

---

---

## МАТЕМАТИЧЕСКАЯ ЛОГИКА, ТЕОРИЯ ИЗМЕРЕНИЙ И СОЦИОЛОГИЧЕСКОЕ “МЫШЛЕНИЕ ПРИЗНАКАМИ”<sup>1</sup>

Ю.Н.Толстова

*(Москва)*

Настоящая публикация преследует следующие цели. 1. Продемонстрировать читателю-гуманитарию значение опубликованной в настоящем номере статьи Е.Е.Витяева и А.Д.Логвиненко, в которой под определённым углом зрения обсуждаются затронутые ниже вопросы. 2. Показать, что свойственное исследователю-социологу “мышление признаками” (термин Э.Ноэль) приводит к естественности использования для поиска эмпирических социологических закономерностей математической логики. 3. Обсудить направление развития репрезентационной теории измерений, связанное с аксиоматическим заданием эмпирической системы; показать, что соответствующий подход позволяет связать измерение с поиском логико-математических закономерностей из п. 2. 4. Рассмотреть ряд аспектов понятия закономерности в эмпирическом социологическом исследовании применительно к закономерностям из п. 2.

*Ключевые слова:* признак, математическая система, исчисление предикатов, логические функции, эмпирическая система, числовая система, шкала, репрезентационная теория измерений, аксиоматика.

---

<sup>1</sup> Работа над статьей осуществлялась автором при поддержке международного фонда RSS Contract No.: 854/1997.

## **I. Некоторые определения репрезентационной теории измерений. Переход к мышлению признаками как основа шкалирования.**

В соответствии со ставшими уже классическими принципами репрезентационной теории измерений понятие шкалы предполагает наличие некоторых ЭСО и ЧСО и означает гомоморфное отображение первой во вторую. Поясним сказанное на примере (нами уже были предприняты попытки пояснить соответствующие принципы для читателя-гуманитария [11, 12]).

Наличие ЭСО означает, что исследователь, во-первых, выделил некоторое множество эмпирических объектов, скажем, совокупность работников какого-либо предприятия (это - множество-носитель ЭСО); и, во-вторых, предположил, что между этими объектами имеется некоторое отношение порядка. Например, исследователь может полагать, что у каждого из рассматриваемых людей имеется определённое эмоциональное состояние, которое можно назвать его удовлетворённостью работой. Если мы полагаем, что имеет смысл только сходство (различие) таких состояний, то говорим о наличии отношения равенства (неравенства) между людьми и можем поставить задачу построения номинальной шкалы. Если же у нас имеются основания считать осмысленной какую-либо упорядоченность рассматриваемых эмоциональных состояний отдельных людей, то мы вводим в рассмотрение ещё и отношение порядка, получая тем самым основание претендовать на построение порядковой шкалы. Для постановки задачи построения интервальной шкалы требуется обоснование содержательной осмысленности отношения порядка для расстояний между объектами. Мы упомянули только три наиболее “любимых” социологами типа шкал. В действительности существует бесконечное количество типов шкал. Конечно, не все они актуальны для социологии, но

не только названные три типа играют существенную роль при решении многих проблем социологического измерения.

Измерение указанных выше отношений означает нахождение способа выяснения для любых двух объектов, равны они или нет (для номинальной шкалы), какой из них больше, а какой меньше (для порядковой), расстояние между какими парами объектов больше, между какими - меньше (для интервальной шкалы) и т.д. Делается это обычно путём сопоставления с каждым объектом некоторого числа (разным респондентам может отвечать одно и то же число) - его шкального значения. При этом предполагается, что известные числовые отношения между шкальными значениями говорят нам о соответствующих эмпирических отношениях между объектами. Для такого измерения, как правило, используется какой-либо из известных методов шкалирования. Например, задаётся “лобовой” вопрос: “Насколько Вы удовлетворены ... ?” с традиционным веером ответов от “Совершенно не удовлетворён” (код 1) до “Вполне удовлетворён” (код 5). Подчеркнём, что, задавая это вопрос и полагая, что используемая шкала - порядковая, мы тем самым выражаем свою уверенность в том, что, скажем, если один человек отметил код 2, а другой - код 4, то рассматриваемые эмоциональные состояния этих людей отличны друг от друга и первый человек менее удовлетворён работой, чем второй. Опишем тот же процесс в несколько иных терминах.

