

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

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

Кафедра социальной информатики

**ИСПОЛЬЗОВАНИЕ МЕТОДА СРАВНЕНИЯ СРЕДНИХ ПРИ
ИЗУЧЕНИИ РОЛИ СОЦИАЛЬНЫХ СЕТЕЙ В ЖИЗНИ
МОЛОДЕЖИ**

(автореферат бакалаврской работы)

Студента 5 курса 531 группы
направления 09.03.03 – Прикладная информатика,
профиль – Прикладная информатика в социологии
Социологического факультета
Вантеева Михаила Петровича

Научный руководитель
кандидат социологических наук,
доцент

_____ С.В. Курганова
подпись, дата

Зав. кафедрой
кандидат социологических наук, доцент

_____ И.Г. Малинский
подпись, дата

Саратов 2019

ВВЕДЕНИЕ

Актуальность проблемы. Современные развитые общества не случайно характеризуют термином «информационные». В настоящее время резко увеличились потоки информации, возросла и ее роль в экономической, политической, социальной сферах, в развитии всего общества в целом¹, а соответственно, приобретают решающую роль получение, хранение, обработка, распространение и использование информации и знания. В свою очередь, происходят кардинальные изменения во всех отраслях знания, в том числе и в социальных науках.

На базе электронных источников и компьютерных технологий развиваются новые методы и направления исследований общественных процессов, основанные на анализе больших массивов данных и документов. С одной стороны, отмечается появление все большего числа новых приемов обработки социальной информации, а также совершенствования старых, уже хорошо знакомых. С другой, даже довольно сложный инструментарий становится все более доступным для широкого числа пользователей, что обуславливается массовой компьютеризацией и технологизацией аналитических приемов.

Развитие современного общества во многом обусловлено развитием компьютерных информационных технологий, которые пронизывают все сферы жизнедеятельности человека, обеспечивая функционирование общественного организма. Это общемировая тенденция, которая в полной мере проявляется и в сфере научных исследований. В настоящее время без использования компьютера крайне сложно представить себе современную социологию, ведь именно благодаря применению компьютерной техники и соответствующего программного обеспечения значительно упростились процедуры обработки социологических данных.

¹ Кастельс, М. Информационная эпоха: экономика, общество и культура / М.Кастельс. М.: ГУ ВШЭ, 2000. 608 с.; Кастельс, М. Власть коммуникации: учеб. пособие / М. Кастельс. М.: ГУ ВШЭ, 2016. 563 с.

Сравнение средних является одним из наиболее популярных способов анализа социологических данных. При этом правильнее было бы говорить не столько о методе, сколько о группе методов, подпадающих под это название. Сюда относятся как самые простые, довольно легко осуществляемые даже с помощью ручной обработки, методы (например, t-тест одной выборки), так и очень сложные, поддающиеся интерпретации только опытных аналитиков (многомерные разновидности дисперсионного анализа). Тем не менее, с помощью новых компьютерных технологий даже рядовые пользователи оказываются способными использовать сравнение средних для решения своих исследовательских задач. В данной работе мы попытаемся проиллюстрировать возможности метода сравнения средних в области исследования социологических данных через построение простых и многомерных моделей.

Степень научной разработанности. Метод сравнения средних был известен еще задолго до повсеместного распространения компьютеров и их вхождения в повседневную практику социальных исследований. До сих пор еще представляют интерес работы, раскрывающие особенности дисперсионного анализа именно как математической процедуры.

Его актуальность и востребованность подтверждается появлением все новых работ, включающих наряду с другими методами анализа данных и сравнение средних.

Однако появление специализированных компьютерных программ статистической обработки эмпирического материала и их популяризация среди специалистов-гуманитариев привели к настоящему буму работ, посвященных анализу данных при помощи компьютерных программ. Особенно много публикаций посвящается возможностям анализа статистической информации с помощью программы SPSS. Таким образом, тема нашего исследования, связанная с особенностями обработки социологических данных при помощи метода сравнения средних в программе статистической обработки SPSS, является на сегодняшний день весьма актуальной. Это и определило выбор объекта и предмета исследования, постановку его цели и задач.

Исходя из вышесказанного, **объектом** настоящего исследования является группа методов сравнения средних, представленная в специализированной компьютерной программе статистической обработки данных SPSS. **Предметом** исследования выступает технология применения и социологическая интерпретация результатов методов сравнения средних на примере изучения роли социальных сетей в жизни молодежи г. Саратова.

Цель исследования – раскрыть познавательный потенциал применения разных методов сравнения средних. Для достижения поставленной цели необходимо решить следующие **задачи**:

1. охарактеризовать группу методов сравнения средних как особые математические процедуры преобразования данных;
2. выявить возможности статистического пакета SPSS в реализации группы методов сравнения средних;
3. проанализировать обозначенными методами роль социальных сетей в жизни молодежи и выявить их дифференциацию внутри молодежной группы.

Эмпирической базой исследования являются материалы авторского социологического исследования на тему «Роль социальных сетей в жизни саратовской молодежи»¹.

Научная новизна заключается в обосновании группы методов сравнения средних как способа раскрытия сложных нелинейных связей между переменными с разными типами шкал.

