2. Описание неопределенностей в теории принятия решений

Одна из основных проблем в теории принятия решений – необходимость учета неопределенностей, оценки и управления рисками. Для описания неопределенностей применяют различные подходы.

Прежде всего, необходимо разобраться с проблемами измерения различных величин, используемых в процессе принятия решения. Они могут быть измерены в тех или иных количественных или качественных шкалах. Поскольку в выборе конкретной шкалы имеется некоторый произвол (например, расстояние можно измерять в аршинах, саженях, верстах, метрах или парсеках), то естественно потребовать, чтобы принимаемое решение не зависело от этого произвола (например, от того, в каких единицах измерено расстояние). Теории измерений посвящена первая глава настоящей части.

Для описания неопределенностей и рисков чаще всего используется вероятностно-статистический подход. Во второй главе рассматриваются основные понятия и результаты в области теории вероятностей и статистики. Изложение нацелено на удовлетворение потребностей менеджеров, экономистов, инженеров, связанных с теорией принятия решений. Приводятся все необходимые понятия и результаты современных вероятностно-статистических методов, при этом исключен ряд излишних подробностей, традиционно включаемых в математические курсы.

Изложение во второй главе доведено до современного уровня. Одним из наиболее интересных и продуктивных современных направлений является статистика интервальных данных. В ней исходные данные – не числа, а интервалы. Таким образом, неопределенность величин, используемых в процессе принятия решения, моделируется путем замены конкретных численных значений на интервалы, в которых содержатся рассматриваемые величины. Статистике интервальных данных посвящена третья глава настоящей части.

Интервальные данные – это частный случай нечетких данных. В последнее время теория нечеткости все чаще используется в экономических исследованиях [1]. Нечеткость, расплывчатость, размытость понятий и величин – типичная черта многих задач принятия решений. Основам теории нечеткости посвящена четвертая глава. В ней рассмотрена, в частности, система принципиально важных утверждений, согласно которой теория нечеткости в определенном смысле сводится к теории случайных множеств – одной из частей вероятностно-статистической теории.

Материалы второй части во многом основываются на результатах, приведенных в монографиях [2,3].

Литература

1. Карминский А.М., Оленев Н.И., Примак А.Г., Фалько С.Г. Контроллинг в бизнесе. Методологические и практические основы построения контроллинга в организациях. – М.: Финансы и статистика, 1998. – 256 с.

2. Орлов А.И. Устойчивость в социально-экономических моделях. – М.: Наука, 1979. -296 с.

3. Орлов А.И. Эконометрика. – М.: Экзамен, 2002. – 576 с.

2.1. Шкалы измерения и инвариантные алгоритмы

2.1.1. Основные шкалы измерения

Почему необходима теория измерений? Теория измерений (в дальнейшем сокращенно ТИ) является одной из составных частей эконометрики. Она входит в состав *статистики объектов нечисловой природы*. Необходимость использования ТИ в теории принятия решений рассмотрим на примере экспертного оценивания, в частности, в связи с агрегированием мнений экспертов, построением обобщенных показателей и рейтингов.

Использование чисел в жизни и хозяйственной деятельности людей отнюдь не всегда предполагает, что эти числа можно складывать и умножать, производить иные арифметические действия. Что бы вы сказали о человеке, который занимается умножением телефонных номеров? И отнюдь не всегда 2+2=4. Если вы вечером поместите в клетку двух животных, а потом еще двух, то отнюдь не всегда можно утром найти в этой клетке четырех животных. Их может быть и много больше - если вечером вы загнали в клетку овцематок или беременных кошек. Их может быть и меньше - если к двум волкам вы поместили двух ягнят. Числа используются гораздо шире, чем арифметика.

Так, например, мнения экспертов часто выражены в *порядковой шкале* (подробнее о шкалах говорится ниже), т.е. эксперт может сказать (и обосновать), что один показатель качества продукции более важен, чем другой, первый технологический объект более опасен, чем второй, и т.д. Но он не в состоянии сказать, *во сколько раз* или *на сколько* более важен, соответственно, более опасен. Экспертов часто просят дать ранжировку (упорядочение) объектов экспертизы, т.е. расположить их в порядке возрастания (или убывания) интенсивности интересующей организаторов экспертизы характеристики. Ранг - это номер (объекта экспертизы) в упорядоченном ряду значений характеристики у различных объектов. Такой ряд в статистике называется вариационным. Формально ранги выражаются числами 1, 2, 3, ..., но с этими числами нельзя делать привычные арифметические операции. Например, хотя в арифметике 1 + 2 = 3, но нельзя утверждать, что для объекта, стоящем на третьем месте в упорядочении, интенсивность изучаемой характеристики равна сумме интенсивностей объектов с рангами 1 и 2. Так, один из видов экспертного оценивания - оценки учащихся. Вряд ли кто-либо будет утверждать, что знания отличника равны сумме знаний двоечника и троечника (хотя 5 = 2 + 3), хорошист соответствует двум двоечникам (2 + 2 = 4), а между отличником и троечником такая же разница, как между хорошистом и двоечником (5 - 3 = 4 - 2). Поэтому очевидно, что для анализа подобного рода качественных данных необходима не всем известная арифметика, а другая теория, дающая базу для разработки, изучения и применения конкретных методов расчета. Это и есть ТИ.

При чтении литературы надо иметь в виду, что в настоящее время термин "теория измерений" применяется для обозначения целого ряда научных дисциплин. А именно, классической метрологии (науки об измерениях физических величин), рассматриваемой здесь ТИ, некоторых других направлений, например, алгоритмической теории измерений. Обычно из контекста понятно, о какой конкретно теории идет речь.

Краткая история теории измерений. Сначала ТИ развивалась как теория психофизических измерений. В послевоенных публикациях американский психолог С.С. Стивенс основное внимание уделял шкалам измерения. Во второй половине XX в. сфера применения ТИ стремительно расширяется. Посмотрим, как это происходило. Один из томов выпущенной в

США в 1950-х годах "Энциклопедии психологических наук" назывался "Психологические измерения". Значит, составители этого тома расширили сферу применения РТИ с психофизики на психологию в целом. А в основной статье в этом сборнике под названием, обратите внимание, "Основы теории измерений", изложение шло на абстрактно-математическом уровне, без привязки к какой-либо конкретной области применения. В этой статье [1] упор был сделан на "гомоморфизмах эмпирических систем с отношениями в числовые" (в эти математические термины здесь вдаваться нет необходимости), и математическая сложность изложения возросла по сравнению с работами С.С. Стивенса.

Уже в одной из первых отечественных статей по РТИ (конец 1960-х годов) было установлено, что баллы, присваиваемые экспертами при оценке объектов экспертизы, как правило, измерены в порядковой шкале. Отечественные работы, появившиеся в начале 1970-х годов, привели к существенному расширению области использования РТИ. Ее применяли к педагогической квалиметрии (измерению качества знаний учащихся), в системных исследованиях, в различных задачах теории экспертных оценок, для агрегирования показателей качества продукции, в социологических исследованиях, и др.

Итоги этого этапа были подведены в монографии [2]. В качестве двух основных проблем РТИ наряду с *установлением типа шкалы* измерения конкретных данных был выдвинут поиск алгоритмов анализа данных, результат работы которых не меняется при любом допустимом преобразовании шкалы (т.е. является *инвариантным* относительно этого преобразования).

Метрологи вначале резко возражали против использования термина "измерение" для качественных признаков. Однако постепенно возражения сошли на нет, и к концу XX в. ТИ стала рассматриваться как общенаучная теория.

Шесть типов шкал. В соответствии с ТИ при математическом моделировании реального явления или процесса следует прежде всего установить *типы шкал*, в которых измерены те или иные переменные. Тип шкалы задает *группу допустимых преобразований шкалы*. Допустимые преобразования не меняют соотношений между объектами измерения. Например, при измерении длины переход от аршин к метрам не меняет соотношений между длинами рассматриваемых объектов - если первый объект длиннее второго, то это будет установлено и при измерении в аршинах, и при измерении в метрах. Обратите внимание, что при этом численное значение длины в аршинах отличается от численного значения длины в метрах - не меняется лишь результат сравнения длин двух объектов.

Укажем основные виды шкал измерения и соответствующие группы допустимых преобразований.

В *шкале наименований* (другое название этой шкалы - *номинальная;* это - переписанное русскими буквами английское название шкалы) **допустимыми** являются все взаимно-однозначные преобразования. В этой шкале числа используются лишь как метки. Примерно так же, как при сдаче белья в прачечную, т.е. лишь для различения объектов. В шкале наименований измерены, например, номера телефонов, автомашин, паспортов, студенческих билетов. Номера страховых свидетельств государственного пенсионного страхования, медицинского страхования, ИНН (индивидуальный номер налогоплательщика) измерены в шкале наименований. Пол людей тоже измерен в шкале наименований, результат измерения принимает два значения - мужской, женский. Раса, национальность, цвет глаз, волос - номинальные признаки. Номера букв в алфавите - тоже измерения в шкале наименований. Никому в здравом уме не придет в голову складывать или умножать номера телефонов, такие операции не имеют смысла. Сравнивать буквы и говорить, например, что буква П лучше буквы С, также никто не будет. Единственное, для чего годятся измерения в шкале наименований - это различать объекты. Во многих случаях только это от них и требуется. Например, шкафчики в раздевалках для взрослых различают по номерам, т.е. числам, а в детских садах используют рисунки, поскольку дети еще не знают чисел.

В *порядковой шкале* числа используются не только для различения объектов, но и для установления порядка между объектами. Простейшим примером являются оценки знаний учащихся. Символично, что в средней школе применяются оценки 2, 3, 4, 5, а в высшей школе ровно тот же смысл выражается словесно - неудовлетворительно, удовлетворительно, хорошо, отлично. Этим подчеркивается "нечисловой" характер оценок знаний учащихся. В порядковой шкале **допустимыми** являются все строго возрастающие преобразования.

Установление типа шкалы, т.е. задания группы допустимых преобразований шкалы измерения - дело специалистов соответствующей прикладной области. Так, оценки привлекательности профессий мы в монографии [2], выступая в качестве социологов, считали измеренными в порядковой шкале. Однако отдельные социологи не соглашались с нами, полагая, что выпускники школ пользуются шкалой с более узкой группой допустимых преобразований, например, интервальной шкалой. Очевидно, эта проблема относится не к математике, а к наукам о человеке. Для ее решения может быть поставлен достаточно трудоемкий эксперимент. Пока же он не поставлен, целесообразно принимать порядковую шкалу, так как это гарантирует от возможных ошибок.

Оценки экспертов, как уже отмечалось, часто следует считать измеренными в порядковой шкале. Типичным примером являются задачи ранжирования и классификации промышленных объектов, подлежащих экологическому страхованию.

Почему мнения экспертов естественно выражать именно в порядковой шкале? **Как показали многочисленные опыты, человек более правильно (и с меньшими затруднениями) отвечает на вопросы качественного, например, сравнительного, характера, чем количественного.** Так, ему легче сказать, какая из двух гирь тяжелее, чем указать их примерный вес в граммах.

В различных областях человеческой деятельности применяется много других видов порядковых шкал. Так, например, в минералогии используется шкала Мооса, по которому минералы классифицируются согласно критерию твердости. А именно: тальк имеет балл 1, гипс - 2, кальций - 3, флюорит - 4, апатит - 5, ортоклаз - 6, кварц - 7, топаз - 8, корунд - 9, алмаз - 10. Минерал с большим номером является более твердым, чем минерал с меньшим номером, при нажатии царапает его.

Порядковыми шкалами в географии являются - бофортова шкала ветров ("штиль", "слабый ветер", "умеренный ветер" и т.д.), шкала силы землетрясений. Очевидно, нельзя утверждать, что землетрясение в 2 балла (лампа качнулась под потолком - такое бывает и в Москве) ровно в 5 раз слабее, чем землетрясение в 10 баллов (полное разрушение всего на поверхности земли).

В медицине порядковыми шкалами являются - шкала стадий гипертонической болезни (по Мясникову), шкала степеней сердечной недостаточности (по Стражеско-Василенко-Лангу), шкала степени выраженности коронарной недостаточности (по Фогельсону), и т.д. Все эти шкалы построены по схеме: заболевание не обнаружено; первая стадия заболевания; вторая стадия; третья стадия… Иногда выделяют стадии 1а, 1б и др. Каждая стадия имеет свойственную только ей медицинскую характеристику. При описании групп инвалидности числа используются в противоположном порядке: самая тяжелая - первая группа инвалидности, затем - вторая, самая легкая - третья.

Номера домов также измерены в порядковой шкале - они показывают, в каком порядке стоят дома вдоль улицы. Номера томов в собрании сочинений писателя или номера дел в архиве предприятия обычно связаны с хронологическим порядком их создания.

При оценке качества продукции и услуг, в т.н. квалиметрии (буквальный перевод: измерение качества) популярны порядковые шкалы. А именно, единица продукции оценивается как годная или не годная. При более тщательном анализе используется шкала с тремя градациями: есть значительные дефекты - присутствуют только незначительные дефекты - нет дефектов. Иногда применяют четыре градации: имеются критические дефекты (делающие невозможным использование) - есть значительные дефекты - присутствуют только незначительные дефекты - нет дефектов. Аналогичный смысл имеет сортность продукции - высший сорт, первый сорт, второй сорт,…

При оценке экологических воздействий первая, наиболее обобщенная оценка - обычно порядковая, например: природная среда стабильна - природная среда угнетена (деградирует). Аналогично в эколого-медицинской шкале: нет выраженного воздействия на здоровье людей - отмечается отрицательное воздействие на здоровье.

Порядковая шкала используется и во многих иных областях. В эконометрике это прежде всего различные методы экспертных оценок. (см. посвященную им главу в части 3).

Все шкалы измерения делят на две группы - шкалы качественных признаков и шкалы количественных признаков.

Порядковая шкала и шкала наименований - основные шкалы качественных признаков. Поэтому во многих конкретных областях результаты качественного анализа можно рассматривать как измерения по этим шкалам.

Шкалы количественных признаков - это шкалы интервалов, отношений, разностей, абсолютная. По шкале *интервалов* измеряют величину потенциальной энергии или координату точки на прямой. В этих случаях на шкале нельзя отметить ни естественное начало отсчета, ни естественную единицу измерения. Исследователь должен сам задать точку отсчета и сам выбрать единицу измерения. Допустимыми преобразованиями в шкале интервалов являются линейные возрастающие преобразования, т.е. линейные функции. Температурные шкалы Цельсия и Фаренгейта связаны именно такой зависимостью: $^0C = 5/9\ (^0F - 32)$, где ^0C - температура (в градусах) по шкале Цельсия, а ^0F - температура по шкале Фаренгейта.

Из количественных шкал наиболее распространенными в науке и практике являются шкалы *отношений*. В них есть естественное начало отсчета - нуль, т.е. отсутствие величины, но нет естественной единицы измерения. По шкале отношений измерены большинство физических единиц: масса тела, длина, заряд, а также цены в экономике. Допустимыми преобразованиями шкале отношений являются подобные (изменяющие только масштаб). Другими словами, линейные возрастающие преобразования без свободного члена. Примером является пересчет цен из одной валюты в другую по фиксированному курсу. Предположим, мы сравниваем экономическую эффективность двух инвестиционных проектов, используя цены в рублях. Пусть первый проект оказался лучше второго. Теперь перейдем на валюту самой экономически мощной державы мира - юани, используя фиксированный курс пересчета. Очевидно, первый проект должен опять оказаться более выгодным, чем второй. Это очевидно из общих соображений. Однако алгоритмы расчета не обеспечивают автоматически выполнения этого очевидного условия. Надо проверять, что оно выполнено. Результаты подобной проверки для средних величин описаны ниже.

В шкале разностей есть естественная единица измерения, но нет естественного начала отсчета. Время измеряется по шкале *разностей*, если год (или сутки - от полудня до полудня) принимаем естественной единицей измерения, и по шкале интервалов в общем случае. На современном уровне знаний естественного начала отсчета указать нельзя. Дату сотворения мира различные авторы рассчитывают по-разному, равно как и момент рождества Христова. Так, согласно новой статистической хронологии [3], разработанной группой известного историка акад. РАН А.Т.Фоменко, Господь Иисус Христос родился примерно в 1054 г. по принятому ныне летоисчислению в Стамбуле (он же - Царьград, Византия, Троя, Иерусалим, Рим).

Только для *абсолютной* шкалы результаты измерений - числа в обычном смысле слова. Примером является число людей в комнате. Для абсолютной шкалы допустимым является только тождественное преобразование.

В процессе развития соответствующей области знания тип шкалы может меняться. Так, сначала температура измерялась по *порядковой* шкале (холоднее - теплее). Затем - по *интервальной* (шкалы Цельсия, Фаренгейта, Реомюра). Наконец, после открытия абсолютного нуля температуру можно считать измеренной по шкале *отношений* (шкала Кельвина). Надо отметить, что среди специалистов иногда имеются разногласия по поводу того, по каким шкалам следует считать измеренными те или иные реальные величины. Другими словами, процесс измерения включает в себя и определение типа шкалы (вместе с обоснованием выбора определенного типа шкалы). Кроме перечисленных шести основных типов шкал, иногда используют и иные шкалы.

2.1.2. Инвариантные алгоритмы и средние величины

Основное требование к алгоритмам анализа данных формулируется в ТИ так: **выводы, сделанные на основе данных, измеренных в шкале определенного типа, не должны меняться при допустимом преобразовании шкалы измерения этих данных.** Другими словами, выводы должны быть *инвариантны* по отношению к допустимым преобразованиям шкалы.

Таким образом, одна из основных целей теории измерений - борьба с субъективизмом исследователя при приписывании численных значений реальным объектам. Так, расстояния можно измерять в аршинах, метрах, микронах, милях, парсеках и других единицах измерения. Массу (вес) - в пудах, килограммах, фунтах и др. Цены на товары и услуги можно указывать в юанях, рублях, тенге, гривнах, латах, кронах, марках, долларах США и других валютах (при условии заданных курсов пересчета). Подчеркнем очень важное, хотя и вполне очевидное обстоятельство: выбор единиц измерения зависит от исследователя, т.е. субъективен. *Статистические выводы могут быть адекватны реальности только тогда, когда они не зависят от того, какую единицу измерения предпочтет исследователь, т.е. когда они инвариантны относительно допустимого преобразования шкалы.*

Оказывается, сформулированное условие является достаточно сильным. Из многих алгоритмов эконометрического анализа данных ему удовлетворяют лишь некоторые. Покажем это на примере сравнения средних величин.

Пусть X_1 , X_2 ,…, X_n - выборка объема n. Часто используют среднее арифметическое

$$X_{ср} = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Использование среднего арифметического настолько привычно, что второе слово в термине часто опускают. И говорят о средней зарплате, среднем доходе и других средних для конкретных экономических данных, подразумевая под "средним" среднее арифметическое. Такая традиция может приводить к ошибочным выводам. Покажем это на примере расчета средней заработной платы (среднего дохода) работников условного предприятия (табл.1).

Табл.1. Численность работников различных категорий, их заработная плата и доходы (в условных единицах).

№ п/п	Категория работников	Число работников	Заработная плата	Суммарные доходы
1	Низкоквалифицированные рабочие	40	100	4000
2	Высококвалифицированные рабочие	30	200	6000
3	Инженеры и служащие	25	300	7500
4	Менеджеры	4	1000	4000
5	Генеральный директор (владелец)	1	18500	18500
6	Всего	100		40000

Первые три строки в табл.1 вряд ли требуют пояснений. Менеджеры - это директора по направлениям, а именно, по производству (главный инженер), по финансам, по маркетингу и сбыту, по персоналу (по кадрам). Владелец сам руководит предприятием в качестве генерального директора. В столбце "заработная плата" указаны доходы одного работника соответствующей категории, а в столбце "суммарные доходы" - доходы всех работников соответствующей категории.

Фонд оплаты труда составляет 40000 единиц, работников всего 100, следовательно, средняя заработная плата составляет 40000/100 = 400 единиц. Однако эта средняя арифметическая величина явно не соответствует интуитивному представлению о "средней зарплате". Из 100 работников лишь 5 имеют заработную плату, ее превышающую, а зарплата остальных 95 существенно меньше средней арифметической. Причина очевидна - заработная плата одного человека - генерального директора - превышает заработную плату 95 работников - низкоквалифицированных и высококвалифицированных рабочих, инженеров и служащих.

Ситуация напоминает описанную в известном рассказе о больнице, в которой 10 больных, из них у 9 температура 40 ^{0}С, а один уже отмучился, лежи в морге с температурой 0 ^{0}С. Между тем средняя температура по больнице равна 36 ^{0}С - лучше не бывает!

Сказанное показывает, что среднее арифметическое можно использовать лишь для достаточно однородных совокупностей (без больших выбросов в ту или иную сторону). А какие средние использовать для описания заработной платы? Вполне естественно использовать медиану. Для данных табл.1 медиана - среднее арифметическое 50-го и 51-го работника, если их заработные платы расположены в порядке неубывания. Сначала идут зарплаты 40 низкоквалифицированных рабочих, а затем - с 41-го до 70-го работника - заработные платы высококвалифицированных рабочих. Следовательно, медиана попадает именно на них и равна 200. У 50-ти работников заработная плата не превосходит 200, и у 50-ти - не менее 200, поэтому медиана показывает "центр", около которого группируется основная масса исследуемых величин. Еще одна средняя величина - мода, наиболее часто встречающееся значение. В рассматриваемом случае это заработная плата низкоквалифицированных рабочих, т.е. 100. Таким образом, для описания зарплаты имеем три средние величины - моду (100 единиц), медиану (200 единиц) и среднее арифметическое (400 единиц). Для наблюдающихся в реальной жизни распределений доходов и заработной платы справедлива та же закономерность: мода меньше медианы, а медиана меньше среднего арифметического.

