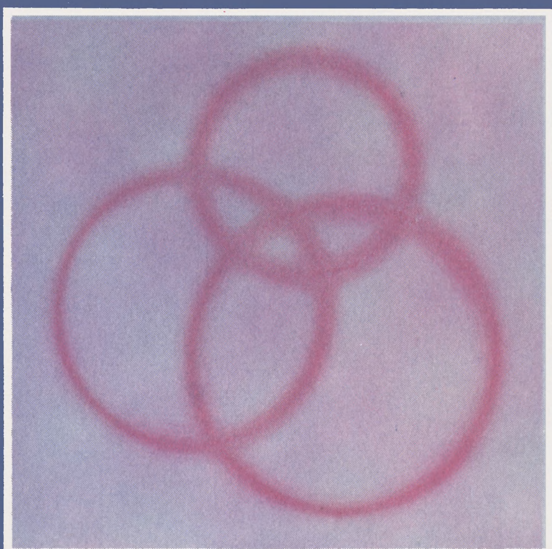
НОВОЕ
В ЖИЗНИ,
НАУКЕ,
ТЕХНИКЕ

СЕРИЯ
МАТЕМАТИКА,
КИБЕРНЕТИКА

8'80

А. И. Орлов

ЗАДАЧИ ОПТИМИЗАЦИИ И НЕЧЕТКИЕ ПЕРЕМЕННЫЕ



НОВОЕ
В ЖИЗНИ,
НАУКЕ,
ТЕХНИКЕ

Серия
«Математика,
кибернетика»
№ 8, 1980 г.

Издается
ежемесячно
с 1967 г.

А. И. Орлов,
кандидат физико-математических наук

ЗАДАЧИ ОПТИМИЗАЦИИ И НЕЧЕТКИЕ ПЕРЕМЕННЫЕ

Издательство
«Знание»
Москва
1980

ББК 22.1
О 66

Орлов А. И.

О 66 Задачи оптимизации и нечеткие переменные.— М.: Знание, 1980.— 64 с. (Новое в жизни, науке, технике. Сер. «Математика, кибернетика»; № 8).

11 к.

Теория нечетких множеств Л. А. Заде, прикладная математическая статистика и задачи оптимизации — три области математики, взаимосвязи которых посвящена брошюра. Концепция нечеткости позволяет учитывать неопределенность реальных явлений в математических моделях, а оптимизационный подход — развивать статистику объектов нечисловой природы, в частности статистику нечетких множеств. Идеи, результаты и задачи обсуждаются с позиции математика, занимающегося прикладными вопросами.

Рассчитано на широкий круг читателей, интересующихся математикой, кибернетикой и их приложениями.

202 204

ББК 22.1
51

ВВЕДЕНИЕ

В последние десятилетия бурное развитие прикладных исследований выдвинуло на первый план три области математики — математическую статистику, теорию экстремальных задач и теорию нечеткости.

История теории нечеткости началась в 1965 г. со статьи Л. А. Заде [1] (а предыстория тянется от древнегреческих философов). За 15 лет вышли тысячи статей и книг по этой теории и ее приложениям, возник специальный журнал «Нечеткие множества и системы» *. Но бурные споры о необходимости использования нечетких множеств, их роли и месте в математике не утихли.

Целесообразность развития теории вероятностей и математической статистики не вызывала в этот период сомнений. В реферативном журнале «Математика» за год рассматривалось около пяти тысяч публикаций по этой тематике. В. Н. Тутубалин темпераментно обсуждал прикладные вопросы [2, 3, 4]. Специалисты многих областей применяли статистические методы, хотя в написанных ими учебных пособиях зачастую имелись существенные ошибки.

Теория экстремальных задач стала основой математических методов в экономике. Впрочем, термин «оптимизация» настолько известен, что вряд ли имеет смысл его растолковывать.

Три разные теории? Задачи брошюры — показать их связь между собой, обосновать необходимость использования концепции нечеткости при моделировании реальных явлений, продемонстрировать целесообразность применения экстремально-статистических подходов при обработке нечетких данных.

---

* International Journal of Fuzzy Sets and Systems.

Что ждет читателя? Сначала мы показываем нечеткость как реальных явлений, так и отражающих их понятий естественного языка, а также обсуждаем способы учета нечеткости в математических моделях: теории устойчивости, использование случайных величин, нечетких переменных в смысле Л. А. Заде. Вслед за основными определениями и обзором итогов пятнадцатилетнего развития мы устанавливаем связь между теорией нечеткости и случайными множествами.

Как известно, математическая статистика — это часть математики, занимающаяся методами анализа статистических данных. А статистические данные могут быть разными: числами, векторами, кривыми. В последние десятилетия выяснилась необходимость анализа данных нечисловой природы: бинарных отношений (например, ранжировок, разбиений), множеств, а также и *нечетких множеств*. Чтобы охватить все эти частные случаи, возникла теория — статистика в пространствах общей природы. Обратим также внимание на (репрезентативную) теорию измерений, которая занимается проблемами обработки статистических данных, определенных с точностью до допустимого преобразования шкалы.

Оптимизационный подход давно использовался в математической статистике: метод наименьших квадратов был предложен Гауссом в 1795 г. Основные задачи прикладного многомерного статистического анализа, как показал С. А. Айвазян [5], являются экстремальными. При переходе к статистике объектов нечисловой природы роль теории оптимизации еще более возрастает: даже для определения среднего и математического ожидания необходимо решать экстремальные задачи (см. далее п. 9).

Задачи обработки нечетких данных мы рассматриваем как часть математической статистики. Выделяем три основные задачи: описание материала, оценивание, проверку гипотез. После краткого разбора ряда методик получения нечетких данных мы обсуждаем аксиоматическое введение расстояний между нечеткими множествами, оптимизационный подход к их классификации. Под оцениванием мы имеем в виду в сcновном получение группового мнения совокупности экспертов (ответ каждого из которых нечеток) и выяснение поведения этого группового мнения при увеличении числа экспертов. В заключение рассматривается проверка гипотезы о совпадении нечетких множеств, описывающих «истинные» мнения экспертов.

Читатель уже встретил «ключевые слова», не отраженные в названии брошюры: «прикладная математическая статистика», «экспертные оценки». Они соответствуют «незримому коллективу», к которому принадлежит автор. Этот коллектив состоит из математиков, доказывающих теоремы в области математической статистики и одновременно занимающихся приложениями статистических методов (в основном в экономике и медицине). «Стержнем» является основанный в 1969 г. всесоюзный научно-исследовательский семинар «Многомерный статистический анализ и вероятностное моделирование реальных процессов» (руководители — С. А. Айвазян, Л. Д. Мешалкин). Основные публикации [6—9].

Часть этого коллектива совместно с прикладниками организовала в 1974 г. комиссию «Экспертные решения» Научного совета АН СССР по комплексной проблеме «Кибернетика», которая занимается в основном разработкой математических методов анализа экспертных оценок. Читателю придется поверить автору на слово, что экспертные оценки заслуживают того, чтобы ими заниматься. Впрочем, с течением времени стало ясно, что они являются лишь одной из областей приложения математических методов анализа нечисловых данных — именно этим мы на самом-то деле занимаемся (о нечисловых данных подробнее рассказано в п. 8). Каковы же другие области? Больше всего мы связаны с экономикой, социологией, психологией, медициной, анализом явлений культуры, теорией принятия решений, распознаванием образов и искусственным интеллектом. Основные публикации комиссии это — [7, 9—13], семинары — «Математические методы в экспертных оценках», «Статистика объектов нечисловой природы» (руководители — Б. Г. Литвак, А. И. Орлов, Ю. Н. Тюрин).

Теория нечеткости также входит в круг наших интересов. Написать эту популярную брошюру меня побудила возможность изложить общий подход к обработке данных, нечетко отражающих реальность. В научных публикациях принято приводить методы, теоремы с подробными доказательствами, результаты применений, а методологическим вопросам обычно уделяют мало внимания. Здесь же изложен сложившийся в нашем незримом коллективе подход, а технические подробности опущены.

Изложение идет, естественно, с моей персональной точки зрения. На ее формирование оказали существенное влияние многолетние научные контакты с С. А. Айвазяном, Л. Д. Мешалкиным, Ю. Н. Тюриным, с которыми я должен разделить достоинства брошюры. Недостатки же —мои.

Мы обсуждаем основные идеи рассматриваемых областей, а за подробностями отсылаем к литературе, список которой приведен в конце брошюры. Список этот довольно большой — для удобства читателя, пожелавшего углубиться в какой-либо специальный вопрос. Изложение идет на математическом уровне — а это значит, что необходимо пользоваться строгими определениями, а ряд результатов формулировать в виде теорем. Разделы 2 и 5, более трудные для восприятия, чем остальные, при первом чтении можно пропустить.

Глава 1. КОНЦЕПЦИЯ НЕЧЕТКОСТИ

1. Введение в нечеткость мира

Философы Древней Греции обсуждали софизм «Куча». В современной формулировке он звучит так: «Одно зерно не составляет кучу. Если к тому, что не составляет кучи, добавить одно зерно, то куча не получится. Следовательно, никакое количество зерен не составляет кучу». Между тем каждый согласится, что 100 000 000 000 зерен — довольно большая куча.

О чем этот софизм? В нем обсуждаются два понятия: «несколько зерен» и «куча» и показывается, что граница между ними в мышлении людей и в отражающем это мышление естественном языке является нечеткой. Разные люди укажут ее по-разному. Более того, один и тот же человек в зависимости от ситуации * может признавать или нет одну и ту же совокупность зерен кучей. Даже если он априори задастся некоторой точной границей для числа зерен, ответы будут колебаться из-за ошибок в определении на глаз числа зерен в предъявленной ему совокупности. Если же много раз предлагать человеку принять решение относительно различных наборов зерен, то испытуемый при малом числе зерен будет уверенно отвечать: «это

* В известном психофизическом опыте испытуемый опускает левую руку в емкость с холодной водой, правую — с горячей, а затем обе — в сосуд с теплой. Тогда левой руке теплая вода кажется горячей, а правой — холодной.

не куча», при большом: «это куча», а в промежуточных случаях его ответы будут колебаться между первым и вторым.

Описанная ситуация типична. Понятия естественного языка, отражающие (и формирующие!) соответствующие элементы мышления человека, как правило, не могут быть эксплицированы * с помощью математического понятия множества. Границы понятия всегда размыты, что непримиримо противоречит математическому определению, в соответствии с которым мы должны в ответ на вопрос: «Принадлежит ли этот элемент данному множеству?» ответить совершенно определенно: либо «да», либо «нет». Можно ли точно указать, при какой длине волны красный цвет переходит в оранжевый? Каждый человек, имеющий нормальное зрение, сможет назвать длину волны, соответствующую красному цвету, и длину, соответствующую оранжевому. Но всегда найдется промежуточная область, цвет которой можно с некоторой степенью уверенности отнести к красному, а с некоторой — к оранжевому. У одних промежуточная область окажется шире, у других — у́же. Художник, наверно, выделит самостоятельный оттенок и на вопрос экспериментатора: «Это красный цвет или оранжевый?» ответит: «Не тот и не другой».

Нечеткость свойственна не только естественному языку, но и языкам науки. Так, из-за атомной структуры вещества принципиально нельзя определить длину предмета (в сантиметрах) с точностью до тысячи знаков после запятой. Это означает, что длину следует описывать не одним действительным числом, а совокупностью чисел.

Может показаться, что язык математики является четким. Однако это не так. Мы уже употребляли понятие множества. Но теория множеств может быть аксиоматизирована многими различными способами. При этом в одном случае так называемая аксиома континуума верна, а в другом — верно ее отрицание. С точки зрения конструктивного подхода к основаниям математики понятие «предел функции в точке» может быть уточнено восемью различными способами.

Между тем математики, не являющиеся специалистами по основаниям этой науки, употребляют понятия «множество», «предел функции в точке» весьма широко, не заботясь о том, какой аксиоматической теории они придержи-

---

* Т. е. определены на формальном языке.

ваются. И обычно довольно успешно — подобная практика не приводит к ошибкам. Точно так же естественный язык используется по назначению, несмотря на нечеткость своих понятий.

Выделим две стороны нечеткости мира: первую — возникающую при построении математических моделей реальных явлений, вторую — присущую мышлению и восприятию человека.

При построении моделей реальных явлений исследователь, желая пользоваться математическими методами, зачастую вносит определенность там, где ее нет по существу. Это может привести к иллюзиям и ошибкам. Автору брошюры представляется, что при многих социологических опросах интересуются мнениями опрашиваемых там, где эти мнения являются весьма нечеткими или, скажем, несформировавшимися ввиду отсутствия необходимости размышлять о соответствующем предмете. Поскольку организатор опроса желает иметь четкие мнения (именно так составлена анкета), то он их получает, но эти мнения имеют довольно слабую связь с размытыми представлениями опрашиваемых, что и приводит иногда к существенным ошибкам в прогнозировании на основании подобных социологических опросов (ср. [14]).

Особенно опасным является использование «искусственной определенности» при управлении. Так, всем хорошо известно, к каким затруднениям приводило применение «валовой продукции» в качестве основного показателя деятельности предприятия. Трудно сравнивать методы лечения больных, если врачи различных центров под одними и теми же терминами понимают разное (так, по определению одной группы медиков «затяжное течение острой пневмонии» имеет место в 6% случаев, а по мнению другой — в 60%). Можно представить себе «крен» научной жизни, если в качестве основного критерия эффективности работы будет использоваться «число публикаций» или «число цитирований».

Прежде чем перейти к обсуждению различных путей учета неопределенностей реальных явлений в математических моделях, рассмотрим нечеткость, связанную с психическими процессами человека. Сам основатель теории нечеткости Л. А. Заде рассматривает ее как аппарат анализа и моделирования гуманистических систем, т. е. систем, в которых участвует человек [15, с. 6]. Его подход «опирается на предпосылку о том, что элементами мышления человека являются не числа, а элементы некоторых нечетких множеств, или классов объектов, для которых переход от «принадлежности» к «непринадлежности» не скачкообразен, а непрерывен» [15, с. 7]. Таким образом, теорию нечеткости необходимо использовать в различных психологических областях, в частности при моделировании реакции опрашиваемого в социологических и экспертных опросах. Представители этих отраслей деятельности разрабатывают методы получения информации, представленной в виде нечетких множеств, а потому более адекватной, чем в категорической числовой форме.

Другая область использования аппарата нечеткости — психофизика, в которой рассматривается реакция человека как прибора, измеряющего факторы физической природы. Характеристики этого прибора естественно описывать в терминах теории нечеткости (см., например, статью П. Б. Шошина в [10]). Используемые в психофизике пороговый подход и модели типа Терстоуна являются частными случаями подобного описания.

Концепция нечеткости полезна и при рассмотрении проблем искусственного интеллекта, в частности при обучении робота пониманию человеческой речи. Хорошо, если бы робот мог выполнять инструкцию типа «варить суп до готовности». Само понятие «готовность супа» не является строго определенным. Более того, в одни дни хозяйка варит суп несколько лучше, в другие — хуже, но его можно есть. Что значит: «можно есть» — тоже не определено четко. И вряд ли можно определить... Ясно ведь, что раз естественный язык нечеток, о чем мы уже очень много говорили, то для общения с роботом люди не будут переучиваться, и роботу придется научиться выполнять нечеткие приказания.

Если, забегая вперед, сказать о результатах применения аппарата нечеткости в реальных задачах, то оказывается, что в ряде случаев с его помощью удается получить больше, чем с помощью других методов, или же получить то же, но с меньшими затратами. Так, группа химиков во главе с академиком В. В. Кафаровым для изучения процессов, протекающих в ванной стекловаренной печи при производстве листового стекла, применила метод Фурье и алгоритм, основанный на использовании нечетких множеств [16]. Последний потребовал в 5 раз меньше, чем при «обычных» методах, машинного времени при одинаковой точности найденного с помощью ЭВМ решения.

2. О методах учета нечеткости в математических моделях

Выделим три основных метода подобного учета. В первом, применяемом в теории устойчивости (общий обзор дан в [9]), сначала находят точные решения, а затем оценивают их вариацию при колебаниях исходных данных в границах допустимых ошибок.

В стохастических моделях (второй метод) в качестве экспликаций исходных понятий, соответствующих реальным объектам или явлениям, рассматриваются случайные величины, а затем вероятностные соображения используют на всех этапах получения принимаемого решения, которое также оказывается случайным.

Третий метод — описание исходных понятий нечеткими множествами и прослеживание этой нечеткости вплоть до окончательного решения.

Конечно, можно указать разнообразные связи между перечисленными методами. Так, мы вскоре покажем, что третий в некотором смысле сводится ко второму, а сейчас коротко рассмотрим общие подходы к решению задач всех трех указанных выше методов.

Общий подход к анализу устойчивости. Пусть X — пространство исходных данных, Y — пространство принимаемых решений. Пусть решение принимается с помощью функции $f: X \to Y$. Поскольку исходные данные известны лишь с некоторой точностью, то вместо $x \in X$ решение принимается по $x' \in X$. Надо ввести функцию потерь: $\rho(y, y')$ ($y \in Y$, $y' \in Y$) показывает величину ущерба, понесенного из-за принятия решения y' вместо y. В качестве ρ можно использовать метрику * в Y. Потери, вызванные неточностью исходных данных, есть $\rho(f(x), f(x'))$.

Дальнейшие рассуждения зависят от предположений о характере неточности. В стиле классической теории устойчивости дифференциальных уравнений с каждой точкой $x \in X$ можно связать «область допустимых колебаний» $G(x)$, в которую входят все те x', которые могут рассматриваться вместо x в силу неопределенности исходных дан-

* Метрикой в некотором множестве Y называется функция $\rho: Y \times Y \to R^1$, такая, что при любых $x, y, z \in Y$ а) $\rho(x, y) \geqslant 0$; б) $\rho(x, y) = 0$ тогда и только тогда, когда $x = y$; в) $\rho(x, y) = \rho(y, x)$; г) $\rho(x, y) + \rho(y, z) \geqslant \rho(x, z)$.

ных. Тогда максимально возможный ущерб есть точная верхняя грань *:

$$\gamma(x, G(x)) = \sup_{x' \in G(x)} \rho(f(x), f(x')). \qquad (1)$$

Что обычно понимают под «областью допустимых колебаний»? Пусть τ — метрика в X. В качестве $G(x)$ естественно рассматривать шар $\{x' : \tau(x, x') \leqslant r\}$ некоторого радиуса r с центром в x. Чем радиус шара r меньше, тем исходные данные точнее.

Представляется естественным изучить поведение модели при различных r и особенно при $r \to 0$. Несколько обобщая, будем считать, что с каждой точкой $x \in X$ связано семейство областей допустимых колебаний $G_r(x)$, где r — произвольное положительное число. Наряду с максимально возможным ущербом $\gamma(x, G(x))$, определенным в (1), рассмотрим:

$$\gamma(x) = \lim_{r \to 0} \gamma(x, G_r(x)). \qquad (2)$$

Пример. Пусть X и Y—совокупности действительных чисел, $\rho(y, y') = |y - y'|$, $\tau(x, x') = |x - x'|$. В этом случае $\gamma(x) = 0$ тогда и только тогда, когда функция f непрерывна в точке x. Таким образом, понятие непрерывности есть частный случай обсуждаемого нами понятия устойчивости.

Определение 1. Функция f называется устойчивой в точке x, если $\gamma(x) = 0$. Функция f называется ε-устойчивой, если $\gamma(x) \leqslant \varepsilon$.

Таким образом, в теории устойчивости имеются два основных элемента — функция f и область допустимых колебаний G. Простейшая задача — изучение устойчивости данной f при данной G. Напрашиваются два ее обобщения: при заданной G найти все устойчивые f и, наоборот, по заданной f описать все G, при которых f является устойчивой. Оба эти обобщения оказываются плодотворными (некоторые примеры будут приведены далее). Подробное обсуждение общей схемы устойчивости имеется в [9].

Стохастические модели. В них нечеткость проявляется в том, что вместо исходных данных x прихо-

---

* Число α называется точной верхней гранью некоторого числового множества X ($\sup_X$), если: а) α не меньше любого $x \in X$ и б) для любого $\beta < \alpha$ существует по крайней мере одно $x(\beta) \in X$, такое, что $x(\beta) > \beta$.

дится иметь дело со случайными величинами $x' = x'(\omega)$, приводящими к случайным потерям $\rho(f(x), f(x'(\omega)))$. Чтобы характеризовать потери одним числом, рассматривают либо математическое ожидание

$$\gamma_{\text{ср}}(x|x') = E\rho(f(x), f(x')), \tag{3}$$

либо квантиль $\gamma_\alpha(x|x')$, близкого к 1 уровня α, определяемый из условия:

$$P\{\rho(f(x), f(x')) \leqslant \gamma_\alpha(x|x')\} = \alpha. \tag{4}$$

Величины, задаваемые формулами (3) и (4), являются аналогами ущерба $\gamma(x, G(x))$ из (1). Хотелось бы подробнее поговорить о различных способах изучения неопределенности с помощью вероятностных моделей, но для этого надо вводить новые понятия и обсуждать конкретные постановки задач, т. е. все дальше отклоняться от основной темы брошюры. Поэтому остается отослать читателя к монографии [9] и указанной там литературе.

Напрашивается ехидный вопрос: «Для чего вообще автор ведет разговор про разные описания устойчивости, а не берет быка за рога — не начинает брошюру словами: „Нечетким множеством называется...“?» Действительно, сейчас сложилась традиция последовательного изложения отдельной математической теории без связи ее с другими частями науки и с приложениями. Так писать легче (в своей области являешься специалистом) и безопаснее (никто не обвинит в методологических ошибках). В результате получаются произведения, одновременно и безукоризненные, и малополезные. Мне же в этой брошюре хочется подчеркнуть методологическое единство различных областей математики и ее приложений.

Теперь перейдем к третьему методу — использованию нечетких множеств для моделирования реальных явлений. Но сначала необходимо, наконец-то, определить понятие нечеткого множества.

3. Элементы теории нечетких множеств

Вернемся к софизму «Куча». Он обсуждался Э. Борелем в начале XX в. в связи с дискуссией с А. Пуанкаре по поводу понятия физической непрерывности (см. [18]). Борель предложил описывать понятие «куча» последовательностью неотрицательных чисел p_n, $n = 1, 2, \ldots$, где p_n — доля тех, кто назвал бы n зерен «кучей», среди всех говорящих на русском языке. В этом предложении выделим два момента. Во-первых, понятие «куча» предлагается описывать функцией, отображающей натуральный ряд в отрезок [0,1]. Во-вторых, указан способ нахождения этой функции с помощью экспертных оценок (или, если угодно, социологического опроса).

О мыслях Э. Бореля по поводу софизма «Куча» я узнал из его популярной книги [18], первое издание которой вышло в 1950 г. Трудно сказать, почему эти мысли не получили тогда развития. Во всяком случае теория нечеткости ведет свое начало с первой статьи Л. А. Заде [1] 1965 г., в которой появился термин «fuzzy» — «нечеткий» *. В отличие от «голой» идеи Бореля в [1] предлагался аппарат расчетов с помощью нечетких множеств. О достоинствах и недостатках этого весьма развитого к настоящему времени аппарата мы подробно поговорим. Здесь сыграл свою роль и удачный выбор термина «fuzzy», давшего имя теории, и подчеркнутое противопоставление нового подхода классическим. В результате — поток публикаций и споров.

Определение 2. (Л. А. Заде). «Нечеткое подмножество A универсального множества U характеризуется функцией принадлежности $\mu_A : U \to [0,1]$, которая ставит в соответствие каждому элементу $u \in U$ число $\mu_A(u)$ из отрезка [0,1], характеризующее степень принадлежности элемента u подмножеству A» [19, с. 32].

Таким образом, понятие «куча» описывается, по Борелю, нечетким множеством K: универсальным множеством является натуральный ряд N, а функция принадлежности имеет вид: $\mu_K(n) = p_n$, $n = 1, 2, \ldots$

Обычные множества являются частными случаями нечетких — для них $\mu_A(u) = 1$ (если $u \in A$) или $\mu_A(u) = 0$ (если $u \notin A$).

Отметим нечеткость определения Л. А. Заде. Говорится, что нечеткое множество «характеризуется» функцией принадлежности, но само по себе оно не определяется. Поскольку в дальнейшем используется только понятие функции принадлежности, то мы вправе сказать, что нечеткое множество и есть функция принадлежности. Однако

* Термин «fuzzy» (англ.), «floue» (франц.) в различных публикациях переводится на русский язык как «нечеткий», «размытый», «расплывчатый», «туманный» и т. д. Мы придерживаемся первого варианта — «нечеткий».

для интуиции более подходит первый термин, чем второй. Мы будем использовать их как синонимы.

Приведем привычный для публикаций Л. А. Заде пример нечеткого множества [19, с. 32—33].

«Пусть универсальное множество U представляет собой отрезок [0,100], и переменная u, принимающая значения из этого интервала, интерпретируется как «возраст». Нечеткое подмножество универсального множества U, обозначаемое термином «старый», можно определить функцией принадлежности вида:

$$\mu_A(u) = \begin{cases} 0 & \text{при } 0 \leqslant u \leqslant 50, \\ \left(1 + \left(\dfrac{u-50}{5}\right)^{-2}\right)^{-1} & \text{при } 50 < u \leqslant 100». \end{cases}$$

Этот пример носит явно искусственный характер, поскольку выбор вида функции $\mu_A(u)$ ничем не мотивирован. Ясно, однако, что в результате опроса достаточно большой группы испытуемых можно оценить функцию принадлежности понятия «старый». Ясно также, что она будет меняться в зависимости от возраста испытуемых, а возможно, и от каких-либо иных их показателей. Л. А. Заде предпочитает не касаться сложных проблем статистического исследования, развивая формальный аппарат.

Описание этого аппарата занимает много страниц. К настоящему времени он достаточно хорошо и подробно изложен на русском языке (см. монографию [19], статьи Л. А. Заде [15], Р. Беллмана и Л. Заде [20], обзор [21] и т. д.). В настоящей брошюре мы не можем претендовать на развитие аппарата теории нечеткости в сколько-нибудь полном виде и потому ограничимся только теми определениями, которые нам понадобятся в дальнейшем.

Определение 3. Пусть A и B — два нечетких подмножества универсального множества U с функциями принадлежности $\mu_A(u)$ и $\mu_B(u)$ соответственно. Пересечением $A \cap B$, произведением AB, объединением $A \cup B$, суммой $A + B$, отрицанием $\bar{A}$ называются нечеткие подмножества с функциями принадлежности

$$\mu_{A \cap B}(u) = \min(\mu_A(u), \ \mu_B(u)),$$
$$\mu_{AB}(u) = \mu_A(u)\, \mu_B(u),$$
$$\mu_{A \cup B}(u) = \max(\mu_A(u), \mu_B(u)),$$
$$\mu_{A+B}(u) = \mu_A(u) + \mu_B(u) - \mu_A(u)\, \mu_B(u),$$
$$\mu_{\bar{A}}(u) = 1 - \mu_A(u), \ u \in U$$

соответственно.

Если A и B являются обычными множествами, т. е. их функции принадлежности принимают только значения 0 или 1, то это определение приводит к обычным понятиям пересечения, объединения и отрицания множеств. Вместо одного понятия «пересечение» в теории нечетких множеств рассматриваются два — «пересечение» и «произведение», а вместо «объединения» — также два: «объединение» и «сумма».

Некоторые из обычных свойств операций над множествами сохраняются и в теории нечетких множеств, другие же нет. Приведем примеры тех и других [9, § 4.1].

Как известно, законами де Моргана называются тождества

$$\overline{A \cup B} = \bar{A} \cap \bar{B}, \quad \overline{A \cap B} = \bar{A} \cup \bar{B}.$$

Теорема 1. Законы де Моргана верны для нечетких множеств, а именно справедливы тождества (6) и, кроме того, такие

$$\overline{A + B} = \bar{A}\bar{B}, \ \overline{AB} = \bar{A} + \bar{B}.$$

Теорема 2. Равенство

$$A + A = A \tag{7}$$

верно тогда и только тогда, когда A — четкое множество (т. е. функция принадлежности принимает значения только 0 и 1).

Теорема 3. Для любых нечетких множеств A, B и C

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C). \tag{8}$$

В то же время равенство

$$A\,(B + C) = AB + AC \tag{9}$$

имеет место тогда и только тогда, когда при всех $u \in U$

$$(\mu_A^2\,(u) - \mu_A\,(u))\,\mu_B\,(u)\,\mu_C\,(u) = 0.$$

Сказанного достаточно, чтобы констатировать, что понятие нечеткого множества является обобщением понятия множества, причем нетривиальным. Вместе с тем видны и некоторые недостатки рассматриваемого математического аппарата. Так, почти невозможно учитывать зависимость реалий, моделируемых нечеткими множествами,— в качестве «общей части» можно использовать либо пересечение, либо произведение, в то время как видов зависимостей явно больше чем два. Поэтому имеет смысл предложить обобщение аппарата Заде, позволяющее избавиться от указанного недостатка. В качестве подобного обобщения автор в 1975 г. предложил рассматривать теорию случайных множеств (см. ниже пп. 5, 6).

Отметим еще «принцип обобщения» для нечетких множеств [19, с. 47]. Пусть A — нечеткое подмножество U с функцией принадлежности $\mu_A(u)$, $u \in U$. Пусть f — отображение из U в V. Тогда нечеткое подмножество $f(A)$ универсального множества V определяется формулой:

$$\mu_{f(A)}(f(u)) = \mu_A(u), \; u \in U.$$

Свойственная Л. А. Заде нечеткость изложения привела к тому, что он забыл рассмотреть случай, когда $f(u_1) = f(u_2)$ при $u_1 \neq u_2$. В соответствии с общим духом этой теории можно предложить следующую модификацию определения Л. А. Заде:

$$\mu_{f(A)}(v) = \begin{cases} 0, & v \notin f(U), \\ \max\limits_{u \in f^{-1}(v)} \mu_A(u), & v \in f(U). \end{cases}$$

Принцип обобщения позволяет рассматривать функции от нечетких переменных, в частности, изучать устойчивость моделей в стиле предыдущего пункта 2. Разберем этот вопрос несколько подробнее.

В силу неопределенности реальных явлений вместо исходных данных x имеем нечеткое множество x' с функцией принадлежности $\mu_{x'}(y), y \in X$. В соответствии с принципом обобщения решение представляется нечетким множеством $f(x')$ в Y. Можно ввести ряд показателей устойчивости. Так, потери характеризуются нечетким множеством $\rho(f(x), f(x'))$. Чтобы характеризовать потери одним числом в стиле (3)—(4), введем в Y множества уровня α:

$$Y_\alpha = \{y \in Y : \mu_{f(x')}(y) \geqslant \alpha\},$$

где $0 \leqslant \alpha \leqslant 1$. В качестве показателя устойчивости можно использовать диаметр Y_α при некотором α:

$$d_\alpha = \sup_{y,\, y' \in Y_\alpha} \rho(y, y').$$

4. Пятнадцать лет теории нечеткости

По приписываемому Н. Бурбаки утверждению, к которому автор присоединяется, «логически говоря, возможно вывести почти всю современную математику из единого источника — теории множеств». Использование концепции нечеткости открывает путь к «удвоению» математики: заменяя обычные множества нечеткими, мы можем каждому объекту в математике поставить в соответствие его нечеткий аналог. Рассматривают нечеткие отношения, теоремы, алгоритмы, топологии, классификации и т. д. При этом определение производных понятий требует некоторой математической техники, которая и развивается во многих опубликованных к настоящему времени статьях и книгах.

Теория нечеткости молода. Это проявляется, например, в том, что основное внимание многих авторов направлено на развитие языка, а не на выяснение его свойств, статьи представляют собой цепочки определений, теоремы встречаются редко, и они просты. Зато обсуждаются и сравниваются между собой различные определения одного и того же понятия, например, «нечеткого отношения предпочтения».

Целесообразность использования теории нечеткости в приложениях подробно обосновывалась в п. 1. Тем не менее большинство авторов изучают теоретические вопросы, примеры носят иллюстративный характер. Есть ряд работ, связанных с анализом реальных данных, примером которых является исследование профессора Понсара из Дижонского университета, о котором он рассказывал на одном из упомянутых в предисловии семинаров. С помощью нечетких множеств (точнее, нечетких графов) Понсар описал структуру связей различных частей Франции между собой, а затем, применив соответствующий алгоритм классификации (использующий, естественно, понятие нечеткого множества), выделил экономические районы своей страны. Полученное им районирование отличалось, естественно, от официально принятого.

Критика работы Понсара может идти по разным направлениям. Нельзя ли было применить какой-либо классический метод? Каковы качество, надежность полученного решения? Дело в том, что методы математической статистики позволяют оценить, скажем, точность полученных значений параметров, если для анализа данных применяются вероятностные методы. Теория нечеткости такими возможностями не располагает. Наконец, нельзя ли подобрать какой-либо метод, который даст лучший результат? (Конечно, понятие «лучший» должно быть определено.) Наши вопросы носят методологический характер, тематику экономической географии мы при этом не затрагиваем.

Ясно, что сформулированные в предыдущем абзаце вопросы можно задать по поводу любого исследования, основанного на использовании реальных данных. Новизна используемых методов (теория нечеткости) их обостряет.

Работа является более обоснованной, если в ней применены не один, а несколько методов и проведено сравнение полученных результатов, как это сделано в [16]. Это соответствует известной рекомендации теории исследования операций о целесообразности решения одной и той же задачи с использованием различных моделей явлений, различных методов расчетов (см., например, [9, § 3.7], [22]).

Как уже отмечалось в п. 1, методы, основанные на теории нечеткости, оказались в [16] лучше, чем другие. Однако остается возражение такого сорта: «Не проводилось ли сравнение с заведомо слабым соперником?»

Иногда от энтузиастов теории нечеткости требуют: «Приведите пример убедительной прикладной значимости теории нечеткости, пример, когда ее методы позволяют получить результат, а любые другие либо не позволяют, либо дают худшие результаты». Требование кажется естественным. Но оно некорректно. Ни один прикладной метод не удовлетворяет такому требованию.

Действительно, всегда найдутся содержательные возражения, основанные на особенностях области приложений. Даже оставляя их в стороне (а это означает, что мы отказываемся констатировать наличие или отсутствие «результата»), мы оказываемся перед совокупностью «любых других методов». Она необозрима даже теоретически — ведь любой метод может быть развит дальше! Тем более практически провести расчеты по всем известным методам для какой-либо задачи «стоит слишком дорого», если только она не является сверхважной. Значит, результаты любого практического применения можно оспаривать, а на вопрос: «А нельзя ли сделать лучше?» есть стандартный ответ: «Скорее всего, можно».

Не надо обвинять автора в агностицизме. Практическая рекомендация сводится к тому, что в пакет прикладных программ на ЭВМ целесообразно включать ряд методов для решения одних и тех же задач и проводить для реальных данных расчеты по всем этим программам. И некоторые из этих программ должны быть основаны на использовании нечетких множеств.

Иногда разрабатываются методы подбора «наилучшего» способа расчетов в конкретной задаче. Поиск «наилучшего» противоречит плюралистической позиции автора, выраженной в предыдущем абзаце, а потому рассматривается им лишь как вынужденная временная мера.

Не могу удержаться от одного примера. Для сравнения различных методов оценивания параметров в математической статистике используют дисперсии оценок. Считают, что наблюдения извлечены из генеральной совокупности с распределением (теперь чаще говорят просто — «из распределения», опуская бессмысленные слова «генеральная совокупность»), принадлежащим некоторому параметрическому семейству (ср. [11, § 1]). Лучшей оценкой считается та, у которой наименьшая дисперсия. Она называется эффективной (подробнее см., например, [23]). Ее и рекомендуют использовать.

Так вот, эта рекомендация является псевдонаучной. Дело в том, что параметрическое семейство распределений — это абстракция, к которой при обработке реальных данных можно приближаться, но нельзя достигнуть. Рассмотрим, например, модель, в которой предполагается, что наблюдения извлечены из нормального распределения с неизвестным средним и единичной дисперсией. Тогда, как известно, эффективной оценкой среднего является среднее арифметическое наблюдений [23, с. 524]. Однако в реальных условиях надо считаться с возможностью «засорения» выборки — оператор может неправильно записать показание прибора, наблюденный объект может относиться к другому классу (не тому, что изучается), могут возникнуть помехи в приборе и т. д. Поэтому более реально предполагать, что наблюдения извлекаются из описанного выше распределения с вероятностью $1 - \varepsilon$, а с вероятностью ε — из некоторого другого распределения, описывающего всевозможные помехи. При этом положительное число ε — мера «засорения». Оно достаточно мало, если влияние помех не слишком велико. Точнее говоря, помехи оказывают действие примерно в $100\varepsilon\,\%$ случаев.

Что же получается? Если распределение, индуцированное помехами, является, например, распределением Коши, то, как нетрудно показать, среднее арифметическое наблюдений не будет приближаться к среднему исходной нормальной совокупности. То есть среднее арифметическое будет настолько плохой оценкой (несостоятельной), что ее даже называть оценкой неудобно. И это при любом положительном ε. В то же время выборочная медиана остается хорошей оценкой, при малых ε она мало отличается от оцениваемого параметра (точнее, приближается к нему при увеличении числа наблюдений и уменьшении ε, а при фиксированном ε и большом объеме наблюдений отличается от оцениваемого среднего на величину порядка ε). Значит, целесообразнее использовать выборочную медиану, чем среднее арифметическое наблюдений. Однако при $\varepsilon = 0$ дисперсия выборочной медианы больше, чем у среднего арифметического, она не является эффективной оценкой.

Таким образом, оптимальность и даже «разумность» оценки могут быстро теряться при малейших отклонениях от исходных предположений. Т. е. первое требование — устойчивость, а уж потом — оптимальность.

Вернемся к примеру. В случае «засорений в $100\,\varepsilon$ % случаев» статистическая теория развита Тьюки (G. W. Tukey) и Хубером (P. J. Huber), а также их последователями (ссылки см. в [9, § 2.1]). Она является параметрической, что, конечно, большой недостаток. От искусственного предположения о принадлежности распределения к параметрическому семейству удается отказаться в непараметрической статистике (см. [11]). Что же касается оценок среднего, линии регрессии и т. д., то наиболее перспективными являются методы взвешенных моментов, развиваемые Л. Д. Мешалкиным и его учениками (см. [24], [25]).

Помимо дискуссий о пользе применения аппарата теории нечеткости в конкретных приложениях, можно поставить общий вопрос о возможности описания реального мира с помощью этого аппарата. И тут мы сталкиваемся с парадоксом. Было приведено много соображений в пользу того, что адекватным образом объекта является нечеткое множество A, характеризующееся функцией принадлежности $\mu_A(u)$. При этом u — вполне определенный элемент множества U. Может ли он соответствовать чему-то реальному? Разберем оба возможных ответа.

Если «да», то в соответствии со сказанным выше надо вместо элемента u рассматривать нечеткое множество u с функцией принадлежности $\mu_u(w)$, где w принадлежит некоторому другому универсальному множеству W. Однако относительно w можно задать тот же самый вопрос, что и относительно u. И если ответ опять «да», то мы получаем, что w также не определенный элемент определенного множества, а нечеткое множество и т. д.

Если же на каком-то этапе будет ответ «нет», то положение не лучше — для описания реального явления привлечены фиктивные элементы, которым ничто в мире не соответствует.

Таким образом, принимая тезис «все в мире нечетко», мы вынуждены выбирать между двумя возможностями: описанием явлений с помощью бесконечной последовательности нечетких множеств и описанием тех же явлений с помощью фиктивных элементов. А тезис этот достаточно хорошо обоснован (см. п. 1).

Аналогичные рассуждения можно провести, разбирая не нечеткость элемента универсального множества, а нечеткость значения функции принадлежности. Л. А. Заде [19, с. 52—57] разработал аппарат нечетких множеств с нечеткими функциями принадлежности, благоразумно не вдаваясь при этом в рассуждения о том, на какой же итерации считать функцию принадлежности четкой.

Конечно, описанный нами парадокс не мешает успешно применять теорию нечеткости в конкретных приложениях. Из него вытекает лишь необходимость указания и обсуждения «границ применимости» этой теории. Для вероятностно-статистических методов эта методологическая работа давно ведется (см. [2—4]).

По-видимому, при обсуждении концепции нечеткости один из первых вопросов таков: «Чем понятие нечеткого множества отличается от понятия случайной величины?» Действительно, распределение случайной величины — один из примеров нечеткого множества. Но при этом сумма значений функции принадлежности по всем элементам универсального множества U (или интеграл, если U — интервал) равна 1, что совсем не является обязательным для произвольного нечеткого множества. Второй аргумент — различие аппаратов этих двух теорий. Поэтому в публикациях по нечетким множествам обычно подчеркивается, что они не сводятся к случайным величинам (см., например, [20]). Но при этом делается неправомерный вывод, что теория нечеткости не имеет отношения к теории вероятностей. На самом деле первую из них можно включить во вторую, но используя при этом не случайные величины, а случайные множества.

Теория случайных множеств — достаточно развитая часть теории вероятностей. Ее применение оказывается полезным в математической экономике, экспертных оценках, при моделировании распространения лесных пожаров и т. д. (см. [9, гл. 4]).

Описанию связи между нечеткими и случайными множествами посвящен один из следующих разделов. Автор надеется, что эта связь позволит применить в теории нечеткости результаты и методы, накопленные в теории случайных множеств, и, наоборот, даст возможность перенести понятия и постановки задач из первой теории во вторую, что послужит прогрессу в обеих теориях и их приложениях.

5. Вспомним определения

При первом чтении этот раздел можно пропустить и обращаться к нему при возникновении вопросов. Необходимость же его основана на горьком опыте недоразумений, возникающих, когда собеседники под одним и тем же термином понимают разное.

Целесообразно сделать несколько замечаний о том, что в этой брошюре понимается под теорией вероятностей вообще и под случайными величинами и случайными множествами в частности.

Теорию вероятностей мы понимаем как аксиоматическую математическую теорию (используется аксиоматика А. Н. Колмогорова [26]). В качестве таковой она не имеет никакого непосредственного отношения к реальному миру. В частности, выражение «вероятность попадания в цель равна 0,6» не имеет смысла внутри нее.

Чтобы связать теорию вероятностей с прикладной областью, необходимо построить вероятностную модель. Например, результат выстрела будем описывать случайной величиной, принимающей значения 1 (попадание) и 0 (промах) с вероятностями 0,6 и 0,4 соответственно. Предыдущая фраза и задает модель. При рассуждениях на языке этой модели, допуская некоторую вольность речи, можно говорить о вероятности попадания в цель.

Правильно ли модель описывает реальность? Этот вопрс не входит в компетенцию теории вероятностей. Она — только инструмент. А инструмент не несет ответственности за то, как им пользуются.

Вопрос о применимости вероятностной модели должен решаться, видимо, с тех же методологических позиций, что и вопрос о применимости модели, использующей понятия из любой другой области математики. В частности, достаточно наличия статистической устойчивости [2—4]. Мы не будем развивать эту сложную тему подробнее.

Напомним, что исходным объектом в теории вероятностей является вероятностное пространство. Пусть задано произвольное множество Ω, элементы которого будем обозначать ω; в Ω выделена совокупность его подмножеств $\mathscr{F}$, элементы $\mathscr{F}$ называются событиями. Предполагается, что $\mathscr{F}$ удовлетворяет некоторым естественным условиям (на профессиональном языке: $\mathscr{F}$ есть σ-алгебра). В частности, для любых множеств A и B из $\mathscr{F}$ их пересечение $A \cap B$, объединение $A \cup B$ и разность $A \setminus B$ входят в $\mathscr{F}$. Кроме того, в Ω задана вероятностная мера $\mathbf{P}$, т. е. для любого события $A \in \mathscr{F}$ определена его вероятность $\mathbf{P}(A)$. Тройка $\{\Omega, \mathscr{F}, \mathbf{P}\}$ и называется вероятностным пространством, или пространством элементарных событий.

Случайная величина $\xi = \xi(\omega)$ со значениями в множестве X — это функция из Ω в X, т. е. в стандартных обозначениях $\xi : \Omega \to X$. Предполагается, что в X задана σ-алгебра $\mathfrak{a}$ и для любого $C \in \mathfrak{a}$ выполнено включение $\xi^{-1}(C) \in \mathscr{F}$, т. е. ξ является $(\mathscr{F}, \mathfrak{a})$ — измеримой функцией. Распределением вероятностей случайной величины ξ называется функция $P : \mathfrak{a} \to [0, 1]$, определенная так:

$$P(C) = \mathbf{P}(\xi^{-1}(C)).$$

Предыдущие два абзаца соответствуют определениям, данным в справочнике Ю. В. Прохорова и Ю. А. Розанова [27, с. 132—133]. Иногда термин «случайная величина» используют лишь в случае, когда X — действительная прямая, а введенный выше объект называют «случайным элементом». Эта несогласованность не должна удивлять. Подобная нечеткость терминологии встречается и для более простых объектов. Пустое множество одни обозначают $\varnothing$, а другие — Λ. Даже в арифметике разнобой: в десятичной дроби целую часть от дробной одни отделяют запятой: 4,12, а другие — точкой: 4.12.

Мы напомнили основные определения. Рассказывать о том, что такое теория вероятностей, не входит в нашу задачу (см. для этого, например, [2]). Отметим только, что понятие вероятностного пространства вполне доступно школьникам: опираясь на него, автор изложил ученикам 8—9-х классов наряду с элементарной теорией доказательство закона больших чисел (с помощью неравенства Чебышева) и методологию проверки статистических гипотез всего за 14 часов [28]!

Что же такое случайное множество? Мы ограничимся рассмотрением подмножеств фиксированного конечного множества $Y = \{y_1, y_2, \ldots, y_k\}$. Множество Y имеет 2^k подмножеств. Их совокупность обозначают 2^Y.

Определение 4. Случайная величина со значениями в 2^Y называется случайным подмножеством Y.

Предполагается, естественно, что в упомянутую в общем определении понятия случайной величины σ-алгебру $\mathfrak{a}$ входят все подмножества 2^Y, т. е. $\mathfrak{a}=2^{(2^Y)}$.

Пусть $A = A(\omega)$ — случайное подмножество Y. Как описать распределение A? Поскольку все подмножества 2^Y состоят из конечного числа элементов, прообразы которых (при отображении A) не пересекаются, то распределение A полностью описывается совокупностью 2^k чисел:

$$P(A = X) = \mathbf{P}(\{\omega : A(\omega) = X\}),\ X \subseteq Y.$$

Числа $P(A = X)$ неотрицательны и в сумме составляют 1. Обратно, по совокупности 2^k неотрицательных чисел, в сумме составляющих 1, можно построить такие вероятностное пространство и случайное множество, что эти числа будут представлять собой распределение построенного случайного множества (см. [9, с. 171]).

С помощью распределения выражаются различные вероятности, связанные со случайным множеством. Рассмотрим, например, вероятность накрытия точки $y \in Y$ случайным подмножеством A множества Y. Накрытие имеет место при тех и только тех $\omega \in \Omega$, при которых $y \in A(\omega)$. Значит, вероятность накрытия есть $\mathbf{P}(\{\omega : y \in A(\omega)\})$. Эта вероятность равна сумме $P(A = X)$ по всем $X \subseteq Y$, содержащим y, поскольку $\{\omega : y \in A(\omega)\}$ есть объединение $\{\omega : A(\omega) = X\}$ по всем $X \subseteq Y$, содержащим y, а события $\{\omega : A(\omega) = X_1\}$ и $\{\omega : A(\omega) = X_2\}$ имеют пустое пересечение при $X_1 \neq X_2$. Для простоты записи введем обозначение

$$P(y \in A) = \mathbf{P}(\{\omega : y \in A(\omega)\}).$$

Как мы показали, для вероятности накрытия верно равенство

$$P(y \in A) = \sum_{y \in X,\ X \subseteq Y} P(A = X).$$

Как известно, понятие распределения вероятностей зачастую является более «работающим», чем исходное понятие вероятностного пространства. Так, моменты действительной случайной величины выражаются через ее распределение. Нам в дальнейшем понадобится еще одна подобная связь.

Пусть f — действительная функция от подмножества Y, т. е. $f : 2^Y \to R^1$. Тогда для любого случайного подмножества $A = A(\omega)$ множества Y можно рассмотреть действительную случайную величину $f(A)$. Она принимает конечное число значений, а потому ограничена и имеет все моменты. Распределение и моменты $f(A)$ выражаются через распределение случайного множества A. Так, например, для математического ожидания имеем:

$$Ef(A) = \sum_{X \subseteq Y} f(X) P(A = X).$$

Рассмотрим несколько случайных множеств $A(\omega)$, $B(\omega)$, $C(\omega)$, определенных на одном и том же вероятностном пространстве и являющихся подмножествами Y. Операции над случайными множествами вводятся естественным образом — при каждом $\omega \in \Omega$ должно быть выполнено соответствующее соотношение. Так, случайное множество C называется объединением случайных множеств A и B и обозначается $C = A \cup B$, если при каждом $\omega \in \Omega$ имеем $C(\omega) = A(\omega) \cup B(\omega)$. Естественно, обычные свойства операций над множествами сохраняются и для случайных множеств.

Понятие совместного распределения случайных множеств является частным случаем понятия совместного распределения случайных величин [27, с. 133]. Именно, совместным распределением случайных множеств $A_1, A_2, \ldots, A_n$ называется совокупность 2^{kn} чисел

$$P(A_1 = X_1, A_2 = X_2, \ldots, A_n = X_n) = \mathbf{P}(\{\omega : A_1(\omega) = X_1,$$

$$A_2(\omega) = X_2, \ldots, A_n(\omega) = X_n\}), \ X_i \subseteq Y, \ i = 1, \ldots, n.$$

Распределение результата операции можно выразить через совместное распределение участвующих в операции множеств. Например,

$$P(A \cap B = X) = \sum_{X_1, X_2 : X_1 \cap X_2 = X} P(A = X_1, B = X_2).$$

Для доказательства этого соотношения достаточно заметить, что $(A \cap B)(\omega) = X$ тогда и только тогда, когда существует пара подмножеств X_1 и X_2 множества Y такая, что $A(\omega) = X_1$, $B(\omega) = X_2$ и $X_1 \cap X_2 = X$.

Понятие независимости случайных множеств также не требует специального определения: в соответствии со справочником Ю. В. Прохорова и Ю. А. Розанова [27] случайные множества $A_1, A_2, \ldots, A_n$ являются независимыми, если

$$P(A_1 = X_1, A_2 = X_2, \ldots, A_n = X_n) = \\ = P(A_1 = X_1) P(A_2 = X_2) \ldots P(A_n = X_n)$$

для любых $X_i \subseteq Y$, $i = 1, \ldots, n$.

Читателя теперь уже не должна смущать фраза: «Пусть A_1 $A_2, \ldots, A_n$ — независимые, одинаково распределенные случайные множества ...» Впрочем, он вправе спросить: «Что в рассказанном специфично для случайных множеств?» Вопрос законный — многое справедливо для случайных величин со значениями в произвольном конечном множестве. Специфическими же являются: нахождение вероятности накрытия точки случайным множеством, вычисление распределения пересечения $A \cap B$ через совместное распределение A и B. Повторим кратко основные определения для случайных величин со значениями в некотором конечном множестве $X = \{x\} = \{x_1, \ldots, x_m\}$.

Случайная величина $\xi = \xi(\omega)$ — это функция $\xi: \Omega \to X$, такая, что прообразы элементов X, т. е. $\xi^{-1}(x_i)$, $i = 1, \ldots, m$, являются событиями: $\xi^{-1}(x) \in \mathcal{F}$. Распределение вероятностей ξ описывается набором m неотрицательных чисел, в сумме дающих 1:

$$P(\xi = x) = \mathrm{P}(\{\omega : \xi(\omega) = x\}), \quad x \in X.$$

Тогда для любого $Y \subseteq X$ имеем:

$$P(\xi \in Y) = \sum_{x \in Y} P(\xi = x).$$

Пусть $f: X \to R^1$ — произвольная действительная функция. Тогда $f(\xi)$ — действительная случайная величина, все значения которой принадлежат конечному множеству $\{f(x), x \in X\}$. Распределение и моменты $f(\xi)$ выражаются через распределение ξ, в частности:

$$Ef(\xi) = \sum_{x \in X} f(x) P(\xi = x).$$

Пусть $\xi_1, \xi_2, \ldots, \xi_n$ — случайные величины на одном и том же пространстве элементарных событий $\{\Omega, \mathcal{F}, \mathrm{P}\}$ со значениями в конечных множествах $X_1, X_2, \ldots, X_n$ соответственно. Их совместное распределение описывается набором чисел:

$$P(\xi_1 = x^1, \xi_2 = x^2, \ldots, \xi_n = x^n) = P(\{\omega : \xi_1(\omega) = x^1, \xi_2(\omega) = x^2, \ldots, \xi_n(\omega) = x^n\}), \; x^i \in X_i, \; i = 1, \ldots, n.$$

Случайные величины $\xi_1, \xi_2, \ldots, \xi_n$ являются независимыми, если

$$P(\xi_1 = x^1, \xi_2 = x^2, \ldots, \xi_n = x^n) = P(\xi_1 = x^1) P(\xi_2 = x^2) \ldots P(\xi_n = x^n)$$

для любых $x^i \in X_i$, $i = 1, \ldots, n$.

Если множества значений не являются конечными, то приходится возиться с понятием измеримости. А надо ли? Все, что рассматривается в математических моделях реальных явлений, можно считать состоящим из конечного числа элементов. Хотя бы потому, что по теориям некоторых современных физиков число атомов во Вселенной конечно. А раз так, то использование бесконечных множеств не является необходимым. Оно иногда полезно — вычислить интеграл от $g(x) = x^2$ гораздо легче, чем соответствующую интегральную сумму. А иногда вредно — затрудняет работу.

Сам жанр популярной брошюры дает автору право игнорировать технические сложности измеримости, рассматривая лишь конечные множества. Но и в реальной работе излишнее внимание к подобным вопросам является, по мнению автора, симптомом схоластического перерождения математики, о котором красочно писал В. Н. Тутубалин [4].

Сделаем еще одно замечание о случайных подмножествах конечного множества $Y = \{y_1, y_2, \ldots, y_k\}$. Рассмотрим совокупность 2^k чисел:

$$P(X \subseteq A) = \mathbf{P}(\{\omega : X \subseteq A(\omega)\}),\ X \subseteq Y.$$

Среди них находятся и числа $P(\{y_j\} \subseteq A) = P(y_j \in A)$, $j = 1, \ldots, k$, которые мы рассматривали ранее.

Теорема 4. Наборы чисел $P(A = X)$, $X \subseteq Y$, и $P(X \subseteq A)$, $X \subseteq Y$, выражаются один через другой.

6. Нечеткие множества как проекции случайных

Покажем, что теория нечетких множеств в определенном смысле сводится к теории случайных множеств и тем самым включается в теорию вероятностей.

Определение 5. Пусть $A = A(\omega)$ — случайное подмножество конечного множества Y. Нечеткое множество B, определенное на Y, называется проекцией A и обозначается $\operatorname{Proj} A$, если

$$\mu_B(y) = P(y \in A)$$

при всех $y \in Y$.

Термин «проекция» оправдывается тем, что $P(y \in A) = P(\{y\} \subseteq A\}$ и вектор $\{P(\{y\} \subseteq A),\ y \in Y\}$, задающий в соответствии с определением 5 функцию принадлежности нечеткого множества, является подвектором вектора $\{P(X \subseteq A),\ X \subseteq Y\}$, который определяет распределение случайного множества A (см. теорему 4).

В соответствии с определением 5 каждому случайному множеству A можно поставить в соответствие нечеткое множество $\operatorname{Proj} A$. Как показывает следующая теорема, обратное отображение определено для каждого нечеткого множества (но, вообще говоря, не является однозначным).

Теорема 5. Для любого нечеткого подмножества B множества Y существует случайное подмножество A множества Y, такое, что $B = \mathrm{Proj}\, A$.

Оказывается, соответствие между нечеткими и случайными множествами можно установить так, чтобы результаты операций также соответствовали друг другу. Приведем соответствующие результаты, доказательства которых содержатся в [9, § 4.2].

Теорема 6. Если $\mathrm{Proj}\, A = B$, то $\mathrm{Proj}\, \bar{A} = \bar{B}$.

Следующие две теоремы показывают, каким образом операция пересечения случайных множеств «раздваивается» на две операции над нечеткими множествами — пересечение и произведение (см. определение 3).

Теорема 7. Если случайные подмножества A_1 и A_2 множества Y независимы, то $\mathrm{Proj}\, (A_1 \cap A_2)$ является произведением $\mathrm{Proj}\, A_1$ и $\mathrm{Proj}\, A_2$.

Определение 6. Носителем случайного множества A называется совокупность всех тех $y \in Y$, для которых $P\, (y \in A) > 0$.

Теорема 8. Равенство

$$\mathrm{Proj}\, (A_1 \cap A_2) = (\mathrm{Proj}\, A_1) \cap (\mathrm{Proj}\, A_2)$$

верно тогда и только тогда, когда пересечение носителей случайных множеств $\bar{A}_1 \cap A_2$ и $A_1 \cap \bar{A}_2$ пусто.

Предоставляем читателю сформулировать соответствующие теоремы для операции объединения, используя справедливость законов де Моргана для нечетких множеств (теорема 1) и теорему 6.

Теорема 9. Пусть $B_1, B_2, B_3, \ldots, B_t$ — некоторые нечеткие подмножества конечного множества Y. Пусть

$$B^m = ((\ldots ((B_1 \bigcirc B_2) \bigcirc B_3) \bigcirc \ldots) \bigcirc B_{m-1}) \bigcirc B_m,$$
$$m = 1, \ldots, t,$$

где $\bigcirc$ — символ одной из следующих операций: пересечение, произведение, объединение, сумма. При этом на разных местах могут стоять разные знаки, но на фиксированном месте при всех m стоит один и тот же знак. Тогда существуют случайные подмножества $A_1, A_2, A_3, \ldots, A_t$ множества Y, такие, что

$$\mathrm{Proj}\, A_i = B_i, \ i = 1, 2, 3, \ldots, t \quad \text{и}$$

$$\mathrm{Proj}\, \{((\ldots ((A_1 * A_2) * A_3) * \ldots) * A_{m-1}) * A_m\} =$$
$$= B^m, \ m = 1, \ldots, t,$$

где $*$ обозначает операцию пересечения, если в определении B^m на соответствующем месте стоит знак пересечения или произведения, и операцию объединения, если в B^m стоит символ объединения или суммы.

По-видимому, теорема 9 верна и при произвольной расстановке скобок при определении B^m, однако эта гипотеза пока не доказана и не опровергнута.

Как представляется автору, теоремы 5—9 дают основания полагать, что любое понятие или утверждение из теории нечетких множеств может быть «промоделировано» на языке случайных множеств. Теоремы 7 и 8 показывают, что использование пересечения и произведения нечетких множеств означает справедливость вполне определенных предположений о зависимости между моделируемыми понятиями, в то время как применение случайных множеств дает возможность гораздо более гибко учитывать указанную зависимость.

Имеется и другая, более важная связь теории нечеткости с вероятностно-статистическими методами. Именно, нечеткие множества можно рассматривать как один из видов статистических данных, к обработке которых применяется весь набор статистических методов. Чтобы обсудить эту связь (глава 3 брошюры), нам понадобится сначала рассмотреть современные подходы к обработке статистических данных общей природы — как числовой, так и нечисловой.

Глава 2. МАТЕМАТИЧЕСКАЯ СТАТИСТИКА И ЗАДАЧИ ОПТИМИЗАЦИИ

7. О развитии математической статистики

Как пишут А. Н. Колмогоров и Ю. В. Прохоров в Большой Советской Энциклопедии, «математическая статистика — раздел математики, посвященный математическим методам систематизации, обработки и использования статистических данных для научных и практических выводов. При этом статистическими данными называются сведения о числе объектов в какой-либо более или менее обширной совокупности, обладающих теми или иными признаками...» [30]. Далее выделяется часть математической статистики, опирающаяся на вероятностные модели. Эта часть не исчерпывает все содержание рассматриваемого раздела математики.

Привести выписку из Большой Советской Энциклопедии было уместно, поскольку в настоящее время известно несколько сотен различных пониманий термина «статистика». В [31] собрано около 200 соответствующих цитат. В частности, широко распространена, особенно среди математиков-преподавателей, тенденция сводить математическую статистику к задачам, сформулированным на вероятностном языке. При этом за ее пределами остается так называемый анализ данных, важными частями которого являются кластер-анализ *, многомерное шкалирование (метрическое и неметрическое), репрезентативная теория измерений ** и др. Некоторое представление об этих частях математической статистики можно получить из [6—13], где есть и дальнейшие ссылки.

В математической статистике можно выделить две стороны — решаемые задачи и природа рассматриваемых объектов, которые часто используются как основания для внутреннего деления этой области. Чисто прагматически выделим три вида задач, между которыми, разумеется, имеются многообразные связи — описание материала, оценивание, проверка гипотез.

Чтобы описать совокупность наблюдений, используют разнообразные таблицы, вычисляют средние и показатели разброса, строят гистограммы, проводят группировку наблюдений, в частности, формальными методами (кластер-анализ). Цель состоит в том, чтобы дать предварительное представление о статистическом материале, использовав меньшее количество чисел, чем в нем имеется, т. е. «сжав информацию».

Роль задач описания материала по-разному оценивается в практической деятельности и в существующих учебниках. Каждому, кто вел реальную статистическую работу совместно со специалистами соответствующей прикладной

---

* Близкое содержание вкладывается в термины «распознавание образов без учителя», «автоматическая классификация», «таксономия». Речь идет о методах разбиения совокупности объектов на группы (кластеры) схожих между собой объектов.

** Репрезентативная — связанная с представлениями. Не путать с алгоритмической теорией измерения [32]! Репрезентативная теория измерений, о которой пойдет речь в следующем параграфе, не имеет также прямого отношения и к классической теории ошибок измерения.

области, хорошо известно, что эти задачи весьма важны. Во всяком случае при работе автора с медиками они являются основными. Содержательно осмысленное описание материала позволяет сформулировать гипотезы для статистической проверки и указать параметры, оценить которые желательно. Соотношение затрат времени на описание материала и на остальные виды статистических задач в исследованиях, в которых участвовал автор, составляло примерно 10 : 1.

В учебниках же для математиков задачи описания материала почти полностью игнорируются (исключение составляют формальные методы кластер-анализа, описанные в соответствующих руководствах, см. [8]). Курсы по математической статистике для лиц иных специальностей ненамного лучше в этом отношении.

В настоящее время отсутствует система подготовки специалистов по прикладной математической статистике. Как следствие продукция деятелей этой области страдает разнообразными дефектами. Математики разрабатывают отдельные методы, не желая думать о системе анализа данных в целом. Лица с прикладной подготовкой, не зная математики, допускают разнообразные ошибки, из-за чего их продукция не вызывает доверия у математика, но, увы, широко распространяется среди тех, кто не может самостоятельно обнаружить «ляпы».

Как источник «ляпов» особенно выделю критерий Колмогорова (может быть, потому, что он относится к бывшей узкой специальности автора этой брошюры). Этот критерий позволяет проверить гипотезу о том, что выборка взята из данного теоретического распределения (подробнее см. [11, с. 10—19]). Распределение статистики критерия Колмогорова при большом объеме выборки хорошо известно. Возникает искушение использовать этот критерий в случае, когда теоретическое распределение известно с точностью до параметров (а не полностью). Параметры оценивают по выборке, а затем в критерии Колмогорова в качестве теоретического распределения используют распределение, в котором истинные значения параметров заменены на оцененные по выборке. Такая замена приводит к уменьшению обычно используемых процентных точек примерно в 1,5 раза (ср. таблицы на с. 15 и 18 в [11]), если выборка взята из нормального распределения. Если же пользоваться таблицами статистики Колмогорова в исходной форме, то гипотеза нормальности будет приниматься гораздо чаще, чем следовало бы.

Математикам описанный факт хорошо известен. А вот прикладникам, берущимся писать о математической статистике, увы, не всегда. В учебнике [33] для студентов экономических вузов по курсу «Общая теория статистики» подробно описано неправильное применение критерия Колмогорова, причем авторы не подозревают, что оно неправильное. Бедные студенты! В одной из лучших книг по применению статистических методов в медицине [34] допущена та же ошибка.

От вопросов обучения и проблем контроля качества научной продукции вернемся к оставшимся видам статистических задач.

Оценивание — это и оценивание параметров, одномерных и многомерных, и оценивание функций — линии регрессии, плотности распределения. Эта часть математической статистики подробно изложена в руководствах (см., в частности, [23, 35]). Мы вернемся к оцениванию, когда будем обсуждать экстремальный подход к статистике и обработку данных общей природы.

Проверку гипотез можно рассматривать как часть теории оценивания, когда параметром является номер гипотезы, которая справедлива, и его-то и надо оценить. Однако специальный вид задачи оценивания (параметр дискретен) порождает специфические проблемы и методы, поэтому проверку гипотез рассматривают как самостоятельную часть математической статистики.

В ряде ситуаций необходимо проверить гипотезу о том, что две выборки извлечены из одного и того же распределения. Например, изучается продолжительность заболевания у пациентов, которых лечили лекарством А (первая выборка) и лекарством Б (вторая выборка). Имеется великое множество критериев для проверки этой гипотезы однородности: критерий Стьюдента, хи-квадрат, Колмогорова, Смирнова, омега-квадрат, Вилкоксона, Ван-дер Вардена и т. д. Каким из них пользоваться? С теоретических позиций для каждого из них имеется ситуация, когда он оптимален (по сравнению со всеми другими). Так, критерий Стьюдента оптимален, если наблюдения имеют нормальное распределение, а критерий Вилкоксона* — когда они

---

* Он оптимален среди ранговых критериев и асимптотически оптимален среди всех.

имеют логистическое распределение. При этом хорошие свойства критерия Стьюдента быстро теряются при отклонении от нормальности, а критерий Вилкоксона остается весьма достойным при любом распределении наблюдений.

Вот и решайте, каким критерием пользоваться (автор сейчас внедряет критерий Вилкоксона среди медиков). Может быть, применять все критерии, включенные в пакет статистических программ, как вытекает из общих соображений п. 4? А как тогда трактовать различия выводов, полученных с помощью этих критериев?

Указанная проблема связана с *системностью* использования статистического аппарата. Она порождена необходимостью применения комплекса методов при описании материала, оценивании параметров и проверке гипотез. Между тем существующая теория ориентирована на рассмотрение отдельных методов. Образно выражаясь, их можно использовать как кирпичи при построении здания статистического исследования. Но науки о том, как строить это здание, пока нет. Есть эмпирические рекомендации — дома-то строят!

Аналогию можно продолжить. Где строить дом? Статистическое исследование какого процесса предпринять? Какие предварительные сведения нужны для принятия решения о строительстве? Это — сфера архитектора, т. е. организатора статистического исследования.

Мы выделили три вида деятельности: изготовление кирпичей, строительство дома, архитектуру. Кто же ими занимается? Кирпичи изготовляют математики-профессионалы. Навалена необозримая гора (см. выше список критериев для проверки гипотезы однородности). Разбившись на небольшие группы, математики тщательно полируют кирпичи на своем участке, не интересуясь соседями и потребителями. Дома строят инженеры, продукцией которых являются «методики обработки данных» соответствующей области и их программное обеспечение. У них попросту нет времени, чтобы тщательно разглядывать каждый кирпич— собрать бы дом к сроку. Тем более нет возможности посмотреть, как кирпичи «стыкуются» друг с другом. Внимание сосредоточивается обычно на программном обеспечении. В результате система что-то считает. Но профессиональный математик, познакомившись с ней, хватается за голову: чем обоснованы те или иные шаги? Как правило, оказывается, что никто не изучал соответствующий вопрос с математической точки зрения.

Рассмотрим один пример. Известно, что для интерпретации результатов регрессионного анализа необходима достаточная однородность совокупности наблюдений. (Предыдущая фраза написана не на математическом языке, а на «содержательном».) Поэтому обычно рекомендуют разбить наблюдения на однородные группы и лишь затем в каждой из групп применять регрессионный анализ. Применять же его можно лишь при некоторых условиях. Обычно предполагают независимость и нормальную распределенность наблюдений. Однако в результате применения процедуры разбиения наблюдения в выделенной группе оказываются зависимыми, а их распределение отнюдь не нормально и не приближается к нормальному при увеличении объема выборки. (Последнее утверждение — нестрогая формулировка теоремы [37]). Как же «стыковать» два кирпича—алгоритмы классификации и регрессии? Для ответа на этот вопрос требуется математическая работа. В настоящее время наиболее обоснованным представляется применение λ-регрессии Л. Д. Мешалкина и А. И. Курочкиной [25].

Было бы целесообразно математикам-профессионалам, кроме полировки кирпичей, заняться вопросами строительства. Для этого нужна посредническая служба — кто-то должен выделить нерешенные проблемы прикладной статистики и сформулировать их как математические задачи. Автор попытался это сделать в [36], где приведена, разумеется, лишь небольшая часть нерешенных проблем.

Что же касается деятельности архитектора — организации статистического исследования, то и сейчас и в ближайшем будущем о формализации вряд ли можно серьезно говорить. Организаторами и методологами статистических исследований являются сейчас люди разных специальностей, их деятельность слишком тесно связана с потребностями той сферы деятельности, в которой предпринимается работа.

Мы разобрали виды статистических задач. Перейдем теперь к природе статистических данных. По традиции статистика ассоциируется с обработкой чисел. И по сей день понятие выборки действительных чисел является основным в большинстве статистических публикаций. Однако уже при изучении корреляции (например, работы А. А. Чупрова в начале XX в.) исходные данные — выборка двухмерных векторов, а не чисел. Далее появился многомерный статистический анализ, в самом названии которого подчеркивается, что выборка состоит из векторов, лежащих в многомерном (но конечномерном) пространстве. Независимо развивалась статистическая техника анализа процессов, текущих во времени, — статистика случайных процессов. Поскольку процессы описываются функциями, а функции можно складывать и умножать, то как обобщение появилась статистика в гильбертовом пространстве. Есть еще статистика на группах.

Прикладные потребности последних десятилетий привели к необходимости статистической обработки таких данных нечисловой природы, как бинарные отношения, измерения в шкалах, отличных от абсолютной, множества, в том числе нечеткие. При попытке рассмотреть их с единой точки зрения мы вынуждены строить статистическую теорию для обработки данных произвольной природы, лежащих в пространстве, которое не является линейным (конечномерным или гильбертовым), не имеет структуры группы и вообще не обладает какими-либо специфическими свойствами. Чтобы доказывать теоремы, приходится все-таки кое-что предполагать, например, что входящие в выборку наблюдения являются элементами топологического пространства. (Как уже отмечалось, при чтении этой брошюры можно считать, что наблюдения берутся из конечного множества (п. 5), что оправдывается, помимо прочего, близостью свойств компакта и конечного множества.)

Направление математической статистики, занимающееся перечисленными выше объектами, называем «статистикой объектов нечисловой природы». В ней классические задачи описания материала, оценивания и проверки гипотез рассматриваются для неклассических объектов.

8. Объекты нечисловой природы

Начнем с теории измерений. У автора интерес к ней возник при попытке осмыслить противоречие в экспериментальных данных.

Коллектив новосибирских социологов изучал привлекательность профессий для выпускников новосибирских школ [38]. Был составлен список из 80 профессий. Опрашиваемых просили оценить каждую из этих профессий одним из баллов 1, 2, ..., 10. Чем больше нравится, тем выше балл. В качестве единой оценки привлекательности профессии использовали среднее арифметическое выставленных ей баллов. В частности, физика получила средний балл 7,69, а математика — 7,50. Значит, выпускники предпочитают математике физику.

Однако по данным доктора педагогических наук Г. И. Щукиной [39], ленинградские школьники средних классов больше любят математику, чем физику. Налицо противоречие. Чем оно вызвано? Различным возрастом опрашиваемых? Тем, что пристрастия школьников меняются от города к городу? Или же применяемой методикой сбора и обработки данных? Зависимость научного вывода от методики означает просто-напросто ошибочность этой методики...

Займемся тем, что в нашей власти, — математическими методами обработки данных. Баллы, названные школьниками, — действительно ли это числа? Одна профессия оценена в 10 баллов, а другая — в 2, значит ли это, что первая в 5 раз привлекательнее второй? Вряд ли.

Естественно считать, что школьник действительно может сказать, какая из двух профессий ему больше нравится, или же они для него одинаково привлекательны. Но трудно признать, что он может сказать, во сколько раз одна из них лучше другой.

Приходим к следующей модели поведения опрашиваемого. Имеется десять групп профессий, упорядоченных по степени привлекательности. Каждая из включенных в список профессий относится к одной из десяти групп.

Казалось бы, повторена начальная формулировка. «Небольшое» отличие в том, что в соответствии с этой моделью школьнику все равно, какой именно системой баллов с 10 градациями пользоваться: 1,2, ..., 10, или 1, 4, 9, ..., 100, или —10, —7, —5, —3, —2, —1, 0, 1, 5, 25. А если так, то методы обработки должны обеспечивать инвариантность выводов относительно перехода от одной системы баллов к другой.

Формализуем сказанное в математических терминах на примере единой оценки привлекательности профессии. Пусть $x_1, x_2, \ldots, x_n$ — ответы n школьников, касающиеся математики, а $y_1, y_2, \ldots, y_n$ — физики. Пусть единая оценка вычисляется с помощью функции f, т. е. $f(x_1, x_2, \ldots, x_n)$— оценка математики, $f(y_1, y_2, \ldots, y_n)$ — физики. Пусть оказалось, что

$$f(x_1, x_2, \ldots, x_n) < f(y_1, y_2, \ldots, y_n). \qquad (10)$$

Переход к другой системе баллов опишем с помощью строго возрастающей функции φ: $R^1 \to R^1$, отображающей действительную прямую R^1 в себя. При этом балл z заменяется на $\varphi(z)$. В соответствии с описанной выше моделью поведения опрашиваемого ответы относительно математики будут иметь вид $\varphi(x_1)$, $\varphi(x_2)$, ..., $\varphi(x_n)$, а относительно физики — $\varphi(y_1)$, $\varphi(y_2)$, ..., $\varphi(y_n)$. Способ вычисления единой оценки f можно применять, если из (10) следует, что

$$f(\varphi(x_1),\ \varphi(x_2),\ \ldots,\ \varphi(x_n))<f(\varphi(y_1),\ \varphi(y_2),\ \ldots,\ \varphi(y_n)). \quad (11)$$

Оказывается, что для этого функция f должна иметь вполне определенный вид. Точнее, потребуем, чтобы для любых x_i, y_i, $i=1, 2, \ldots, n$ и любой строго возрастающей функции φ неравенства (10) и (11) были либо одновременно выполнены, либо одновременно не выполнены. Можно сказать немного по-иному: при любых возможных ответах школьников упорядоченность единых оценок не меняется при переходе к (произвольной) новой системе баллов. Пусть выполнены еще два технических требования: функция f непрерывна по совокупности переменных и симметрична, т. е. не меняется при любой перестановке аргументов (последнее означает, что мнения всех школьников учитываются в одинаковой мере). Пусть, наконец, f — средняя величина, т. е.

$$\min_{1\leqslant i\leqslant n} x_i \leqslant f(x_1, x_2,\ldots,x_n) \leqslant \max_{1\leqslant i\leqslant n} x_i.$$

Чтобы описать вид f, вытекающий из только что сформулированных условий, напомним определение порядковых статистик. Числа x_1, x_2, ..., x_n упорядочим в порядке возрастания (точнее, неубывания), получим $x(1)\leqslant x(2)\leqslant\ldots\leqslant x(n)$. Совокупность (упорядоченная) чисел $x(1)$, $x(2)$, ..., $x(n)$ называется вариационным рядом, а его члены — порядковыми статистиками: i-я порядковая статистика —это $x(i)$ (о порядковых статистиках см. [40—41]).

Теорема 10. [9, § 3.3]. Пусть выполнены перечисленные выше условия. Тогда найдется натуральное число m, $1\leqslant m\leqslant n$, такое, что $f(x_1, x_2, \ldots, x_n)$ есть m-я порядковая статистика:

$$f(x_1,\ x_2,\ \ldots,\ x_n)\equiv x(m).$$

Из теоремы 10 вытекает, что брать среднее арифметическое баллов в качестве единой оценки нельзя — при переходе к другой системе баллов упорядоченность средних арифметических может измениться на противоположную (например, (10) будет выполнено, а (11) — нет). А что же можно использовать? Например, медиану x ($[n/2]+1$) — член вариационного ряда с номером $[n/2]+1$, где квадратными скобками обозначена целая часть числа. Поэтому теорема 10 названа в [9] «теоремой о медиане».

Теперь опишем общие постановки репрезентативной теории измерений. Основное понятие в ней — группа допустимых преобразований Φ прямой в себя. Теория измерений— это теория инвариантов относительно групп допустимых преобразований. Функция g называется инвариантом, если

$$g(x_1, x_2, \ldots, x_n)=g(\varphi(x_1), \varphi(x_2), \ldots, \varphi(x_n))$$

для любых чисел x_i, $i=1,2, \ldots, n$, и любого допустимого преобразования $\varphi \in \Phi$.

Говорят, что данные измерены в шкале с группой допустимых преобразований Φ, если с содержательной точки зрения наборы чисел $(x_1, x_2, \ldots, x_n)$ и $(\varphi(x_1), \varphi(x_2), \ldots, \varphi(x_n))$, $\varphi \in \Phi$ несут одну и ту же информацию. Тогда выводы можно делать лишь с помощью инвариантов относительно группы Φ.

Типичные постановки задач в теории измерений таковы. Является ли g инвариантом относительно Φ? По данной Φ найти все инварианты g, удовлетворяющие некоторому условию. Для данного g найти наиболее широкую группу Φ, относительно которой g является инвариантом.

Свяжем теорию измерений с задачей о сравнении средних. Функцию $g: R^{2n} \to R^1$ определим следующим образом:

$$g(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_n) = \begin{cases} 0, & f(x_1, \ldots, x_n) < f(y_1, \ldots, y_n); \\ 1, & f(x_1, \ldots, x_n) \geqslant f(y_1, \ldots, y_n). \end{cases}$$

Пусть Φ_0 — группа всех строго возрастающих преобразований, непрерывных во всех точках (проверьте, что это действительно группа!). Тогда g является инвариантом относительно Φ_0, если и только если (10) и (11) либо одновременно выполнены, либо одновременно не выполнены. Теорема 10 дает описание всех инвариантов относительно Φ_0, построенных по средним f.

Отметим, что теория измерений вкладывается в общую схему устойчивости п. 2. В качестве пространства исходных данных возьмем $X=R^n$, в качестве пространства принимаемых решений—множество Y значений g. Отображение из X в Y задается функцией g. В качестве ρ используется произвольная метрика в Y. С каждой точкой $(x_1, x_2, \ldots, x_n) \in X$ связывается область допустимых колебаний:

$$G(x_1, \ldots, x_n)=\{(x'_1, \ldots, x'_n):(x'_1, \ldots, x'_n)=(\varphi(x_1), \ldots, \varphi(x_n)), \varphi \in \Phi\}.$$

Инвариантность g соответствует устойчивости g в смысле определения 1 из п. 2 (считая $G_r \equiv G$ при всех r).

Шкала с группой Φ_0 всех строго возрастающих непрерывных преобразований прямой в себя называется порядковой. По этой шкале измерены, в частности, мнения школьников о привлекательности профессий. В медицине часто используют понятие «группа тяжести заболевания»; значения группы тяжести также следует считать измеренными в порядковой шкале. То же можно сказать о шкале твердости Мосса, по которой минералы классифицируются согласно критерию твердости, о бофортовой шкале ветров («штиль», «слабый ветер», «умеренный ветер» и т. д.). По мнению автора, в порядковой шкале измерены и трудности освоения «элемента культуры» [42].

Наиболее широкая группа Φ у шкалы наименований, называемой также номинальной или классификационной шкалой. У этой шкалы группа Φ — группа всех взаимно-однозначных отображений прямой на себя. Она соответствует простейшему типу измерения, в котором числа используются лишь как имена объектов. По шкале наименований измерены, скажем, номера телефонов или автомашин.

Если Φ — группа линейных преобразований $\varphi(x)=ax+b$, $a>0$, то измерения проводятся в шкале интервалов. Поскольку, например, температура C, измеренная по шкале Цельсия, выражается через температуру F, измеренную по шкале Фаренгейта, следующим образом: $C=5(F-32)/9$, то температуру естественно считать измеряющейся по шкале интервалов. Термин «шкала интервалов» поясняется следующей теоремой.

Теорема 11. Пусть равенство

$$\frac{\varphi(x_1)-\varphi(x_2)}{\varphi(x_3)-\varphi(x_4)}=\frac{x_1-x_2}{x_3-x_4} \tag{12}$$

справедливо при всех x_1, x_2, x_3, x_4, таких, что $x_3 \neq x_4$. Тогда $\varphi(x)$ — линейное преобразование. Обратно, равенство (12) справедливо для всех линейных преобразований. Значит, отношение интервалов — инвариант в шкале интервалов.

Измерение многих величин — длины, веса, стоимости — производится лишь с точностью до выбора единицы измерения. Допустимыми преобразованиями шкалы являются $\varphi(x)=ax$, $a>0$; это означает, что смена единицы измерения приводит к умножению всех результатов измерения на некоторое положительное число и, наоборот, для любого $a>0$ можно подобрать такую новую единицу измерения, что при переходе к ней все результаты измерения умножаются на a. Шкала с группой допустимых преобразований $\Phi=\{ax, a>0\}$ называется шкалой отношений. Название объясняется тем, что отношение x_1/x_2 является инвариантом в этой шкале.

Видимо, в настоящее время шкала отношений является наиболее распространенной, во всяком случае гораздо более распространенной, чем абсолютная шкала, в которой группа допустимых преобразований состоит из одного-единственного преобразования — тождественного. Абсолютная шкала применяется, когда имеется естественная единица измерения и естественное начало отсчета, например при подсчете числа людей.

В ряде задач полезны шкалы с иными группами Φ [9, 29, 43]. Мы перечислили лишь наиболее распространенные. Удается указывать инвариантные расстояния, меры близости, показатели связи [9, гл. 3] относительно соответствующих шкал. В стиле теоремы 10 удается находить инвариантные средние и в иных шкалах, чем порядковая, но из более узкого класса средних.

Естественная система аксиом приводит к так называемым ассоциативным средним (формулировки см. в [44]). Из работ А. Н. Колмогорова, М. Нагумо и Б. де Финетти начала тридцатых годов следует, что общий вид ассоциативных средних таков:

$$f_F(x_1,\ldots,x_n)=F^{-1}\left(\frac{1}{n}\sum_{1\leqslant i\leqslant n}F(x_i)\right), \tag{13}$$

где $F(z)$ — строго монотонная функция (т. е. строго возрастающая или строго убывающая), $F^{-1}(z)$ — обратная к ней. При $F(z)$, равной z, $\ln z$, z^{-1}, z^2, формула (13) определяет среднее арифметическое, среднее геометрическое, среднее гармоническое, среднее квадратическое соответственно, т. е. наиболее часто используемые средние.

Потребуем, чтобы неравенства (10) и (11) были либо одновременно выполнены, либо одновременно не выполнены при всех x_i, y_i, $i=1, \ldots, n$ и $\varphi\in\Phi$, где $f=f_F$ из (13). Если данные измерены в шкале интервалов, то этому условию удовлетворяет только среднее арифметическое. Если же они измерены в шкале отношений, то только степенные средние (получаются при $F(z)=z^c$, $c\neq0$ и $F(z)=\ln z$). Сформулированные результаты верны при некоторых предположениях регулярности [9, §3.4). Об их применении рассказано в [9, гл. 3].

Остановимся несколько подробнее на репрезентативной теории измерений. Эта теория прошла долгий путь. Вначале она развивалась как часть математических методов психологии [43]. В качестве основных понятий в результате долгого развития стали рассматривать «эмпирическую систему с отношениями», «числовую систему с отношениями» и гомоморфизмы из первой во вторую. Для используемых в приложениях шкал совокупность гомоморфизмов удавалось описать как $\{\varphi \bigcirc h, \varphi \in \Phi\}$, где h — один из гомоморфизмов, Φ — группа преобразований, называемых допустимыми, прямой в себя. Математическая техника этого описания иногда достаточно сложна [29, 43].

Советские исследователи (см. обзор в [9]) в течение семидесятых годов стали применять теорию измерений к социологическим, экономическим, медицинским данным, данным экспертных оценок и контроля качества, короче говоря, столь же широко, как и прочие статистические методы.

Постепенно выяснилось, что в новой, более широкой области приложений указать «эмпирическую систему с отношениями» отнюдь не легче, чем непосредственно привести группу допустимых преобразований. Поэтому в [9] за основу взято именно последнее понятие. Это позволяет охватить обычно используемые шкалы, избежав трудностей, продемонстрированных в [29]. На мой взгляд, аксиоматизация понятия шкалы, используемая в [29, 43], неудачна, поскольку излишне сложна, т. е. приводит к необходимости трудоемких математических исследований. Возможно, она отражает потребности психологов, но нет оснований сохранять ее при переходе на другие виды данных. Поэтому в ее изложении в [29] чувствуется налет схоластики, а само классическое определение шкалы имеет, во всяком случае вне психологии, лишь исторический интерес.

Опишем инварианты в шкале наименований и порядковой шкале.

Теорема 12. Пусть g — инвариант в шкале наименований. Тогда $g(x_1, x_2, \ldots, x_n)=g(B)$, где B — матрица порядка $n \times n$ с элементами

$$b_{ij} = \begin{cases} 1, & x_i = x_j, \\ 0, & x_i \neq x_j. \end{cases}$$

Теорема 13. Пусть g — инвариант в шкале порядка. Тогда $g(x_1, x_2, \ldots, x_n)=g(C)$, где C — матрица порядка $n \times n$ с элементами

$$c_{ij}=\begin{cases}1, & x_i\leqslant x_j,\\ 0, & x_i>x_j.\end{cases}$$

Всмотримся в заключения теорем 12 и 13. Задать матрицу B — это значит задать разбиение совокупности $\{x_i\}$ на группы, в каждую из которых входят совпадающие элементы, т. е. разбиение на классы эквивалентности. С помощью матрицы C также задается разбиение на классы эквивалентности (x_i и x_j входят в один класс тогда и только тогда, когда $c_{ij}=1$ и $c_{ji}=1$), и, кроме того, между этими классами вводится отношение порядка (строгого).

Мы перешли к следующему виду объектов нечисловой природы — бинарным отношениям (популярное введение в их теорию дано в [45]). Наиболее интересны для приложений описанные выше отношения, порожденные разбиениями, между классами которых в одном случае вводится «совершенный строгий порядок» [45, с. 120], а в другом — не вводится. Если каждый из классов, между которыми вводится этот порядок, состоит из одного элемента, то получаем объект, который в статистике принято называть ранжировкой. Если классы могут состоять более чем из одного элемента, то говорят о ранжировке со связями. Для описания сходства, помимо разбиений, представляют интерес толерантности (рефлексивные симметричные отношения). В отличие от разбиений, транзитивность не обязательно выполнена для толерантностей...

Автор поймал себя на желании переписать ряд определений и фактов о бинарных отношениях из различных учебников. Если писать толстую монографию, то так стоило бы сделать для удобства читателя. Но в краткой брошюре остается отослать к [45] и соответствующим разделам [9—13]. Нам понадобится, однако, следующее определение.

Определение 7. Бинарным отношением на конечном множестве X называется (произвольное) подмножество его декартова квадрата X^2.

Если на это подмножество наложить соответствующие условия, то получаются отношения описанных выше типов. Так, бинарное отношение $Q\subseteq X^2$ является эквивалентностью, если $(x,x)\in Q$ при всех $x\in X$ (рефлексивность); из $(x, y)\in Q$ вытекает $(y, x)\in Q$ (симметричность); из $(x, y)\in Q$ и $(y, z)\in Q$ вытекает $(x, z)\in Q$ (транзитивность) при любых x, y, z из X.

Каждое бинарное отношение может быть описано матрицей типов B и C, состоящей из элементов 0 и 1. Пусть для определенности X состоит из k элементов x_1, x_2, ..., x_k. Бинарное отношение Q полностью описывается матрицей $||q_{ij}||$ порядка $k \times k$, где $q_{ij}=1$, если $(x_i, \ x_j) \in Q$, и $q_{ij}=0$ в противном случае.

Бинарные отношения сведены к множествам, которые сами по себе представляют важный вид статистических данных [9, гл. 4]. Естественно, возникает мысль определить нечеткие отношения как нечеткие подмножества декартова квадрата исходного универсального множества. Эта мысль плодотворна (см. [15, 19, 21], статьи В. Б. Кузьмина и С. В. Овчинникова в [12, 13]).

Нечеткие множества, которым была посвящена первая глава, также представляют собой объекты нечисловой природы. Какие же статистические методы применимы к нечисловым данным*?

9. Оптимизационный подход в статистике

Чтобы описать статистический материал, первым делом вводят «показатели центральной тенденции» — математическое ожидание, медиану. Это — тогда, когда обрабатывают числа. А как описать центр эмпирического распределения, скажем, для совокупности множеств?

Обратим внимание на то, что математическое ожидание и медиана, равно как и их эмпирические аналоги (среднее арифметическое наблюдений и выборочная медиана), являются решениями задач оптимизации.

Рассмотрим сначала действительную случайную величину ξ. Будем использовать обозначения

$$M(f, X) = M(f(x), x \in X) = \operatorname*{Arg\,min}_{x \in X} f(x) \tag{14}$$

для совокупности точек, в которых достигается глобальный минимум функции $f(x)$ на множестве X. Таким образом, если $x_0 \in M(f, X)$, то $f(x) > f(x_0)$ для всех $x \in X/M(f, X)$. (Обозначение A/B означает разность множеств A и B — совокупность всех элементов множества A, не входящих в B.) Конечно, в каждом конкретном случае следует доказывать непустоту $M(f, X)$.

---

* Под объектами нечисловой природы мы понимаем те, которые нецелесообразно рассматривать как определяемые числами. Разумеется, в п. 8 перечислены не все подобные объекты.

Теорема 14. Пусть * $f_0(x)=E(\xi-x)^2$. Тогда

$$M(f_0, R^1)=\{E\xi\}.$$

Доказательство. Справедливо равенство

$$(\xi-x)^2=((\xi-E\xi)+(E\xi-x))^2=(\xi-E\xi)^2+2(E\xi-x)(\xi-E\xi)+(E\xi-x)^2.$$

Поскольку математическое ожидание суммы случайных величин равно сумме их математических ожиданий и постоянный множитель можно выносить за знак математического ожидания, то

$$f_0(x)=D\xi+(E\xi-x)^2,$$

откуда и следует требуемое.

Теорема 15. Для произвольных чисел $x_1, x_2, \ldots, x_n$ положим

$$f_n(x)=\sum_{1\leqslant i\leqslant n}(x_i-x)^2.$$

Тогда

$$M(f_n, R^1)=\{\bar{x}\}=\left\{\frac{x_1+x_2+\ldots+x_n}{n}\right\}.$$

Обратим внимание на то, что $M(f, X)$ — множество. В теоремах 14 и 15 каждое из этих множеств состоит из одной точки. В общем случае это, конечно, не так.

Определение 8. Медианой случайной величины ξ называется множество

$$\mathrm{Med}\,\xi=\{x : F(x)\leqslant{}^1/_2,\ F(x+)\geqslant{}^1/_2\}, \tag{15}$$

где $F(x)=P(\xi<x)$.

В (15) через $F(x+)$ обозначен предел справа функции распределения $F(x)$. Множество Med ξ состоит либо из одной точки, либо из отрезка; последнее соответствует существованию отрезка, для всех точек которого $F(x)={}^1/_2$.

Теорема 16. Для произвольных чисел $x_1, x_2, \ldots, x_n$ положим

$$g_n(x)=\sum_{1\leqslant i\leqslant n}|x_i-x|\cdot \tag{16}$$

---

* Для обозначения математического ожидания случайной величины ξ мы используем символ $E\xi$ (от expectation — ожидание), а не $M\xi$ (от mean — среднее), поскольку существует много видов средних (см. (13)). Дисперсию ξ обозначаем $D\xi$. В формулировках теорем существование используемых объектов специально не оговаривается.

Если n нечетно, то

$$M(g_n,\ R^1)=\{x([n/2]+1)\},$$

если же n четно, то

$$M(g_n,\ R^1)=[x([n/2]),\ x([n/2]+1)],$$

где $x(i)$ — i-я порядковая статистика совокупности $\{x_1,\ x_2,\ \ldots,\ x_n\}$.

Теорема 17. Пусть $g(x)=E\,|\xi-x|$. Тогда $M(g,R^1)=\operatorname{Med}\xi$.

В соответствии с определением 8 эмпирической медианой естественно называть $M(g_n,R^1)$. Если n четно, то это — отрезок. При обсуждении теории измерений мы называли медианой правый конец этого отрезка. В учебниках по общей теории статистики [33] медиана — середина отрезка. Таким образом, имеются три определения медианы выборки, за каждым из которых стоит некоторое теоретическое представление.

Важное значение в методологии статистики имеет закон больших чисел. Приведем его строгую формулировку.

Теорема 18. (А. Я. Хинчин, [46, с. 208]). Если случайные величины $\xi_1,\ \xi_2,\ \ldots$ независимы, одинаково распределены и имеют конечные математические ожидания $(a=E\,\xi_n)$, то при $n\to\infty$ для любого фиксированного $\varepsilon>0$

$$P\left\{\left|\frac{1}{n}\sum_{1\leqslant k\leqslant n}\xi_k-a\right|<\varepsilon\right\}\longrightarrow 1. \qquad (17)$$

Обозначая среднее арифметическое наблюдений $\bar{\xi}$ (в стиле теоремы 15), вместо (17) пишут: $\bar{\xi}\to E\,\xi_1$ (по вероятности). Говорят, что последовательность случайных величин η_n, $n=1,\ 2,\ \ldots$, сходится к константе a по вероятности, если для любого $\varepsilon>0$ при $n\to\infty$

$$P\{|\eta_n-a|<\varepsilon\}\to 1. \qquad (18)$$

В силу теорем 14 и 15 закон больших чисел можно переписать по-иному: при $n\to\infty$ по вероятности

$$M(f_n,\ R^1)\to M(f_0,\ R^1). \qquad (19)$$

Хотелось бы аналогичным образом сформулировать и доказать сходимость эмпирической медианы к теоретической:

$$M(g_n,\ R^1)\to M(g,\ R^1). \qquad (20)$$

Однако в (19) и (20) используются не числа, а множества! В (19) это не очень мешает: множества состоят каждое из одной точки. А в (20) каждое из участвующих множеств может быть отрезком.

Чтобы придать смысл формулам (19) и (20), переформулируем закон больших чисел. Прежде всего заметим, что неравенство, стоящее в (17) под знаком вероятности, эквивалентно включению

$$\bar{\xi}\in(a-\varepsilon,\ a+\varepsilon). \tag{21}$$

Вспомнив теорему 15, перепишем (21) в виде (f_n определяется по ξ_i, а не по x_i)

$$M(f_n,\ R^1)\subseteq(a-\varepsilon,\ a+\varepsilon). \tag{22}$$

Займемся правой частью (22). Наряду с (14) будем использовать обозначение

$$M_\varepsilon(f,\ X)=\left\{x\in X : f(x)<\inf_{y\in X} f(y)+\varepsilon\right\}. \tag{23}$$

Из доказательства теоремы 14 вытекает, что

$$(a-\varepsilon,\ a+\varepsilon)=M_{\varepsilon^2}(f_0,\ R^1). \tag{24}$$

Теперь переформулируем закон больших чисел, применяя замены (22) и (24).

Теорема 19. Если случайные величины ξ_1, ξ_2, ... независимы, одинаково распределены и имеют конечные дисперсии, то для любого фиксированного $\varepsilon>0$

$$\lim_{n\to\infty} P\{M(f_n,\ R^1)\subseteq M_\varepsilon(f_0,\ R^1)\}=1. \tag{25}$$

Теорема 19 немедленно обобщается, в частности, остается верной при замене f_0 на g, а f_n на g_n (тогда она означает сходимость эмпирической медианы к теоретической). Но за общность приходится платить — кроме математических ожиданий, как в теореме 18, случайные величины должны иметь дисперсии. Впрочем, с точки зрения прикладника эта плата ничтожна.

Оптимизационный подход позволяет определять понятия математического ожидания и эмпирического среднего в пространствах общей природы. Но что такое, скажем, математическое ожидание для случайной ранжировки?

Определение 9. Пусть ξ — случайная величина со значениями в пространстве X, а ρ — некоторая функция на X^2. Математическим ожиданием ξ относительно ρ назовем $M(E\rho(\xi,\ x),\ x\in X)$.

Определение 10. Пусть x_1, x_2, ..., x_n — произвольные элементы X, а ρ — некоторая функция на X^2.

Эмпирическим средним относительно ρ назовем $M(h_n, X)$, где

$$h_n(x) = \sum_{1 \leqslant i \leqslant n} \rho(x_i, x). \tag{26}$$

Если X — конечно, то решения оптимизационных задач, упомянутых в определениях 9 и 10, всегда существуют. Общий случай рассмотрен в [47, 48].

Для пространств произвольной природы справедлив закон больших чисел (аналог теоремы 19).

Теорема 20. Пусть случайные величины ξ_1, ξ_2, ..., со значениями в пространстве X независимы и одинаково распределены. Пусть ρ — некоторая действительная функция на X^2. Пусть выполнены некоторые условия регулярности, наложенные на X, ρ и распределение ξ_1 [47, 48]. Тогда для любого фиксированного $\varepsilon > 0$

$$\lim_{n \to \infty} P\{M(h_n, X) \subseteq M_\varepsilon(h, X)\} = 1, \tag{27}$$

где h_n определено в (26), а $h(x) = E\rho(\xi_1, x)$.

Справедлив также и усиленный закон больших чисел, в котором рассматривается сходимость с вероятностью 1, а не по вероятности, как в теореме 20. Доказательство теоремы 20 для конечного X фактически содержится в [9, §4.4].

Зачем понадобились переформулировки классического закона больших чисел и доказательство его обобщения для произвольного пространства? Исходным импульсом служило желание разобраться в способах обработки ранжировок, полученных от экспертов. Один из распространенных способов нахождения группового мнения экспертов — вычисление так называемой медианы Кемени. (О его месте в общей теории анализа нечисловой экспертной информации см. обзорный доклад нашего «незримого коллектива» [49].)

Определение 11. Пусть X — пространство бинарных отношений определенного вида (ранжировок, разбиений и т. д.) на конечном множестве. Медианой Кемени бинарных отношений x_1, x_2, ..., x_n называется их эмпирическое среднее относительно расстояния Кемени—Снелла.

Появилось новое понятие «расстояние Кемени — Снелла» между бинарными отношениями. Напомним (см. п. 8), что бинарное отношение на множестве из k элементов взаимно-однозначно описывается матрицей порядка $k \times k$, состоящей из элементов 0 и 1.

Определение 12. Пусть Q_1 и Q_2—бинарные отношения на множестве из k элементов, описываемые матрицами $||s_{ij}||$ и $||t_{ij}||$ соответственно. Расстоянием Кемени—Снелла между Q_1 и Q_2 называется

$$d(Q_1, Q_2) = \sum_{1 \leqslant i,j \leqslant k} |s_{ij} - t_{ij}|.$$

Название расстояния d объясняется тем, что для пространства ранжировок оно выведено из некоторой системы аксиом в книге Кемени и Снелла [50], вышедшей в США в начале шестидесятых годов. Далее аналогичный вывод был осуществлен для бинарных отношений других типов (обзор дан в [9, §4.3]). Большую роль в пропаганде аксиоматического подхода к введению расстояния между бинарными отношениями сыграла группа новосибирцев во главе с Б. Г. Миркиным [51, 52]. Аналогичный подход может быть применен и для множеств. Оказалось, что системе аксиом типа Кемени — Снелла удовлетворяют те и только те расстояния между множествами, что представляются в виде меры симметрической разности [9, §4.3].

В соответствии с теоремой 20 медиана Кемени при $n \to \infty$ сходится к математическому ожиданию случайной ранжировки относительно расстояния Кемени—Снелла, более того, начиная с некоторого (случайного) n, совпадает с ним. Если же распределение монотонно относительно некоторого центра x_0 (это означает, что при удалении от x_0 вероятность уменьшается [9, §4.6]), то x_0 и является этим математическим ожиданием. Приняв, что распределение ответа эксперта является монотонным относительно «истины», приходим к тому, что медиана Кемени сходится к «истине». Вот оно — вероятностное обоснование целесообразности применения медианы Кемени при обработке экспертных данных.

Читатель проделал в обратном порядке путь автора — от необходимости установить асимптотическое поведение медианы Кемени к законам больших чисел в пространствах общей природы и соответствующей переформулировке обычных законов больших чисел. Точка? Нет, продолжение.

Автор десять лет занимался пропагандой математики среди школьников [53], а в качестве теоретической основы использовал «эвристический подход» Д. Пойа [54—56], в котором рекомендуется постоянно задавать себе вопрос: «Все ли данные Вами использованы?» [54, с. 63]. Такой же вопрос задал я себе и по поводу теоремы 20 (увы, не сразу). Что же оказалось? В доказательстве фактически не использовалось то, что оба аргумента функции ρ лежат в одном и том же пространстве. Так откажемся от этого лишнего требования. Получилась следующая теорема [48].

Теорема 21. Пусть случайные величины ξ_1, ξ_2, ... со значениями в пространстве X независимы и одинаково распределены, пусть Y — некоторое пространство, ρ — действительная функция на $X \times Y$. Пусть выполнены некоторые условия регулярности, наложенные на X, Y, ρ и распределение ξ_1. Тогда для любого фиксированного $\varepsilon > 0$

$$\lim_{n \to \infty} P\{M(h_n, Y) \subseteq M_\varepsilon(h, Y)\} = 1, \tag{28}$$

где

$$h_n(y) = \sum_{1 \leqslant i \leqslant n} \rho(\xi_i, y),\ h(y) = E\rho(\xi_1, y),\ y \in Y. \tag{29}$$

Теорема получена, «эвристический подход» Д. Пойа оправдал себя. А зачем она нужна? Оказывается, она объединяет в себе довольно много — состоятельность оценок максимального правдоподобия и минимального контраста, в частности, устойчивых* оценок Хубера-Тьюки (см. п. 4), оценки параметров регрессии, полученные методами наименьших квадратов и наименьших модулей [57], и, конечно, законы больших чисел, рассмотренные в теореме 20.

Действительно, вспомним, что оценкой максимального правдоподобия параметра $\delta \in \Delta$, построенной по независимым одинаково распределенным наблюдениям ξ_1, ξ_2, ..., $\xi_n \in X$, извлеченным из совокупности с плотностью, принадлежащей параметрическому семейству $\{l(x, \delta),\ x \in X,\ \delta \in \Delta\}$, называется решение задачи

$$\prod_{1 \leqslant i \leqslant n} l(\xi_i, \delta) \longrightarrow \max_{\delta \in \Delta}. \tag{30}$$

Если положить $Y = \Delta$, $\rho(\xi, y) = -\ln l(\xi, y)$, то от (30) переходим к (29). Чтобы констатировать состоятельность оценки максимального правдоподобия, т. е. то, что она стремится по вероятности к «истинному» значению параметра, достаточно показать, что величина $E_{\delta(0)} \ln l(\xi_1, \delta)$, где математическое ожидание берется по распределению при $\delta = \delta(0)$, достигает максимума при $\delta = \delta(0)$. Предоставляем это читателю.

---

* В оригинале — «robust»; Ю. В. Линник предлагал переводить это слово как «крепкий». Используемый иногда термин «робастный» не соответствует духу русского языка.

При $Y=\Delta$ оценка $\delta_n \in M(h_n, \Delta)$ «истинного» значения параметра называется оценкой минимального контраста. Введенные в (14) обозначения позволяют не размышлять об единственности подобной оценки.

Задачи регрессии рассмотрим в наиболее простой постановке (общий случай разобран в статье автора в [12]). Пусть ξ_i двухмерный вектор, $\xi_i=(\xi_i^1, \xi_i^2)$. Желательно вторую координату ξ_i^2 приблизить наилучшим образом с помощью линейной функции от первой, т. е. с помощью $a\xi_i^1+b$, подобрав оптимальные a и b. Положим $X=Y=R^2$, $\rho(\xi_i, (a, b))=\varphi(|\xi_i^2-a\xi_i^1-b|)$, где φ — неубывающая функция на $[0, +\infty)$. При $n\to\infty$ с помощью теоремы 21 получаем асимптотику решений соответствующих задач оптимизации. Если $\varphi(x)=x^2$, то речь идет о методе наименьших квадратов [23, гл. 37], если же $\varphi(x)=x$, то о методе наименьших модулей. Видно также, что полученные с помощью этих методов линии регрессии отнюдь не всегда совпадают, даже асимптотически (а именно, асимптотическое совпадение имеет место тогда и только тогда, когда совпадает условное математическое ожидание и условная медиана ξ_i^2 при фиксированном ξ_i^1).

Задачи оптимизации играют важнейшую роль и в других областях прикладной статистики. Их обзор дан С. А. Айвазяном [5]. А автору, увы, остается воскликнуть: «Никто не обнимет необъятного!» [17]. Сошлюсь только еще раз на [49].

Вспомним совет Д. Пойа и взглянем на теорему 21. От каких лишних условий осталось избавиться? В (28) предполагается, что задача оптимизации имеет решение (иначе (28) тривиально). А нужно ли это предположение? Нет!

Теорема 22. Теорема 21 остается справедливой при замене (28) на следующее утверждение: для любых $\varepsilon>0$ и $C>1$

$$\lim_{n\to\infty} P\{M_\varepsilon(h_n, Y)\subseteq M_{C\varepsilon}(h, Y)\}=1.$$

Не будем утомлять читателя дальнейшими обобщениями (см. [48]). Воздержимся также и от обсуждения теории экстремальных задач [58]. Отметим только, что для решения задач оптимизации оказывается весьма полезным метод случайного поиска [59]. Таким образом, связь между оптимизацией и статистикой оказывается полезной для обеих сторон.

Обратите внимание: до сих пор ни слова не сказано о доверительных вероятностях, в которых Ю. И. Алимов [60] считает нужным видеть суть математической статистики. Это понятие появляется в более продвинутых вопросах, чем разбирались. И, увы, соответствующие результаты в статистике объектов нечисловой природы зачастую еще не получены. Так, например, мы не умеем указывать доверительную область для математического ожидания случайной ранжировки относительно расстояния Кемени — Снелла, если используем медиану Кемени в качестве его оценки.

В статистическом оценивании имеется общий метод, использующий задачи оптимизации, а в проверке гипотез общего метода нет (пока?). Можно применять статистики интегрального типа (см. статью автора в [12]); ряд задач проверки гипотез в конкретных пространствах рассматривается в [9—13] и ниже в п. 12.

При рассмотрении статистических задач мы исходим из аксиоматики теории вероятностей А. Н. Колмогорова [26]. В [60] сделана попытка разбирать эти задачи, опираясь на частотный подход Р. Мизеса [61—64]. По моему мнению, из-за стремления приблизиться к реальности подобный метод приводит к чрезмерному усложнению аппарата и рассуждений (сравните, например, [2] и [60]).

Глава 3. СТАТИСТИКА НЕЧЕТКИХ МНОЖЕСТВ

10. Сбор и описание нечетких данных

В этой главе мы рассматриваем нечеткие множества как один из видов объектов нечисловой природы и применяем к их анализу методы математической статистики, прежде всего оптимизационный подход.

Выделим два способа получения нечетких данных — непосредственный и как результат обработки четких экспертных мнений или, более общо, четких данных. Второй из них рекомендовал применять Борель для определения нечеткого множества зерен, соответствующего понятию «куча» (см. п. 1, 3). С общей точки зрения этот способ не требует специального обсуждения — речь идет об оценке параметра — функции принадлежности нечеткого множества, для чего могут быть использованы стандартные методы, в частности, основанные на оптимизационном подходе.

Непосредственное получение нечетких данных требует разработки специальных методик (см., например, статью П. Б. Шошина в [10]). Иногда задачу экспертам облегчают, предлагая им выдать в качестве ответа «ступеньку», т. е. нечеткое множество, функция принадлежности которого равна 0 вне некоторого отрезка и 1 на нем. «Ступенька»— промежуточный этап между «классикой» (когда ответ — одно число) и нечетким множеством произвольного вида. Вместе с тем постановка задачи, целью которой является выдача интервала возможных значений, не вызывает затруднений для понимания.

Существует и иная методика — когда значения функции принадлежности сами нечетки, а именно представляют собой описанные выше «ступеньки». Ее использовал профессор Понсар (см. п. 4). Соответствующий аппарат описан в [19, с. 52—57].

В целом можно констатировать, что в настоящее время имеется ряд способов получения нечетких данных и нарочитая условность примеров в публикациях Л. А. Заде должна рассматриваться лишь как педагогический прием, обоснованность которого мы не будем здесь обсуждать.

Для описания данных — нахождения эмпирического среднего, классификации наблюдений — весьма полезно использование расстояния между нечеткими множествами. Как и в случае бинарных отношений, расстояния можно вводить аксиоматически. Так, в заметке И. З. Батыршина в [7] показано, что все удовлетворяющие естественной системе аксиом расстояния между нечеткими подмножествами универсального множества $U=\{u_1, u_2, \ldots, U_k\}$ из k элементов имеет вид:

$$\rho(A, B) = \sum_{1 \leqslant j \leqslant k} t_j(\mu_A(u_j), \mu_B(u_j)),$$

где $t_j(x, y)$ — расстояние на $[0,1]$, $j=1, \ldots, k$. Наиболее простой вид $\rho(A, B)$ имеет при $t_j(x, y) = |x-y|$, а именно

$$\rho_1(A, B) = \sum_{1 \leqslant j \leqslant k} |\mu_A(u_j) - \mu_B(u_j)|.$$

Для расстояния ρ_1 легко записывается эмпирическое среднее (относительно ρ_1) для совокупности n нечетких множеств (см. определение 10). Достаточно при каждом u_j взять указанную в теореме 16 эмпирическую медиану значений функций принадлежности рассматриваемых нечетких множеств для этого u_j.

Если же рассматривать ρ, а не ρ_1, то задача распадается на подзадачи нахождения эмпирических средних относительно t_j значений функций принадлежности в точке u_j, но ответ столь просто не записывается. Кроме того, может оказаться целесообразным искать минимум не на всем пространстве нечетких множеств, а лишь на некоторой его части, например, среди тех, у которых функция принадлежности имеет единственный максимум (что естественно, если все наблюдения таковы); при этом еще труднее явно указать ответ.

Обсудим несколько подробнее использование расстояния между нечеткими множествами (или лишь чуть более общо — меры близости*) для классификации наблюдений. Обычно пишут [8], что речь идет о классификации конечномерных векторов. Однако ряд алгоритмов использует лишь попарные расстояния (или меры близости). Ясно, что их можно применять для классификации объектов произвольной природы, лишь бы имелись нужные расстояния. В качестве примера рассмотрим алгоритм кластер-анализа «Дендрограмма», известный также под названием «агломеративный иерархический алгоритм средней связи» [8, с. 102]. На первом шагу алгоритма каждое наблюдение рассматривается как отдельный кластер. Далее на каждом шагу происходит объединение двух самых близких кластеров. Выдача ЭВМ напоминает перевернутое дерево, каждая ветвь которого соответствует кластеру, появляющемуся на каком-либо шаге работы алгоритма; слияние ветвей соответствует объединению кластеров, а ствол — заключительному шагу, когда все наблюдения оказываются объединенными в один кластер. По форме выдачи результата алгоритм и получил свое название. Для его работы необходимо определить расстояние между кластерами. Естественно использовать обобщенные средние по Колмогорову (см. (13)) всевозможных попарных расстояний между элементами рассматриваемых кластеров. Расстояние между кластерами C и D, состоящими из n_1 и n_2 элементов, соответственно, определяется по формуле

$$\tau(C, D) = F^{-1}\left(\frac{1}{n_1 n_2} \sum_{i \in C} \sum_{j \in D} F(\rho(X_i, X_j))\right),$$

---

* Понятие меры близости отличается от понятия расстояния (другими словами, метрики) только тем, что аксиома треугольника не предполагается обязательно справедливой.

где ρ — рассмотренное выше расстояние между нечеткими множествами, F — строго монотонная функция. Соображения теории измерений позволяют ограничить круг возможных F: годятся только степенные средние, т. е. $F(z)=z^\alpha$ при $\alpha\neq0$ или $F(z)=\ln z$ [9, § 4.5]. Чтобы получить разбиение, надо «разрезать» дерево на определенной высоте, т. е. объединять кластеры лишь до тех пор, пока расстояние между ними не превосходит заранее выбранной константы.

Алгоритм кластер-анализа «Дендрограмма», пригодный и для классификации нечетких данных, полностью описан (по поводу «реальности» получающихся при этом кластеров см. статью автора в [65]).

11. Оптимизационные задачи с нечеткими переменными

Конечно, множество таких задач возникает в различных приложениях. Мы остановимся лишь на тех, что появляются в статистике нечетких данных, как обычно, не рассматривая способов их решения.

Продолжим обсуждение алгоритмов классификации. Некоторые из них включают в себя решение задачи оптимизации на каждом шагу. Для примера опишем аналог хорошо известной «Форели» [8], предложенной Н. Г. Загоруйко и его коллективом много лет назад. На пространстве, в котором лежат наблюдения, заданы две меры близости — ρ и τ. Берется одно из наблюдений и вокруг него описывается шар радиуса R, определяемый мерой близости ρ (напомним, что шаром с центром в x относительно ρ называется множество всех элементов y рассматриваемого пространства, таких, что $\rho(x, y)\leqslant R$). Берутся наблюдения, попавшие в этот шар, и находится их эмпирическое среднее относительно τ. Оно берется за новый центр, вокруг которого снова описывается шар радиуса R относительно ρ, и процедура повторяется. (Чтобы алгоритм был полностью определен, необходимо сформулировать правило выбора элемента эмпирического среднего в качестве нового центра, если эмпирическое среднее состоит более чем из одного элемента.)

Когда центр шара зафиксируется, попавшие в этот шар элементы объявляются первым кластером и исключаются из дальнейшего рассмотрения. Алгоритм применяется к совокупности оставшихся наблюдений, выделяет из нее второй кластер и т. д. В исходной «Форели» классифицировались точки конечномерного пространства, в качестве мер близости ρ и τ использовались евклидово расстояние и его квадрат соответственно*.

Всегда ли центр тяжести «остановится»? При реальных расчетах в течение многих лет так было всегда. Но доказательство было получено лишь в 1977 г. — для исходной «Форели» П. М. Блехером и М. Я. Кельбертом [7, с. 358—361], а для описанного выше ее обобщения — автором [7, с. 361—364]. Дальнейшее изучение провел десятиклассник московской школы № 2 Миша Севрюк. Он установил, что остановка происходит не из-за конечности совокупности классифицируемых наблюдений, а из-за того, что в каждом шаре единичного радиуса число наблюдений не превышает заранее фиксированного числа. Через это число удается оценить число шагов до остановки. К сожалению, доказательство Севрюка не удается пока перенести на общий случай, оно проведено лишь для классической «Форели» (см. его заметку в [66]).

Любопытно, что только что рассмотренный алгоритм классификации тесно связан с упомянутыми выше в п. 4, 7 λ-моментами Л. Д. Мешалкина [67].

Перейдем к задачам оценивания, решаемым с помощью оптимизационного подхода п. 9. Основной объект здесь— последовательность независимых одинаково распределенных наблюдений ξ_1, ξ_2, ..., каждое из которых представляет собой случайную величину в пространстве нечетких множеств. Таким образом, предполагается, что реальные данные являются реализациями «случайного нечеткого множества».

Сейчас следовало бы подробно объяснить смысл этого понятия, способы задания распределения вероятностей, связь и отличие его от понятия случайного множества (см. определения в п. 5), соотнесение с одноименным понятием, используемым проф. Р. Фероном (Лион, Франция), и т. д. Но, увы, читателю здесь предлагается популярная брошюра, а не толстая монография. Третья глава в целом представляет собой описание синтеза идей, рассмотренных в первых двух главах, происходит как бы умножение трудностей восприятия материала этих глав, а потому описание результатов синтеза приходится делать «с птичьего полета».

* Тогда эмпирическое среднее есть центр тяжести (теорема 15)

Отметим лишь несколько моментов. Закон больших чисел (теорема 20) позволяет установить асимптотику эмпирических средних совокупности нечетких множеств. Поскольку расстояния, используемые в предыдущем параграфе, представляются в виде суммы по элементам универсального множества (универсума), то математическое ожидание случайного нечеткого множества есть нечеткое множество, функция принадлежности которого для каждого элемента универсума есть математическое ожидание (в классическом смысле) случайной величины — функции принадлежности случайного нечеткого множества (если в задаче оптимизации нет дополнительных ограничений). Читателю рекомендуется ввести подходящие обозначения и записать предыдущую фразу в виде формулы.

Необходимость оценивания параметров возникает, например, при попытках описать применяемый людьми оператор «очень», переводящий нечеткое множество «высокий человек» в нечеткое множество ««очень высокий человек», «старый» — в «очень старый» и т. д. [19]. Простейшее предположение — оператор «очень» приводит к сдвигу функции принадлежности вдоль универсального множества. Возникает задача оценки величины этого сдвига на основе опроса ряда испытуемых, проверки гипотез о том, что у всех людей параметр сдвига одинаков, что данные допускают адекватное описание с помощью этого простейшего предположения и т.д. Теоремы п. 9 устанавливают состоятельность соответствующих оценок.

Стоит упомянуть также о регрессионных постановках. Пусть каждый из n испытуемых выдает в ответ на вопрос два нечетких множества A_i, B_i, $i=1, 2, \ldots, n$. Необходимо приблизить B_i с помощью A_i. Положим

$$\mu_{\alpha A+\beta}(x)=\mu_A(\alpha x+\beta),\quad x\in R^1.$$

Тогда в качестве приближения B_i естественно взять $\alpha^* A_i+\beta^*$, где $(\alpha^*, \beta^*)\in M(h_n, R^2)$, а

$$h_n(\alpha, \beta)=\sum_{1\leqslant i\leqslant n}\rho(\alpha A_i+\beta, B_i),$$

здесь ρ — расстояние между нечеткими множествами (п. 10). При обычных предположениях оценки α^*, β^* являются состоятельными (п. 9). Естественно называть $\alpha^* A_i+\beta^*$ регрессией B_i на A_i.

12. Проверка гипотез о нечетких множествах

Пусть ответ эксперта — нечеткое множество. Естественно считать, что его ответ, как показание любого прибора, содержит погрешности. Если есть несколько экспертов, то в качестве единой оценки группового мнения естественно взять эмпирическое среднее их ответов. Но возникает коварный вопрос: действительно ли все эксперты измеряют одно и то же? Может быть, глядя на данный предмет, они оценивают его с разных сторон? Например, на научную статью можно смотреть как с теоретической точки зрения, так и с прикладной, и оценки будут скорее всего различны (если они совпадают, то работа либо никуда не годится, либо является выдающейся).