Структура работы. Работа состоит из введения, двух разделов (1 раздел «Группа методов сравнения средних: общая характеристика и типология», 2 раздел «Особенности применения различных типов метода сравнения средних

¹ Социологическое исследование на тему «Роль социальных сетей в жизни саратовской молодежи» было проведено в 2017 г. Методом сбора информации было выбрано формализованное интервью. Объектом исследования выступала молодежь в возрасте от 20 до 30 лет. Выборочная совокупность конструировалась по принципу квотно-стратифицированной, критерии отбора – пол и возраст респондентов. В общей сложности в исследовании приняли участие 100 респондентов. Электронная база данных была создана в программе SPSS 19.0.

для проверки гомогенности социологических данных»), заключения, списка использованных источников и приложения.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

В первом разделе «Группа методов сравнения средних: общая характеристика и типология» рассматриваются основные принципы и логика реализации анализа методами сравнения средних, дается их классификация, общая характеристика и особенности реализации в программе SPSS.

Сравнение средних представляет собой большую и разветвленную группу параметрических методов исследования данных, объединяет которые сравнение средних разных выборок. Наиболее простым методом из них является t-тест, поскольку дает ответ на вопрос о статистическом различии двух выборок по уровню выраженности измеряемого признака.

Всего выделяют три варианта t-теста. Т-тест для одной выборки предназначен для сопоставления среднего значения одной выборочной совокупности с некой эталонной величиной, задаваемой, как правило, генеральной совокупностью. В последнем случае исследователь вынужден прибегать к неким внешним источникам статистической информации, что иногда может вызвать определенные затруднения. Т-тест для двух независимых выборок нацелен на сравнение средних арифметических двух выборок. Причем для сравниваемых выборок должны быть заданы значения одной зависимой переменной с количественной шкалой. Т-тест для двух зависимых (или связанных) выборок дает возможность сравнить средние арифметические двух измерений одной и той же переменной для одной и той же выборки. Таким образом, в данной случае обязательным условием выступает повторное измерение для одной выборки, что в практике социологических исследований является довольно редким.

Дисперсионный анализ — это статистический метод, позволяющий анализировать влияние различных факторов на исследуемую переменную.

Впервые данный метод был разработан в 1925 году биологом Рональдом Фишером.

Главной целью дисперсионного анализа является оценка степени значимости различий между средними с помощью сравнения дисперсий. Дисперсию изучаемого признака разбивают на независимые слагаемые, причем каждое характеризует влияние определенного фактора или же их взаимодействия. Сравнение таких слагаемых дает возможность оценить значимость каждого исследуемого фактора, а также их комбинаций.

Дисперсионный анализ является целым комплексом методов, которые выделяются в зависимости от того, сколько факторов принимает участие в рассмотрении, сколько зависимых переменных исследуется и по тому, как соотносятся друг с другом выборки значений.

Самым распространенным и наиболее простым является метод однофакторного дисперсионного анализа. В этом случае анализируется влияние одного фактора на дисперсию выборок, заданных одной зависимой переменной. Случай же, когда исследуется одновременное влияние двух и более факторов, называется многофакторным дисперсионным анализом. Когда же в анализе участвуют несколько зависимых переменных, тогда обращаются к многомерному дисперсионному анализу. Наконец, когда исследуется воздействие факторов на одну и ту же группу объектов, но «замерянных» в разные временные отрезки, используют дисперсионный анализ с повторными измерениями.

Обращаясь к дисперсионному анализу, исследователь стремится выяснить, как изменяется конкретный признак в изменяющихся условиях действия переменной-фактора. Идея дисперсионного анализа заключается в сопоставлении «факторной дисперсии», или рассеяние, задаваемое изменением уровня фактора, и «остаточной дисперсии», определяемой случайными причинами. Если их различие является статистически значимым, то делается вывод о том, что фактор влияет на X и при изменении его уровня групповые средние различаются статистически значимо. Когда определили, что фактор

оказывает статистически значимое влияние на X , тогда надо уточнить, какие именно уровни оказывают значимое воздействие, и для этого проводят попарное сравнение средних. С этой целью можно обратиться к t-тесту для двух выборок, что является несколько трудоемким. Или же использовать апостериорные множественные сравнения.

Второй разновидностью дисперсионного анализа является *многофакторный дисперсионный анализ*. Основное отличие многофакторного анализа состоит в том, что в данном случае берут не одну, а несколько независимых переменных-факторов, имеющих номинальную шкалу. Сразу же отметим, что принципиальной разницы между данными разновидностями дисперсионного анализа не существует. Многофакторный анализ не отменяет общие принципы проведения дисперсионного анализа, а только делает ее более сложной, так как кроме анализа влияния на зависимую переменную каждого из факторов по отдельности, надо еще исследовать и их совместное действие. Взаимодействие двух факторов означает, что влияние первого фактора реализуется особым образом на разных уровнях второго фактора. Наиболее простым и полезным в процессе интерпретации взаимодействий выступает такой инструмент как диаграммы средних значений для каждой ячейки таблицы сопряженности независимых переменных.

Третьей разновидностью дисперсионного анализа является *многомерный дисперсионный анализ*, который нацелен на решение задач, где надо учитывать не одну, а две и более зависимых переменных. Представим, что нам необходимо сравнить мужчин и женщин одновременно по двум зависимым переменным. В подобной ситуации одним из возможных подходов является двукратное применение t-теста. Очевидным достоинством такого решения является простота и ясность, однако нельзя не заметить и двух недостатков. