

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»**

Кафедра социальной информатики

**СПЕЦИФИКА ИСПОЛЬЗОВАНИЯ МАТЕМАТИЧЕСКИХ
МОДЕЛЕЙ ПРИ АНАЛИЗЕ ОБЩЕСТВЕННОГО МНЕНИЯ**

(автореферат бакалаврской работы)

Студента 4 курса 451 группы
направления 09.03.03 Прикладная информатика
профиль Прикладная информатика в социологии
Социологического факультета
Крашенинникова Романа Георгиевича

Научный руководитель
профессор, доктор социологических наук

подпись, дата

С.В. Ситникова

Заведующий кафедрой
кандидат социологических наук, доцент

подпись, дата

И.Г. Малинский

Саратов 2023

ВВЕДЕНИЕ

Актуальность проблемы. При проведении исследований в социальных и гуманитарных науках зачастую появляется необходимость провести анализ набора данных. В том числе, при изучении общественного мнения зачастую невозможно установить всеобъемлющий набор фактов по поверхностному частотному анализу измеряемых признаков¹. Для более углубленного анализа в исследованиях применяются математические модели.

Математические методы позволяют социологам проводить исследования с высокой точностью и достоверностью, а также предсказывать различные тенденции и события. В целом, использование математических методов помогает улучшить качество исследований и сделать более точные выводы о социальных процессах и явлениях.

Актуально применение математических методов и при изучении предпочтений населения в вопросе исчисления времени. Только при помощи математических методов можно установить взаимосвязь между значением социально-демографической характеристики респондента и предпочитаемым вариантом исчисления времени. Это позволяет делать предположения, как будут меняться предпочтения населения со временем, когда номинальные численности и процентные соотношения этих демографических страт будут меняться по заранее прогнозируемой траектории.

Изменения исчисления времени проводились как по экономическим причинам, так и по причинам существования социального запроса на установление того или иного варианта исчисления времени. Исследования влияния исчисления времени на экономические процессы противоречат друг другу, проводимые как на разных этапах исторического развития России, так и разными исследовательскими группами.

Конкретно в Саратове, постоянное поясное исчисление времени может быть обосновано: фактом его совпадения с исчислением времени,

¹ Семенов В. А. Математические методы в гуманитарных исследованиях: учебное пособие для вузов / В. А. Семенов, В. А. Макаридина. – Москва: Издательство Юрайт, 2022. – 250 с.

установленным в столице – городе Москва, упрощением построения логистических цепочек, пользой для здоровья большинства работников большей продолжительностью утреннего сна, повышением производительности труда, и многим другим.

Также возможно обоснование постоянного исчисления времени, опережающего поясное: сохранением исторически сложившейся традиции, стимулированием занятий спортом и здорового образа жизни, поддержанием региональной идентичности, и многим другим.

Сезонные переводы изначально вводились, как попытка заимствования успешных практик других стран, они заимствуют многие успешные аспекты разных вариантов исчисления времени, между которыми осуществляются сезонные переводы, при однозначном вреде, который оказывает на здоровье граждан сам момент перевода часов.

Именно наличие большого числа факторов, влияющих на предпочтения населения в вопросе исчисления времени определяет актуальность и необходимость применения математических методов, при изучении общественного мнения в вопросе исчисления времени.

Степень научной разработанности данной проблемы.

Применение математических моделей в исследованиях зачастую тогда имеет место, когда в его процессе возникает необходимость в анализе набора данных. Существует множество исследований, статей, научных работ, книг, где положения доказываются со ссылкой на авторитетные источники, умозрительный анализ, без опоры на конкретные наборы значений, или проводится качественный анализ, где идет упор на детализацию фактов, а не количественное сравнение².

Такие исследования могут иметь место в малоизученных, относительно новых, или областях, контакт с которыми редок и немногочислен. Однако, зачастую исследования проводятся количественным подходом, когда область

² Давыдов А. А. Математическая социология: обзор зарубежного опыта // Социологические исследования. 2018. Вып.4. С. 105-111.