Говоря об интересующем нас отношении порядка, мы делаем шаг по направлению к “мышлению признаками” [7]: полагаем, что рассматриваемые объекты (люди, потенциальные респонденты) обладают каким-то общим свойством (в данном случае - неким эмоциональным настроением, который можно назвать удовлетворённостью), по-разному проявляющимся у разных людей. То, что является общим для всех объектов, и называется признаком. Конкретные проявления этого общего посредством реализации процесса измерения выражаются в виде значений признака. Таким образом, мы полагаем, что, хотя само понятие

признака зарождается в эмпирической системе, значения его являются элементами математической системы.

Если в ЭСО мы выделяем лишь некоторые “смутно” понимаемые отношения между рассматриваемыми объектами, то, моделируя ЭСО посредством измерения, мы превращаем эти отношения во всем знакомые числовые отношения между значениями рассматриваемого признака.

Всё сказанное относится к тому случаю, когда и ЭСО, и ЧСО отвечают одному признаку. Однако в действительности практически всегда социолог использует многомерное признаковое пространство. Чтобы говорить об этом в “измерительном” ракурсе необходимо вспомнить, что задачи практики уже давно потребовали от теории измерений определённых обобщений тех классических представлений, о которых шла речь выше. В [12] отмечалось, что жизнь заставляет социолога отказаться и от принятия в расчёт только тех эмпирических отношений, которые значимы для трёх упомянутых типов шкал, и вообще от задания эмпирической системы (ЭС) в виде системы с отношениями, и от понимания шкалы как гомоморфизма, и от трактовки измерения как отображения реальности именно в числовую систему, а не в произвольную математическую (с отношениями или без - соответственно, МСО и МС) (в частности, понятие признака тоже может быть связано с нечисловыми измерениями, скажем, это понятие может использоваться при построении модели малой группы в виде графа). Развиваясь, репрезентационная теория измерений заставляет нас выделить главное в ней - понимание измерения как построения модели ЭС (ЭС нельзя назвать “чистым” фрагментом реальности; при её формировании уже присутствуют аспекты моделирования, например, в процессе выделения в реальности интересующих нас отношений между объектами; впрочем, и выделение самих объектов тоже может рассматриваться как фрагмент процесса моделирования) с помощью элементов некой МС.

В настоящей статье нас будет интересовать лишь один вид нечисловой МС - так называемое признаковое пространство, точки которого отвечают объектам (например, респондентам), а координаты этих точек определяются значениями рассматриваемых признаков для этих объектов (скажем, ответам респондентов на вопросы анкеты; одна координатная ось пространства в таком случае отвечает одному анкетному вопросу). Коснёмся ниже мы и одной из причин, мешающей представлению ЭС как ЭСО и, следовательно, - построению шкалы на основе прямого гомоморфного отображения ЭСО в какую-нибудь МСО. Эта причина - аксиоматическое задание ЭС.

Итак, осуществление измерения - это окончательный переход к мышлению признаками (это положение имеет место, даже если мы используем какие-либо сложные и тонкие методы измерения, опираясь на данные, не являющиеся непосредственным отображением каждого респондента в точку признакового пространства; так, для сбора данных можно использовать метод парных сравнений; но в итоге мы всё равно приходим к признаковому пространству). Далее, анализируя полученные в результате измерения данные (ведь измерение, как правило, не является самоцелью, а требуется для анализа полученных с его помощью результатов, для поиска интересующих исследователя закономерностей, “скрывающихся” в исходных данных), мы будем получать “содержательные” выводы, формулируемые в терминах признаков: “такие-то два признака имеют сильную статистическую связь”, “второй признак линейно зависит от седьмого” и т.д.

Как известно, получение такого рода утверждений для социолога практически всегда представляет собой довольно сложную проблему. Так, для решения любой задачи всегда “на вооружении” у исследователя имеется множество методов. Разные методы, вообще говоря, приводят к разным результатам (как известно, только для измерения связи между двумя признаками существует более сотни коэффициентов). Встаёт во-

прос о выборе конкретного метода, либо о сравнении результатов применения разных методов друг с другом и т.д. Или другой аспект проблемы: условия применимости конкретных методов очень часто не соблюдаются. Актуальной становится задача выявления того, в каких масштабах соответствующие нарушения могут иметь место, не приводя при этом к слишком сильной неадекватности результатов. И т.д., и т.д.

В силу сказанного, представляется полезной постановка вопроса о том, нельзя ли каким-либо конструктивным способом описать всё множество содержательных выводов, которые могут быть получены для конкретной совокупности полученных социологом данных. И оказывается, что этот вопрос не бессмыслен, на него существует положительный ответ. Именно об этом по существу и идёт речь в статье Е.Е.Витяева и А.Д.Логвиненко. Рассмотрим более подробно некоторые её аспекты, попытавшись параллельно показать, как эти аспекты вписываются в некоторую более широкую картину, описывающую современное положение дел с анализом социологических данных.

Остановимся на номинальных данных - наиболее распространённых в социологии и актуальных для неё.

Обоснуем возможность описания основных интересующих социолога закономерностей на языке исчисления предикатов первого порядка.

## **II. Общий вид утверждений, получаемых с помощью известных методов анализа номинальных данных.**

Прежде всего уточним, что значит делать какие-то выводы в терминах рассматриваемых признаков. Вероятно, исходя из здравого смысла, подобные выводы должны иметь вид:

“5-е значение 8-го признака часто встречается с 3-м значением 14-го и 1-м значением 2-го”, “из того, что 3-й признак принимает 2-е значение одновременно с тем, что 4-й принимает 5-е значение, как правило, следует, что 6-й признак принимает либо 2-е, либо 3-е”, “из того, что 3-й признак принимает какое-либо значение, кроме 2-го, следует, что 7-й признак принимает 4-е значение” и т.д. (надеюсь, что для понимания сказанного не требуется более конкретно формулировать подобные утверждения: скажем, указывать, что 3-й признак - это возраст, его 5-е значение - указание того, что возраст конкретного респондента заключён в интервале от 35 до 40 лет и т.д.).

Действительно, именно такой вид по существу имеют выводы, получающиеся с помощью известных способов анализа номинальных данных. Покажем это. А именно, попытаемся продемонстрировать, что практически все методы направлены на поиск того, что можно назвать взаимодействием, по аналогии с тем, как этот термин используется в математической статистике (в частности, в дисперсионном анализе).

Напомним, что использование этого термина предполагает выделение среди всех признаков главного (зависимого, выходного, признака - функции, признака-следствия) и группы детерминирующих его признаков (независимых, входных, признаков - аргументов, признаков-причин, предикторов). “Взаимодействие” в дисперсионном анализе (описание дисперсионного анализа можно найти, например, в [9]) означает сочетание значений независимых признаков, определяющих тот или иной уровень зависимого. При этом в дисперсионном анализе зависимый признак предполагается количественным, то есть таким, значения которого, получены, по крайней мере, по интервальной шкале.

Роль поиска взаимодействий в эмпирической социологии вряд ли можно преувеличить. Однако представляется, что потребность практики делает целесообразным расширение этого понятия:

- во-первых, будем считать, что таковым может быть не только конъюнкция выражений типа “значение второго признака равно 3”, а любые логические функции от таких выражений (предполагаем, что читатель знает определение основных логических функций - конъюнкции, дизъюнкции, импликации, отрицания);

- во-вторых, будем полагать, что наше взаимодействие может определять не только значение непрерывного признака (как в дисперсионном анализе), но и значение, либо сочетание значений любых дискретных признаков, частоту в таблице сопряжённости, а может и ничего не определять, но тогда естественно требовать просто истинность взаимодействия как логической функции.

При таком понимании взаимодействия можно сказать, что поиск взаимодействий разного рода служит основой всех известных методов анализа номинальных данных. Приведем примеры.

(Для систематизации всех методов анализа номинальных данных можно прибегнуть к их классификации, предложенной в [13]. В её основе по существу лежит именно представление о разных видах взаимодействий. Она основывается на том, что мы как бы “рассыпаем” все рассматриваемые признаки на отдельные альтернативы, а далее выделяем группы методов в зависимости от того, каким образом эти альтернативы “склеиваются” в тех закономерностях, которые ищутся при реализации метода.)

Детерминационный анализ [15]) позволяет говорить о влиянии какого-либо одного значения (альтернативы) рассматриваемого признака на произвольное значение некоторого другого признака (даются ответы на вопросы типа: насколько интенсивно учителя читают Учительскую газету? В какой мере аудитория Учительской газеты состоит из учителей?).

Анализ фрагментов таблиц сопряжённости [3, 18] дает возможность в данной двумерной таблице сопряженности

выделить такие подтаблицы, которые как бы “оттягивают” на себя основную долю общей связи исходных признаков (то есть найти такие сочетания значений рассматриваемых признаков, которые в определенном смысле детерминируют друг друга). (Скажем в [3] показывается, что наличие сильной связи между типом трехпоколенной семьи и желательным вариантом расселения объясняется тем, что молодые матери-одиночки в гораздо меньшей мере, чем другие молодые семьи, желают разъезжаться с бабушками-дедушками, а брачные пары в возрасте от 30 до 40 лет - наоборот, гораздо в большей степени, чем другие, желают расселиться).

Методы последовательных разбиений (типа AID: AID3, THAID, CHAID) [3, 17, 19, 20] позволяют находить сочетания значений разных признаков, детерминирующих по-разному трактующее поведение респондента: оно может быть задано одним или несколькими значениями рассматриваемых признаков, долей, уровнем числового признака и т.д. (решаются задачи типа: какими социально-демографическими характеристиками можно описать респондентов, проголосовавших на выборах за кандидата N; выявить сочетание ответов на вопросы анкеты, характерные для людей с ежемесячным доходом свыше 100.000 рублей).

Логлинейный анализ [1] дает возможность находить такие сочетания значений рассматриваемых признаков, которые определяют отклонение частот таблицы сопряженности от “средних”, отвечающих равномерному распределению.

Методы поиска логических закономерностей [4] позволяют выявлять логические функции от сочетаний значений разных признаков, истинные для рассматриваемых совокупностей объектов.

Заметим, что упомянутые в п. I типичные для “количественной” статистики утверждения вида “такие-то два признака имеют сильную статистическую связь”, “второй признак линейно зависит от седьмого” и т.д. в случае номинальных дан-

ных тоже имеют описанный выше вид. Скажем, наличие сильной статистической связи между двумя признаками всегда означает, что те или иные значения одного признака всегда встречаются в сочетании с определёнными значениями другого. Другими словами, связь всегда опирается на прогнозную модель: при наличии связи между двумя признаками знание значения одного признака даёт в то же время информацию о возможных значениях другого. Это верно даже при том понимании связи, которое лежит в основании коэффициентов, опирающихся на известный критерий “Хи-квадрат” [6] (это утверждение неявно использовалось нами при обсуждении в предыдущем абзаце методов анализа фрагментов таблицы сопряжённости).

Можно говорить и о форме зависимости между номинальными признаками. Как известно, так называемая дихотомизация номинальных данных позволяет применять к ним любой количественный метод (правда, этот подход имеет только эвристическое обоснование, но, тем не менее, весьма популярен; практика подтверждает его целесообразность). Скажем, можно использовать регрессионную технику со всеми присущими ей особенностями. Возможен поиск как линейной, так и нелинейной зависимости. И поиск нелинейной зависимости по существу является поиском тех же взаимодействий [2].

Таким образом, можно сказать, что большинство известных методов анализа номинальных данных приводит нас к утверждениям того типа, о котором шла речь в начале параграфа. Нетрудно видеть, что подобные утверждения могут быть легко переведены на язык исчисления предикатов первого порядка (об этом исчислении см., например, [6]). Поговорим об этом подробнее, но сначала напомним язык такого исчисления.

### **III. Описание языка узкого исчисления предикатов.**

Опишем соответствующий алфавит для рассматриваемого случая. Прежде всего - о нелогических символах формализованного языка.

Предметные (индивидные) константы: конкретные номера респондентов, для обозначения которых могут использоваться буквы  $a, b, c, \dots$ . Предметные (индивидные) переменные - обозначения произвольных номеров респондентов:  $x, y, z, \dots$ .

$p$  - местные предикатные константы: одноместные - "для респондента  $x$  рассматриваемый признак принимает такое-то значение". Примеры: "возраст человека  $x$  лежит в интервале от 35 до 40 лет"; "возраст человека  $x$  лежит в интервале от 15 до 20 лет"; "профессия респондента  $x$  - врач"; "профессия респондента  $x$  - учитель" и т.д.; двуместные - "профессия респондента  $x$  не совпадает с профессией респондента  $y$ ", "респондент  $x$  читает те же газеты, что и респондент  $y$ ".

Понятие формулы определяется рекурсивно:

1) любая предикатная константа  $P(x), P(x,y), P(x,y,z), \dots$  является формулой;

2) если  $A$  - формула, то  $\neg A$  - тоже формула;

3) если  $A$  и  $B$  - формулы, то  $A \& B, A \vee B, A \supset B$  - тоже формулы;

4) если  $A$  - формула и  $a$  - предметная переменная, то  $\forall a$  и  $\exists a A$  - формулы;

5) ничто иное, кроме перечисленного в п.п. 1)-4), формулой не является.

Будем считать, что читателю известно, как определяется истинность логических формул в обычной классической двузначной логике, понятием которой мы будем далее пользоваться.

**IV. Возможность рассмотрения результатов, получаемых с помощью известных методов анализа номинальных данных, как формул в языке узкого исчисления предикатов.**

Нетрудно показать, что большинство утверждений, подобных сформулированным в начале п. II, означают истинность определённых формул в нашем исчислении предикатов. Пусть, например, предикат (предикатная константа)  $P(x)$  означает “респондент  $x$  отметил 5-е значение 8-го признака”, предикат  $Q(y)$  - “респондент  $y$  отметил 3-е значение 14-го признака”, а предикат  $R(z)$  - “респондент  $z$  отметил 1-е значение 2-го признака. Тогда приведённое выше утверждение “5-е значение 8-го признака, как правило, встречается либо с 3-м значением 14-го, либо с 1-м значением 2-го” будет означать, что почти для всех  $x$  будет истинной формула  $(P(x) \& (Q(x) \vee R(x)))$ .

Теперь предположим, что  $P(x)$  означает “респонденту  $x$  отвечает 2-е значение 3-го признака”,  $Q(x)$  - “респонденту отвечает 5-е значение 4-го признака”,  $R(x)$  - предикат “значение 6-го признака для респондента  $x$  равно или 2, или 3”. Тогда выражение “из того, что 3-й признак принимает 2-е значение одновременно с тем, что 4-й принимает 5-е значение, как правило, следует, что 6-й признак принимает либо 2-е, либо 3-е,” и т.д., означает, что почти для всех  $x$  будет истинно выражение  $((P(x) \& Q(x)) \supset R(x))$ .

Пусть  $S(x)$  - “значение 23-го признака для респондента  $x$  равно 2”,  $T(x)$  - “значение 7-го признака для респондента  $x$  равно 4”. Тогда утверждение “из того, что 23-й признак принимает какое-либо значение, кроме 2-го, следует, что 7-й признак

принимает 4-е значение” будет эквивалентно утверждению истинности формулы  $(\neg(S(x)) \supset T(x))$ .

( $\supset$  - импликация,  $\&$  - конъюнкция,  $\vee$  - дизъюнкция,  $\neg$  - отрицание.)

Нетрудно видеть, что таким образом в виде формул узкого исчисления предикатов действительно можно выразить очень многие интересующие социолога “закономерности”, “скрывающиеся” в эмпирических данных. А если учесть, что большинство методов анализа номинальных данных, как было показано в предыдущем параграфе, позволяет выявлять “закономерности” именно такого вида, то можно сказать, что практически все интересующие социолога закономерности выражаются на языке формул исчисления предикатов первого порядка.

Обсудим вопрос о том, что характеризуют наши истинные формулы: ЭС или МС. В расширенной теории измерений считается, что ЭС. Именно это утверждается в рассматриваемой статье Е.Е.Витяева и А.Д.Логвиненко. Но нам представляется, что такой подход нецелесообразен. Поясним это.

Выше мы связали переход от ЭС к МС с переходом к “мышлению признаками”. Этот процесс для социологии является сложным, многоступенчатым и в то же время весьма принципиальным. Он состоит из ряда шагов, включающих формирование понятий (идеализированных объектов, идеальных типов), переход к моделирующим эти понятия показателям и индикаторам (так называемой операционализации понятий) и т.д.

Все недостатки анкетных опросов, по большому счёту, связаны с некорректностью упомянутого перехода. Представляется, что адекватное его сложности описание шагов построения ЭС и МС пока ещё не разработано. Здесь мы не будем говорить об этом более подробно. Однако заметим, что целесообразно отнести к области построения ЭС не весь процесс формирования представлений о признаке. Ту его часть, которая касается непосредственного отображения в МС, то есть непосредственного указания того, какое шкальное значение должно

отвечать тому или иному содержательно понимаемому состоянию шкалируемого объекта, вероятно, целесообразно считать связанной с построением МС.

Но вряд ли можно связывать истинность наших формул и с самой МС. Эта система - фрагмент математики. И мы не должны эту систему “затуманивать”, приписывая ей качества, не предполагаемые известными математическими положениями. Наши предикаты не принадлежат используемой для выводов МС.

Представляется целесообразным помимо ЭС и МС ввести в рассмотрение ещё одну конструкцию - ММСЭД: математическую модель структуры эмпирических данных. Она была введена нами в [5]. ММСЭД - это в точности тот фрагмент используемой МС, который является моделью ЭС. Его элементы могут обладать не всеми свойствами МС и в то же время обладать такими свойствами, которые отражают структуру ЭС, но не обусловлены свойствами МС (последнее как раз и иллюстрируется справедливостью для ММСЭД какого-либо логического утверждения из числа определенных выше).

## **V. Конструктивный подход к описанию всех закономерностей рассматриваемого вида.**

Естественным представляется желание конструктивно описать множество всех “закономерностей”, неявно содержащихся в собранной статистической информации. Но для этого необходимо ответить, по крайней мере, на два вопроса.

Первый - о систематизации упомянутого множества. Перебрать все формулы указанного вида не под силу даже современным ЭВМ.

Второй вопрос - о понимании того, что означает истинность формулы не для всех, а лишь “почти для всех” переменных.

Ответы на эти вопросы можно получить из статьи Е.Е.Витяева и А.Д.Логвиненко. Они базируются на двух моментах, к краткому обсуждению которых мы и перейдём (попытаемся не повторять содержание статьи, а дополнить его некоторыми соображениями, представляющимися, на наш взгляд, полезными для социолога).

Расширение понятия эмпирической системы.

Выше мы уже упоминали о том, что практика заставила исследователей, работающих в области теории измерений, отойти от тех строгих определений шкалы, которые были введены в первые десятилетия существования названной теории и отражённые, например, в [10]). Не последнюю роль в соответствующем процессе сыграло развитие представлений об ЭС, связанных с её аксиоматизацией [16]. В рамках этого направления и лежит статья Е.Е.Витяева и А.Д.Логвиненко. Опишем более подробно суть этого направления.

В первые годы после институционализации интересующей нас теории измерений данный подход к измерению называли репрезентационным (поскольку, в соответствии с положениями этой теории, под измерением понимается некая репрезентация, представление ЭС в МС). В настоящее время говорят об аксиоматическо - репрезентационном подходе: предполагается, что ЭС устанавливается не только определённым набором отношений, заданных на элементах её множества - носителя, но и некоторой совокупностью утверждений (аксиом), которым эти отношения должны удовлетворять. Предполагается, что при формулировке этих аксиом могут использоваться логические языки. Например, язык узкого исчисления предикатов. Один из возможных примеров такого использования был приведён выше.