Итак, возник вопрос: как проверить согласованность мнений экспертов? Сначала надо определить понятие согласованности. Пусть A — нечеткий ответ эксперта. Будем считать, что

$$\mu_A(u)=\mu_{N(A)}(u)+\xi_A(u), \qquad (31)$$

где $N(A)$ — «истинное» нечеткое множество, а $\xi_A(u)$ — «погрешность» эксперта как прибора*. Рассмотрим две постановки.

Мнения экспертов $A(1), \ldots, A(m)$ будем считать согласованными, если

$$N(A(1))=N(A(2))=\ldots=N(A(m)). \qquad (32)$$

Рассмотрим две группы экспертов. В первой у всех «истинное» мнение $N(A)$, а во второй у всех—$N(B)$. Две группы будем считать согласованными по мнениям, если

$$N(A)=N(B). \qquad (33)$$

Согласованность определена. Как же ее проверить? Если экспертов достаточно много, то эти гипотезы можно проверять отдельно для каждого элемента универсального множества. Мы попадаем в область классической статистики. Правда, пока доберешься до конкретного способа проверки, приходится разобрать ряд вопросов, например: для всех ли экспертов погрешности одинаково распределены? независимы ли погрешности? известно ли их распределение? известна ли дисперсия?

---

* Об экспертных оценках и соответствующих математических моделях см. [9—13], а также [67—68].

Посмотрим повнимательнее на гипотезу (33). Фактически спрашивается: совпадают ли «показатели центральной тенденции» у совокупностей, из которых извлечены выборки? Однако какие «показатели» взять? Математическое ожидание? Медиану? Может быть, расширить понятие согласованности, добавив требование одинаковой распределенности погрешностей? Тогда вместо (33) получаем проверку однородности, т. е. того, что две выборки извлечены из одной и той же совокупности. С гипотезой однородности связан свой клубок проблем (п. 7), однако можно без опаски рекомендовать критерий Вилкоксона, он же Уилкоксона, он же Манна-Уитни или Манна-Уитнея*.

Вспомним теперь, что проверять гипотезы нам надо не для одного избранного элемента универсума, а для всех. Результаты проверки для отдельных элементов, скорее всего, не являются независимыми (как случайные величины), поэтому неясно, как свести их вместе. Это — нерешенная проблема [36].

Зачем это нагромождение неизвестностей? Ведь профессионал-статистик, перед которым будет приказом начальства поставлена задача проверки обсуждаемых гипотез, предложит соответствующую методику. Правда, она окажется довольно длинной, и ни одного образца мы здесь не можем привести. А цель предыдущего изложения — показать, что методик может быть весьма много, и подбор подходящей— творческий трудоемкий процесс, который должен осуществляться совместно статистиком и специалистом прикладной области, ответственным за содержательную постановку задачи. По оценке известного польского математика Г. Штейнгауза на подобные обсуждения должно уходить не менее 50 часов, если, конечно, рассчитывать на плодотворность совместной работы статистика и прикладника [70, с. 24].

Дополнительные сложности возникают, если число экспертов мало. Иногда их можно преодолеть за счет увеличения числа элементов универсума. Если оно стремится к бесконечности, то число неизвестных параметров растет пропорционально объему выборки (а также объему имеющейся информации). Подобная ситуация (ее называют асимптотикой Колмогорова, который предложил рассматривать по-

* Разнобой в транскрипции фамилий весьма мешает: попробуй догадайся, что Кендалл и Кендэл — один и тот же Kendall (см. также [69, с. 217]). Не писать ли по-старинному фамилию на языке оригинала, как это принято и сейчас в медицинских журналах?

добные постановки в дискриминантном анализе) не является стандартной для статистики — обычно число параметров фиксировано, а объем выборки стремится к бесконечности. За последние годы в асимптотике Колмогорова получен ряд важных результатов, прежде всего Л. Д. Мешалкиным и его сотрудниками.

Предположим дополнительно независимость случайных величин, относящихся к различным элементам универсума. Тогда речь идет о проверке гипотез по совокупности малых выборок. Автором предложен прием, позволяющий проверять гипотезы (32) и (33) в асимптотике Колмогорова (число выборок растет, объемы их остаются ограниченными, каждой выборке соответствует свой параметр), за описанием которого отошлем к [12, 13, 66].

Конкретные алгоритмы разработаны в случае, когда ответы экспертов — четкие множества, т. е. представляют собой случайные множества с независимыми элементами. Соответствующие публикации — это [9, § 4.5, 13, 66,67]. Приложения — модель независимых парных сравнений в экспертных оценках, некоторые медицинские вопросы.

В результате читатель озадачен — много нерешенных задач и мало конкретных рекомендаций. Что и требовалось. Призываю: действуйте сами!

ЛИТЕРАТУРА

1. Zadeh L. A. Fuzzy Sets. — Inf. Control., 1965, v. 8, p. 338—353.

2. Тутубалин В. Н. Теория вероятностей в естествознании. — М.: Знание, 1972 (серия «Математика, кибернетика», 1972, № 2).

3. Тутубалин В. Н. Статистическая обработка рядов наблюдений. — М.: Знание, 1973 (серия «Математика, кибернетика», 1973, № 3).

4. Тутубалин В. Н. Границы применимости (вероятностно-статистические методы и их возможности). — М.: Знание, 1977 (серия «Математика, кибернетика», 1977, № 1).

5. Айвазян С. А. Экстремальная формулировка основных проблем прикладной статистики. —В кн.: Всесоюзная школа «Программно-алгоритмическое обеспечение прикладного многомерного статистического анализа» (18—28 сентября 1979 г., пос. Цахкадзор АрмССР). Тезисы докладов. — Ереван, 1979, с. 24—49.

6. Многомерный статистический анализ в социально-экономических исследованиях. — Ученые записки по статистике, т. 26. — М.: Наука, 1974.

7. Прикладной многомерный статистический анализ. — Ученые записки по статистике, т. 33. — М.: Наука, 1978.

8. Айвазян С. А., Бежаева З. И., Староверов О. В. Классификация многомерных наблюдений. — М.: Статистика, 1974.

9. Орлов А. И. Устойчивость в социально-экономических моделях. — М.: Наука, 1979.

10. Статистические методы анализа экспертных оценок. — Ученые записки по статистике, т. 29. — М.: Наука, 1977.

11. Тюрин Ю. Н. Непараметрические методы статистики. — М.: Знание, 1978 (серия «Математика, кибернетика», 1978, № 4).

12. Экспертные оценки. — Вопросы кибернетики, вып. 58. — М.: Научный совет АН СССР по комплексной проблеме «Кибернетика», 1979.

13. Экспертные методы в системном анализе. — Труды Всесоюзного научно-исследовательского института системных исследований ГКНТ и АН СССР, вып. 4. — М.: ВНИИСИ, 1979.

14. Ротарь В. И. О симметрической зависимости и одной модели выборочного обследования. — Экономика и математические методы, 1973, т. 9, № 6, с. 1150—1156.

15. Заде Л. А. Основы нового подхода к анализу сложных систем и процессов принятия решений. — В кн.: Математика сегодня. — М.: Знание, 1974, с. 5—49 (серия «Математика, кибернетика», 1974, № 7).

16. Кафаров В. В., Дорохов И. Н., Марков Е. П. Принцип описания химико-технологических процессов с помощью нечетких множеств. — Доклады Академии наук СССР, 1978, т. 243, № 1, с. 159—162.

17. Сочинения Козьмы Пруткова. — М.: Художественная литература, 1976.

18. Борель Э. Вероятность и достоверность. — М.: Государственное издательство физико-математической литературы, 1961.

19. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений. — М.: Мир, 1976.

20. Беллман Р., Заде Л. Принятие решений в расплывчатых условиях. — В кн.: Вопросы анализа и процедуры принятия решений. — М.: Мир, 1976, с. 172—215.

21. Гусев Л. А., Смирнова И. М. Размытые множества. Теория и приложения. (Обзор). — Автоматика и телемеханика, 1973, № 5, с. 66—85.

22. Вентцель Е. С. Исследование операций. — М.: Знание, 1976 (серия «Математика, кибернетика», 1976, № 1).

23. Крамер Г. Математические методы статистики. — М.: Мир, 1975.

24. Мешалкин Л. Д. Параметризация многомерных распределений. — В сб.: Прикладной многомерный статистический анализ. — М.: Наука, 1978, с. 11—18.

25. Мешалкин Л. Д., Курочкина А. И. Новый подход к параметризации регрессионных зависимостей. — В сб.: Исследования по математической статистике, 3. Записки научных семинаров Ленинградского отделения Математического института АН СССР, т. 87. — Л.: 1979, с. 79—86.

26. Колмогоров А. Н. Основные понятия теории вероятностей. — М.: Наука, 1974.

27. Прохоров Ю. В., Розанов Ю. А. Теория вероятностей. (Основные понятия. Предельные теоремы. Случайные

процессы.) — М.: Наука, 1973 (серия «Справочная математическая библиотека»).

28. Орлов А. И. Вероятностное пространство, неравенство Чебышева и закон больших чисел - — основа курса теории вероятностей для школьников. — В сб.: Подготовка студентов педагогических институтов к внеурочной работе по математике, вып. 2. — Вологда, 1976, с. 13—29.

29. Пфанцагль И. Теория измерений. — М.: Мир, 1976.

30. Колмогоров А. Н., Прохоров Ю. В. Математическая статистика. — Большая Советская Энциклопедия, т. 15 (третье издание). — М.: Советская энциклопедия, 1974, с. 1428—1438.

31. Налимов В. В. Вероятностная модель языка. — М.: Наука, 1974.

32. Стахов А. П. Алгоритмическая теория измерения. — М.: Знание, 1979 (серия «Математика, кибернетика», 1979, № 6).

33. Козлов В. С., Эрлих Я. М., Долгушевский Ф. Г., Полушин П. И. Общая теория статистики. — М.: Статистика, 1975.

34. Гублер Е. В. Вычислительные методы анализа и распознавания патологических процессов. — Л.: Медицина, 1978.

35. Тюрин Ю. Н. Что такое математическая статистика. — М.: Знание, 1975 (серия «Математика, кибернетика», 1975, № 5).

36. Орлов А. И. Некоторые проблемы асимптотической теории статистик. — В сб.: указанном в [5], с. 104—113.

37. Орлов А. И. Алгоритмические аспекты статистики объектов нечисловой природы. — В сб., указанном в [5], с. 261—264.

38. Шубкин В. Н. Социологические опыты. — М.: Мысль, 1970.

39. Щукина Г. И. Проблема познавательного интереса в педагогике. — М.: Педагогика, 1971.

40. Дэйвид Г. Порядковые статистики. — М.: Наука, 1979.

41. Ефимов А. Н. Порядковые статистики — их свойства и приложения. — М.: Знание, 1980 (серия «Математика, кибернетика», 1980, № 2).

42. Петров В. М., Яблонский А. И. Математика и социальные процессы. - М.: Знание, 1980 (серия «Математика, кибернетика, 1980, № 1).

43. Суппес П., Зинес Дж. Основы теории измерений. — В кн.: Психологические измерения. — М.: Мир, 1967, с. 9—110.

44. Джини К. Средние величины. — М.: Статистика, 1970.

45. Шрейдер Ю. А. Равенство, сходство, порядок. — М.: Наука, 1971.

46. Гнеденко Б. В. Курс теории вероятностей. — М.: Наука, 1969.

47. Орлов А. И. Случайные множества: законы больших чисел, проверка статистических гипотез. — Теория вероятностей и ее применения, т. 23, № 2 (1978), с. 462—464.

48. Орлов А. И. Статистика объектов нечисловой природы. — Теория вероятностей и ее применения, т. 25, № 3 (1980).

49. Тюрин Ю. Н., Литвак Б. Г., Орлов А. И., Сатаров Г. А., Шмерлинг Д. С. Анализ нечисловой информации. — В сб.: указанном в [5], с. 231—243.

50. Кемени Дж., Снелл Дж. Кибернетическое моделирование. Некоторые приложения. — М.: Советское радио, 1972.
51. Миркин Б. Г. Проблема группового выбора. — М.: Наука, 1974.
52. Миркин Б. Г. Анализ качественных признаков. — М.: Статистика, 1976.
53. Гусев В. А., Орлов А. И., Розенталь А. Л. Внеклассная работа по математике в 6—8 классах. — М.: Просвещение, 1977.
54. Пойа Д. Как решать задачу. — М.: Учпедгиз, 1961.
55. Пойа Д. Математика и правдоподобные рассуждения. — М.: Наука, 1975.
56. Пойа Д. Математическое открытие. Решение задач: основные понятия, изучение и преподавание. — М.: Наука, 1970.
57. Мудров В. И., Кушко В. Л. Метод наименьших модулей. — М.: Знание, 1971 (серия «Математика, кибернетика», 1971, № 7).
58. Иоффе А. Д., Тихомиров В. М. Теория экстремальных задач. — М.: Наука, 1974.
59. Растригин Л. А. Случайный поиск. — М.: Знание, 1979 (серия «Математика, кибернетика», 1979, № 1).
60. Алимов Ю. И. Альтернатива методу математической статистики. — М.: Знание, 1980 (серия «Математика, кибернетика», 1980, № 3).
61. Мизес Р. Вероятность и статистика. — М. — Л.: 1930.
62. Хинчин А. Я. Частотная теория Р. Мизеса и современные идеи теории вероятностей. — Вопросы философии, 1961, № 1, 2.
63. Сачков Ю. В. Введение в вероятностный мир. — М.: Наука, 1971.
64. Купцов В. И. Детерминизм и вероятность. — М.: Политиздат, 1976.
65. Прикладная статистика. — Ученые записки по статистике, т. 41. — М.: Наука, 1980.
66. Алгоритмическое и программное обеспечение прикладного статистического анализа.— Ученые записки по статистике, т. 36.— М.: Наука, 1980.
67. Экспертные оценки в задачах управления. — М.: Институт проблем управления (автоматики и телемеханики), 1980.
68. Китаев Н. Н. Групповые экспертные оценки. — М.: Знание, 1975 (серия «Математика, кибернетика», 1975, № 3).
69. Рыбакин А. И. (составитель). Словарь английских личных имен. — М.: Советская энциклопедия, 1973.
70. Штейнгауз Г. Задачи и размышления. — М.: Мир, 1974.

СОДЕРЖАНИЕ

Александр Иванович ОРЛОВ

ЗАДАЧИ ОПТИМИЗАЦИИ И НЕЧЕТКИЕ ПЕРЕМЕННЫЕ

Гл. отраслевой редактор Л. А. Ерлыкин
Редактор Г. Г. Карвовский
Мл. редактор Т. Г. Иншакова
Обложка Л. П. Ромасенко
Худож. редактор М. А. Бабичева
Техн. редактор А. М. Красавина
Корректор В. В. Каночкина
ИБ № 2939

Сдано в набор 17.06.80 г. Подписано к печати 14.07.80 г. Т—13861. Формат бумаги $84 \times 108^{1}/_{32}$. Бумага тип. № 3. Гарнитура литературная. Печать высокая. Усл. печ. л. 3,36. Уч.-изд. л. 3,56. Тираж 35470 экз. Заказ № 1335. Цена 11 коп.
Издательство «Знание». 101835, ГСП, Москва, Центр, проезд Серова, д. 4. Индекс заказа 804308.
Чеховский полиграфический комбинат Союзполиграфпрома Государственного комитета СССР по делам издательств, полиграфии и книжной торговли г. Чехов Московской области

ВНИМАНИЮ ЧИТАТЕЛЕЙ!

Издательство «Знание» выпускает в свет ежегодную серию подписных научно-популярных брошюр «Математика, кибернетика». В ее издании принимают участие крупнейшие советские и зарубежные ученые, которые на страницах брошюр рассказывают о современных достижениях и основных направлениях развития математики и кибернетики. Наряду с теоретическими вопросами большое внимание уделяется практическим приложениям к решению важнейших научных и народнохозяйственных проблем. Обсуждаются методологические и философские проблемы развития математического знания, вопросы математического образования. Много места в брошюрах уделяется рассказу о наиболее видных ученых-математиках, об истории развития математической науки.

Серия рассчитана на специалистов математиков, инженеров, учащихся вузов, преподавателей и учителей. Отдельные брошюры будут полезны и школьникам.

В 1981 г. подписчики получат ряд брошюр, среди которых:

Белый Ю. А., кандидат физико-математических наук. Электронные миникалькуляторы и техника вычислений.

Дашевский Л. Н., Шкабара Е. А., кандидаты технических наук. Как это начиналось (О создании первой отечественной электронной вычислительной машины МЭСМ).

Колмогоров А. Н., академик АН СССР, Вавилов В. В., кандидат физико-математических наук, Тропин И. Т., директор школы. Физико-математическая школа при МГУ.

Гнеденко Б. В., академик АН УССР. Из истории науки о случайном.

Данилов Ю. А., кандидат физико-математических наук. Джон фон Нейман.

Осипов Л. А., кандидат технических наук. Программирование на алгоритмических языках.

Брошюры серии «Математика, кибернетика» в продажу не поступают, распространяются они только по подписке. Подписная цена на год — 1 руб. 32 коп. Индекс серии в каталоге «Союзпечати» — 70096.

11 коп.

Индекс 70096