исследований хорошо изучена, её представителей достаточно много, а те, кто не принадлежит к исследуемым субъектам, зачастую имеет то или иное о них представление.

Именно в таком случае возникает необходимость проведения сравнений, расчетов, поиска закономерностей, связей – то, для чего применимо и необходимо построение математических моделей. Например, с помощью математических моделей можно изучать социальные сети, прогнозировать поведение людей и предсказывать тенденции развития общества в целом.

Одним из первых предложенных коэффициентов корреляции был коэффициент корреляции гамма. Он был изобретен Карлом Фридрихом Гауссом, в начале XIX века. Он предложил использовать коэффициент корреляции для измерения связи между двумя переменными, которые могут быть измерены только на номинальной и порядковой шкале. Коэффициент корреляции гамма широко используется в социологии и других социальных науках для изучения связи между различными переменными³.

Карл Пирсон разработал и использовал метод корреляции Пирсона в своих работах по статистике в конце XIX века. Этот метод используется для измерения степени линейной связи между двумя переменными. Карл Пирсон применил этот метод в своих исследованиях различных научных областей, включая биологию, астрономию и статистику.

Коэффициент корреляции фи также был предложен Карлом Пирсоном в 1900 году и является мерой связи между двумя переменными.

Коэффициент V Крамера был предложен в 1946 году Харальдом Крамером и является более точной мерой связи между двумя переменными, чем фи-коэффициент. Он учитывает размер выборки и количество категорий в каждой переменной. Коэффициент V Крамера широко используется в социологии для изучения связи между различными социальными переменными.

³ Шахнович В. В. Корреляционный анализ. Теория и практика / В. В. Шахнович. — М.: КНОРУС, 2018. — 248 с.

Коэффициент корреляции лямбда, был впервые предложен Моррисом Вилкоксом в 1945 году. Моррис Вилкокс, был профессором математической статистики в Принстонском университете, впервые применил коэффициент корреляции лямбда в своих исследованиях по социальной статистике. В одном из своих первых исследований, опубликованном в 1945 году, Вилкокс использовал коэффициент корреляции лямбда для изучения взаимосвязи между предпочтениями в музыке и полом. Он также применил этот коэффициент для изучения взаимосвязи между уровнем образования и политическими взглядами⁴.

В целом, коэффициенты корреляции являются важными инструментами для изучения социальных явлений и помогают социологам выявлять связи и зависимости между различными номинальными и порядковыми переменными.

Помимо математических методов поиска корреляции между двумя переменными, существует также факторный анализ. Факторный анализ был изобретен Чарльзом Спирменом в начале XX века. Чарльз Спирмен использовал факторный анализ для изучения статистических связей между различными показателями интеллекта, которые он назвал факторами общей интеллектуальной способности. Он предположил, что интеллект может быть разделен на несколько компонентов, которые он назвал факторами.

С помощью факторного анализа Спирмен исследовал большой набор тестов на интеллект, чтобы определить, какие из них наиболее сильно связаны друг с другом. Он обнаружил, что большинство тестов на интеллект имеют сильную корреляцию между собой, что подтвердило его гипотезу о наличии общей интеллектуальной способности.

Таким образом, Чарльз Спирмен использовал факторный анализ для создания теории об общей интеллектуальной способности, которая стала известна как "теория г-фактора". Эта теория стала одной из основных теорий интеллекта и оказала значительное влияние на развитие психометрии и психологии в целом. А факторный анализ в дальнейшем использовался

⁴ Гашкова Е. В. Исследование связи признаков методами корреляционного анализа / Е. В. Гашкова, С. И. Ларин. — М. : Изд-во АСВ, 2018. — 160 с.

социологами в исследованиях с целью снижения размерности, объединяя разрозненные атрибуты, имеющие между собой тесную связь в факторы, рассматриваемые, как единое целое.

Первые обоснования к изменению установленного исчисления времени в СССР приводились в сборнике декретов Советской власти от 1957 года⁵, где представлен документ декрета от 1930 года, когда для более оптимального использования светлой части суток было введено так называемой декретное время – время опережающее поясное на 1 час.

В 1980 году были представлены научные обоснования дальнейшего изменения исчисления времени в СССР академиком астрономом Александром Михайловым и председателем Государственной комиссии единого времени и эталонных частот СССР, председателем госстандарта Василием Бойцовым в статье «За стрелками часов»⁶, журнала «Наука и жизнь», где необходимость дальнейшего изменения исчисления времени обосновывалась предполагаемым положительным экономическим эффектом.

Исследования влияния установленного исчисления времени на экономику страны проводил Рауль Яруллин в статье «О влиянии смены часовых поясов и изменения времени на экономику России»⁷, где приводит доводы о неоднозначности различных оценок, которые на разном историческом периоде дают экономическому значению тех или иных изменений исчисления времени в России.

Современные представления об исчислении времени и его значении в вопросе организации трудовой деятельности обозначены в статье Кристофера Барнса «Идеальный рабочий график, определяемый циркадными ритмами»⁸, где

⁵ Валк С.Н., Обичкин Г.Д. Декреты Советской власти [Текст] / Г.Д. Обичкин // Гос. изд. политической литературы, 1957. – 2457 с

⁶ Михайлов, А.А. За стрелками часов [Текст] / А.А. Михайлов, В.В. Бойцов // Наука и жизнь. – 1980. - № 3. – С. 8-10.

⁷ Яруллин Р.Р. О влиянии смены часовых поясов и изменения времени на экономику России [Текст] / Р.Р. Яруллин // Журнал Вестник Башкирского Государственного Университета. – 2014. - № 3. – С. 142-147

⁸ Barnes, С.М. The Ideal Work schedule, as Determined by Circadian Rhythms [Текст] / С.М. Barnes // Harvard Business Review. – 2015. - № 2. – С. 2-3.

автор обозначает необходимость учёта индивидуальных потребностей работников при планировании рабочего времени, и приходит к выводу о необходимости введения гибкого рабочего графика, вместо директивных всеобщих изменений исчисления времени в ту или иную сторону.

О социальном значении установленного исчисления времени разъясняет Тиль Реннеберг в книге «Внутреннее время: хронотипы, социальный джетлаг, и почему ты уставший»⁹. В этой книге автор приводит ряд положений о влиянии установленного исчисления времени на социальную активность и продуктивность совершаемых действия людьми, принадлежащих к различным хронотипам, в процессе проживания при различных вариантах исчисления времени относительно астрономического времени, определённого для данной территории.

О влиянии регулярных изменений исчисления времени в сезонном порядке на безопасность на автомобильных дорогах, и его недостатках в сравнении с постоянным исчислением времени проводит исследование¹⁰ Рейчел Кери, научный сотрудник Университета Колледжа Лондона.

Таким образом представления населения об оптимальном варианте исчисления времени исследовались в Российской, Советской и мировой науке. В процессе изучения сферы исчисления времени, применялись различные подходы, исходя из критериев определения оптимального варианта исчисления времени, с точки зрения экономики, безопасности, географии и других смежных наук. Отдельным критерием считались предпочтения населения. Социологические исследования в этой области зачастую подкрепляли предложения об изменении исчисления времени на более эффективное, оптимальное и предпочитаемое как в России, так и за рубежом.

Для углубленного анализа социологических данных, взаимосвязи между компонентами представлений населения в вопросе исчисления времени,

⁹ Roenneberg T. Internal Time: Chronotypes, Social Jet Lag, and Why You're So Tired [Текст] / T. Roenneberg // Harvard University Press 2012. – 288 с

¹⁰ Carey, R.N. Impact of daylight-saving time on road traffic collision risk: a systematic review [Текст] / R.N. Carey // British Medical Journal. – 2016. - № 2. – С. 7-9.

