

Государственный Университет -
Высшая Школа Экономики

Факультет социологии

Кафедра сбора и анализа социологической информации

Выпускная квалификационная работа

Тема: Возможности использования прямого ДСМ – метода автоматического порождения гипотез для анализа социологических данных

Выполнил студент 532 гр.
Земсков Александр Васильевич

Научный руководитель: доктор технических наук, профессор
Финн Виктор Константинович

Москва 2006

ВВЕДЕНИЕ.....	3
РАЗДЕЛ 1. ТЕОРИЯ И МЕТОДОЛОГИЯ.....	6
ПРЕДПОСЫЛКИ ФОРМИРОВАНИЯ МЕТОДОЛОГИИ.....	6
<i>Эволюционная эпистемология Карла Поппера.....</i>	8
<i>Идея трех миров Карла Поппера.....</i>	9
<i>Закон всеобщей причинной связи по Д.С. Миллю.....</i>	11
<i>Основные методы индуктивного исследования по Д.С. Миллю.....</i>	13
<i>А.С. Есенин-Вольпин и философия неологизма.....</i>	16
НЕКОТОРЫЕ МЕТОДОЛОГИЧЕСКИЕ ПРОБЛЕМЫ СОЦИАЛЬНЫХ НАУК.....	19
<i>Проблема формализации идей в социальной науке.....</i>	19
<i>Проблема измерения и проблема познания.....</i>	22
<i>Синтез познавательных процедур.....</i>	24
РАЗДЕЛ 2. НЕФОРМАЛЬНОЕ ИЗЛОЖЕНИЕ ДСМ-МЕТОДА АПГ. СПЕЦИФИКА И	
ВОЗМОЖНОСТИ ЕГО ИСПОЛЬЗОВАНИЯ.....	25
ДСМ МЕТОД АПГ В НЕФОРМАЛЬНОМ ИЗЛОЖЕНИИ.....	25
<i>Специфика представления данных для работы с ДСМ методом АПГ.....</i>	25
<i>Общие принципы работы ДСМ метода АПГ.....</i>	27
<i>Возможности анализа социального поведения с помощью ДСМ-метода АПГ.....</i>	35
ПРОГРАММНАЯ РЕАЛИЗАЦИЯ ДСМ-МЕТОДА АПГ.....	38
<i>Обзор программы JSM-explorer.....</i>	38
<i>Особенности и ограничения метода.....</i>	45
РАЗДЕЛ 3. ВОЗМОЖНОСТИ ИСПОЛЬЗОВАНИЯ ДСМ-МЕТОДА АПГ В СОЧЕТАНИИ С	
МЕТОДОМ ВИНЬЕТОК (НА ПРИМЕРЕ АНАЛИЗА ВОСТРИЯТИЯ РЕЛИГИОЗНОСТИ).....	46
МЕТОД ВИНЬЕТОК КАК МЕТОД СБОРА ДАННЫХ.....	46
<i>Общее представление о методе виньеток.....</i>	46
<i>Описание условий эксперимента.....</i>	48
АНАЛИЗ ДАННЫХ.....	51
<i>Задача поиска существенных взаимодействий.....</i>	51
<i>Процедуры и результаты анализа данных.....</i>	52
<i>Содержательная интерпретация результатов.....</i>	60
МЕТОДОЛОГИЧЕСКИЕ ВЫВОДЫ.....	64
СПИСОК ЛИТЕРАТУРЫ:.....	65
ПРИЛОЖЕНИЯ.....	67

Введение

Одна из задач науки – поиск скрытых закономерностей, связывающих явления действительности. Научный метод представляет собой набор принципов, позволяющих выявлять эти закономерности в процессе наблюдения за окружающей действительностью и в процессе взаимодействия с ней. Под окружающей действительностью в данном случае можно понимать любую сферу деятельности человека, а также любую часть физического и социального мира.

Говоря о процессе познания окружающей действительности, следует обратить внимание на взаимосвязь трех философских подсистем: онтологии, эпистемологии и логики. Онтология - это учение об общих категориях и закономерностях бытия. Эпистемология является учением о процессе познания этих закономерностей, в результате которого появляется независимое от самой природы знание как особая форма бытия. Логика же представляет собой учение о законах, принципах и формах мышления, с помощью которых осуществляется процесс познания. Таким образом, логика является связующим звеном между онтологией и эпистемологией, а также представляет собой основу, на которой строится методология науки.

Методология – учение о методе научного познания.

Метод – это набор техник и логических процедур, указывающих на то, как следует наблюдать, соотносить и сравнивать факты в разных сферах окружающей действительности для получения знания о них.

Методология включает в себя набор общих принципов познавательной деятельности, и построенную на основе этих принципов систему различных методов и техник, соответствующих различным целям познания и ситуациям, в которых осуществляется познание. Основная задача методологии состоит в решении проблемы адекватности научных методов исследуемой действительности, а именно, в поиске и разработке наиболее подходящих методов для решения тех или иных задач в различных предметных областях.

Данная работа методологическая. Ее цель – продемонстрировать некоторые возможности применения ДСМ метода Автоматического Порождения Гипотез (АПГ)¹ для анализа социологических данных разного рода. ДСМ - метод АПГ представляет собой

¹ Метод впервые был предложен В.К. Финном в конце 80-х годов XX века. Метод был назван в честь Джона Стюарта Милля (ДСМ).

логики-комбинаторный метод поиска причинно-следственных зависимостей в базах данных с неполной информацией.²

Описываемая в данной работе методология способна существенно дополнить традиционно используемую в социологии статистическую методологию, особенно в отношении предварительного анализа социологических данных, решения задач классификации и поиска причинно-следственных зависимостей.

Данная работа представляется нам актуальной в силу следующих причин:

1) На данный момент было проведено достаточно мало экспериментов, связанных с применением ДСМ – метода к анализу социологических данных, что само собой подтверждает актуальность новых экспериментов в данном направлении.

2) Ранее метод применялся для выявления детерминант социального поведения, анализа электорального поведения населения и автоматизированного анализа мнений.³ В данной работе представлены результаты работы по изучению восприятия религиозности с помощью ДСМ метода АПГ.

3) В качестве метода сбора информации в ранее проведенных экспериментах использовался метод формализованного интервью. При этом анкета состояла преимущественно из закрытых вопросов. В данной работе представлены результаты работы ДСМ - метода на данных, собранных с помощью метода виньеток.

4) Имеющиеся на данный момент работы по ДСМ–методу были написаны в большинстве своем специалистами в области искусственного интеллекта. Это приводит к большим трудностям интерпретации данных работ социологами. Поэтому на наш взгляд более «социологическое» изложение данной методологии и наглядная демонстрация работы метода на социологических примерах также является достаточно актуальной задачей.

Работа содержит три основные главы:

1. В первой главе подробно рассматриваются теоретические предпосылки формирования методологии ДСМ - рассуждений. Также проводится анализ некоторых методологических проблем изучения социальной реальности и возможности решения некоторых из них с помощью данной методологии.

² Базой данных (БД) с неполной информацией можно считать любую базу данных о фактах какой либо предметной области, в которой представлены не все факты, подлежащие изучению, а также нет исчерпывающего описания этих фактов. Примером такой БД может служить выборочное социологическое исследование, в котором представлена лишь часть генеральной совокупности объекта исследования лишь по некоторым, наиболее важным для исследования параметрам.

³ Гусакова С.М., Михеенкова М.А., Финн В.К. О логических средствах автоматизированного анализа мнений // НТИ. Сер. 2. – 2001. – № 5. – с. 4-24.

2. Вторая глава посвящена непосредственно описанию ДСМ метода АПГ, его неформальному изложению, описанию особенностей и ограничений метода. В данной главе также приведены наглядные примеры работы метода.

3. Третья глава содержит описание и результаты эксперимента по анализу восприятия религиозности с помощью ДСМ - метода АПГ на данных, собранных с помощью метода виньеток.

Так как работа методологическая, мы уделяем значительно больше внимания описанию методологических и технических аспектов поведенных экспериментов, нежели содержательному анализу результатов. Более подробный содержательный анализ результатов можно найти в дипломной работе Добрыниной Стеллы⁴, являющейся автором программы по исследованию восприятия религиозности, а также осуществлявшей сбор данных для анализа.

⁴ Добрынина Стелла, Использование методов виньеток и репертуарных решеток в исследовании обыденного восприятия традиционных и новых форм религиозности // выпускная квалификационная работа, рукопись, ГУ-ВШЭ 2006

Раздел 1. Теория и методология

Предпосылки формирования методологии

ДСМ-метод автоматического порождения гипотез был разработан и реализован в рамках программы под названием «искусственный интеллект» и является средством, существенно усиливающим интеллектуальные возможности исследователя. ДСМ система родственна интеллектуальным системам, обладающим средствами индуктивного обучения. Среди таких систем можно выделить системы машинного обучения, системы распознавания образов и т.п. Наиболее популярными системами подобного рода являются системы DENDRAL, MYCIN, PROSPECTOR, PLANT.⁵

В приложении к социальным наукам интеллектуальные системы призваны решать следующие типы задач:

- Логическая систематизация знания,
- Формализация рассуждений,
- Уточнение понятий,
- Формулирование концепций⁶ и их коррекция.⁷

При этом интеллектуальные системы предполагают тесное взаимодействие с исследователем, ибо они являются лишь средствами, усиливающими и дополняющими интеллектуальную работу человека в плане автоматизации сложных логических рассуждений.

ДСМ-метод в контексте анализа данных представляет собой метод построения общей теоретической концепции относительно изучаемой базы данных, что включает в себя сравнение объектов по всем имеющимся в базе параметрам, их классификацию, поиск причинно-следственных зависимостей и выдвижение гипотез относительно причин (в социологии уместнее применять термин «детерминанты социальных действий, мнений и установок»). Таким образом, применение ДСМ-метода полезно на этапе построения

⁵ Подробнее см. *Финн В.К.* Гуманитарное знание и искусственный интеллект. / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 38 - 57

⁶ Под концепцией в данном контексте следует понимать систему аксиом и логических выводов, объясняющих некоторый набор фактов.

⁷ *Финн В.К.* Гуманитарное знание и искусственный интеллект. / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 39

общей концепции социологического исследования и на этапе выдвижения основных гипотез.⁸

В целом, в основе методологии лежат:

Во-первых, идея поиска сходства событий и явлений посредством соотношения их между собой.

Во-вторых, идея расчленения объектов на множества характеристик, представляющих собой описания объектов, и множества свойств объектов.

В-третьих, логические процедуры, с помощью которых выводятся и проверяются общие закономерности – индукция, аналогия, абдукция, дедукция.

В-четвертых, предположение о существовании причинно-следственных отношений между событиями.

ДСМ-метод в теоретическом аспекте является частичным решением таких проблем, как

- Проблема формализации идей,
- Проблема поиска существенных характеристик (существенных по отношению к изучаемым свойствам)⁹.
- Проблема синтеза познавательных процедур (индукции, аналогии, абдукции).

К анализу этих проблем и их возможных решений мы вернемся чуть позже. А сейчас следует рассмотреть теоретические основы самой методологии. Среди них можно выделить:

- ✓ идею фальсификации К.Поппера, лежащую в основе его «эволюционной эпистемологии»,
- ✓ идею трех миров К. Поппера как онтологическую основу эпистемологии,
- ✓ Теорию постоянства Джона Стюарта Милля,
- ✓ Понятие индукции и индуктивные методы Д.С. Милля (в частности метод сходства),
- ✓ критерий достаточного основания, предложенный А.С. Есениным - Вольпиным и лежащий в основе философии неологицизма,
- ✓ Абдукцию Ч.С. Пирса.

⁸ Особенно это актуально в предметных областях, имеющих массу собранных ранее данных различного характера. Анализ таких данных позволяет лучше выделить признаки и переменные, необходимые для новых исследований.

⁹ Так, например, при изучении готовности рабочих принимать участие в забастовке для нас существенными характеристиками будут психологические черты личности (агрессивность, импульсивность, подверженность мнению окружающих, лояльность), а несущественными, например, - наличие домашних животных, музыкальные предпочтения и т.п.

Эволюционная эпистемология Карла Поппера

Эпистемология переводится с английского языка как теория научного познания, объясняющая развитие науки и научного знания. Эпистемологию К. Поппера называют эволюционной в силу того, что он смотрит на науку как на продукт «биологической» эволюции, а именно, дарвинистской эволюции, в основе которой лежит принцип естественного отбора.

По Попперу все живые организмы являются решателями *проблем*, которые «рождаются вместе с возникновением жизни». Для решения проблем выдвигаются гипотезы, строятся теории. В процессе эволюции «лучшие» теории вытесняют «худшие» путем естественного отбора. Лучшая теория наиболее устойчива к критике и наименее подвержена фальсификации. Иными словами, лучшая теория более устойчива в мире фактов.

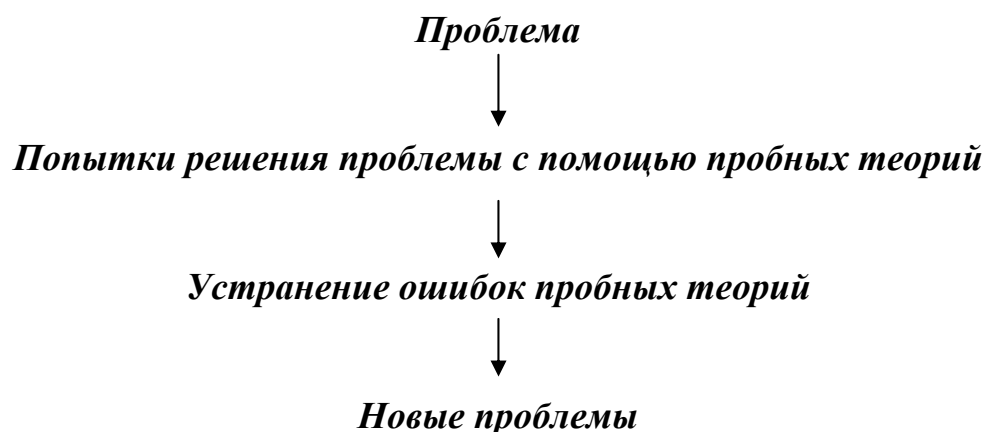
Основной критерий оценки теорий - критерий демаркации, в основе которого лежит принцип опровержимости и фальсифицируемости теорий. В соответствии с этим критерием «любое организованное множество высказываний, такое, что оно не допускает фальсифицируемости, не может считаться наукой»¹⁰. Согласно позиции Поппера, «некоторую систему можно считать научной только в том случае, если ее утверждения способны столкнуться с наблюдениями. Проверка системы состоит в том, чтобы обнаружить такие столкновения, т.е. чтобы опровергнуть ее. Поэтому проверяемость системы означает ее опровержимость, следовательно, она тоже может быть принята в качестве критерия демаркации»¹¹. Таким образом, научной теорией может считаться некоторое обобщение относительно фактов, которое в принципе проверяемо и может быть опровергнуто в том случае, если нам встретится факт, не согласующийся с этим обобщением.

В реальности мы не можем произвести исчерпывающее количество наблюдений, мы не можем оценить все факты действительности, и, как следствие, не можем построить абсолютно истинную обобщающую теорию. Поэтому любые теории должны рассматриваться лишь как пробные теории, а точнее как теоретические гипотезы. А, следовательно, они должны предполагать возможность фальсификации. Лучшие теории – это те теории, которые сложнее опровергнуть. То есть, сложнее найти факты, противоречащие им.

¹⁰ Финн В.К. Эволюционная эпистемология Карла Поппера и эпистемология синтеза познавательных процедур // Эволюционная эпистемология и логика социальных наук. - М.: Эдиториал УРСС, 2000, стр. 367

¹¹ Поппер К. Предположения и опровержения, М.: "Издательство АСТ", 2004, стр. 427

Процесс эволюции теорий можно выразить следующей схемой:



Процесс устранения ошибок предполагает фальсификацию, а затем, с учетом ее результатов, построение новой, более хорошей теории.

Познание в соответствии с данными положениями осуществляется по принципу «гипотезы – фальсификация – новые гипотезы».

Напомним, что наша исследовательская задача состоит в том, чтобы найти на основе фактов наиболее общие и устойчивые закономерности, описывающие изучаемую предметную область. Мы описали общую схему процесса познания.

Теперь нужно ответить на следующие вопросы:

- Что считать фактами, на основе которых строятся гипотезы?
- Как получить гипотезы относительно фактов, и как их проверить?

Для этого перейдем к анализу следующих предпосылок.

Идея трех миров Карла Поппера

Развитие человеческого интеллекта непосредственно связано с развитием языка, что приводит к отделению субъекта (производителя теорий), от объекта (самих теорий). Человек в отличие от более примитивных существ, может «вывести свои теории из организма» и смотреть на них, как на объекты своего интеллектуального творчества. При этом он может уничтожать свои теории при их несостоятельности, не уничтожая самого себя. Таким образом, именно развитие языка способствует развитию знания. Вследствие

этого происходит постепенное усложнение мира, который Поппер разделяет на 3 подмира:

Мир 1 - мир всех тел, силовых полей, организмов, всех физических объектов.

Мир 2 - мир ментальных состояний.

Мир 3 - мир продуктов человеческого сознания, абстрактных теорий и объектов, таких, как доказательство теорем, содержание мыслей и т.п.

При этом мир 3 включает в себя также объекты из мира 1, но уже не их физические формы, а их символическое значение. В отличие от мира 2 (субъективного знания, конкретной информации, содержащийся в наших головах), мир 3 содержит в себе объективное знание. Но вместе с тем, мы видим мир 3 через призму наших предположений об объективности. При этом мы стремимся развивать истинное знание об объективном смысле. С одной стороны мир 3 представляет собой идеальную объективную сущность, но с другой стороны, также включает в себя все теории, посредством которых человек пытается достичь этого идеала.

В качестве фактов для анализа используются объекты мира 1. Будь то идеи, сознания и все другие объекты мира 2 и 3, для формализованного анализа они должны быть представлены в виде некоторых реальных объектов и событий мира 1.

В.К. Финн предложил более содержательный подход к классификации миров, введя классификацию типов «подмиров», являющихся частными случаями мира 1:

Мир 1-1, состоящий из случайных событий, изучаемых вероятностными методами.

Мир 1-2, состоящий из событий, между которыми имеются причинно следственные отношения.

Мир 1-3, объединяющий в себе характеристики миров 1-1 и 1-2, то есть такой мир, в котором имеются как причинно – следственные отношения между событиями, так и случайные события.¹²

Мир 1-1 является адекватной онтологией для вероятностных методов. Мир 1-2 адекватен для методов причинного анализа, в основе которых лежит теория постоянства, предполагающая наличие четкого детерминизма в мире фактов. Но социальная реальность представляет собой скорее мир 1-3, включающий в себя как факты мира 1-1, так и факты мира 1-2. Это приводит к обострению проблемы адекватности методов исследования и ставит вопрос о совмещении двух подходов.

ДСМ метод в чистом виде работает с миром 1-2. Однако возможен ряд допущений и расширений, позволяющих ДСМ методу работать с миром 1-3. О них будет сказано ниже.

¹² Там же, стр. 371

Итак, посредством описания миров дана общая классификация и характеристики событий, что имеет принципиальное значение для выбора адекватной методологии анализа данных.

Теперь более подробно коснемся теории постоянства Д.С. Милля, и понятия причинности по Миллю.

Закон всеобщей причинной связи по Д.С. Миллю

Джон Стюарт Милль в своих работах (в том числе работе «Система логики силлогистической и индуктивной»¹³) развил подход Дэвида Юма к понятию причинности и сформулировал закон единообразия природы, легший в основу описываемой методологии анализа причинно - следственных отношений. Суть этого закона состоит в том, что причинно-следственные отношения постоянны и причина неизменно влечет за собой следствие, поэтому в сходных условиях сходные события ведут себя одинаково.

Теорию причинности Д.С.Милля можно выразить следующими тезисами:

Все явления природы находятся в двух различных отношениях друг к другу: отношении одновременности и отношении последовательности.

Любое явление связано некоторым «правильным» образом с одной стороны с явлениями, существующими одновременно, и с другой стороны, с явлениями предшествующими данному явлению или следующими за ним.

Основными «единообразиями», т.е. неизменными законами являются законы числа и пространства. Единообразие по Миллю и есть «закон причинной связи» - «Та истина, что всякий факт имеет начало, имеет причину, простирается на весь человеческий опыт».¹⁴ Милль доказывает это утверждение следующими образом: «тот факт, что всякое явление зависит от какого либо закона, есть сам закон».¹⁵ Исходя из этого, все события взаимосвязаны и детерминированы некоторыми условиями, содержащими в себе причины. Но, в свою очередь, причины могут быть сложными и разнонаправленными, что влечет за собой проблемы выявления причинной связи. Это означает, что случайностей не бывает. Бывает лишь недостаток информации. Если бы мы имели возможность исчерпывающего наблюдения за всеми фактами, событиями, условиями и

¹³ Милль Д.С. Система логики силлогистической и индуктивной. - М.: Книжное дело, 1900

¹⁴ Там же, стр. 259

¹⁵ Там же, стр. 259

состояниями природы, мы бы смогли построить теорию, описывающую все явления природы в их сложнейше взаимосвязи. Исходя из этих предпосылок, можно заключить, что причинно–следственные отношения между явлениями неизбежно существуют.

Само понятие причины Милль связывает с понятием «необходимости» и «безусловности». То есть, если мы считаем, что А есть причина В, это означает, что А влечет В при любых других условиях. Если есть А, то В неизбежно возникнет. Отношение нельзя считать причинно следственным если оно не безусловно. Для иллюстрации данной идеи Милль приводит в пример последовательность дня и ночи, не являющуюся безусловной. Так день не является причиной ночи, и ночь не является причиной дня. Несмотря на то, что день всегда следует за ночью, а ночь за днем, они не являются причинами друг друга, т.к. есть третье обстоятельство, являющееся причиной данной последовательности – вращение земли вокруг солнца. Если бы земля не вращалась вокруг солнца, то связи между днем и ночью не было бы.

Исходя из данных положений причину Милль определяет как «такое предыдущее и такое сочетание предыдущих, за которым данное явление неизменно и безусловно следует».¹⁶ Однако сложность мира 1-2 два заключается в том, что безусловных последовательностей событий гораздо меньше, чем условных. Сочетание причин может вызывать некоторое следствие, которое, в свою очередь, в сочетаниях с другими событиями влечет другие следствия и т.д. Помимо этого Милль указывает на тот факт, что причина может как предшествовать следствию, так и одновременно сосуществовать с ним, если речь идет о процессах, таких, что один процесс является причиной продолжения другого и т.п. Все эти обстоятельства существенно усложняют процессы поиска причинно-следственных связей среди явлений природы.

Милль также отрицает возможность возникновения случайных событий среди следствий. «Так как все что происходит во вселенной, определяется законами причинной связи и сочетаниями, или распределениями первоначальных причин, то отсюда следует, что наблюдаемые сосуществования между следствиями причин сами не подчинены никаким законам, отличным от законов причинной связи... Между следствиями причин не существует первоначальных и независимых, другими словами, безусловных единообразий сосуществования; и если эти следствия сосуществуют одновременно друг с другом, то лишь потому, что случайно сосуществовали их причины».¹⁷

Милль делает еще три важных допущения относительно причинно – следственных связей.

¹⁶ Там же, стр. 271

¹⁷ Там же, стр. 277

Первое: положение о множественности причин, суть которого состоит в том, что причиной явления могут быть совокупности (суммы) причин.

Второе: причины могут быть как положительными, так и отрицательными, то есть вызывающими наличие определенного эффекта или тормозящие его.

Третье: отсутствие некоторого явления наряду с другими сопричинами также может быть одной из причин наличия или отсутствия эффекта. Например, АВ и не-С всегда влекут за собой Х и т.п.

На основе этих предположений Д.С. Милль сделал попытку создать логический аппарат для формализации рассуждений о причинно – следственных зависимостях. Этот аппарат представляет собой группу методов индуктивного исследования:

- ✓ Метод содства,
- ✓ Метод различия,
- ✓ Косвенный метод,
- ✓ Метод остатков,
- ✓ Метод сопутствующих изменений.

Рассмотрим каждый из методов более подробно.

Основные методы индуктивного исследования по Д.С. Миллю

Исследования причинно-следственных закономерностей бывают двух видов: исследования с помощью опытов и исследования с помощью наблюдения. Этим типам соответствуют разные методы.

Милль выделил два основных способа выявления причины события из числа предшествующих ему обстоятельств. Первый заключается в сопоставлении тех отличных один от другого случаев, в которых данное явление имеет место. Второй состоит в сравнении таких случаев, где изучаемое явление присутствует, со сходными в других отношениях случаями, но в которых этого явления нет. Эти два способа фактически представляют собой метод сходства и метод различия.¹⁸ Приведем формулировки данных методов.¹⁹

Метод сходства Милль формулирует следующим образом:

¹⁸ Милль Д.С. Система логики силлогистической и индуктивной. - М.: Книжное дело, 1900, с. 313

¹⁹ Формулировки приведены там же, стр. 311 - 325

Если два и более случая подлежащие исследованию явления имеют общим лишь одно обстоятельство, то это обстоятельство, в котором только и согласуются все эти случаи, есть причина (или следствие) данного явления.

Так, например, если ABC влечет abc и ADE влечет за собой ade , то очевидно, что причиной a является A .

Данный метод опирается на следующую аксиому: ни одно обстоятельство, которое можно исключить без ущерба для явления или которое может отсутствовать не смотря на присутствие этого явления, не связано причинной связью.²⁰ Таким образом, после исключения случайных обстоятельств останется лишь одно или несколько обстоятельств, которые нельзя будет исключить без ущерба для явления, а, следовательно, которые являются причинами данного явления.

Формулировка метода различия звучит так:

Если случай, в котором исследуемое явление наступает, и случай, в котором оно не наступает, сходны во всех обстоятельствах, кроме одного, встречающегося лишь в первом случае, то это обстоятельство, в котором только и разнятся эти два случая, есть следствие, или причина, или необходимая часть причины явления.

Так, если ABC влечет abc , в то время как BC влечет bc , то очевидно, что A является причиной a .

Следует отметить, что метод различия требует сходства многих обстоятельств, что приводит к сложностям его применения в опытных исследованиях. Как правило, данный метод применяется в исследованиях, проводимых с помощью наблюдения за реальными событиями, т.к. в реальных событиях редко встречаются сильно схожие обстоятельства.

Усложнение и логическая формализация метода сходства и метода различий лежит в основе ДСМ-метода АПГ.

Косвенный (смешанный) метод представляет собой сочетание метода сходства и метода различия. Формулируется он следующим образом:

Если два и более случая возникновения явления имеют общим лишь одно обстоятельство, и два и более случая невозникновения того же явления имеют общими только отсутствие того же самого обстоятельства, то это обстоятельство, в котором только и разнятся оба ряда случаев, есть или следствие, или причина, или необходимая часть причины изучаемого явления.

²⁰ Там же, стр. 312

Так, если ABC влечет abc и ADE влечет за собой ade , и при этом BC влечет bc , то очевидно, что A является причиной a .

Метод остатков:

Если из явления вычтешь ту его часть, которая, как известно из прежних индукций, есть следствие некоторых определенных предыдущих, то остаток данного явления должен быть следствием остальных предыдущих.

Этот метод представляет собой разновидность метода сходства с той лишь разницей, что у нас уже имеется некоторая информация о причинах сопутствующих явлений. Так, если мы знаем, что AB влечет ab , то если из ABC , влекущем abc вычтешь AB , и если остатком следствия будет c , значит C причина наличия c .

Метод сопутствующих изменений:

Всякое явление, изменяющееся определенным образом всякий раз, когда некоторым особенным образом изменяется другое явление, есть либо причина, либо следствие этого явления, либо соединено с ним какой либо причинной связью.

Этот метод хорош для изучения процессов. Видоизменение в данном случае предполагает изменение предыдущего до момента его исчезновения. Так, например, если трение предшествует выделению тепла и при усилении и ослаблении трения количество тепла изменяется, то это говорит о наличии причинно-следственной связи между этими явлениями.

Данные методы представляют собой логический аппарат для построения правдоподобных (не истинных, но пока не отвергнутых) рассуждений. Теперь более подробно коснемся вопроса проверки и фальсификации гипотез, а также идеи правдоподобного вывода.

Для этого обратимся к и идеям философии неологицизма (Есенин-Вольпин) и идеи абдуктивных заключений Ч.С. Пирса, которые являются существенными для формирования рассматриваемой методологии.

А.С. Есенин-Вольпин и философия неологицизма

Идеи Есенина-Вольпина, касающиеся логико-методологических основ науки изложены в следующих работах: «Свободный философский трактат»²¹, «Об антитрадиционной (ультраинтуитивалистской) программе оснований математики и естественно-научном мышлении»²² и др.

Есенин-Вольпин характеризует метод познания не как дедуктивный, а как дискриптивно-аналитический. Данный метод использует следующие логико-методологические средства для построения онтологических характеристик и эпистемологии изучаемых «миров»: построение определений, средств принятия и фальсификации утверждений, критика источников доверия, установление сомнительности утверждений, формулирование правильных рассуждений, выделение рассуждений, удовлетворяющих канонам доказательств.²³

Основа философии неологицизма заключается в «изгнании веры» из области математики и логики, ибо вера нарушает закон достаточного основания. «Верой» Есенин-Вольпин называет тот факт, что некоторые суждения считаются истинными без предъявления доказательств. В соответствии с принципом изгнания веры, истинными суждениями могут считаться только результаты доказательств. (За исключением аксиом, принимаемых в качестве основополагающих утверждений, не требующих доказательств. Любое суждение, не являющееся аксиомой требует доказательства.) В качестве основных критериев истинности знания в дискриптивно-аналитической методологии является *неоспоримость* и *«не-недоказанность»*. Фактически эти выражения являются синонимами.

Неологицизм можно определить как философию обоснованного знания, где в качестве приемов обоснования используются логические рассуждения. Обоснованные рассуждения являются неоспоримыми при «честной критике в диспутах». Иными словами, обоснованные рассуждения должны полностью удовлетворять канонам научного знания (доказательств), относительно которых ведутся научные диспуты. Честная критика в данном случае означает критику в рамках установленных канонов, т.е. нахождение критикующего и критикуемого в равных условиях. В качестве этих условий могут выступать как раз те самые правила логических выводов.

²¹ А.С. Есенин-Вольпин, Свободный философский трактат / Логика, философия, поэзия, защита прав человека. Под редакцией В.К. Финна и А.Ю. Даниэля, М.: РГГУ, 1999

²² А.С. Есенин-Вольпин, Об антитрадиционной (ультраинтуитивалистской) программе оснований математики и естественно-научном мышлении / Логика, философия, поэзия, защита прав человека. Под редакцией В.К. Финна и А.Ю. Даниэля, М.: РГГУ, 1999

²³ Финн В.К. Гуманитарное знание и искусственный интеллект. / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 23

Дискриптивно-аналитический подход дает основание для применения автоматизированных средств логических рассуждений при построении концептуальных схем исследований и анализе данных в социальных науках, в которых до сих пор большая часть выводов делается на основе интуиции исследователя, то есть на вере в истинность предположений. В соответствии с данным подходом, формализация и автоматизация рассуждений способна в некоторой степени избавить исследователя от заблуждений и повысить степень обоснованности выводов.

Абдукция Ч.С. Пирса

Абдукция в Ч.С. Пирса есть вывод, результатом которого является принятие гипотез, объясняющих изучаемое множество фактов.²⁴ Суть ее заключается в том, что предсказываемые на основе гипотез следствия проверяются соответствующими наблюдениями и после этого гипотеза либо опровергается, либо остается в качестве пробной. Абдукцию в смысле Пирса можно также определить как переход от необъясненных фактов к теории, объясняющей их.²⁵

В традиционной логике абдукция понимается как вывод от следствий к посылкам, а не наоборот. Приведем типичный пример абдуктивного умозаключения: «Сократ смертен => Сократ человек, Все смертные – люди». Так, допустим, по имеющимся данным мы заключили, что все люди определенной категории обладают неким свойством, мы выдвигаем гипотезу о том, что это свойство может детерминировать попадание человека в данную категорию. Но если затем мы обнаружим, что этим свойством обладают и некоторые другие люди, не входящие в данную категорию, то гипотеза опровергается. Таким образом, абдукция является логической процедурой, с помощью которой возможно осуществление проверки и фальсификации гипотез. Так, например, если мы обнаружили, что все живые существа, являющиеся людьми – смертны, мы можем сделать заключение о том, что данное свойство присуще именно категории людей и определяет попадание некоторого живого существа именно в эту категорию. Однако если в дальнейшем мы вдруг обнаружим, что собака тоже смертна, и при этом собака не является человеком, то наша гипотеза о том, что свойство быть смертным позволяет отнести некое живое существо к категории людей, опровергается. Однако мы можем

²⁴ Финн В.К. Синтез познавательных процедур и проблема индукции. / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 160 – 167.

²⁵ ФирменЮ., Сколимовский Г. Поиск объективности у Пирса и Поппера. // Эволюционная эпистемология и логика социальных наук. - М.: Эдиториал УРСС, 2000, стр. 227

выдвинуть новую гипотезу о том, что все живые существа смертны. И если некоторое существо смертно, это позволяет нам отнести данное существо к категории живых существ.

Таким образом, абдукция Ч.С. Пирса позволяет нам использовать идею фальсификации К. Поппера.

Итак, мы описали и проанализировали основные методологические предпосылки ДСМ метода АПГ как средства формализованного логического анализа данных. В теоретическом плане можно считать решенной стоящую перед нами проблему поиска скрытых закономерностей и причинно-следственных связей. Однако есть еще масса проблем и вопросов, связанных с онтологией изучаемой предметной области. А именно, что считать фактами, подлежащими социальному анализу; каким образом соотносить их между собой и как определить, насколько данные факты схожи или различны; по каким параметрам соотносить изучаемые факты и какие параметры считать существенными для анализа? Для более глубокого погружения в данную проблематику проанализируем некоторые методологические проблемы, с которыми сталкиваются социальные науки при построении теорий и анализе фактов.

Некоторые методологические проблемы социальных наук

Проблема формализации идей в социальной науке

На сегодняшний день точные методы исследований все глубже проникают в сферы гуманитарных и общественных наук. Значительное развитие знания и его эмпирическая проверка в этих сферах становятся возможными во многом благодаря точным методам, в основе которых лежат рассуждения, выраженные в специализированных научных языках, и эксперименты, позволяющие устанавливать эмпирические зависимости. Такие методы предполагают наличие в каждой науке специального понятийного аппарата, являющегося языком данной науки и описание условий применимости данного аппарата. Иными словами, точные научные методы предполагают четкое определение объекта и предмета науки, в которой они применяются.

Особенностью общества как предмета изучения является целостность взаимодействия различных факторов - политических, экономических, религиозных и т. д. - образующих контекст и причины социального поведения. Целостность и многофакторность также являются главными принципами изучения общества. Но при попытке применения точных методов к изучению общества возникает ряд проблем, кроющихся в специфике самого предмета изучения. Во-первых, это необходимость упорядоченности факторов, определяющих социальное поведение. Данная проблема осложняется, прежде всего, невозможностью четкого разграничения факторов и определения степени влияния каждого фактора на социальные процессы, а также динамическим характером структуры факторов. Во-вторых, это специфика самого человека, являющегося, если так можно выразиться, атомом общества. Если в точных и естественных науках предметом изучения - научными фактами являются объективные события и процессы (к примеру, замерзание воды), то в социальных науках источником знания является сам человек со своим пониманием реальности, причем, человек выступает и как в качестве исследователя и как в качестве исследуемого. Здесь, с одной стороны, наука сталкивается с проблемой сходства и индивидуальности. Так, если в физике в качестве единицы изучения могут выступать элементарные частицы, идентичные друг другу, но порождающие различные структуры с различными свойствами, то в социальной реальности единица анализа - человек или социальная группа, которые в принципе не могут быть идентичными друг другу. Это приводит к сложностям разграничения и классификации социальных явлений и объектов, необходимых для

точного анализа и моделирования социальной реальности. С другой стороны, наука сталкивается с самой спецификой человеческого мышления, основой которого является конструирование идей на базе рассуждений. В данном случае идею можно определить как приблизительное, обоснованное здравым смыслом высказывание относительно какого-либо явления или предмета. Но точные научные методы требуют четко определенного и структурированного языка - понятийного аппарата. Встает вполне естественный вопрос: а в чем разница между идеей и понятием и как преобразовать идею в понятие?

Обычно, в рамках сложившейся традиции, как в обыденном разговорном языке, так и в научном языке термины *идея* и *понятие* четко не отделяются друг от друга. "В традиционной логике под понятием имеют в виду мыслимое содержание, представленное совокупностью существенных признаков, которое относится к некоторой совокупности предметов, называемой объемом понятия".²⁶ Также в традиционной логике рассматриваются определения через род и видовое отличие. В математической логике используются индуктивные определения, абстракция (упрощение модели) и символические определения - введение нового символа через известные символы. Также математики используют неявные определения - аксиомы. Вышеперечисленные виды определений составляют понятия в истинном смысле этого слова.

В гуманитарных науках феномены "идея" и "понятие" часто отождествляются и сливаются. Это утверждение на первый взгляд может показаться упреком гуманитарным наукам со стороны точных наук, однако, это не так. Гуманитарные, в частности общественные науки, изучают одни виды фактов - социальные факты, которые основываются на идеях и объясняются идеями. Точные науки изучают свои факты - объективные явления природы (химия, физика), абстрактные модели объективной реальности (математика). Каждая научная область имеет свои методы, наиболее адекватные для своего предмета изучения. Другое дело, если совмещение этих методов достаточно продуктивно в практическом применении. Тогда возникает необходимость перевода идей с одного языка на другой. Перевод какого-либо понятия с языка математики на язык идей, то есть язык рассуждений и здравого смысла, достаточно прост, так как язык математики фактически является упорядоченным и уточненным языком здравого смысла. А вот перевод человеческих идей и рассуждений на язык математики - уже проблема.

Социологи работают, как правило, с большими массивами данных, где необходимо применение точных методов анализа. И от того, в какой степени язык анализа

²⁶ Финн В.К. Интеллектуальные системы и общество: идеи и понятия // НТИ. Сер. 2. - 1999. - № 10. - с. 6-20

соответствует "языку общества", зависит адекватность и масштабы применения данного метода к анализу социологических данных.

Как показывает опыт последних лет, массивы данных содержат колоссальное количество информации, недоступной глазу исследователя, пользующегося стандартными статистическими методами для проверки своих интуитивно выдвинутых гипотез. Статистический анализ - метод во многом упрощает решение некоторых социологических задач. Во-первых, он оперирует средними величинами и средними разбросами, как базовыми средствами языка статистики. А средние величины уничтожают "индивидуальность", в которой, исходя из теории причинности Милля, кроются причинно-следственные связи. С другой стороны статистический анализ усредняет случайные выбросы, мешающие выявлению причинно-следственных связей. Таким образом, социолог должен в некотором роде постоянно искать компромисс между статистикой и логикой, сходством и индивидуальностью, объективностью и субъективностью, количественным и качественным подходом к анализу данных.

Методология, рассматриваемая в данной работе, способна в некотором роде решить данную проблему - породить гипотезы о причинно-следственных зависимостях с помощью «объективных», автоматизированных рассуждений и анализа достаточно больших массивов данных логическими средствами. Таким образом, мы не теряем индивидуальности, но сталкиваемся с проблемой случайных выбросов, которые могут в результате привести к так называемому «комбинаторному взрыву». В конце данной работы мы предложим один вариант возможного решения данной проблемы.

В целом, ДСМ - системы на данный момент достаточно слабо адаптированы к анализу социологической информации. Адаптация систем к той или иной предметной области требует детального и многоаспектного ее изучения на предмет применения метода. Это как раз и является одной из основных задач данной работы. И основные вопросы, возникающий на первом этапе адаптации метода к той или иной предметной области, звучит так:

С какой информацией мы работаем, как должна быть представлена информация в базах данных? Каким образом она должна быть формализована? Если мы не можем четко определить, что мы изучаем, мы не можем это адекватно измерить. Так, контекст, смысловое содержание и эмоциональная окраска вопросов анкет социологических исследований может восприниматься респондентами по-разному. Таким образом, проблема сводится к отсутствию единого формального языка, на котором исследователь и респондент могли бы обмениваться точными данными. Данная проблема приводит к

потере информации на пути от респондента к исследователю, то есть, потере информации при ее переносе из реальности в базу данных.

Проблема формализации идей и понятий упирается в проблему измерения, которая, выражается в несоответствии индикаторов, полученных в результате операционализации понятий, реальному содержанию понятий.

Проблема измерения тесно связана с проблемой познания.

Проблема измерения и проблема познания

Каким образом проблема познания (поиска объективности) соотносится с проблемой измерения? Для ответа на данный вопрос проанализируем типичную схему социологического исследования в рамках концепции Поппера.

Итак, основные этапы социологического исследования:

- 1) **Осознание некоторой проблемы на основе некоторого опыта.** Это осознание можно кратко сформулировать, как осознание того, что некоторые социальные механизмы работают по законам, которых мы не знаем.
- 2) **Формирование концептуальной схемы исследования.** На данном этапе четко определяется проблема, цели и задачи исследования, проводится анализ имеющихся знаний по данной проблеме или схожим проблемам. Если среди имеющихся знаний есть попытки объяснения проблемной ситуации, и они не соответствуют нашему опыту, мы пытаемся отвергнуть (или уточнить) старые теории посредством создания новых. Выдвигаются основные гипотезы, разрабатывается понятийный аппарат, проводится операционализация понятий. Посредством этого мы пытаемся определить, какие факты из всего множества фактов относятся к делу, и как нам преобразовать эти факты в некую базу данных для дальнейшего анализа.
- 3) **Сбор данных, анализ, проверка гипотез.** Проводим сбор информации, на наш взгляд нужной для решения проблемы. Определенным образом структурируем эту информацию и представляем в виде базы данных (БД) На данном этапе происходит контакт исследователь с респондентом в рамках разработанной на предыдущем этапе пробной теории.
- 4) **Выводы, корректировка имеющегося знания.** Гипотезы, явно не подтверждающиеся и противоречащие фактам, отбрасываются. Остальные гипотезы остаются открытыми за не имением фактов, их фальсифицирующих, и

как бы условно подтверждаются, что на самом деле означает: "на данный момент мы не знаем ничего о фактах, которые могли бы фальсифицировать принимаемые гипотезы, так как эти факты, если они есть в реальности, отсутствуют в нашей концептуальной схеме".

В итоге, проблема измерения сводится к вопросам "Что измерять?" и "Как измерять?". Первая проблема возникает на этапе формирования концептуальной схемы исследования (в частности операционализации). Вторая - на этапе сбора информации, ее кодирования и создании базы данных. А эти проблемы изначально неразрешимы в силу не возможности достижения идеального объективного знания (в идеале достижения объективного знания предполагает возможность исчерпывающего анализа взаимовлияния всех без исключения явлений, событий и фактов).

Таким образом, актуализируется проблема неполноты информации, с которой неизбежно сталкивается социолог в процессе проведения исследования и анализа результатов. В итоге мы: имеем данные о фактах, якобы относящихся к делу, и возможно не имеем данных о каких-то фактах, реально относящихся к делу. Именно эту проблему призван частично решать логико-комбинаторный ДСМ метод АПГ, специально предназначенный для работы в условиях неполноты информации.

По сути, есть два решения проблемы: подбор фактов под теорию, или подбор теории под факты. В первом случае мы изначально формулируем гипотезы о причинно-следственных связях и различных отношениях между переменными, и в то же время по большому счету умозрительно задаем критерии отбора фактов. Во втором случае мы подбираем все возможные гипотезы, объясняющие имеющиеся данные, ищем причинно-следственные связи исходя непосредственно из массива данных. Этот метод в некотором роде является решением двух вышеописанных проблем. Но во втором случае предполагается сложная интеллектуальная работы с базами данных. Что неизбежно приводит к необходимости разработки автоматизированных средств интеллектуальной работы с информацией.

Синтез познавательных процедур

Синтез познавательных процедур (индукции, аналогии, абдукции, дедукции) приводит к частичному решению вышеизложенных проблем. Синтез познавательных процедур посредством формализации рассуждений в некотором роде является способом получения максимально объективного знания из имеющихся фактов.

Проблема синтеза познавательных процедур решается посредством построения логико-математического аппарата и создания специального языка представления фактов и знаний для предметной области. В предыдущих главах были описаны все основные теоретические предпосылки, необходимые для обоснованного построения такого аппарата. Таким образом, остается решить ряд конкретных задач. А именно:

- - Описать язык представления фактов
- -Описать основные логические процедуры и алгоритмы их использования для анализа данных

Решение данных задач нашло свое воплощение в ДСМ методе АПГ, к детальному изложению которого мы переходим в данный момент.

Раздел 2. Неформальное изложение ДСМ-метода АПГ. Специфика и возможности его использования

ДСМ метод АПГ в неформальном изложении

В предыдущей главе было сказано о специфике трех подмиров мира фактов: мира 1-1, состоящего из случайных событий, мира 1-2, состоящего из событий, между которыми имеются причинно – следственные связи и мира 1-3, в котором есть как события из первого мира, так и события из второго мира. Как и все методы Милля, так и ДСМ метод в чистом виде предназначен для работы с миром 1-2. Работа с миром 1-3 должна осуществляться посредством комбинирования логических и статистических средств. В данном разделе рассматриваются логические средства изучения мира 1-2, а также возможности применения этих средств к миру 1-3. В данном разделе будут также рассмотрены возможности комбинирования логических и статистических средств для анализа фактов мира 1-3 (в частности, социологических данных).

Специфика представления данных для работы с ДСМ методом АПГ.

Для формализации фактов мира 1-2 необходимо задать некоторые рамки представления фактов в базах данных. В.К. Финн предложил сделать это посредством введения следующей системы формальных ограничений:

1. Мир состоит из фактов имеющих структуру вида «объект X обладает множеством свойств Y». При этом на множестве всех объектов должна быть задана алгебраическая операция нахождения сходства двух объектов. Факт есть элемент отношения «обладать множеством свойств». Это отношение может либо выполняться, либо не выполняться для каждого из объектов. Поэтому среди множества фактов присутствуют как позитивные факты (объект обладает множеством свойств), так и негативные факты (объект не обладает множеством свойств).

2. Объекты или подобъекты (части объектов) связаны отношением «быть причиной наличия множества свойств». Данное отношение понимается как направленное влияние. То есть оно может быть как позитивным, так и негативным. Влияние позитивно, если оно вызывает наличие изучаемого эффекта, влияние негативно, если оно вызывает отсутствие изучаемого эффекта. Положительные и отрицательные влияния являются, соответственно, (+) причинами и (-) причинами.

3. Учитывая вышесказанное можно разбить мир 1-2 (по отношению к изучаемому эффекту) на четыре непересекающихся подмножества фактов. Позитивные факты – обладающие (+) причинами, негативные факты, содержащие в себе (-) причины, противоречивые факты, содержащие в себе как (+), так и (-) причины и неопределенные факты, относительно которых не зафиксировано наличия ни положительных, ни отрицательных причин.

4. Множество фактов является открытым. Это означает, что изначальная база данных неполна и может расширяться посредством добавления новых фактов. Эти факты могут как соответствовать выдвигаемым гипотезам, так и противоречить им.²⁷

Далее мир, удовлетворяющий этим четырем конкретизациям будем называть (+/-)-миром, или универсумом. Вышеизложенные конкретизации касаются онтологии изучаемого мира.

Относительно эпистемологии мира 1-2 предполагается, что сходство объектов универсума выразимо посредством логических высказываний. То есть относительно некоторого множества объектов мы можем заключить, что, к примеру, «объекты А, В, С сходны в том, что все они обладают множеством свойств Y». Данное высказывание легко формализуется и переводится на язык логики. Далее предполагается, что причины являются источниками эффектов (множеств свойств) и выразимы посредством сходств объектов (то есть нахождением среди описаний объектов таких подобъектов, которыми обладают объекты, схожие в свойствах). Так, если различные объекты А, В, С сходны по какому-либо параметру и все обладают множеством свойств Y, то можно считать, что «найденный параметр - подобъект -X является причиной наличия множества свойств Y». Фактически это и есть метод сходства Д.С.Милля.

Далее следует сделать некоторые замечания относительно онтологии и эпистемологии изучаемого мира.

²⁷ Финн В.К. Эволюционная эпистемология Карла Поппера и эпистемология синтеза познавательных процедур // Эволюционная эпистемология и логика социальных наук. - М.: Эдиториал УРСС, 2000, стр. 372-374.

То, что в онтологии является фактом, в эпистемологии представляет собой высказывание о факте. В свою очередь, тип факта в онтологии (положительный, отрицательный, противоречивый, неопределенный) соответствует оценке высказывания о факте. В.К. Финн предложил в своих работах следующую систему оценок:

- +1 – фактическая или эмпирическая истина,
- 1 – фактическая или эмпирическая ложь,
- 0 – фактическое или эмпирическое противоречие,
- t – неопределенность.²⁸

Эти значения применяются как для изначальной (эмпирической) оценки фактов в базе данных (собранные и преобразованные данные), так и для оценки высказываний о фактах.

Далее для описания фактов и высказываний будем использовать следующие виды констант и переменных:

Переменные для объектов и подобъектов – X, Z, V

Переменные для множеств свойств – Y, U, W

Индивидуальные константы для объектов и подобъектов – $C1, C2, C3 \dots Cn$

Индивидуальные константы для множеств свойств – $A1, A2, A3 \dots An$.

Данные для ДСМ метода могут быть представлены в виде структурного описания субъектов (в социологии - респондентов) как множества дифференциальных признаков. На основе предварительной интерпретации и обработки данных логико-комбинаторными средствами каждому респонденту приписываются оценки признаков. Сходство субъектов (респондентов) определяется как алгебраическая операция. Операция сходства может быть представлена различным образом, в зависимости от варианта, в котором представлены данные. Представление данных и их первичная обработка для соответствия ДСМ системе является очень важным этапом анализа. Далее данный этап будет рассмотрен подробнее.

Общие принципы работы ДСМ метода АПГ

Фактически ДСМ метод представляет собой вариант точной эпистемологии. Теория, разработанная В.К. Финном и легшая в основу ДСМ метода, включает в себя емкий и специфический логико-математический аппарат. Однако, подробное изучение данного аппарата необходимо лишь для его компьютерной реализации и работы над

²⁸ Там же, стр. 375

программными средствами, связанными с методом. Социологу для использования возможностей ДСМ метода АПГ вполне достаточно понимание сути и специфики самого метода, а также знание алгоритма работы ДСМ - системы. Поэтому в задачу данной работы не входит подробное описание формальных средств реализации метода, а приводится лишь общее описание алгоритма работы метода с использованием некоторых элементов логики предикатов и теории множеств. Большое внимание уделяется неформальному описанию метода, а также интерпретации того, что мы имеем на входе и выходе метода.

ДСМ метод АПГ является «теорией автоматизированных рассуждений и способом представления знаний для решения задач в условиях неполноты информации».²⁹ Метод применим к данным, а также знаниям, для которых выполняются следующие условия:

1. Знания должны быть слабо формализованы, а данные хорошо структурированы, что означает, соответственно, неполноту описания предметной области и возможность лишь частичной ее аксиоматизации и определенность сходства объектов как алгебраической операции. Иными словами, база данных может представлять (а так оно обычно и бывает) далеко не полное описание объектов, и при этом мы не имеем представления о каких-либо четко определенных связях и зависимостях, которые могут скрываться в данных. Если соотнести это требование с традиционной техникой проведения социологического исследования, то у нас имеется некая база данных, в которой содержится масса индикаторов и показателей, построенных на основе гипотез, выдвинутых ранее исследователем. Гипотезы эти строились на основе результатов предыдущих исследований (некоторых аксиом, описывающих данную предметную область) и относительно новых проблем, решение которых требует получения новых знаний. Но показатели, которые нужно измерить для решения новых проблем формируются исследователем зачастую на основе интуиции. Поэтому знания в данном случае являются слабо формализованными. В свою очередь, хорошо структурированные данные предполагают: во-первых, выделение некоторых признаков, относительно которых описываются объекты и, во-вторых, приписывание некоторых численных или символических оценок, фиксирующих наличие или отсутствие признака, или же степень его выраженности, что дает возможность поиска сходства объектов.

2. Предметная область, описываемая базой данных должна содержать позитивные и негативные примеры изучаемого явления, а также примеры неопределенности, для которых система автоматически будет строить прогнозы. Фактом в данном случае

²⁹ Там же, стр. 378

является высказывание «объект C обладает множеством свойств A » с оценкой фактически истинно, фактически ложно, фактически противоречиво или фактически неопределенно.

3. В базе данных в неявном виде должны содержаться причинно-следственные зависимости типа «подобъект B есть причина наличия (отсутствия) множества свойств A ». Это означает, что некоторые признаки (переменные) из базы данных должны являться причинами изучаемого явления. По крайней мере, предполагается, что признаки взаимозависимы, и их распределения относительно друг друга не случайны.

В работе метода можно выделить три этапа:

На первом этапе на основе анализа сходства событий порождаются гипотезы о позитивных и негативных причинах изучаемого явления, а также гипотезы о фактической противоречивости или конфликтах. Гипотезы о причинах эффектов порождаются посредством применения правил правдоподобного вывода (ППВ) 1-ого рода.

Правила правдоподобного вывода 1-ого рода формулируются на основе специальных предикатов $M(+)$ и $M(-)$. Предикаты $M(+)$ используются для нахождения сходства для положительных примеров изучаемого явления. Предикаты $M(-)$ используются для нахождения сходства на отрицательных примерах. Эти предикаты состоят из пары элементов (подобъект, множество свойств). Подобъектом в данном случае является общая часть положительных (в случае $M(+)$) и отрицательных (в случае $M(-)$) примеров.

Так, если C считать подобъектом, а A – множеством свойств, то:

Если для пары (C, A) $M(+)$ выполняется, а $M(-)$ не выполняется, то порождается гипотеза о том, что C является причиной наличия A .

Если для пары (C, A) $M(+)$ не выполняется, а $M(-)$ выполняется, то порождается гипотеза о том, что C является причиной отсутствия A .

Если для пары (C, A) $M(+)$ выполняется, и $M(-)$ выполняется, то порождается гипотеза о том, что утверждения « C является причиной A » фактически противоречиво.

Если для пары (C, A) $M(+)$ не выполняется, и $M(-)$ не выполняется, то порождается гипотеза о том, что утверждения « C является причиной наличия A » фактически не определено.

Предикаты $M(+)$ и $M(-)$ представляют собой формализацию индуктивных методов Д.С. Милля. На основе отношения «обладать множеством свойств» данные предикаты порождают новое отношение - «подобъект X является причиной множества свойств Y ».

Гипотезы о причинах образуют новый фрагмент базы знаний,³⁰ который на следующем этапе используется для порождения предсказаний.

На втором этапе работы метода порождаются гипотезы, являющиеся предсказанием наличия или отсутствия явления или конфликта для фактов, изначально имеющих оценку «фактически неопределенно». Они порождаются с помощью применения правил правдоподобного вывода (ППВ) 2-ого рода, представляющих собой выводы по аналогии. Случаи с неопределенным исходом анализируются на предмет наличия или отсутствия положительных и отрицательных причин.

Если объект содержит в себе положительные причины и не содержит отрицательных, то порождается гипотеза о наличии у объекта изучаемого свойства.

Если объект не содержит в себе положительных причин и содержит отрицательные, то порождается гипотеза об отсутствии у объекта изучаемого свойства.

Если объект содержит в себе и положительные и отрицательные причины, то порождается гипотеза о том, что утверждение «объект X обладает множеством свойств Y» фактически противоречиво.

Если объект не содержит в себе ни положительных, ни отрицательных причин, то порождается гипотеза о том, что утверждение «объект X обладает множеством свойств Y» фактически неопределенно.

На третьем этапе проверяется так называемая аксиома каузальной полноты (АКП). АКП выполняется, если каждый факт изначального состояния базы данных объясняется полученными гипотезами. Это означает, что каждый положительный факт в своем описании содержит положительную причину, а каждый отрицательный факт – отрицательную причинную. То есть отсутствуют конфликты. Если же АКП не выполняется, то все три этапа запускаются снова, но уже с учетом изменений в базе данных, сделанных на основе предсказаний, а также с учетом фактов, добавленных в базу данных. АКП представляет собой формализацию критерия достаточного основания и на его основе реализуется абдуктивное принятие гипотез.

³⁰ В базе знаний, в отличие от базы данных содержатся гипотезы о причинах и предсказаниях, представляющие собой высказывания о данных, полученные посредством применения к данным эмпирической индукции.

После применения всех этапов метода может быть три исхода:

1. Все факты начального состояния базы данных объясняются полученными гипотезами. Тогда гипотезы и предсказания принимаются, т.к. удовлетворяют критерию достаточного основания.
2. Некоторые факты начального состояния базы данных не объясняются, но большинство объясняется. Тогда гипотезы и предсказания принимаются, но на «квазидостаточном»³¹ основании.
3. После расширения базы данных появляются необъясненные факты. В таком случае полученные гипотезы нельзя считать обоснованными.³²

Здесь следует сделать замечание относительно поэтапного применения метода в социологических исследованиях. Проведение анализа в несколько этапов (циклов) актуально для этой предметной области лишь при добавлении новых данных в базу. На социологических данных, особенно полученных с помощью опросов вряд ли можно достичь идеального состояния базы знаний. Опросы, как правило, являются разовыми, и добавить данные в базу зачастую не представляется возможным. Поэтому наиболее интересными для социолога являются первый и второй этапы.

Теперь попытаемся изложить алгоритм порождения гипотез в ДСМ - методе, используя некоторые формальные средства.³³ Алгоритм постараемся изложить с учетом специфики социологических данных.

В социологии виде объектов, как правило, выступают респонденты, а в виде описаний респондентов выступают переменные, объединяемые в признаки. Среди описаний выделяются характеристики. Характеристики могут быть: психологическими, социальными, демографическими, биографическими. В качестве психологических характеристик, как правило, выступают индивидуальные черты личности (к примеру, черты характера), социальные характеристики представляют собой социальные установки, выражающиеся в совокупности аттитюдов (отношений) и мнений респондентов. Демографические характеристики, описывают фактическое демографическое положение респондента на момент опроса, а биографические данные представляют собой конкретные

³¹ Грубо говоря, здесь уже идет речь о статистической проверке гипотез. Так, к примеру, можно задать критерий, относительно которого можно принимать или отвергать полученные гипотезы. Так, к примеру, если гипотезы объясняют более 80% фактов изначального состояния базы данных, то ее можно принять, но учитывая вероятность ее ложности, или недообоснованность. Это можно назвать квазидостаточным основанием.

³² Финн В.К. Синтез познавательных процедур и проблема индукции. / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 133 – 138.

³³ Помимо работ В.К. Финна воспользуемся работой Аншакова О.М. «Об одной интерпретации ДСМ метода АПГ», опубликованной в журнале НТИ, 1999 г. №1-2, стр. 45 – 53. В данной работе дается нестандартное описание формального алгоритма работы ДСМ-метода с помощью булевой алгебры множеств.

факты из жизни респондента. В качестве множества изучаемых свойств рассматриваются определенные акты поведения респондентов и поведенческие готовности.³⁴

Данные для анализа представлены в виде матриц следующего вида:

	<i>П1</i>	<i>П2</i>	<i>Пm</i>
<i>С1</i>	<i>V11</i>	<i>V12</i>	<i>V1m</i>
<i>С2</i>	<i>V21</i>	<i>V22</i>	<i>V2m</i>
...	<i>Vij</i>
<i>Сn</i>	<i>Vn1</i>	<i>Vn2</i>	<i>Vnm</i>

Где *С* – респонденты, *П* – признаки, *Vij* – оценки наличия или отсутствия признака у респондента. *Vij* принадлежат множеству {1, -1, 0, т}. Для социологических данных это множество может быть расширено до {1, 1/2, -1, -1/2, 0, т}, где оценки соответствуют следующим ответам на вопросы в анкете:

1 -Да;

1/2 -Скорее да;

0 - Ни да, ни нет;

-1/2 -Скорее нет;

-1 - Нет;

т - Нет данных (нет ответа).

Если для анализа используются не вопросы, а признаки, сформированные на основе нескольких вопросов, то можно использовать логико-комбинаторное преобразование данных. Например, для признака, сформированного на основе 4-х вопросов, эти комбинации могут быть следующими (если рассматривать исходную шкалу как +,0,-):

все 4 плюса	+1
3 плюса и нуль	
2 плюса и 2 нуля	+ 1/2
есть и + и -	0
2 минуса и 2 нуля	- 1/2
3 минуса и нуль	-1
все 4 минуса	
все нули	т
3 нуля и +,	
3 нуля и -,	
есть missing (хотябы 1)	

³⁴ Помимо поведенческих готовностей, анализируемых с помощью прямого ДСМ метода, можно анализировать мнения, используя обратный ДСМ-метод. В данной работе мы коснемся только прямого метода.

Аналогичным образом задается матрица для поведенческих готовностей:

	<i>A1</i>	<i>A2</i>	<i>Ak</i>
<i>C1</i>	<i>V11</i>	<i>V12</i>	<i>V1k</i>
<i>C2</i>	<i>V21</i>	<i>V22</i>	<i>V2k</i>
....	<i>Vij</i>
<i>Cn</i>	<i>Vn1</i>	<i>Vn2</i>	<i>Vnk</i>

Где *C* – респонденты, *A* – поведенческие готовности, *Vij* – оценки наличия или отсутствия признака у респондента. *Vij* принадлежит множеству {1, -1, 0, т} или {1, 1/2 -1, -1/2, 0, т}.

Для поиска сходства объектов задается определение операции сходства:

Сходство:	+1	+ ½	0	- ½	-1	т
+1	+1	+ ½	т	0	0	т
+ ½	+ ½	+ ½	т	0	0	т
0	т	т	0	т	т	т
- ½	0	0	т	- ½	- ½	т
-1	0	0	т	- 1/2	-1	т
т	т	т	т	т	т	т

Или же, если шкала не расширенная:

Сходство:	+1	0	-1	т
+1	+1	т	0	т
0	т	0	т	т
-1	0	т	-1	т
т	т	т	т	т

После подобного описания данных и определения операция сходства, можно применить алгоритм прямого ДСМ метода.

Чтобы понять алгоритм, в соответствии с которым работают правила 1-ого рода и производится поиск сходства объектов, необходимо ввести несколько определений.³⁵

О – множество всех объектов. Допустим объект (его идентификационный знак – номер респондента) отождествляется с его описанием. Фактически, **О** – это множество описаний всех объектов. В нашем случае под описанием понимается совокупность оценок признаков, соответствующих каждому объекту.

С – множество всех подмножеств характеристик, претендующих на роль возможных причин. Характеристика может состоять из совокупности значений по нескольким признакам. Если у нас, к примеру, 3 переменных, то множество характеристик

³⁵ В данном случае используется интерпретация ДСМ метода, предложенная Аншаковым О.М в статье «Об одной интерпретации ДСМ метода АПГ», НТИ, 1999 г. №1-2, стр. 45 – 53.

состоит из всех возможных наборов значений этих переменных по всем возможным подмножествам исходного множества из 3 переменных. Так, если у нас, к примеру, признаки представлены в виде дихотомических переменных, то все возможные элементы множества S можно задать следующей таблицей (см. таблица 1):

Таблица 1

1	2	3
+	+	+
+	+	-
+	-	+
+	-	-
-	+	+
-	+	-
-	-	+
-	-	-
+	+	
+	-	
-	+	
-	-	
	+	+
	+	-
	-	+
	-	-
+		+
+		-
-		+
-		-
+		
-		
	+	
	-	
		+
		-

В данной таблице представлены все возможные комбинации признаков, по которым мы ищем сходство, а также все возможные комбинации значений этих признаков в подгруппах характеристик. Подгруппами в данном случае называются комбинации самих признаков. ((1, 2, 3); (1,3); (1,2); (2,3); (1); (2); (3))

Далее задается еще два отношения:

Отношение $c < o$, означающее, что характеристика c извлекается из объекта o , выполняющееся тогда, когда мы можем зафиксировать относительно данного объекта факт наличия или отсутствия данного свойства. Также задается отношение частичного порядка $c1 \leq c2$, означающее, что характеристика $c1$ скупер характеристике $c2$. Так, например, если $c1$ – это пол, а $c2$ – это пол и возраст, то данное отношение выполняется.

Пусть множество O является подмножеством O . Тогда, если $c < O$ (c извлекается из O), то c можно считать сходством объектов подмножества O . Допустим, что c , являющееся сходством объектов подмножества O , наиболее емкая характеристика, то есть $ci \leq c$, где ci – это любая характеристика, извлекающаяся из O , являющаяся менее емкой, чем характеристика c , и при этом также являющаяся сходством объектов из O . Это значит, что c является наиболее емким сходством на множестве объектов O .

Каждый из выделенных подобъектов изначально может претендовать на роль гипотезы о причине изучаемого эффекта (свойства). Алгоритм ДСМ – метода проверяет все подобъекты на предмет выполнимости относительно них предикатов $M(+)$ и $M(-)$. $M(+)$ выполняется, если данный подобъект (сочетание признаков) одновременно встречаются среди двух и более плюс - примеров (т.е. представляют собой сходства этих примеров). $M(-)$ выполняется, если данный подобъект (сочетание свойств) одновременно встречаются среди двух и более минус – примеров. Далее, для порождения гипотез о причинах используются правила правдоподобного вывода первого рода, описанные выше. (стр.29)

В содержательном смысле гипотезы дают нам ключ к пониманию того, какие признаки (сочетания признаков) формируют пространство детерминации того или иного поведения объекта.³⁶

Возможности анализа социального поведения с помощью ДСМ-метода АПГ.

В данной главе рассматривается один из примеров применения прямого ДСМ-метода для анализа социологических данных.

В социологии одной из главных проблем является проблема несоответствия установок (аттитюдов, мнений) реальному поведению людей. А точнее вопрос звучит так: как установить взаимосвязи между аттитюдами и поведением, как взвесить вклад многочисленных социальных и индивидуально - личностных переменных, потенциально обуславливающих социальные действия?³⁷ Эту проблему возможно частично решить с помощью ДСМ - метода в контексте изучения социального поведения.

Поведенческие готовности обычно представляют собой предмет социологических и маркетинговых исследований. К примеру, вопросы заказчика обычно формулируется следующим образом:

- Купит человек товар данной марки, или не купит?
- Если купит, то почему?
- Кто этот человек – кто потенциальный потребитель данного товара?

Условно говоря, многие используемые сейчас методы направлены на решение именно этих задач. Обычно используется сегментация по различным параметрам (демографическим, психографическим и т.п.). ДСМ метод также подходит для решения этих задач. Так, на первом этапе мы получаем ответ на вопрос, почему человек купит товар, а почему нет. На втором этапе мы получаем ответ на вопрос о том, купит ли какой-либо конкретный человек с тем или иным набором характеристик данный товар. Таким же образом мы можем получить ответ на вопрос о том, кто является потребителем данного товара. Проще говоря, с помощью ДСМ метода мы получаем набор существенных характеристик, влияющих на поведенческие готовности. Далее, эти характеристики можно анализировать более подробно с помощью других методов.

³⁶ Данилова Е.Н., Михеенкова М.А., Московский С.С. Об одной модели детерминации социального поведения / Интеллектуальные системы и общество, РГТУ, М.: 2001, стр. 23

³⁷ Данилова Е.Н., Михеенкова М.А., Климова С.Г. Возможности применения логико-комбинаторных методов для анализа социальной информации // Социология: 4М. – 1999. – № 11. – с. 142-160.

Метод успешно применялся для выявления детерминант социального поведения, а именно: изучения поведенческих установок рабочих относительно участия / не участия в забастовке. Данные для эксперимента были получены путем опроса рабочих двух заводов в г. Санкт-Петербурге и в г. Ельце.³⁸ Измерялись социальные установки, а также психологические и демографические характеристики. На основе анализа и логико-комбинаторного преобразования множества переменных был получен ряд признаков, таких как индивидуализм, признание закона, уровень материального положения, лояльность директору, ощущение себя совладельцем предприятия, установка на переговоры и т.п. Данные о поведенческих готовностях формировались на основе вопроса о том, готов ли рабочий участвовать в забастовке. Рабочие, готовые принять участие в забастовке относились к множеству положительных примеров. Рабочие, не готовые участвовать в забастовке относились к отрицательным примерам. На выходе первого этапа применения метода (порождения гипотез) в качестве гипотез для каждой из этих двух подгрупп были получены несколько наборов признаков и значений этих признаков, представляющие собой как раз те самые детерминационные множества, которые в социологическом контексте можно назвать кластерами, типами социальных субъектов.

Так, например, Среди субъектов, не участвовавших в забастовке, выделились три подгруппы, приведенные в таблице 2:

Таблица 2

<i>Выделенные совокупности признаков</i>		<i>Метафоры</i>
<i>Индивидуализм</i>	+	"обиженный бедняк"
<i>Признание закона</i>	-	
<i>Ощущение враждебного окружения</i>	+	
<i>Низкий уровень материального положения</i>	+	
<i>Зависимость от предприятия</i>	+	"Советский рабочий"
<i>Ощущение себя совладельцем</i>	+	
<i>Лояльность к директору</i>	+	
<i>Низкий уровень материального положения</i>	+	
<i>Зависимость от предприятия</i>	+	"Крепосной директора"
<i>Лояльность к директору</i>	+	
<i>Установка на переговоры</i>	+	
<i>Потребность в рабочих организациях</i>	-	

Данные наборы признаков можно интерпретировать как причины отсутствия изучаемого эффекта, то есть причины нежелания участвовать в забастовке. Следует отметить, что было выявлено несколько причин, а точнее детерминант, влекущих не

³⁸ Климova С.Г., Михеевкова М.А., Панкратов Д.В. ДСМ-метод как метод выявления детерминант социального поведения // НТИ. Сер. 2. – 1999. – № 12. – с. 3-14.

желание участвовать в забастовке. Каждый из наборов можно содержательно интерпретировать как подтип рабочих, демонстрирующих не желание бастовать (отрицательные примеры изучаемого явления).

Итак, на выходе мы получаем наборы признаков и значений этих признаков, определяющих тот или иной тип поведения. Иными словами, мы находим пространство, относительно которого следует далее анализировать изучаемое явление. Статистические методы в силу своей специфики не способны дать подобные результаты. В статистике, как правило, анализируются связи между конкретными признаками и поведенческими готовностями. Логико-комбинаторный подход позволяет нам найти различные сочетания различных признаков и их значений, соответствующие разным типам поведения. Например, статистический вывод звучит так: «признак С влияет на поведение А. Чем выше значение по признаку С, тем вероятнее, что человек поведет себя в соответствии с А». Логико-комбинаторный подход говорит: если у человека выражен признак С и при этом не выражен признак Д, то человек ведет себя в соответствии с А, а если признак Р выражен, то человек не ведет себя в соответствии с А и т.п. Иными словами, мы пытаем подобными логическими высказываниями описать нашу совокупность респондентов.

На втором этапе мы можем породить гипотезы относительно поведения субъектов, которые не ответили на вопрос о готовности участвовать в забастовке. Также мы можем просто делать прогнозы, имея схожие по структуре наборы данных.

Итак, мы охарактеризовали суть и принцип работы ДСМ метода АПГ. Метод не только полезен для анализа данных, но также стимулирует логическую систематизацию данных и моделей, в соответствии с которыми проводятся исследования.

Программная реализация ДСМ-метода АПГ

Обзор программы JSM-explorer

Пакет JSM-explorer представляет собой программную реализацию ДСМ-метода, предназначенную для работы с методом в различных предметных областях. В данной главе мы попытаемся познакомить читателя с программой на примере решения некоторой гипотетической задачи.

Как уже было сказано ранее, важными задачами социологии являются задачи поиска детерминант социального поведения, мнений, установок, а также задачи поиска причин социальных событий. Предположим, что мы изучаем склонность к преступлениям и криминальным действиям, и у нас имеется некоторая база данных, описывающая группу респондентов по некоторым признакам, относящимся к изучаемому предмету. В этой базе данных представлены как преступники, так и добропорядочные граждане. Мы разделяем совокупность на эти две группы, фиксируя факт совершения / не совершения человеком тяжкого преступления. Также для каждого респондента фиксируются следующие номинальные признаки:

P1 – из неблагополучной семьи / из благополучной семьи.

P2 – низкий уровень материального положения / высокий уровень материального положения.

P3 – импульсивный / спокойный.

P4 – не имеет высшего образования / имеет высшее образование.

U – судим за совершение тяжкого преступления / не судим.

Все данные представлены в виде дихотомических переменных. Всего в базе 15 респондентов. Все мужчины. Из них 6 человек имеют судимость (их мы отнесем к (+) примерам), 7 человек не имеют судимости (их отнесем к (-) примерам) и относительно двух человек нет данных относительно судимостей. Изначально данные представлены в виде файла SPSS. (Рис1.) Переменные свойств P1, P2, P3, P4 закодированы следующим образом (1-1-я альтернатива относительно свойства, 2 – 2-я альтернатива относительно свойства, 9 наличие или отсутствие свойства не определено). Кодам переменной U соответствуют значения (1-судим, 2 – не судим, 3 – нет данных относительно судимости).

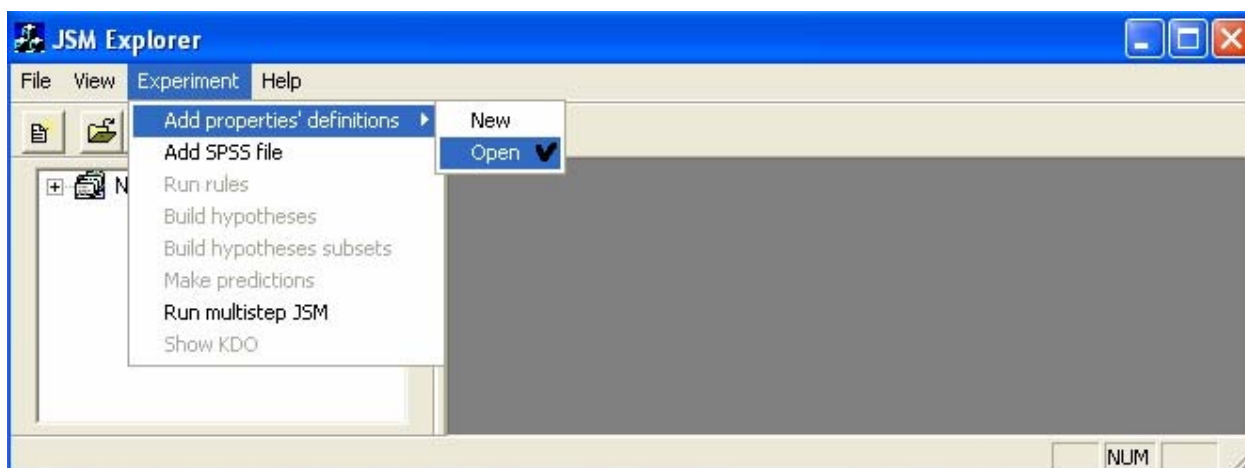
Рис.1. Данные в Spss

	P1	P2	P3	P4	U	var	var	var
1	1,00	2,00	1,00	1,00	1,00			
2	1,00	1,00	2,00	1,00	1,00			
3	1,00	1,00	1,00	2,00	1,00			
4	1,00	2,00	1,00	2,00	1,00			
5	2,00	9,00	1,00	2,00	1,00			
6	9,00	1,00	9,00	2,00	1,00			
7	1,00	2,00	1,00	1,00	2,00			

Программа JSM-explorer воспринимает данные в формате tab-delimited (dat). Поэтому сохраняем исходный файл SPSS в соответствующем формате под именем **example_1**.

Далее следует открыть этот файл из приложения JSM-explorer, однако, прежде чем мы откроем файл, необходимо задать правила интерпретации данных и функцию сходства, в соответствии с которой, будет работать алгоритм метода. Для этого заходим в команду *Experiment / Add properties definitions / Open*.

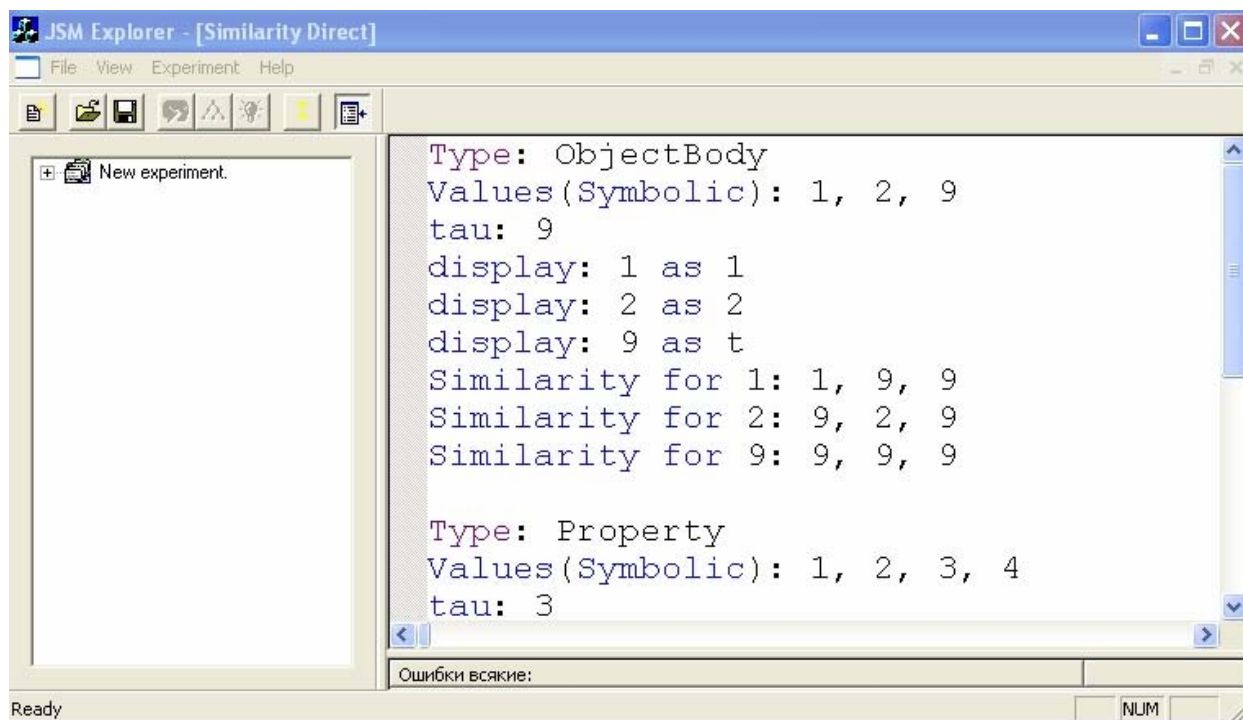
Рис.2 Окно программы JSM-explorer.



Перед нами появляется окно, предлагающее выбрать файл, содержащий необходимые описания объектов и функцию сходства, написанный в соответствии с правилами специального языка, созданного для программы JSM-explorer. Обычно этот файл создается заранее, но программа позволяет редактировать правила непосредственно в режиме работы. В нашем случае такой файл был создан заранее и сохранен под названием **Similarity Direct**. Название говорит о том, что правила, описанные в файле, предназначены для прямого ДСМ – метода.

После открытия файла в окне появляется описание необходимых свойств и функций (см. рис.3).

Рис 3: Пример описания свойств



Далее приведен листинг всего файла с описанием свойств. Разберем подробно, что означают следующие команды.

```

Type: ObjectBody ‘команда означает, что следующие строки относятся к телу объекта
Values(Symbolic): 1, 2, 9 ‘используемые коды
tau: 9 ‘определение внутренних значений
display: 1 as 1 ‘ и соответствующих им кодов
display: 2 as 2
display: 9 as t
Similarity for 1: 1, 9, 9 ‘определение функции сходства
Similarity for 2: 9, 2, 9
Similarity for 9: 9, 9, 9
  
```

```

Type: Property ‘описание изучаемого эффекта (свойства)
Values(Symbolic): 1, 2, 3, 4
tau: 3 ‘определение внутренних значений
plus: 1 ‘ и соответствующих им кодов для свойств
minus: 2
zero: 4
display: 1 as +1
display: 2 as -1
display: 3 as t
display: 4 as 0
  
```

В данном файле особенно важно корректно задать функцию сходства, т.к. она должна отражать наши содержательные представления относительно изучаемых признаков.

В нашем случае имеет смысл сходство по первому и второму значению каждого из признаков, однако сходство на неопределенности смысла не имеет. Так, например, если для двух респондентов неизвестно, какими из возможных значений признака P1 они обладают, то сходны они лишь в том, что исследователю не удалось зафиксировать значения данного признака относительно этих респондентов.

Нижеприведенной матрицей задается функция сходства. 9 (тау) означает, что сходство для данного сочетания значений признаков не определяется.

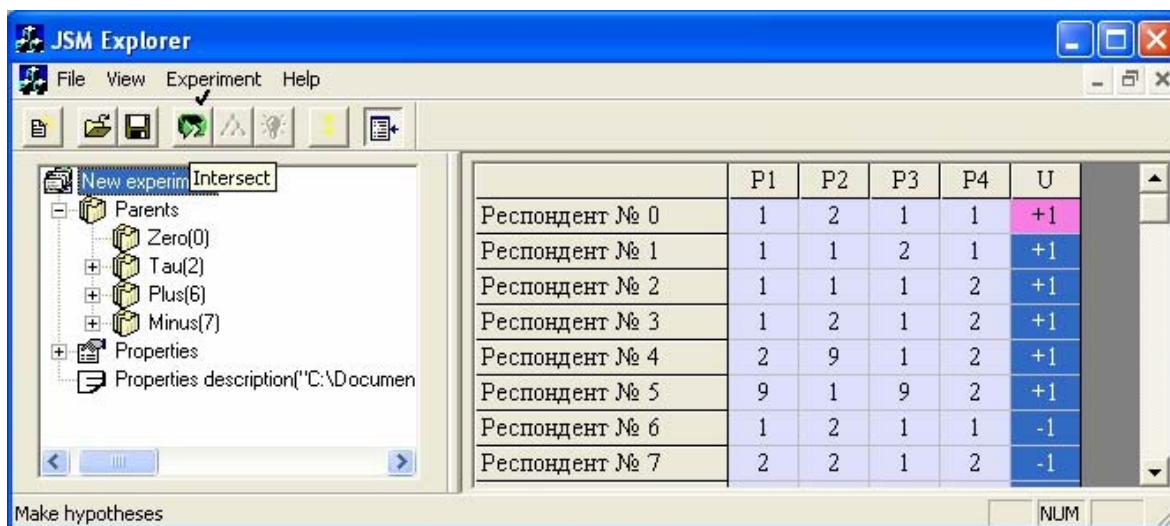
Similarity for 1: 1, 9, 9

Similarity for 2: 9, 2, 9

Similarity for 9: 9, 9, 9

Теперь мы можем открыть сохраненный файл example_1, посредством процедуры **File / Create from tab-delimited**. Появляется диалоговое окно, в котором предлагается выбрать переменные, описывающие тело объекта, а также переменную свойства. После того, как мы поместили переменные в соответствующие колонки, появляется окно эксперимента, состоящее из окна с данными, а также окна проекта.


Рис 4. Окно проекта, данные



The screenshot shows the JSM Explorer software interface. On the left, there is a tree view under 'New experim Intersect' with folders for 'Parents' (containing Zero(0), Tau(2), Plus(6), Minus(7)) and 'Properties' (containing Properties description('C:\Documen...)). The main area displays a data table with the following content:

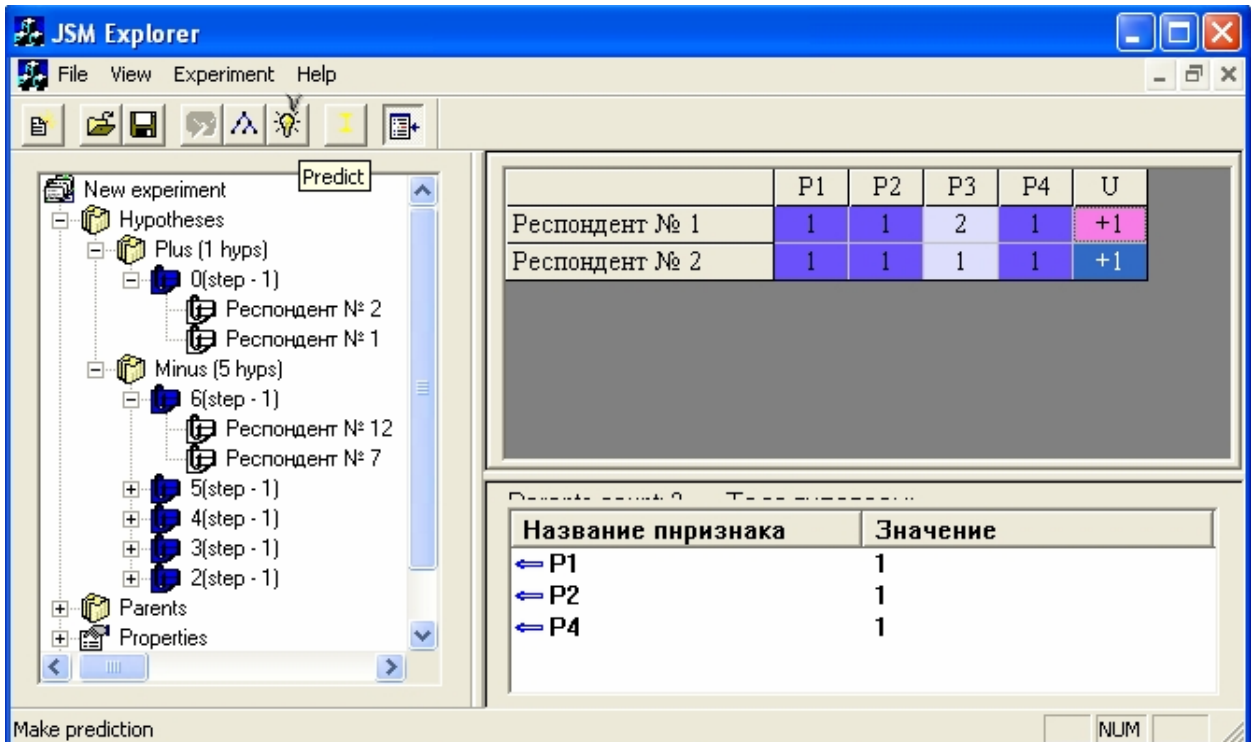
	P1	P2	P3	P4	U
Респондент № 0	1	2	1	1	+1
Респондент № 1	1	1	2	1	+1
Респондент № 2	1	1	1	2	+1
Респондент № 3	1	2	1	2	+1
Респондент № 4	2	9	1	2	+1
Респондент № 5	9	1	9	2	+1
Респондент № 6	1	2	1	1	-1
Респондент № 7	2	2	1	2	-1

At the bottom of the window, there is a 'Make hypotheses' button and a 'NUM' field.

В окне проекта представлена информация о родителях ((+) примерах, (-) примерах и неопределенных примерах). В левой части окна расположена матрица данных. Теперь с помощью команды **Experiment / Make hypothesis**, обозначаемой иконкой , запускаем

процедуру построения гипотез (ППВ-1). Результаты первого этапа выглядят в виде множества гипотез на (+) и на (-) примерах. (см. рис. 5)

Рис.5 Множество гипотез



На выходе мы имеем одну плюс-гипотезу и 5 минус-гипотез относительно детерминант склонности к преступлениям. Гипотеза представлена в виде множества респондентов, на которых было найдено соответствующее сходство, а также в виде тела гипотезы (сочетания признаков). Для интерпретации все множество гипотез можно представить в виде следующей таблицы:

Таблица 3: гипотезы

Гип.	Число родителей	Тело гипотезы		
+	2	Из неблагополучной семьи	импульсивный	не имеет высшего образования
	2	низкий уровень материального положения	спокойный	имеет высшее образование.
-	3	из благополучной семьи.	высокий уровень материального положения.	
	2	из благополучной семьи.	спокойный	
	2	высокий уровень материального положения	спокойный	
	2	из благополучной семьи	высокий уровень материального положения	имеет высшее образование
	2	из благополучной семьи	высокий уровень материального положения	имеет высшее образование

Теперь мы можем попытаться содержательно проинтерпретировать получившиеся гипотезы.

Так, мы видим, что по нашим данным, склонность к преступлениям детерминируется сочетанием трех показателей:

- не благополучная семья.
- импульсивность.
- отсутствие высшего образования.

Данная гипотеза представляет для нас ценность, прежде всего как результат, помогающий на качественном уровне понять характер причинно-следственных связей и закономерностей, существующих в изучаемой предметной области. Основываясь на содержательной стороне гипотезы, мы можем построить небольшую концептуальную схему.

Предположим, что склонность к преступлению является результатом установки, которая в определенной ситуации толкает человека на преступные действия или, по крайней мере, не препятствует им. Неблагополучная семья воспитывает в человеке склонности и ценности, ориентированные на принятие силы, унижения, агрессии и т.п. Это в дальнейшем может привести к определенным фрустрациям. Импульсивность может усиливать возникающие фрустрации и влечь противоправное поведение. При этом отсутствие высшего образования, как фактора во многом формирующего определенные культурные, моральные и нравственные нормы, не тормозит преступные действия человека. Аналогичным образом мы можем «домыслить» модель с учетом минус - гипотез.

Таким образом, главное предназначение ДСМ - метода на данном этапе - усиление интеллектуальных возможностей исследователя. Метод помогает нам сделать на основе имеющихся фактов некоторые обобщения в виде логических высказываний относительно причинно-следственных связей. Так, если мы посмотрим на данные, то путем мысленных сравнений и сопоставлений мы можем сами прийти к некоторым из полученных гипотез. На рисунке 6 наглядно изображены полученные гипотезы. Желтым цветом показана (+) гипотеза, другими цветами (-) минус гипотезы. Однако человеческие интеллектуальные способности вряд ли позволят выявить подобные сходства и провести подобные сопоставления даже на таком маленьком массиве данных. В реальных же исследованиях признаков и респондентов значительно больше.

Далее метод позволяет нам построить предсказание относительно случаев с неопределенным исходом. Предсказание осуществляется с помощью команды **Experiment / Make Predictions**. В результате работы данного этапа метода респонденту №13 было

предсказано значение (-1) по гипотезам 6 и 3. Однако не удалось получить предсказание относительно Респондента №14. В его описании не было обнаружено ни полюс - гипотез ни минус - гипотез (Рис. 7).

Рис.6 Гипотезы

P1	P2	P3	P4	U
1,00	2,00	1,00	1,00	1,00
1,00	1,00	2,00	1,00	1,00
1,00	1,00	1,00	1,00	1,00
1,00	2,00	1,00	2,00	1,00
2,00	12,00	1,00	2,00	1,00
12,00	1,00	12,00	2,00	1,00
1,00	2,00	1,00	1,00	2,00
2,00	2,00	1,00	2,00	2,00
1,00	1,00	2,00	2,00	2,00
2,00	1,00	2,00	2,00	2,00
2,00	2,00	2,00	12,00	2,00
1,00	2,00	2,00	1,00	2,00
2,00	2,00	1,00	1,00	2,00
2,00	2,00	1,00	1,00	3,00
2,00	1,00	1,00	1,00	3,00

Рис.7 Предсказания

	P1	P2	P3	P4	U
Респондент № 6	1	2	1	1	-1
Респондент № 7	2	2	1	2	-1
Респондент № 8	1	1	2	2	-1
Респондент № 9	2	1	2	2	-1
Респондент № 10	2	2	2	t	-1
Респондент № 11	1	2	2	1	-1
Респондент № 12	2	2	1	1	-1
Респондент № 13	2	2	1	1	-1p(1)
Респондент № 14	2	1	1	1	t

Итак, нами был проведен анализ гипотетических данных с помощью первого и второго этапов метода. Были получены гипотезы и построен прогноз. Помимо этого в программе JSM-explorer реализована возможность пошагового применения метода и проверки аксиомы каузальной полноты (АКП) на каждом шаге эксперимента. Как уже было сказано ранее, АКП выполняется в том случае, если после осуществления предсказаний (ППВ-2) каждый положительный факт содержит в себе положительные

причины, каждый отрицательный факт содержит в себе отрицательные причины и отсутствуют противоречия. Если АКП не выполняется, то необходимо расширение базы данных.

На данный момент мы рассмотрели общие принципы работы программы JSM-explorer, продемонстрировали работу первого и второго этапов метода на простом гипотетическом примере. Мы не будем более подробно касаться многоэтапного ДСМ - метода и процедур проверки АКП, так как, во-первых, это напрямую не связано с экспериментальными задачами, решаемыми в данной работе (см. следующую главу), а, во-вторых, специфика социологических исследований практически не позволяет нам до конца объяснить поведение объекта исследования с помощью причинно-следственных связей. Большинство социологических выводов делаются на квазидостаточном основании.

Теперь суммируем основные положения, касающиеся особенностей и ограничений ДСМ-метода АПГ относительно анализа социологической информации.

Особенности и ограничения метода

Главная особенность метода заключается в том, что в ДСМ-методе реализуются автоматизированные рассуждения, на базе которых формулируются гипотезы о причинно-следственных зависимостях, неявно содержащихся в данных. Обычно данный этап исследования предшествует этапу, на котором применяются численные методы анализа данных. ДСМ-метод является средством, усиливающим интеллектуальные возможности исследователя в плане поиска неявных взаимосвязей, содержащихся в данных. ДСМ-метод АПГ можно отнести к эвристическим методам.

Метод имеет ряд ограничений, связанных со следующими пунктами:

Во-первых, ДСМ-метод чувствителен к случайным колебаниям в данных. Так, если несколько респондентов каким-то образом ошиблись, это может привести к опровержению истинной гипотезы или порождению ложной гипотезы. Здесь статистика имеет значительное преимущество.

Во-вторых, ДСМ-метод нельзя назвать универсальным методом работы с данными. Он требует тщательной адаптации к любой области знаний, в которой применяется. Поэтому реализация ДСМ – метода в рамках универсального программного пакета является сложной задачей, но с сегодняшними темпами развития информационных технологий она становится все более реальной.

Раздел 3. Возможности использования ДСМ-метода АТГ в сочетании с методом виньеток (на примере анализа восприятия религиозности)

Метод виньеток как метод сбора данных

Общее представление о методе виньеток

На сегодняшний день в социологии достаточно часто применяются многофакторные эксперименты. Под экспериментом понимается опытное исследование воздействия группы факторов на зависимую переменную, интересующую исследователя. В последнее время был сделан ряд открытий и инноваций в области планирования эксперимента. Одной из таких инноваций является метод виньеток, позволивший совместить экспериментальные методы с широкомасштабными выборочными исследованиями. Метод виньеток позволяет получить достаточно валидные и надежные данные и представляет собой хороший инструмент для измерения мнений и установок.

Использование метода виньеток в экспериментальных планах приобрело популярность после выхода статьи К. Тёрмана, П. Росси и Дж. Лама «Факторный эксперимент в изучении обыденного восприятия психических заболеваний»³⁹. В данной статье предлагается новый подход к изучению обыденного восприятия поведенческих симптомов психических заболеваний. Авторы продемонстрировали возможности применения факторных экспериментов для изучения установок людей по отношению к сложным социальным объектам. В роли таких объектов выступают виньетки - краткие описания индивидов или социальных ситуаций, в которых систематически варьируются различные признаки, представляющие собой множество факторов, влияющих на формирование социальных установок относительно изучаемых объектов.

К. Тёрман, П. Росси и Дж. Лам уделяли большое внимание изучению обыденного восприятия различных форм и проявлений психических заболеваний. Исследователей интересовал вопрос о том, какие поведенческие факторы влияют на оценку психологического состояния индивида.

³⁹ Thurman Q.C., Lam J.A., Rossi P.H. A factorial survey approach to the study of popular conceptions of mental illness//The Sociological Quarterly, Vol. 29, № 4, 1988, p. 565-588

В процессе исследования авторами было создано большое количество виньеток, представлявших собой описания гипотетических индивидов, обладающих различными социально-демографическими характеристиками и различными симптомами психологических заболеваний. В виньетках было представлено множество вариаций симптомов психологических заболеваний с различной степенью выраженности и различными формами проявлений (уровнями). Общий объем выборки составил 7150 виньеток. К. Тёрман, П. Росси и Дж. Лам провели исследование на выборке из 143 студентов. Каждому респонденту давалась анкета, состоявшая из 50 виньеток, случайным образом отобранных из общего числа виньеток. В качестве зависимой переменной выступала субъективная оценка уровня психического заболевания индивидов, описанных в винюетке. Независимыми переменными являлись факторы, представлявшие собой симптомы психических заболеваний. В результате эксперимента авторы пришли к выводу о том, что именно психологические факторы являются определяющими в оценке психологического состояния индивидов. Также авторы продемонстрировали широкие возможности использования метода виньеток в экспериментальных исследованиях.⁴⁰

Метод виньеток позволяет решить некоторые проблемы, связанные с качеством получаемой от респондента информации. Во-первых, респонденту предлагается оценивать не абстрактные признаки, а комплексные ситуации, что в значительной степени приближает условия формирования оценок к условиям реальной жизни. Во-вторых, метод не принуждает респондента рефлексировать относительно факторов, которые влияют на его решения, а лишь требует от респондента оценки комплексной ситуации. Респонденты, как правило, не осознают того, какие факторы в действительности влияют на их решения. Им также бывает сложно оценить степень влияния различных факторов на принятия решений. Метод виньеток в значительной мере позволяет избежать этих проблем. В-третьих, метод позволяет более точно оценить эффект изменения общей оценки ситуации относительно изменения какого-либо отдельного параметра.

Итак, суммируем основные особенности метода виньеток. Во-первых, единицами анализа являются не респонденты, а винюетки. Таким образом, объем выборки измеряется количеством винюеток. Во-вторых, набор винюеток определяется экспериментальным планом в соответствии с моделью факторного эксперимента. В-третьих, при формировании винюеток исследователем варьируется не один конкретный фактор, а целое

⁴⁰ Подробнее смотреть выпускную квалификационную работу Добрыниной Стелы / Использование методов винюеток и репертуарных решеток в исследовании обыденного восприятия традиционных и новых форм религиозности/ Методическую часть

множество различных факторов, которые вместе описывают многомерную социальную конструкцию.

В задачу данной главы не входит подробное описание метода виньеток и возможностей его применения в социологии.⁴¹ Наша задача заключается в демонстрации аналитических возможностей ДСМ – метода на данных, собранных с помощью метода виньеток. Для этого нам достаточно имеющегося обзора. Теперь перейдем непосредственно к описанию эксперимента и исходной базы данных.

Описание условий эксперимента

Исследование, проведенное Добрыниной Стелой в рамках дипломной работы, было посвящено изучению обыденного восприятия религиозности. Цель исследования заключалась в определении того, каким образом формируется мнение относительно людей, в поведении которых имеются те или иные признаки религиозности.

Предполагалось, что людей в виньетках можно классифицировать по степени религиозности. Степень религиозности выступала в качестве зависимой переменной. Также предполагалось, что на оценку степени религиозности влияют следующие группы факторов:

- внешний вид человека
- поведение человека
- совершаемые человеком религиозные ритуалы
- религиозные убеждения.

Каждый из этих признаков был разбит на группы подпризнаков. Так, например, внешний вид был разделен на одежду и прическу. Религиозные ритуалы были разделены на такие поведенческие характеристики как соблюдение поста, посещение храма, молитву. Каждая виньетка состояла из четырех уровней четырех описанных факторов.

Так как метод виньеток впервые применялся к изучению обыденного восприятия религиозности, а исследование проводилось в условиях ограниченности ресурсов (студенческая работа), эксперимент проводился на малой выборке. В результате эксперимента было опрошено 48 человек (из них 24 студента и 24 человека в возрасте от 40 до 55 лет). При этом использовалась выборка доступных случаев. Таким образом, в данном исследовании не идет речи о какой-либо внешней валидности эксперимента и по полученным данным вряд ли можно делать выводы относительно генеральной совокупности. Но в задачи данной работы не входила оценки внешней валидности, а

⁴¹ Это описание можно найти в дипломной работе Добрыниной Стелы.

целью эксперимента было выявление причинно-следственных взаимосвязей, если таковые существовали. Поэтому процедуру сбора данных можно считать корректной относительно поставленных задач. А данные можно считать корректными для анализа на качественном уровне.

Изначально последовательность виньеток состояла из 116 штук. Однако один респондент не в состоянии оценить такое количество ситуаций. Поэтому данная последовательность случайным образом была разделена на 4 равные части, после чего каждому респонденту предъявлялась анкета (один блок) состоящая из 29 виньеток. Для минимизации эффекта влияния одной винюетки на другую, каждому респонденту винюетки предъявлялись в случайном порядке.

Анкета представляла собой напечатанный на девяти страницах бланк с инструкцией. В инструкции человеку было предложено принять участие в социологическом эксперименте, заполнив предлагаемую анкету, далее описывалась задача, которую предстояло выполнить респонденту, после чего была дана расшифровка баллов шкалы, по которой должна была проводиться оценка.

Весь эксперимент проводился лично исследователем. Анкета давалась респондентам для самостоятельного заполнения, однако всегда существовала возможность посоветоваться с исследователем. Заметим, что такой потребности обычно не возникало, что свидетельствует о том, что в инструкции было вполне достаточно информации для заполнения анкеты. В конце анкеты респонденты должны были ответить на два открытых вопроса о том, по какой причине ими был проставлен тот или иной балл. И в заключения следовали вопросы о поле, возрасте, образовании респондента. По времени заполнения анкета занимала от 15 до 25 минут.

По каждому блоку виньеток было опрошено по 12 человек (6 студентов и 6 не-студентов). Однако в качестве исследуемого объекта были взяты не люди, а винюетки, следовательно, объем выборки составил 1392 винюетки.⁴²

Собранные данные были закодированы в файле SPSS. Изначально в файле было 82 переменных, соответствующие следующим блокам:

- социально-демографические характеристики
- внешний вид человека
- поведение человека
- совершаемые человеком религиозные ритуалы
- религиозные убеждения.

⁴² См. дипломную работу Добрыниной Стелы / Использование методов винюеток и репертуарных решеток в исследовании обыденного восприятия традиционных и новых форм религиозности/ Процедура сбора данных и выборка.

- оценка религиозности.
- социально-демографические характеристики респондентов (подробнее см. приложение 1).

Был проведен предварительный анализ частотных распределений. Затем была построена регрессионная модель, описывающая влияние отдельных характеристик и их уровней на оценку религиозности. Результаты анализа позволили сделать следующие содержательные выводы:⁴³

- социально демографические характеристики не оказывают существенного влияния на оценку религиозности. Их можно исключить из дальнейшего анализа.

- среди таких показателей, как внешний вид человека, поведение человека, совершаемые человеком религиозные ритуалы, религиозные убеждения существуют характеристики, оказывающие значимое влияние на оценку религиозности. Среди них в основном выделяются характеристики православия.

- была сформулирована гипотеза о том, что достаточно сильное влияние на оценку религиозности могут оказывать не столько отдельные характеристики, сколько их взаимодействия. В связи с этим были поставлены следующие задачи: введение в модель взаимодействий переменных, отбор наиболее существенных взаимодействий, претендующих на роль причин восприятия человека как религиозного / нерелигиозного, дальнейший анализ полученных взаимодействий. Для решения этих задач нами было предложено использовать прямой ДСМ – метод АПГ.

⁴³ Подробнее см. там же, Анализ результатов эксперимента по методу виньеток.

Задача поиска существенных взаимодействий

Задача поиска существенных взаимодействий заключается в выделении из всего множества сочетаний характеристик, варьируемых в виньетках, таких сочетаний, которые оказывают достаточно сильное положительное или отрицательное влияние на оценку религиозности. Иными словами, задача заключается в поиске таких частей описаний, которые можно считать причинами высоких / низких оценок. Задача актуализируется в силу следующих обстоятельств:

1) Каждая переменная в отдельности не оказывает сильного влияния на оценку религиозности. Однако некоторые сочетания переменных и соответствующих уровней этих переменных, по мнению исследователей, способны оказывать существенное влияние на оценку.

2) Введение в модель всех возможных взаимодействий переменных для дальнейшего отбора из них наиболее значимых с помощью множественной регрессии в принципе невозможно в силу того, что число переменных в модели будет слишком большим. Так как в модели 57 переменных, только двойных взаимодействий можно составить $57 + 56 + 55 + \dots + 2 + 1 = 57 \cdot (57 + 1) / 2 = 1653$. Это число больше числа респондентов в «растянутой» базе данных (1392). А если ввести в модель еще и тройные взаимодействия, то число переменных значительно превысит число кейсов. В такой ситуации, ни о какой модели множественной регрессионной не может быть и речи. Правда есть вариант построения большого числа парных регрессионных моделей с каждой из независимых переменных. Но эта процедура очень громоздка и не дает возможности использовать огромный плюс факторного эксперимента, реализуемого с помощью метода виньеток, выражающийся в отсутствие корреляций между независимыми переменными. Эта процедура также не дает возможности учитывать совместное влияние всего множества факторов.

В итоге был сделан вывод о том, что традиционные статистические методы не позволяют нам решить задачу поиска существенных взаимодействий в условиях данного эксперимента. В результате этого нами был предложен вариант решения данной задачи с помощью ДСМ – метода АПГ в сочетании с некоторыми статистическими инструментами. Был разработан следующий алгоритм действия:

1) Исходная база данных, состоящая из 1392 респондентов, агрегируется по номерам виньеток таким образом, чтобы в результате были получены средние оценки религиозности относительно каждой виньетки. Так как каждая виньетка оценивается 12-ю респондентами, мы получаем массив данных, состоящий из 116 виньеток ($1392/12 = 116$).

2) Все множество виньеток в агрегированном массиве в зависимости от оценок разбивается на 2 части - религиозные / нерелигиозные (т.е. (+) и (-) примеры). Критерии разбиения определяются в процессе работы с базой данных.

3) Проводятся необходимые процедуры обработки данных, после чего применяется алгоритм прямого ДСМ - метода АПГ. Цель данного этапа работы - поиска существенных сочетаний признаков, соответствующих положительным и отрицательным фактам (т.е. религиозным и нерелигиозным оценкам).

4) Полученные множества сочетаний преобразуются в переменные и подвергаются проверке (внутренней валидации) на исходном массиве из 1392 респондентов. Проверка производится путем подсчета и соотнесения некоторых статистических показателей (подробнее см. ниже).

5) Строится регрессионная модель, в которой в качестве независимых переменных выступают сочетания признаков, а в качестве зависимой переменной – оценка религиозности. Посредством анализа регрессионных коэффициентов и их значимости отбираются наиболее существенные взаимодействия.

Процедуры и результаты анализа данных

Теперь опишем ход и результаты аналитической работы, подробно рассмотрев каждый из вышеописанных пунктов.

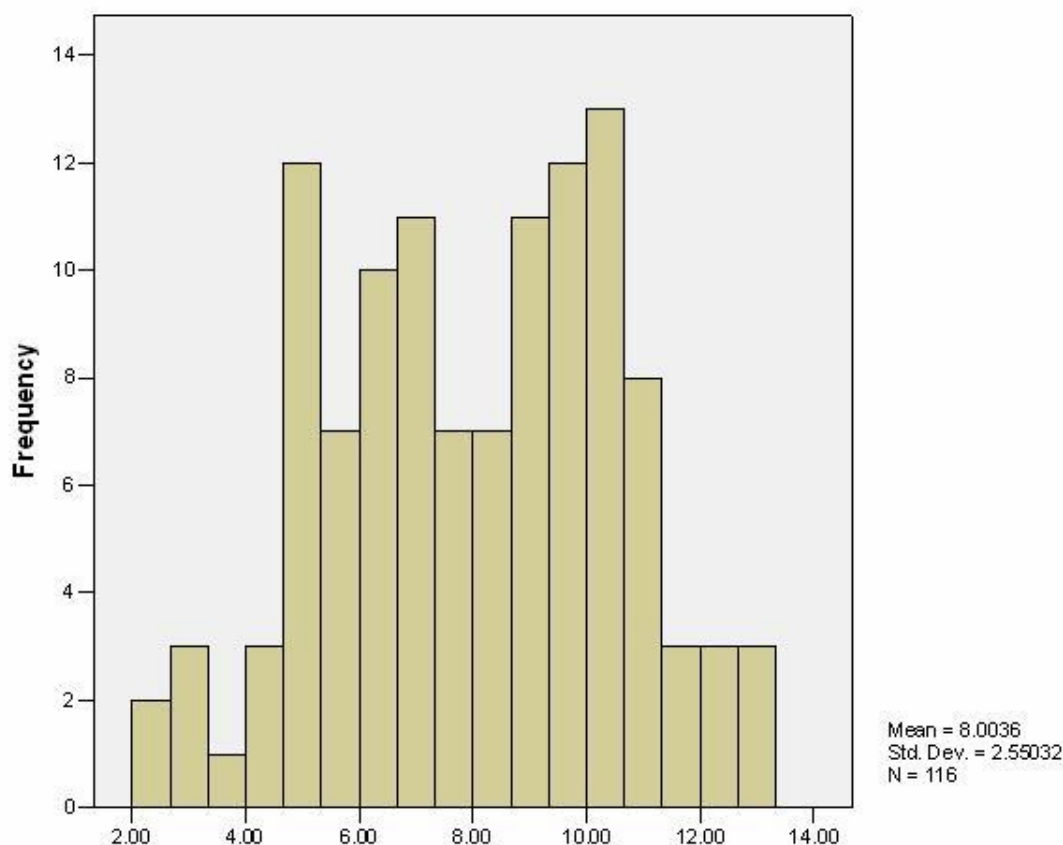
1) Агрегирование. Агрегированный массив включал в себя 59 переменных. Из них 57 переменных отражали характеристики виньеток. Одна переменная представляла собой среднюю оценку степени религиозности человека, описываемого в данной виньетке. Также в массиве была переменная с номером виньетки. Переменные, описывающие характеристики респондента были дихотомическими. Переменная Q, отражавшая среднюю оценку степени религиозности была измерена по интервальной шкале.⁴⁴ В целом, в процессе обработки данных проблем не возникло.

⁴⁴ Изначально виньетки оценивались по шкале от 1 до 15. Предполагалось, что данная шкала интервальная. Поэтому средние оценки виньеток также можно считать интервальными оценками.

2) Разбиение массива на положительные и отрицательные примеры. Анализ распределения средних оценок показал, что среди виньеток можно выделить два подмножества – религиозные и нерелигиозные. Распределение оценок сильно напоминает двугорбое распределение (Рис. 8). Это дало нам основания полагать, что в массиве данных действительно есть положительные и отрицательные примеры изучаемого эффекта (оценки религиозности) И позволяет нам разбить массив на две части.

Также мы предполагали, что среди изучаемых описаний (виньеток) существуют такие сочетания признаков (подобъекты), влекущие положительные оценки и есть сочетания, влекущие отрицательные оценки. Иными словами, мы предполагали, что в данной базе данных в неявном виде содержатся причинно-следственные зависимости. Эти предположения дают нам право воспользоваться прямым ДСМ – методом для поиска причин положительной/отрицательной оценки религиозности людей. (А точнее первым этапом ДСМ – метода, использующим ППВ-1).

Рис. 8 Распределение средних оценок виньеток (Q)



На данном этапе анализа возникла серьезная проблема, касающаяся поиска границы разбиения виньеток на + и – примеры. Проблема заключалась в том, что зависимая переменная представляла собой распределение случайной величины,

складывающееся из распределений средних оценок виньеток, также представляющих собой случайные величины. Каждая средняя оценка в агрегированной базе представляет собой оценку математического ожидания случайной величины, отражающей распределение оценок данной виньетки в генеральной совокупности. Дисперсия оценок каждой виньетки позволяет определить доверительный интервал для математического ожидания оценки данной виньетки. В нашем случае это означает, что каждая виньетка, в действительности оцениваемая в генеральной совокупности как сильно-религиозная, в нашей выборке может получить другую оценку степени религиозности с некоторой долей вероятности. И это оценка с некоторой долей вероятности в реальности может быть нерелигиозной. Из данных статистических соображений вытекают следующие выводы:

- чем выше среднее значение оценок религиозности данной виньетки, тем меньше вероятность того, что эта виньетка в действительности описывает нерелигиозный тип человека. И наоборот.

- Чем ближе оценка к центру шкалы, тем больше вероятность ее ошибочного отнесения к (+) или (-) примерам. В нашем случае это означает, что если мы разобьем шкалу пополам, используя, например, в качестве границы значение 8, то достаточно много реальных плюс - примеров могут оказаться в нашем анализе в минус - примерах и наоборот. При этом мы рискуем потерять много существенных гипотез.

- При установлении не пороговой границы разбиения, а некоего интервала, посредством которого будет отсекается доля примеров, получивших средние оценки (т. е. отсекается доля противоречивых примеров), мы должны учитывать следующее обстоятельство: при расширении интервала уменьшается вероятность ошибочного отнесения (+) примеров к (-) примерам и наоборот. Но при этом уменьшается число кейсов, участвующих в анализе. Это с одной стороны может привести к уменьшению доли ошибочных гипотез, а с другой стороны сократить общее число гипотез. При этом из-за отсутствия достаточного количества примеров могут не быть порождены значимые для дальнейшего анализа гипотезы. Если мы будем двигать только одну границу «области противоречий», то вышеизложенные рассуждения будут верны, отдельно для (+) гипотез и для (-) гипотез, в зависимости от того, с какой стороны меняется граница.

Данные соображения наталкивают на мысль о поиске некоего оптимального решения (разработке и минимизации некоего функционала). Однако эта задача требует массы усилий и не укладывается в рамки данной работы. При этом объем выборки и способ ее построения не позволяют нам провести такого анализа. Проблема усложняется тем, что оценка религиозности описывает некий интервальный конструкт, континуум, на котором нельзя четко отделить религиозных людей от нерелигиозных.

Эти рассуждения привели нас к выводу о том, что статистическими средствами сделать разбиение в данной ситуации не представляется возможным. Поэтому пороговые границы разбиения определялись в процессе нескольких экспериментов таким образом, чтобы как можно больше фактов содержало в себе гипотезы.

3) Процедура порождения гипотез с помощью ДСМ – метода. Было проведено три эксперимента с разными вариантами разбиений. В результате был выбран следующий вариант разбиения:

- К (+) примерам были отнесены виньетки, имеющие оценки выше 8,5.
- К (-) примерам были отнесены виньетки, имеющие оценки ниже 6,5.

При данном разбиении 55 виньеток были отнесены к (+) примерам, 36 виньеток были отнесены к (-) примерам. 25 виньеток были отнесены к противоречивым примерам и не почувствовали в поиске причин. По результатам работы первого этапа ДСМ - метода 49 из 55 положительных фактов содержали в себе (+) причины и 25 из 36 отрицательных фактов содержали в себе (-) причины. Мы не можем утверждать, что такое разбиение является оптимальным, но из трех пробных вариантов этот вариант разбиения оказался наилучшим. Наибольшее число фактов содержало в себе гипотезы.

В результате работы ДСМ – метода были получены 41 плюс - гипотеза и 19 минус – гипотез в виде комбинаций значений признаков. В следующей таблице для наглядности приведены некоторые примеры гипотез (см. Таблица 3). Все гипотезы приводятся в приложении (см. Приложение 2).

Таблица 3 Примеры + гипотез

№ гипотезы	Тело гипотезы		
1	Скромно одевается носит длинную юбку и блузу и голову покрывает платком		
2	Не разговаривает с людьми, которые не придерживаются его религии		
3	Постится один месяц в году	Читает только религиозные тексты и не читает никакой	
4	Уединяясь совершает пятикратную молитву в течении дня в строго отведенное для этого		
5	Он посещает религиозные собрания и богослужения каждый день		
6	Постится один месяц в году	Она носит длинные волосы, никогда не стрижет их и не	Читает только религиозные тексты и не читает никакой другой литературы

Мы не будем подробно останавливаться на процедурах подготовки данных для программы JSM – explorer. Они аналогичны процедурам, описанным в предыдущем разделе.

4) Проверка полученных взаимодействий на исходном массиве. Полученные гипотезы были представлены в исходном массиве данных в виде переменных. Значения переменных рассчитывались как произведения значений переменных, входящих в тело гипотезы. Так как переменные дихотомические, переменная, соответствующая гипотезе, принимает значение 1 в том и только том случае, если описание виньетки содержит в себе все признаки, указанные в гипотезе. В остальных случаях эта переменная принимает значение 0. Так, например, переменная, соответствующая третьей гипотезе, принимает значение 1 только в тех случаях, когда в описании виньетки одновременно встречаются следующие характеристики: постится один месяц в году, читает только религиозные тексты и не читает никакой другой литературы. В остальных случаях переменная принимает значение, равное 0. В результате в массиве были созданы 60 новых переменных (Нур1, Нур2.....Нур60).

На данном этапе встала задача проверки того, являются ли в действительности положительные гипотезы причинами высоких оценок степени религиозности, а отрицательные гипотезы – причинами низких оценок степени религиозности. Проверка производилась на исходном массиве данных путем подсчета средних оценок для каждой гипотезы. За базу брались виньетки, содержащие в себе данную гипотезу. В результате для плюс - гипотез средние оценки оказались достаточно высокими, а для минус – гипотез средние оценки оказались достаточно низкими. (См. табл. 9, 10). При этом результаты Т-теста на равенство средних для двух независимых выборок (множеств гипотез), показали, что средние значения оценок религиозности среди двух групп (плюс - гипотез и минус - гипотез) значительно различаются (см. Табл. 11). Таким образом, сделав проверку на исходных данных, мы убедились в том, что положительные гипотезы действительно являются причинами высоких оценок степени религиозности, а отрицательные гипотезы – причинами низких оценок степени религиозности.

Табл. 10 Средние оценки для минус - гипотез

Нур	mean	N	St.dev	Нур	mean	N	St.dev
42	5,81	36	3,232155	52	4,39	36	2,749314
43	5,63	24	3,680885	53	4,54	24	2,765929
44	5,63	24	3,997961	54	4,42	24	3,562079
45	5,67	24	4,124863	55	4,04	24	3,085438
46	5,42	24	3,133503	56	4,71	24	3,665102
47	5,94	36	3,992453	57	3,42	36	3,516289
48	5,08	36	2,822107	58	4,29	24	2,678078
49	5,08	24	3,119597	59	3,83	24	2,632021
50	4,79	24	2,918891	60	3,58	24	3,693904
51	3,85	48	3,319764	Average	4,74	0,79	

Табл. 9 Средние оценки для плюс-гипотез

Hyp	mean	N	St.dev	Hyp	mean	N	St.dev
1	10,66	96	3,650027	22	10,29	24	3,342469
2	10,08	48	4,485596	23	9,85	60	3,363941
3	11,72	36	3,230313	24	10,13	24	3,054754
4	10,36	72	3,54967	25	10,04	24	3,532202
5	9,85	108	3,795755	26	10,04	24	3,928482
6	11,33	24	3,619592	27	10,13	24	3,825686
7	10,97	36	2,999868	28	9,92	36	3,952395
8	10,86	36	3,570936	29	9,10	96	4,123052
9	10,06	96	3,690136	30	10,79	24	3,400501
10	11,38	24	3,254595	31	10,13	24	3,53015
11	10,67	24	3,031238	32	9,58	24	4,148511
12	10,69	36	3,77828	33	9,92	24	2,636148
13	10,71	24	3,961545	34	9,06	36	3,94204
14	10,71	24	3,141298	35	9,81	48	3,647536
15	10,58	24	3,322082	36	8,94	36	4,18804
16	10,47	36	3,202554	37	10,21	24	3,729776
17	10,33	24	3,002414	38	9,36	36	2,958308
18	11,42	24	2,903022	39	9,33	24	3,818339
19	10,86	36	4,008226	40	10,25	24	3,937004
20	11,21	24	3,078384	41	8,83	24	3,434691
21	10,38	24	3,333786	Average	10,27	0,69	

Табл. 10 Результаты T-теста на равенство средних

Independent Samples Test								
		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Error Differenc
Hyps	variances assumed	1,300255	0,2588537	27,47	58,00	0,00	5,53	0,20
	Equal variances not assumed			26,06	31,06	0,00	5,53	0,21

5) Отбор наиболее значимых взаимодействий. Построение регрессионной модели. На предыдущем этапе были получены все возможные взаимодействия переменных, влекущие положительные и отрицательные оценки религиозности. Однако поиск этих взаимодействий осуществлялся с помощью логико-комбинаторных средств (прямого ДСМ - метода), что не позволяет нам количественно соотнести влияния различных взаимодействий на зависимую переменную для отбора из них наиболее значимых. При этом статистической проверке подвергалось каждое из взаимодействий в отдельности, что не позволяет нам построить единую модель, обладающую предсказательной силой и учитывающую влияние всех гипотез на зависимую переменную. Для преодоления этих трудностей было предложено построить модель множественной линейной регрессии.

В качестве метода построения множественной регрессионной модели мы решили использовать один из пошаговых методов, автоматически включающих переменные в модель в зависимости от их вклада. Изначально был выбран метод Stepwise, поэтапно включающий в модель переменные в зависимости от силы их влияния на зависимую переменную, которая определяется изменением коэффициента детерминации при добавлении каждой переменной в модель. В результате работы метода в модель были включены лишь переменные, значимо влияющие на зависимую переменную (См. табл.11). Далее для проверки полученных результатов были построены еще одна подобная модель, но с использованием метода Forward (См. табл. 12).

Обе модели дали практически одинаковые результаты, за исключением того, что во вторую модель автоматически была включена 7-я гипотеза. Однако коэффициент при этой гипотезе не значим (Sig = 0,96), поэтому нет смысла учитывать эту переменную в модели. Для дальнейшего анализа была использована первая модель (Метод Stepwise).

Табл. 11 Регрессионная модель. Метод Stepwise $R^2 = 0,274$

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
N_hyp	Var	B	Std. Error	Beta		
	(Constant)	6,64	0,15		44,89	0,00
11	Нур11	4,02	0,76	0,12	5,30	0,00
10	Нур10	3,77	0,84	0,12	4,48	0,00
4	Нур4	3,72	0,45	0,19	8,17	0,00
22	Нур22	3,65	0,76	0,11	4,80	0,00
18	Нур18	3,53	0,78	0,11	4,53	0,00
9	Нур9	3,42	0,40	0,20	8,53	0,00
33	Нур33	3,27	0,76	0,10	4,31	0,00
14	Нур14	3,10	0,84	0,10	3,69	0,00
5	Нур5	2,93	0,38	0,18	7,72	0,00
3	Нур3	2,59	0,85	0,10	3,04	0,00
19	Нур19	2,49	0,84	0,09	2,95	0,00
1	Нур1	2,48	0,43	0,15	5,80	0,00
37	Нур37	2,43	0,87	0,07	2,81	0,01
32	Нур32	2,43	0,79	0,07	3,08	0,00
31	Нур31	2,27	0,87	0,07	2,62	0,01
12	Нур12	2,20	0,80	0,08	2,76	0,01
8	Нур8	1,93	0,75	0,07	2,58	0,01
29	Нур29	1,02	0,51	0,06	1,99	0,05
51	Нур51	-2,52	0,76	-0,11	-3,32	0,00
52	Нур52	-2,26	0,63	-0,08	-3,60	0,00
53	Нур53	-2,10	0,76	-0,06	-2,77	0,01
55	Нур55	-2,60	0,76	-0,08	-3,43	0,00
57	Нур57	-3,56	1,06	-0,13	-3,35	0,00
60	Нур60	3,02	1,50	0,09	2,02	0,04

a Dependent Variable: q ОЦЕНКА

Табл.12 Регрессионная модель. Метод Forward $R^2=0,274$

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
N_hyp	Var	B	Std. Error	Beta		
	(Constant)	6,64	0,15		44,82	0,00
11	Hyp11	4,02	0,76	0,12	5,29	0,00
10	Hyp10	3,77	0,84	0,12	4,48	0,00
4	Hyp4	3,72	0,46	0,19	8,17	0,00
22	Hyp22	3,65	0,76	0,11	4,80	0,00
18	Hyp18	3,54	0,78	0,11	4,53	0,00
9	Hyp9	3,42	0,40	0,20	8,52	0,00
33	Hyp33	3,30	0,88	0,10	3,74	0,00
14	Hyp14	3,10	0,84	0,10	3,69	0,00
5	Hyp5	2,93	0,38	0,18	7,72	0,00
3	Hyp3	2,60	0,88	0,10	2,97	0,00
19	Hyp19	2,50	0,86	0,09	2,90	0,00
1	Hyp1	2,48	0,43	0,15	5,79	0,00
37	Hyp37	2,45	0,92	0,08	2,67	0,01
32	Hyp32	2,43	0,79	0,07	3,08	0,00
31	Hyp31	2,28	0,92	0,07	2,49	0,01
12	Hyp12	2,21	0,80	0,08	2,76	0,01
8	Hyp8	1,93	0,75	0,07	2,57	0,01
29	Hyp29	1,02	0,51	0,06	1,99	0,05
7	Hyp7	-0,05	0,88	0,00	-0,05	0,96
51	Hyp51	-2,52	0,76	-0,11	-3,31	0,00
52	Hyp52	-2,25	0,63	-0,08	-3,60	0,00
53	Hyp53	-2,10	0,76	-0,06	-2,77	0,01
55	Hyp55	-2,60	0,76	-0,08	-3,42	0,00
57	Hyp57	-3,56	1,06	-0,13	-3,34	0,00
60	Hyp60	3,02	1,50	0,09	2,02	0,04

a Dependent Variable: q ОЦЕНКА

Прежде всего, стоит отметить, что модель в целом не противоречит проведенной ранее проверке. Все плюс – гипотезы входят в модель с положительными коэффициентами, а минус – гипотезы входят в модель с отрицательными коэффициентами. Исключение составляет лишь гипотеза №60, вошедшая в модель с большим положительным коэффициентом. При этом виньетки, включающие в себя данное сочетание признаков, имеют достаточно низкие оценки, а среднее значение этих оценок равно 3,58. Это самое низкое из всех средних значений по минус – гипотезам. При исключении этого взаимодействия из анализа модель сильно не меняется (См. приложение 3). При этом коэффициент детерминации незначительно увеличивается и становится равным 0,276. Далее был проведен анализ корреляций между гипотезами, показавший, что Нур60 сильно коррелирует с Нур51 ($r=0,7$) и Нур57 ($r=0,81$). По всей видимости, это произошло из-за самой специфики алгоритма порождения гипотез (гипотезы могут частично пересекаться и если эти пересечения достаточно большие, то возможна высокая корреляция между переменными, которые соответствуют гипотезам).

Это также могло произойти вследствие неполного перебора всех возможных сочетаний признаков в виньетках. Но учитывая то, что 57-я и 51-я гипотезы включены в модель со значимыми отрицательными коэффициентами, можно считать высокий положительный коэффициент при 60-й гипотезе «побочным эффектом» корреляций между указанными тремя гипотезами. Учитывая эти положения, было принято решение не исключать из анализа 60-ю гипотезу, и, ориентируясь на гипотезы 51 и 57, частично пересекающиеся с гипотезой 60, считать влияние гипотезы 60 отрицательным.

Итак, мы описали технические процедуры анализа данных и специфику полученных результатов. Теперь коснемся содержательной стороны полученных результатов.⁴⁵

Содержательная интерпретация результатов

Посмотрим с содержательной стороны на отобранные гипотезы (Табл. 13). Их можно рассматривать как типичные элементы религиозности / не религиозности, встречающиеся в описаниях людей и влекущие соответствующие оценки.

Табл. 13 Содержание гипотез

Плюс - гипотезы		
Нур11	Он носит длинную бороду и длинные волосы	Празднует только религиозные праздники, своего дня рождения не празднует
Нур10	Считает, что истинная вера должна восторжествовать, даже если для этого придется прибегнуть к насилию	Постится большую часть года
Нур4	Уединяясь совершает пятикратную молитву в течении дня в строго отведенное для этого время	
Нур22	Верит в существование Бога-творца и считает, человек никогда не сможет стать подобным Богу	Постится большую часть года
Нур18	Считает, что душа человека бессмертна и после смерти тела отправляется в рай или в ад по воле Бога и никогда больше не может вернуться на землю	Постится большую часть года
Нур9	Выбирает в течении дня определенные часы и посвящает их молитве, которую никогда не пропускает	
Нур33	Он иногда посещает религиозные собрания и богослужения	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой
Нур14	Если он знакомится с людьми, настроенными против его религии, то может вступить с ними в противоборство	Постится большую часть года
Нур5	Он посещает религиозные собрания и богослужения каждый день	
Нур3	Постится один месяц в году	Читает только религиозные тексты и не читает никакой другой литературы

⁴⁵ Более подробное описание содержательных результатов можно найти в дипломной работе Добрыниной Стелы: Использование методов виньеток и репертуарных решеток в исследовании обыденного восприятия традиционных и новых форм религиозности.

Нур19	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой	Читает только религиозные тексты и не читает никакой другой литературы	
Нур1	Скромно одевается, носит длинную юбку и блузу и голову покрывает платком		
Нур37	Никогда не смотрит телевизор, не читает газет, не слушает популярную музыку	Верит в существование Бога-творца и считает, человек никогда не сможет стать подобным Богу	
Нур32	Верит в то, что у человека много жизней и что душа его после смерти тела может переселиться в другого человека	Выходя на улицу он даже когда жарко одевает короткие выше щиколоток брюки и куртку защ. цвета	Читает только религиозные тексты и не читает никакой другой литературы
Нур31	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой		
Нур12	Он часто посещает религиозные собрания и богослужения	Любит читать религиозную, философскую и другую литературу	
Нур8	Постится большую часть года	Празднует все праздники как религиозные так и государственные	
Нур29	Он часто посещает религиозные собрания и богослужения		
Минус - гипотезы			
Нур51	Он никогда не посещает религиозные собрания и богослужения	Не ограничивает себя в выборе блюд	
Нур52	Верит в судьбу и высшие силы		
Нур53	Не молится	Не ограничивает себя в выборе блюд	
Нур55	Не совершает молитвы, поскольку считает, что человек не может общаться с Богом		
Нур57	Считает, что душа человека смертна и умирает вместе с телом	Он никогда не посещает религиозные собрания и богослужения	
Нур60	Считает, что душа человека смертна и умирает вместе с телом	Он никогда не посещает религиозные собрания и богослужения	Не ограничивает себя в выборе блюд

Дальнейший анализ можно проводить в двух направлениях. Первое предполагает некую классификацию типов религиозного поведения, выделение некоторых «полюсов» религиозности. Полученные результаты частично позволяют это сделать. Второе направление предполагает более детальный содержательный анализ результатов на качественном уровне, необходимый для более глубокого проникновения в суть проблематики.

Мы попробуем продолжить анализ в первом направлении и попытаемся выявить «полюсы» религиозного / нерелигиозного поведения, используя полученные результаты. Двигаться во втором направлении мы на данный момент не беремся, оставив эту работу экспертам в содержательной области.

Попробуем выделить основные характеристики, встречающиеся отдельно в плюс – гипотезах и в минус – гипотезах.

Среди (+) гипотез выделяются следующие подгруппы характеристик:

- Регулярное соблюдение поста (Постится большую часть года, Постится один месяц в году). Эти характеристики отмечены в таблице желтым цветом.

- Четко сформированные мировоззренческие взгляды в рамках исповедуемой религии (Убежденность в истинности своей веры, представления о боге, о душе человека). Эти характеристики отмечены зеленым цветом.

- Посещение религиозных собраний, участие в религиозных праздниках. Эти характеристики выделены светло-серым цветом.

- Чтение религиозной литературы. Эти характеристики выделены темно-серым цветом.

- Внешний вид, соответствующий религиозным канонам. Эти характеристики выделены темно-зеленым цветом.

- Соблюдение ритуала молитвы. Эти характеристики выделены красным цветом.

При этом среди минус - гипотез выделяются противоположные характеристики

- Не соблюдение поста (Не ограничивает себя в выборе блюд)

- Мировоззренческие взгляды, противостоящие религиозным воззрениям (Верит в судьбу и высшие силы; считает, что душа человека смертна и умирает вместе с телом)

- Не соблюдение ритуала молитвы (не молится, считает, что человек не может общаться с богом)

- Не посещение религиозных собраний и богослужений.

Эти результаты с одной стороны дают нам возможность построить некий вектор религиозности, в котором каждый из выделенных элементов религиозности будет представлять собой ось соответствующего пространства восприятия. С другой стороны данные результаты позволяют нам сделать предположение о том, что восприятие религиозности как сложный социальный конструкт можно измерить с помощью системы бинарных оппозиций, таких, как молится – не молится, ходит в церковь – не ходит в церковь и т.п.

Также следует обратить внимание на тот факт, что среди минус – гипотез не встречается характеристик, описывающих внешний вид человека. При этом на полном множестве минус – гипотез характеристики внешнего вида встречаются достаточно часто. Это может быть как следствием не полного перебора признаков в виньетках, так и следствием анализа, в результате которого мы неизбежно отсеиваем часть информации, изначально содержащейся в данных. Также это может быть отражением реального восприятия религиозности, в котором такой признак, как «внешний вид, соответствующий религиозным канонам» не может быть выражен в качестве бинарной оппозиции. Так, например, «металлист» в «косухе» и соответствующей экипировке может восприниматься окружающими как очень религиозный человек, если он держит в руках библию. При этом его внешний вид может не оказывать существенного значения на оценку степени его

религиозности. С другой стороны, если человек одевается в соответствии с церковными канонами, то это дает людям основания считать его религиозным. Но если человек не носит соответствующей одежды, то это в целом может не давать оснований считать его нерелигиозным, в отличие, например, от «ритуала молитвы», не соблюдение которого может быть причиной восприятия человека как нерелигиозного.

Полученные результаты позволяют нам сделать еще один интересный вывод. Если учитывать количество элементов, составляющих тело гипотез, то можно заметить, что признак соблюдение / не соблюдение молитвы, вне зависимости от того, с какими признаками он сочетается, оказывает достаточно сильное влияние на оценку религиозности. Это может означать, что социальный вес «молитвы» в восприятии религиозности достаточно высок, по сравнению, например, с социальным весом «соблюдением поста». Таким образом, можно прийти к выводу (а точнее предположению) о том, что каждая из выделенных бинарных оппозиций имеет свой вес в социальном пространстве.

На данном этапе мы закончим попытки интерпретации полученных результатов, оставив их для экспертов в области восприятия религиозности. Наша задача заключалась в демонстрации исследовательских возможностей предлагаемой технологии анализа данных. Считая данную задачу выполненной, кратко резюмируем результаты проделанной работы и полученные выводы.

Методологические выводы

1) В результате проделанной работы был проведен ряд экспериментов с системой JSM-explorer, была детально разобрана специфика работы системы, было проделано достаточно много технических экспериментов, связанных с анализом социологических данных. В результате автором был приобретен ценный опыт работы с методом и его программной реализацией.

2) ДСМ – метод АПГ успешно показал себя в решении задачи поиска существенных взаимодействий. Также были предложены варианты внутренней валидации полученных взаимодействий (гипотез) с помощью некоторых статистических средств. Таким образом, была подготовлена почва для дальнейшей работы в направлении совмещения логико-комбинаторных и статистических средств анализа социологических данных.

3) Был предложен вариант использования ДСМ – метода АПГ для анализа данных, собранных с помощью метода виньеток. Метод виньеток позволяет получать более надежные данные, лучше отражающие объективную социальную реальность, нежели традиционные методы сбора данных. Использование метода виньеток во многом позволяет избавиться от «субъективного» шума в данных, затрудняющего поиск причинно-следственных связей с помощью логико-комбинаторных средств. Это позволяет исследователю выйти за рамки традиционной статистической обработки данных и использовать широкий арсенал логико-комбинаторных средств анализа социологической информации.

4) ДСМ – метод АПГ как метод качественного анализа данных, реализующий в себе синтез познавательных процедур, позволяет значительно снизить субъективное влияние исследователя на результаты анализа. Использование ДСМ – метода АПГ в сочетании с методом виньеток позволяет исследователю производить анализ более «объективных» данных более «объективными» средствами.

Список литературы:

- 1) *Аншаков О.М.* Каузальные модели предметных областей // НТИ. Сер. 2. – 2000. – № 3. – с. 3-17.
- 2) *Аншаков О.М.* Об одной интерпретации ДСМ-метода автоматического порождения гипотез // НТИ. Сер. 2. – 1999. – № 3. – с. 4-12.
- 3) *Войшвилло Е.К., Дяттерев М.Г.* Логика - М.: ВЛАДОС ПРЕСС, 2001
- 4) *Гусакова С.М., Михеенкова М.А., Финн В.К.* О логических средствах автоматизированного анализа мнений // НТИ. Сер. 2. – 2001. – № 5. – с. 4-24.
- 5) *Данилова Е.Н., Михеенкова М.А., Климова С.Г.* Возможности применения логико-комбинаторных методов для анализа социальной информации // Социология: 4М. – 1999. – № 11. – с. 142-160.
- 6) *Данилова Е.Н., Михеенкова М.А., Московский С.С.* Об одной модели детерминации социального поведения / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 263-271.
- 7) *Девятко И.Ф.* Методы социологического исследования. М.: УНИВЕРСИТЕТ, 2003
- 8) *Девятко И.Ф.* Диагностическая процедура в социологии. М.: Наука, 1993
- 9) *Ершов Ю.Л., Палютин Е.А.* Математическая логика. Учебное пособие. №-е изд. – СПб.: Издательство «Лань», 2004
- 10) *Добрынина С.* Использование методов виньеток и репертуарных решеток в исследовании обыденного восприятия традиционных и новых форм религиозности // выпускная квалификационная работа, рукопись, ГУ-ВШЭ 2006
- 11) *Есенин-Вольпин А.С.*, Свободный философский трактат / Логика, философия, поэзия, защита прав человека. Под редакцией В.К. Финна и А.Ю. Даниэля, М.: РГГУ, 1999
- 12) *Есенин-Вольпин А.С.*, Об антитрадиционной (ультраинтуитивной) программе оснований математики и естественно-научном мышлении / Логика, философия, поэзия, защита прав человека. Под редакцией В.К. Финна и А.Ю. Даниэля, М.: РГГУ, 1999
- 13) *Климова С.Г., Михеенкова М.А., Панкратов Д.В.* ДСМ-метод как метод выявления детерминант социального поведения // НТИ. Сер. 2. – 1999. – № 12. – с. 3-14.
- 14) *Милль Д.С.* Система логики силлогистической и индуктивной. - М.: Книжное дело, 1900
- 15) *Назаров Б.В.*, Изучение причинно-следственных отношений с помощью причинного анализа и родственных ему методов: сравнение разных подходов // выпускная квалификационная работа, Рукопись, ГУ-ВШЭ, 2004

- 16) *Поппер К.* Предположения и опровержения, М.: "Издательство АСТ", 2004
- 17) *Поппер К.* Эволюционная эпистемология // Эволюционная эпистемология и логика социальных наук. - М.: Эдиториал УРСС, 2000, стр. 57 - 74
- 18) *Решер Н.* Пирс, Поппер и методологический переворот // Эволюционная эпистемология и логика социальных наук. - М.: Эдиториал УРСС, 2000, стр. 210 - 221
- 19) *Толстова Ю.Н.* Анализ социологических данных. М.: Научный мир, 2000
- 20) *Финн В.К.* Интеллектуальные системы и общество: идеи и понятия // НТИ. Сер. 2. - 1999. - № 10. - с. 6-20
- 21) *Финн В.К.* Эволюционная эпистемология Карла Поппера и эпистемология синтеза познавательных процедур // Эволюционная эпистемология и логика социальных наук. - М.: Эдиториал УРСС, 2000, стр. 364 - 424
- 22) *Финн В.К.* Об особенностях ДСМ-метода как средства интеллектуального анализа данных // НТИ. Сер. 2. – 2001. – № 5. – с. 1-4.
- 23) *Финн В.К.* Гуманитарное знание и искусственный интеллект. / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 38 – 57
- 24) *Финн В.К.* Синтез познавательных процедур и проблема индукции. / Интеллектуальные системы и общество, РГГУ, М.: 2001, стр. 133 – 138
- 25) *Финн В.К.* О некоторых проблемах логики и методологии интеллектуальных систем // НТИ. Сер. 2. – 1999. – № 1-2.
- 26) *Фирмен Ю., Сколимовский Г.* Поиск объективности у Пирса и Поппера // Эволюционная эпистемология и логика социальных наук. - М.: Эдиториал УРСС, 2000, стр. 222 – 279
- 27) *Ядов В.А.* Стратегия социологического исследования: описание, объяснение, понимание социальной реальности. – 7-е изд. – М.: «Добросвет», 2003
- 28) *Thurman Q.C., Lam J.A., Rossi P.H.* A factorial survey approach to the study of popular conceptions of mental illness//The Sociological Quarterly, Vol. 29, № 4, 1988, p. 565-588

Приложения

Приложение 1 Список переменных в исходном массиве данных

- x5_1_м Внешний вид мужчины прическа он бреется налысо и не носит бороды
- x5_2_м Он носит короткую стрижку и не носит бороды (са)
- x5_3_м У него короткая стрижка и длинная борода (и,св)
- x5_4_м Он носит длинную бороду и длинные волосы (п)
- x5_1_ж Внешний вид женщины прическа она носит короткую модную стрижку и обычно пользуется косметикой (са)
- x5_2_ж Она носит длинные волосы и пользуется косметикой (и)
- x5_3_ж Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой (п, св)
- x6_1_м Внешний вид ужчины одежда Без упоминания (св,са)
- x6_2_м Модно одевается
- x6_3_м Выходя на улицу он даже когда жарко одевает костюм темного цвета (п)
- x6_4_м Выходя на улицу он даже когда жарко одевает короткие выше щиколоток брюки и куртку защ. цвета (и)
- x6_1_ж Внешний вид женщины одежда без упоминания (са)
- x6_2_ж модно одевается
- x6_3_ж скромно одевается носит длинную юбку и блузу и голову покрывает платком (п,св)
- x6_4_ж Всегда одевает просторный черный балахон и платок закрывающий лицо (и)
- x7_1 Книги Любит читать религиозную, философскую и другую литературу
- x7_2 Читает только религиозные тексты и не читает никакой другой литературы (п, св, са, и)
- x8_1 Праздники Празднует все праздники как религиозные так и государственные (са)
- x8_2 Празднует только религиозные праздники, своего дня рождения не празднует (П, И)
- x8_3 Не празднует никаких праздников, в том числе и свой день рождения (Св)
- x9_1 Еда Не ограничивает себя в выборе блюд (П, Са)
- x9_2 Никогда не употребляет в пищу кровь животных и не курит (Св)
- x9_3 Никогда не употребляет в пищу алкоголь, свинину и не курит (И)
- x11_1 Сми Без упоминания (Са, Св)
- x11_2 Вечерами читает газеты и смотрит телевизор и любит современную музыку
- x11_3 Никогда не смотрит телевизор, не читает газет, не слушает популярную музыку (П, И)
- x12_1 Бог Без упоминания
- x12_2 Верит в судьбу и высшие силы
- x12_3 Верит в существование Бога-творца и считает, что каждый человек может стать Богом (Са)
- x12_4 Верит в существование Бога-творца и считает, человек никогда не сможет стать подобным Богу (П, И, Св)
- x13_1 Разговоры о Боге Он обычно не касается религиозных тем в разговоре
- x13_2 Он любит поговорить о Боге и о душе (П, Св, Са, И)
- x14_1 Реинкарнация Верит в то, что у человека много жизней и что душа его после смерти тела может переселиться в другого человека (Са)
- x14_2 Считает, что душа человека бессмертна и после смерти тела отправляется в рай или в ад по воле Бога и никогда больше не может вернуться на землю (П, И)
- x14_3 Считает, что душа человека смертна и умирает вместе с телом (Св)
- x15_1 Конец света Он не воспринимает разговоры о возможном конце света всерьез
- x15_2 В последнее время он часто говорит своим друзьям о том, что на Землю в скором времени высадятся войска межгалактической конфедерации и наступит конец света.(са)
- x15_3 Он часто говорит своим друзьям о том, что очень скоро произойдет катаклизм космического масштаба и наступит конец света, тогда спастись смогут лишь истинно верующие.(и)
- x15_4 Он часто говорит своим друзьям о том, скоро на землю придет Бог и наступит конец света, тогда спастись смогут лишь истинно верующие (п)
- x15_5 В последнее время он часто говорит своим друзьям о том, что очень скоро на землю придет Бог со своим войском и истребит всех неверующих. (св)
- x16_1 Другие конфессии Не обращает внимания на вероисповедание своих знакомых
- x16_2 Не разговаривает с людьми, которые не придерживаются его религии (Св)
- x16_3 Если оказывается в кругу людей не его религии, то обязательно начинает их поучать и старается преподать им духовные истины (П)
- x16_4 Если он знакомится с людьми, настроенными против его религии, то может вступить с ними в противоборство (Са)

- x16_5 Считает, что истинная вера должна восторжествовать, даже если для этого придется прибегнуть к насилию (И)
- x17_1 Пост Без упоминания (Св, Са)
- x17_2 Постится один месяц в году (И)
- x17_3 Постится большую часть года (П)
- x18_1 Храм Он никогда не посещает религиозные собрания и богослужения
- x18_2 Он иногда посещает религиозные собрания и богослужения
- x18_3 Он часто посещает религиозные собрания и богослужения
- x18_4 Он посещает религиозные собрания и богослужения каждый день
- x19_1 Молитва Не молится
- x19_2 Не совершает молитвы, поскольку считает, что человек не может общаться с Богом (Св)
- x19_3 Молится время от времени (Са)
- x19_4 Выбирает в течении дня определенные часы и посвящает их молитве, которую никогда не пропускает (П)
- x19_5 Уединяясь совершает пятикратную молитву в течении дня в строго отведенное для этого время (И)
- relig Степень религиозности

Приложение 2 Полный набор полученных гипотез

№ комбинации	Комбинация					
1	x6_3ж			Скромно одевается носит длинную юбку и блузу и голову покрывает платком		
2	x16_2			Не разговаривает с людьми, которые не придерживаются его религии		
3	x17_2	x7_2		Постится один месяц в году	Читает только религиозные тексты и не читает никакой другой литературы	
4	x19_5			Уединяясь совершает пятикратную молитву в течении дня в строго отведенное для этого время		
5	x18_4			Он посещает религиозные собрания и богослужения каждый день		
6	x17_2	x5_3ж	x7_2	Постится один месяц в году	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой	Читает только религиозные тексты и не читает никакой другой литературы
7	x12_4	x5_3ж		Верит в существование Бога-творца и считает, человек никогда не сможет стать подобным Богу	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой	
8	x17_3	x8_1		Постится большую часть года	Празднует все праздники как религиозные так и государственные	
9	x19_4			Выбирает в течении дня определенные часы и посвящает их молитве, которую никогда не пропускает		
10	x16_5	x17_3		Считает, что истинная вера должна восторжествовать, даже если для этого придется прибегнуть к насилию	Постится большую часть года	
11	x5_4м	x8_2	x17_1	Он носит длинную бороду и длинные волосы	Празднует только религиозные праздники, своего дня рождения не празднует	
12	x18_3	x7_1		Он часто посещает религиозные собрания и богослужения	Любит читать религиозную, философскую и другую литературу	

13	x19_5	x8_2	Уединяясь совершает пятикратную молитву в течении дня в строго отведенное для этого время	Празднует только религиозные праздники, своего дня рождения не празднует	
14	x16_4	x17_3	Если он знакомится с людьми, настроенными против его религии, то может вступить с ними в противоборство	Постится большую часть года	
15	x12_2	x8_2	Верит в судьбу и высшие силы	Празднует только религиозные праздники, своего дня рождения не празднует	
16	x5_4м	x8_2	Он носит длинную бороду и длинные волосы	Празднует только религиозные праздники, своего дня рождения не празднует	
17	x12_2	x19_4	Верит в судьбу и высшие силы	Выбирает в течении дня определенные часы и посвящает их молитве, которую никогда не пропускает	
18	x14_2	x17_3	Считает, что душа человека бессмертна и после смерти тела отправляется в рай или в ад по воле Бога и никогда больше не может вернуться на землю	Постится большую часть года	
19	x5_3ж	x7_2	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой	Читает только религиозные тексты и не читает никакой другой литературы	
20	x13_1	x6_3ж	Он обычно не касается религиозных тем в разговоре	скромно одевается носит длинную юбку и блузу и голову покрывает платком	
21	x13_1	x8_2	Он обычно не касается религиозных тем в разговоре	Празднует только религиозные праздники, своего дня рождения не празднует	
22	x12_4	x17_3	Верит в существование Бога-творца и считает, человек никогда не сможет стать подобным Богу	Постится большую часть года	
23	x5_4м		Он носит длинную бороду и длинные волосы		
24	x14_3	x5_4м	Считает, что душа человека смертна и умирает вместе с телом	Он носит длинную бороду и длинные волосы	
25	x14_3	x18_4	Считает, что душа человека смертна и умирает вместе с телом		

26	x18_3	x5_3м	x7_1	Он часто посещает религиозные собрания и богослужения	У него короткая стрижка и длинная борода	Любит читать религиозную, философскую и другую литературу
27	x19_4	x7_2		Выбирает в течении дня определенные часы и посвящает их молитве, которую никогда не пропускает	Читает только религиозные тексты и не читает никакой другой литературы	
28	x14_1	x7_2		Верит в то, что у человека много жизней и что душа его после смерти тела может переселиться в другого человека	Читает только религиозные тексты и не читает никакой другой литературы	
29	x18_3			Он часто посещает религиозные собрания и богослужения		
30	x13_2	x18_3		Он любит поговорить о Боге и о душе	Он часто посещает религиозные собрания и богослужения	
31	x17_1	x5_3ж		Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой		
32	x14_1	x6_4м	x7_2	Верит в то, что у человека много жизней и что душа его после смерти тела может переселиться в другого человека	Выходя на улицу он даже когда жарко одевает короткие выше щиколоток брюки и куртку защ. цвета	Читает только религиозные тексты и не читает никакой другой литературы
33	x18_2	x5_3ж		Он иногда посещает религиозные собрания и богослужения	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой	
34	x16_4	x5_3ж		Если он знакомится с людьми, настроенными против его религии, то может вступить с ними в противоборство	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой	
35	x16_5			Считает, что истинная вера должна восторжествовать, даже если для этого придется прибегнуть к насилию		
36	x11_1	x18_4		Он посещает религиозные собрания и богослужения каждый день		
37	x11_3	x12_4		Никогда не смотрит телевизор, не читает газет, не слушает популярную музыку	Верит в существование Бога-творца и считает, человек никогда не сможет стать подобным Богу	

38	x12_4	x18_2		Верит в существование Бога-творца и считает, человек никогда не сможет стать подобным Богу	Он иногда посещает религиозные собрания и богослужения	
39	x19_4	x8_2		Выбирает в течении дня определенные часы и посвящает их молитве, которую никогда не пропускает	Празднует только религиозные праздники, своего дня рождения не празднует	
40	x13_2	x6_3ж		Он любит поговорить о Боге и о душе	скромно одевается носит длинную юбку и блузу и голову покрывает платком	
41	x6_3ж	x8_3		скромно одевается носит длинную юбку и блузу и голову покрывает платком	Не празднует никаких праздников, в том числе и свой день рождения	
42	x19_1	x9_3		Не молится	Никогда не употребляет в пищу алкоголь, свинину и не курит	
43	x18_1	x6_2м	x7_1	Он никогда не посещает религиозные собрания и богослужения	Модно одевается	Любит читать религиозную, философскую и другую литературу
44	x19_1	x6_4м		Не молится	Выходя на улицу он даже когда жарко одевает короткие выше щиколоток брюки и куртку защ. цвета	
45	x17_1	x5_1ж		Внешний вид женщины прическа она носит короткую модную стрижку и обычно пользуется косметикой		
46	x19_1	x5_3ж		Не молится	Она носит длинные волосы, никогда не стрижет их и не пользуется косметикой	
47	x16_4	x19_1		Если он знакомится с людьми, настроенными против его религии, то может вступить с ними в противоборство	Не молится	
48	x5_1м	x9_1		он бреется налысо и не носит бороды	Не ограничивает себя в выборе блюд	
49	x6_4ж	x9_1		сегда одевает просторный черный балахон и платок закрывающий лицо	Не ограничивает себя в выборе блюд	
50	x13_1	x9_1		Он обычно не касается религиозных тем в разговоре	Не ограничивает себя в выборе блюд	

51	x18_1	x9_1		Он никогда не посещает религиозные собрания и богослужения	Не ограничивает себя в выборе блюд	
52	x12_2	x17_1		Верит в судьбу и высшие силы		
53	x19_1	x9_1		Не молится	Не ограничивает себя в выборе блюд	
54	x15_3	x18_1		Он часто говорит своим друзьям о том, что очень скоро произойдет катаклизм космического масштаба и наступит конец света, тогда спастись смогут лишь истинно верующие	Он никогда не посещает религиозные собрания и богослужения	
55	x11_1	x19_2		Не совершает молитвы, поскольку считает, что человек не может общаться с Богом		
56	x11_2	x6_4м		Вечерами читает газеты и смотрит телевизор и любит современную музыку	Выходя на улицу он даже когда жарко одевает короткие выше щиколоток брюки и куртку защ. цвета	
57	x14_3	x18_1		Считает, что душа человека смертна и умирает вместе с телом	Он никогда не посещает религиозные собрания и богослужения	
58	x12_2	x17_1	x9_1	Верит в судьбу и высшие силы	Не ограничивает себя в выборе блюд	
59	x12_2	x17_1	x6_4м	Верит в судьбу и высшие силы	Выходя на улицу он даже когда жарко одевает короткие выше щиколоток брюки и куртку защ. цвета	
60	x14_3	x18_1	x9_1	Считает, что душа человека смертна и умирает вместе с телом	Он никогда не посещает религиозные собрания и богослужения	Не ограничивает себя в выборе блюд

Приложение 3 Регрессионная модель без переменной Нур60 (Метод Stepwise)

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
28	(Constant)	6,708197	0,150811		44,48	0,00
11	Нур11	3,96	0,76	0,12	5,21	0,00
10	Нур10	3,72	0,84	0,11	4,43	0,00
4	Нур4	3,65	0,46	0,19	8,02	0,00
22	Нур22	3,58	0,76	0,11	4,72	0,00
18	Нур18	3,49	0,78	0,11	4,48	0,00
9	Нур9	3,35	0,40	0,20	8,36	0,00
33	Нур33	3,21	0,76	0,10	4,23	0,00
14	Нур14	3,05	0,84	0,09	3,64	0,00
5	Нур5	2,87	0,38	0,18	7,56	0,00
3	Нур3	2,58	0,85	0,10	3,03	0,00
1	Нур1	2,43	0,43	0,15	5,70	0,00
19	Нур19	2,43	0,84	0,09	2,89	0,00
32	Нур32	2,39	0,79	0,07	3,03	0,00
37	Нур37	2,39	0,86	0,07	2,76	0,01
31	Нур31	2,22	0,86	0,07	2,57	0,01
12	Нур12	2,21	0,80	0,08	2,78	0,01
8	Нур8	1,90	0,75	0,07	2,54	0,01
29	Нур29	0,96	0,51	0,06	1,87	0,06
57	Нур57	-3,29	0,63	-0,12	-5,26	0,00
55	Нур55	-2,67	0,76	-0,08	-3,51	0,00
52	Нур52	-2,32	0,63	-0,09	-3,71	0,00
54	Нур54	-2,29	0,76	-0,07	-3,02	0,00
53	Нур53	-2,17	0,76	-0,07	-2,85	0,00
50	Нур50	-1,92	0,76	-0,06	-2,52	0,01

a Dependent Variable: ОЦЕНКА