Для чего в экономике используются средние величины? Обычно для того, чтобы заменить совокупность чисел одним числом, чтобы сравнивать совокупности с помощью средних.

Пусть, например, $Y_1, Y_2,...,Y_n$ - совокупность оценок экспертов, "выставленных" одному объекту экспертизы (например, одному из вариантов стратегического развития фирмы), $Z_1, Z_2,...,Z_n$ - второму (другому варианту такого развития). Как сравнивать эти совокупности? Очевидно, самый простой способ - по средним значениям.

А как вычислять средние? Известны различные виды средних величин: среднее арифметическое, медиана, мода, среднее геометрическое, среднее гармоническое, среднее квадратическое. Напомним, что общее понятие средней величины введено французским математиком первой половины XIX в. академиком О. Коши. Оно таково: средней величиной является любая функция $f(X_1, X_2,...,X_n)$ такая, что при всех возможных значениях аргументов значение этой функции не меньше, чем минимальное из чисел $X_1, X_2,...,X_n$, и не больше, чем максимальное из этих чисел. Все перечисленные выше виды средних являются средними по Коши.

При допустимом преобразовании шкалы значение средней величины, очевидно, меняется. Но выводы о том, для какой совокупности среднее больше, а для какой - меньше, не должны меняться (в соответствии с требованием инвариантности выводов, принятом как основное требование в ТИ). Сформулируем соответствующую математическую задачу поиска вида средних величин, результат сравнения которых устойчив относительно допустимых преобразований шкалы.

Пусть $f(X_1, X_2,\dots,X_n)$ - среднее по Коши. Пусть среднее по первой совокупности меньше среднего по второй совокупности:

$$f(Y_1, Y_2,\dots,Y_n) < f(Z_1, Z_2,\dots,Z_n).$$

Тогда согласно ТИ для устойчивости результата сравнения средних необходимо, чтобы для любого допустимого преобразования g из группы допустимых преобразований в соответствующей шкале было справедливо также неравенство

$$f(g(Y_1), g(Y_2),\dots,g(Y_n)) < f(g(Z_1), g(Z_2),\dots,g(Z_n)).$$

т.е. среднее преобразованных значений из первой совокупности также было меньше среднего преобразованных значений для второй совокупности. Причем сформулированное условие должно быть верно для любых двух совокупностей $Y_1, Y_2,\dots,Y_n$и $Z_1, Z_2,\dots,Z_n$ и, напомним, любого допустимого преобразования. Средние величины, удовлетворяющие сформулированному условию, назовем допустимыми (в соответствующей шкале). Согласно ТИ только такими средними можно пользоваться при анализе мнений экспертов и иных данных, измеренных в рассматриваемой шкале.

С помощью математической теории, развитой в монографии [2], удается описать вид допустимых средних в основных шкалах. Сразу ясно, что для данных, измеренных в шкале наименований, в качестве среднего годится только мода.

2.1.3. Средние величины в порядковой шкале

Рассмотрим обработку мнений экспертов, измеренных в порядковой шкале. Справедливо следующее утверждение.

Теорема 1. *Из всех средних по Коши допустимыми средними в порядковой шкале являются только члены вариационного ряда (порядковые статистики).*

Теорема 1 справедлива при условии, что среднее $f(X_1, X_2,\ldots,X_n)$ является непрерывной (по совокупности переменных) и симметрической функцией. Последнее означает, что при перестановке аргументов значение функции $f(X_1, X_2,\ldots,X_n)$ не меняется. Это условие является вполне естественным, ибо среднюю величину мы находим для *совокупности (множества)*, а не для *последовательности*. Множество не меняется в зависимости от того, в какой последовательности мы перечисляем его элементы.

Согласно теореме 1 в качестве среднего для данных, измеренных в порядковой шкале, можно использовать, в частности, медиану (при нечетном объеме выборки). При четном же объеме следует применять один из двух центральных членов вариационного ряда - как их иногда называют, левую медиану или правую медиану. Моду тоже можно использовать - она всегда является членом вариационного ряда. Но никогда нельзя рассчитывать среднее арифметическое, среднее геометрическое и т.д.

Приведем численный пример, показывающий некорректность использования среднего арифметического $f(X_1, X_2) = (X_1 + X_2)/2$ в порядковой шкале. Пусть $Y_1= 1$, $Y_2 = 11$, $Z_1= 6$, $Z_2= 8$. Тогда $f(Y_1, Y_2) = 6$, что меньше, чем $f(Z_1, Z_2) = 7$. Пусть строго возрастающее преобразование g таково, что $g(1) = 1$, $g(6) = 6$, $g(8) = 8$, $g(11) = 99$. Таких преобразований много. Например, можно положить $g(x) = x$ при x, не превосходящих 8, и $g(x) = 99(x-8)/3 + 8$ для x, больших 8. Тогда $f(g(Y_1), g(Y_2)) = 50$, что больше, чем $f(g(Z_1), g(Z_2)) = 7$. Как видим, в результате допустимого, т.е. строго возрастающего преобразования шкалы упорядоченность средних величин изменилась.

Таким образом, ТИ выносит жесткий приговор среднему арифметическому - использовать его с порядковой шкале нельзя. Однако же те, кто не знает теории измерений, используют его. Всегда ли они ошибаются? Оказывается, можно в какой-то мере реабилитировать среднее арифметическое, если перейти к вероятностной постановке и к тому удовлетвориться результатами для больших объемов выборок. В монографии [2] получено также следующее утверждение.

Теорема 2. *Пусть* Y_1, $Y_2,\ldots,Y_m$ - *независимые одинаково распределенные случайные величины с функцией распределения* $F(x)$, *а* Z_1, $Z_2,\ldots,Z_n$ - *независимые одинаково распределенные случайные величины с функцией распределения* $H(x)$, *причем выборки* Y_1, $Y_2,\ldots,Y_m$ *и* Z_1, $Z_2,\ldots,Z_n$ *независимы между собой и* $MY_1 > MZ_1$. *Для того, чтобы вероятность события*

$$\{w : \frac{g(Y_1)+g(Y_2)+\ldots+g(Y_m)}{m} > \frac{g(Z_1)+g(Z_2)+\ldots+g(Z_n)}{n}\}$$

стремилась к 1 при $\min(m,n)\to\infty$ *для любой строго возрастающей непрерывной функции* g, *удовлетворяющей условию*

$$\overline{\lim_{|x|\to\infty}} \left|\frac{g(x)}{x}\right| < \infty,$$

необходимо и достаточно, чтобы при всех x *выполнялось неравенство* $F(x) \le H(x)$, *причем существовало число* x_0 , *для которого* $F(x_0)<H(x_0)$.

Примечание. *Условие с верхним пределом носит чисто внутриматематический характер. Фактически функция* g - *произвольное допустимое преобразование в порядковой шкале.*

Согласно теореме 2 средним арифметическим можно пользоваться и в порядковой шкале, если сравниваются выборки из двух распределений, удовлетворяющих приведенному в теореме неравенству. Проще говоря, одна из функций распределения должна всегда лежать над другой. Функции распределения не могут пересекаться, им разрешается только касаться друг друга. Это условие выполнено, например, если функции распределения отличаются только сдвигом, т.е.

$$F(x) = H(x+b)$$

при некотором b. Последнее условие выполняется, если два значения некоторой величины измеряются с помощью одного и того же средства измерения, у которого распределение погрешностей не меняется при переходе от измерения одного значения рассматриваемой величины к измерению другого.

2.1.4. Средние по Колмогорову

Естественная система аксиом (требований к средним величинам) приводит к так называемым ассоциативным средним. Их общий вид нашел в 1930 г. А.Н.Колмогоров [4]. Теперь их называют «средними по Колмогорову». Они являются обобщением нескольких из перечисленных выше средних.

Для чисел $X_1, X_2,\ldots,X_n$ среднее по Колмогорову вычисляется по формуле

$$G\{(F(X_1)+F(X_2)+\ldots F(X_n))/n\},$$

где F - строго монотонная функция (т.е. строго возрастающая или строго убывающая), G - функция, обратная к F. Среди средних по Колмогорову - много хорошо известных персонажей. Так, если $F(x) = x$, то среднее по Колмогорову - это среднее арифметическое, если $F(x) = \ln x$, то среднее геометрическое, если $F(x) = 1/x$, то среднее гармоническое, если $F(x) = x^2$, то среднее квадратическое, и т.д. (в последних трех случаях усредняются положительные величины).

Среднее по Колмогорову - частный случай среднего по Коши. С другой стороны, такие популярные средние, как медиана и мода, нельзя представить в виде средних по Колмогорову. В монографии [2] доказаны следующие утверждения.

Теорема 3. *В шкале интервалов из всех средних по Колмогорову допустимым является только среднее арифметическое (при справедливости некоторых внутриматематических условий регулярности).*

Таким образом, среднее геометрическое или среднее квадратическое температур (в шкале Цельсия) или расстояний не имеют смысла. В качестве среднего надо применять среднее арифметическое. А также можно использовать медиану или моду.

Теорема 4. *При справедливости некоторых внутриматематических условий регулярности в шкале отношений из всех средних по Колмогорову допустимыми являются только степенные средние с $F(x) = x^c$, $c \neq 0$, и среднее геометрическое.*

Замечание. *Среднее геометрическое является пределом степенных средних при $c \to 0$.*

Есть ли средние по Колмогорову, которыми нельзя пользоваться в шкале отношений? Конечно. Например, с $F(x) = e^x$.

Аналогично средним величинам могут быть изучены и другие статистические характеристики - показатели разброса, связи, расстояния и др. (см., например, [2]). Так, коэффициент корреляции не меняется при любом допустимом преобразовании в шкале интервалов, как и отношение дисперсий, дисперсия не меняется в шкале разностей, коэффициент вариации - в шкале отношений, и т.д.

Приведенные выше результаты о средних величинах широко применяются, причем не только в экономике, менеджменте, теории экспертных оценок или социологии, но и в инженерном деле, например, для анализа методов агрегирования датчиков в АСУ ТП доменных печей. Велико прикладное значение ТИ в задачах стандартизации и управления качеством, в частности, в квалиметрии. Здесь есть и интересные теоретические результаты. Так, например, любое изменение коэффициентов весомости единичных показателей качества продукции приводит к изменению упорядочения изделий по средневзвешенному показателю (эта теорема доказана проф. В.В. Подиновским).

В теории принятия решений необходимо использовать только инвариантные алгоритмы обработки данных. В настоящей главе показано, что требование инвариантности выделяет из многих алгоритмов усреднения лишь некоторые, соответствующие используемым шкалам измерения. Инвариантные алгоритмы в общем случае рассматриваются в математической теории измерений [5]. Нацеленное на прикладные исследования изложение теории измерений дается в монографиях [2,6].

Литература

1. Суппес П., Зинес Дж. Основы теории измерений. - В сб.: Психологические измерения. - М.: Мир, 1967. С. 9-110.

2. Орлов А.И. Устойчивость в социально-экономических моделях. - М.: Наука, 1979. - 296 с.

3. Носовский Г.В. , Фоменко А.Т. Империя. Русь, Турция, Китай, Европа, Египет. Новая математическая хронология древности. - М.: Изд-во "Факториал", 1996. - 752 с.

4. Колмогоров А.Н. Избранные труды: Математика и механика. - М.: Наука, 1985. С. 136-138.

5. Пфанцагль И. Теория измерений. - М.: Мир, 1976. - 165 с.

6. Орлов А.И. Эконометрика. – М.: Экзамен, 2002. – 576 с.

Контрольные вопросы

1. Всегда ли имеет смысл складывать числа, используемые в той или иной области человеческой деятельности.

2. Приведите примеры величин, измеренных в шкале наименований.

3. Приведите примеры величин, измеренных в порядковой шкале.

4. Приведите примеры величин, измеренных в шкале интервалов.

5. Приведите примеры величин, измеренных в шкале отношений.

6. Постройте пример, показывающий некорректность использования среднего арифметического $f(X_1, X_2) = (X_1 + X_2)/2$ в порядковой шкале, используя допустимое преобразование $g(x) = x^2$ (при положительных усредняемых величинах x).

7. Постройте пример, показывающий некорректность использования среднего геометрического в порядковой шкале.

8. Какие средние величины целесообразно использовать при расчете средней заработной платы (или среднего дохода)?

Темы докладов, рефератов, исследовательских работ

1. Теория измерений как научная дисциплина, посвященная гомоморфизмам эмпирических систем с отношениями в числовые.

2. Показатели разброса, связи, показатели различия (в том числе метрики) в порядковой шкале.

3. Ранговые методы математической статистики как инвариантные методы анализа порядковых данных.

4. Показатели разброса, связи, показатели различия (в том числе метрики) в шкале интервалов.

5. Показатели разброса, связи, показатели различия (в том числе метрики) в шкале отношений.

6. Теорема В.В. Подиновского: любое изменение коэффициентов весомости единичных показателей качества продукции приводит к

изменению упорядочения изделий по средневзвешенному показателю.