определяющими характер предпочтений необходимо применение математических моделей. Высокая степень изученности применения математических моделей в изучении математических данных, собираемых, в том числе при анализе социальных процессов позволяет эффективно их применять при изучении общественного мнения, учитывая специфику использования математических моделей при исследовании предпочитаемых вариантов исчисления времени.

Цель исследования – показать эффективность использования математических методов при изучении общественных процессов, на примере исследования предпочтений населения в вопросе исчисления времени.

Для достижения поставленной цели необходимо решить следующие задачи:

- представить теоретические основания математических моделей корреляционного, факторного анализа;
- обосновать эффективность применения математических моделей при анализе социальных процессов;
- определить отношение жителей города Саратов к различным вариантам исчисления времени;
- выявить факторы, влияющие на отношение жителей города Саратов к различным вариантам исчисления времени.

Объектом исследования являются математические методы.

Предметом исследования является специфика использования математических методов при изучении предпочтений населения в вопросе исчисления времени.

Теоретико-методологической основой данного исследования выступили математические модели, в частности, статистические меры, корреляционные модели, а также регрессионная модель, и модели снижения размерности: факторная модель, кластерная модель. Применяемые для анализа взаимосвязей между измеряемыми признаками, поиска причинно-следственных связей между критериями определения оптимального варианта исчисления времени и

предпочитаемым вариантом исчисления времени с точки зрения жителей города Саратов.

Эмпирической базой выпускной квалификационной работы являются результаты авторского исследования на тему «Предпочтения населения города Саратов в вопросе исчисления времени» (N=100 человек); исследования эффективности использования математических моделей при анализе социальных процессов; социологические исследования отношения жителей города Саратов к различным вариантам исчисления времени. Среди нормативных документов использовалась информация о внесении изменений в федеральный закон «Об исчислении времени».

Практическая значимость исследования заключается в возможности использования основных положений и выводов данной работы для дальнейшего изучения выбранной проблематики.

Структура бакалаврской работы определена задачами исследования, логикой раскрытия темы. Квалификационная работа состоит из введения, трех разделов, заключения, списка использованных источников и приложений.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первом разделе «Теоретические основы формирования и применения математических моделей» описываются теоретические основы применения статистических мер, математических методов корреляционного, регрессионного, факторного, кластерного анализа. Определяется математическая логика их построения, требования, которые предъявляются к набору данных, для их применения, их теоретическая обоснованность, критерии эффективности.

Среди статистических мер описывается среднее, медиана, мода, стандартное отклонение, дисперсия, переменные для которых они могут быть рассчитаны, и возможности их применения в анализе и построении других, более сложных математических моделей.

Характеризуется сущность каждого метода корреляционного анализа для номинальных переменных, таких как хи-квадрат, производных от него:

коэффициента сопряженности Пирсона, ϕ , V Крамера, альтернативы применению хи-квадрата – коэффициента корреляции лямбда.

Описываются методы корреляционного анализа для взаимосвязи порядковых переменных: ρ Спирмена, τ у Кендалла, гаммы Гудмена-Краскала, математические формулы их расчета, возможность применения на различных по структуре наборах данных, преимущества и недостатки их применения в отдельных случаях.

Поясняются математические законы применение теста Стьюдента, однофакторного дисперсионного анализа и метода ранговых сравнений Краскала-Уоллиса, применяемых для поиска взаимосвязи между количественными и качественными признаками, методом сравнения средних.

Также описывается методика построения регрессионной модели, требования, которые она предъявляет к данным: связи независимых переменных с зависимой, отсутствие коллинеарности, гетероскедастичности, очищение данных от выбросов.

Поясняется логика построения факторной и кластерной модели для наборов данных. Специфика их построения и сущность принципа снижения размерности.

По итогам первого раздела сформулирован набор математических методов, и доказывается их эффективность при математическом анализе наборов данных.

Во втором разделе «Эффективность применения математических моделей при изучении социальных процессов» представляются практические методы применения математических моделей при изучении социальных процессов, обосновывается эффективность их применения.

Для каждого математического метода, эффективность которого при анализе наборов данных доказывается в первом разделе, приводится пример его применения при изучении социальных процессов, и на его основе доказывается эффективность и необходимость применять математические модели при анализе социальных процессов.

Обосновывается необходимость применения математических моделей при изучении социальных процессов, в целях углубленного анализа данных, и уточнения выводов.

По итогам второго раздела для сформулированного набора математических методов поясняется методика и доказывается эффективность его применения при изучении социальных процессов

В третьем разделе «Особенности использования математических методов при анализе мнений жителей города Саратова о специфике исчисления времени в регионе» математические модели корреляционного, факторного анализа применяется для выявления факторов, влияющих на отношение жителей города Саратов к различным вариантам исчисления времени. Поясняется логика построения корреляционных и факторных моделей анализа для конкретного набора данных по теме изучения отношения населения города Саратов к различным вариантам исчисления времени.

По итогам третьего раздела доказывается эффективность применения математических методов при изучении предпочтений населения в вопросе исчисления времени.

ЗАКЛЮЧЕНИЕ

В процессе написания бакалаврской работы была показана эффективность использования математических методов при изучении общественных процессов на примере исследования предпочтений населения в вопросе исчисления времени.

Была изучена область исследований об исчислении времени, выявлены различные подходы к установлению оптимального варианта исчисления времени, критерии его определения, исследования фактического действия различных вариантов исчисления времени. Представлены аргументы при установлении того или иного варианта исчисления времени на различных исторических периодах.

Рассмотрена проблематика применения математических моделей при анализе общественных процессов. Обоснована необходимость их применения,

история возникновения данной необходимости и сохраняющаяся традиция применения математических методов в современных исследованиях, в том числе в области социальных наук.

В рамках первого раздела рассмотрены различные математические модели с точки зрения их строгого научного описания. Перечислены такие статистические меры, как среднее, медиана, мода, стандартное отклонение, и дисперсия, их непосредственное значение и применение в анализе, также при расчетах более комплексных показателей. Приведены корреляционные модели для выявления связи между переменными, измеренными на разных типах шкал. Для номинальных шкал: хи-квадрат Пирсона, коэффициент сопряженности Пирсона, лямбда, фи, V Крамера. Для порядковых шкал: ро Спирмена, тау Кендалла, гамма Гудмена-Краскала. Для проверки связи между категориальными и количественными признаками: тест Стьюдента, однофакторный дисперсионный анализ. Для проверки связи между номинальными и порядковыми переменными, в случае равных интервалов между порядками применяют метод дисперсионного анализа ранговых сравнений Краскала-Уоллиса. Также были рассмотрены: регрессионная модель для прогнозирования, факторная модель для снижения размерности признаков и кластерная модель для снижения размерности наблюдений.

В рамках второго раздела была описана эффективность применения рассмотренных математических методов применительно конкретным исследованиям в социальных и гуманитарных науках. Хи-квадрат позволил установить, что с возрастом растет предпочтение ценностей традиционализма над ценностями обновления в исследовании основания ценностей консолидации россиян. Коэффициент сопряженности Пирсона позволил установить, что высокая доля мигрантов к местному населению способствует росту социально-политической нестабильности в Арабских странах. V Крамера позволил установить, что формальное наличие трудового договора существенно влияет на социальные гарантии прекариев. Коэффициент лямбда позволил установить, что насильственные действия чаще свойственны представителям мужского пола

среди учеников 5-11 классов, при изучении темы насилия в школе и проблем семьи. Коэффициент корреляции гамма позволил установить, что основной детерминантой уровня счастья является материальное положение. При помощи теста Стьюдента удалось установить, что возраст обучающихся в очной форме существенно ниже возраста студентов заочной формы. При помощи метода однофакторного дисперсионного анализа удалось установить, что кредитно-заемное поведение существенно отличается для людей с разными доходами, более бедные чаще пользуются услугами микрофинансовых организаций, люди со средним доходом – потребительскими кредитами, а люди с более высоким доходом – ипотечным кредитованием. Метод дисперсионного анализа Краскала-Уоллиса позволил установить разницу в уровне жизни в городах разных стран в оценках их жителей.

Регрессионная модель применялась при изучении взглядов россиян на социальную политику государства, и было установлено, что в представлении россиян ответственность государства за достойный уровень жизни пенсионеров и безработных существенно снизилась за время между двумя волнами исследования, все ещё оставаясь одной из самых высоких в Европе, в то время как непосредственно оценка уровня жизни пенсионеров и безработных значительно выросла.

Факторная модель позволила в исследовании факторов гражданского участия на электронных платформах установить, что: безопасность личных данных, информационная безопасность и безопасность частной жизни определяют фактор безопасности; разнообразие выхода в интернет, электронный документооборот и доверие цифровым платформам определяют фактор доступности; знание возможностей использования, удобство использования и эффективность использования платформ определяют фактор эффективности.

Кластерная модель позволила определить 6 групп населения по характеру потребления алкоголя, он позволил автору прийти к выводу, что для большей части россиян процесс потребления алкоголя является комбинированным, а также преимущественно связан с социальным взаимодействием и совместным

приемом пищи. В высокодоходном социальном слое чаще отдается предпочтение дорогим и легким сортам алкоголя, в среднем сочетаются разные стили, а также популярно производство домашнего алкоголя, в низших слоях свойственно традиционное представление о потреблении алкоголя, фабричный алкоголь, пиво и водка, также ей свойственно чрезмерное потребление алкоголя, его крайние и деструктивные формы.

В рамках третьего раздела были практически применены математические модели в авторском исследовании предпочтений населения в исчислении времени.

При помощи коэффициента корреляции гамма было установлено, что отношение к текущему варианту исчисления времени зависит от возраста, в более младших возрастных группах чаще недовольны текущим вариантом исчисления времени, в то время как в более старших возрастных группах, текущий вариант исчисления времени чаще устраивает.

При помощи коэффициента корреляции лямбда установлено, что отношение к текущему варианту исчисления времени зависит от семейного положения. Состоящих в браке чаще текущий вариант исчисления времени устраивает, в то время как не состоящие в браке чаще недовольны текущим вариантом исчисления времени.

При помощи критерия хи-квадрат было установлено, что предпочитаемый вариант исчисления времени зависит от возраста. Представители старших возрастных групп сторонники сохранения текущего варианта исчисления времени, в то время как представители более младших возрастных групп предпочитают установление московского времени.

При помощи коэффициентов корреляции фи и V Крамера было установлено, что предпочитаемый вариант исчисления времени зависит от значимости фактора личного удобства при определении оптимального варианта исчисления времени. Те, кто считает критерий личного удобства значимым чаще предпочитают московское время, в то время как, считающие критерий личного

удобства не значимым предпочитают сохранение текущего варианта исчисления времени.

Факторный анализ позволил 12 критериев определения оптимального варианта исчисления времени выразить, как 5 факторов: решение проблемы автомобильных аварий путём заимствования успешных мировых практик, превалирование субъективных потребностей в личном удобстве при проведении пассивного или активного отдыха, консервативная компонента сохранения традиционно установленного в регионе варианта исчисления времени при недопущении его дальнейшего изменения, компонента экономической целесообразности, как базового фактора установления внутривластного соответствия исчисления времени географическому месторасположению региона, а также компонент профессиональной заинтересованности в удобстве построения логистики.

Таким образом в социологии применение математических методов способствует повышению эффективности проведения социологических исследований. Такие исследования, помимо того, что характеризуются большей глубиной выводов, также, как правило, имеют большую практическую значимость, и перспективу дальнейшего и ещё более углубленного анализа. Используемые в комплексе, математические модели позволяют преобразовывать исследовательские предположения в научные факты, с конкретной измерительной точностью. Выводы, выстраиваемые из этих фактов, являются не исключительно умозрительными, но также могут быть визуализированы таблично и графически. Это способствует наглядности таких выводов, их применимости в широкой области наук, их продвижению и социальной значимости таких исследований.