Е.Е.Витяев и А.Д.Логвиненко вполне справедливо отмечают, что априорное теоретическое выявление аксиом, задающих ЭС, - дело чаще всего невозможное, и более перспективной

выглядит попытка нахождения этих аксиом эмпирическим путём.

Предполагается, что в используемой ММСЭД определены некоторые предикатные константы (не уточняется, какие именно; примером могут служить предикаты, описанные выше) и определён язык узкого исчисления предикатов. Среди гипотетического множества формул, истинных на совокупности шкальных значений (являющихся элементами МС), выделяется такое их подмножество, элементы которого, собственно говоря, и следует назвать аксиомами: показывается, что остальные истинные формулы из них выводимы. Совокупность аксиом задана достаточно конструктивно (автор называет их аксиомами ЭС; учитывая наше определение ММСЭД, можно сказать, что эти аксиомы действительно характеризуют ЭС, если принимать во внимание их “дочисловой” вариант).

В статье показывается, каким способом аксиомы могут быть проверены для каждой конкретной ситуации. При этом строго, посредством задания определённой теоретико-вероятностной конструкции, определяется, в каких случаях аксиому можно считать выполненной “почти для всех” объектов совокупности.

Таким образом, рассматриваемая статья практически предлагает своеобразный поиск “ядра” всех истинных для изучаемой ММСЭД формул. Все остальные истинные формулы будут из них выводимы.

#### *Некоторые аспекты понятия закономерности.*

Авторы показывают, что рассматриваемые ими аксиомы удовлетворяют некоторым свойствам, которые позволяют считать их законами изучаемой ЭС, и подробно обсуждают целесообразность такого понимания закона. Некоторые аспекты этого понимания значимы не только для рассматриваемой в статье ситуации. Поговорим о них подробнее.

Е.Е.Витяев и А.Д.Логвиненко связывают понятие закона с простотой и неупрощаемостью и показывают, как такое представление о законе реализуется в тех случаях, когда в качестве последнего фигурирует та или иная формула узкого исчисления предикатов. Они строго определяют эти свойства и в качестве аксиом выбирают именно такие формулы, которые можно назвать законами.

Нам представляется, что аналогичные соображения должны использоваться в социологии гораздо чаще: практически всегда, когда мы имеем дело с анализом данных. Поясним это. При этом будем говорить не о законе, а о закономерности как о чём-то менее значимом (здесь не будем обсуждать эти понятия). Результаты применения метода всегда могут рассматриваться как выявление некой закономерности, характеризующей изучаемую совокупность объектов.

Почти каждый метод анализа данных предоставляет исследователю массу вариантов решения поставленной задачи. Окончательный выбор остаётся за социологом. И осуществить его бывает весьма непросто. Какую размерность искомого латентного пространства выбрать при многомерном шкалировании? Какой шаг разбиения выбрать при использовании какого-либо из алгоритмов многомерной классификации? Конечно, в каждом алгоритме предусматриваются некоторые критерии, позволяющие говорить о качестве полученного решения (значение функции стресса в многомерном шкалировании, оптимизируемый функционал в алгоритмах классификации). Тем не менее, мы, как правило, не можем этими критериями ограничиться. В силу вступают другие, неформализованные соображения, связанные с возможностью человека воспринять полученные результаты как некие закономерности. Подобные соображения как раз и связаны со сложностью этих результатов, и, в первую очередь, с количеством описывающих их характеристик.

Отвечая на вопросы типа тех, которые были сформулированы выше, мы стремимся к тому, чтобы количество параметров, характеризующих решение, было бы не очень большим. Если многомерное шкалирование приведёт нас к 70-мерному пространству, мы вряд ли будем такой результат расценивать как нахождение некоторой закономерности, предпочтём искать оную путём снижения размерности пространства. По той же причине мы не воспримем результаты классификации как построение содержательной типологии объектов, если количество типов превысит, скажем, 50. Алгоритмы типа AID всегда предусматривают останов в том случае, если длина описания каждого взаимодействия будет слишком большой и т.д.

По существу мы здесь имеем дело с психологическими аспектами представлений исследователя о том, что такое “закономерность”. Насколько нам известно, в литературе таким аспектам уделяется мало внимания (хотя в когнитологии имеются работы, посвящённые, скажем, психологической стороне восприятия человеком разных видов графического представления статистических данных [8]). Поэтому работа Е.Е.Витяева и А.Д.Логвиненко и с этой точки зрения представляется интересной.

В заключение мы хотим отметить, что конструктивность всех содержащихся в рассматриваемой статье положений является следствием опоры на строгие определения всех рассматриваемых понятий и корректные доказательства всех формулируемых утверждений. Весьма вероятно, что математическая корректность статьи может сделать её чтение тяжёлым для читателя - гуманитария. Тем не менее, мы глубоко убеждены в пользе её публикации именно в предлагаемом виде. Объясняется это, на наш взгляд, двумя причинами. Во-первых, социолог должен учиться понимать, какого вида рассуждения сопряжены с корректным использованием в социологии математического аппарата. Во-вторых, мы полагаем, что читателями нашего

журнала являются в числе прочих и люди с профессиональным математическим образованием, заинтересованные в применении своих знаний к социологии.

#### ЛИТЕРАТУРА

1. *Аптон Г.* Анализ таблиц сопряжённости. М.:Финансы и статистика, 1982.
2. *Аргунова К.Д.* Качественный регрессионный анализ в социологии. М.:ИСАН СССР, 1990.
3. Интерпретация и анализ данных в социологических исследованиях. М.:Наука, 1987
4. *Лбов Г.С.* Методы обработки разнотипных экспериментальных данных. Новосибирск:Наука, 1981.
5. Логика социологического исследования. М.:Наука, 1987. С. 104-138
6. *Миркин Б.Г.* Группировки в социально-экономических исследованиях. М.:Финансы и статистика, 1985.
7. *Ноэль Э.* Массовые опросы. Введение в методику демоскопии. М.:Ава-эстра, 1993.
8. *Плотинский Ю.М.* Визуализация информации. М.:Изд-во МГУ, 1994.
9. Статистические методы анализа социологической информации. М.:Наука, 1979.
10. *Суппес П., Зинес Дж.* Основы теории измерений//Психологические измерения. М.:Мир, 1967.
11. *Толстова Ю.Н.* Методология математического анализа данных//Социологические исследования, 1990. №6. С.77-87.
12. *Толстова Ю.Н.* Математика в социологии: элементарное введение в круг основных понятий (измерение, статистические закономерности, принципы анализа данных). М.:ИСИ АН СССР, 1990
13. *Толстова Ю.Н.* К вопросу о принципах построения учебного курса “Анализ социологических данных”//Социология: методология, методы, математические модели, 1993-1994. №3-4.
14. *Толстова Ю.Н.* Обобщенный подход к определению понятия социологического измерения//Методология и методы социологических исследований. М.:ИСРОСАН, 1996. С.66-95.

15. *Чесноков С.В.* Детерминационный анализ в социально-экономических данных. М.:Наука, 1982.
16. *Krantz D.H., Luce R.D., Suppes P., Tversky A.* Foundation of Measurement. V.3. N.Y. - L.:Acad. Press, 1990.
17. *Sonquist J., Baker E.L., Morgan J.N.* Searching for Structure. Ann arbor:Institute for Social research, 1973.
18. *Ростовцев П.С.* Черно-белый анализ связи переменных// Социология: методология, методы, математические модели, 1998. №10.
19. *Agresti A.* Categorical data analysis. N.-Y.:John Wiley and sons, 1990.
20. *Derrick F., Magidson J.* Using CHAID with the gains chart option//Proceedings of the 1992 annual meeting of American stat.Ass. Business and Economics Section, 1992.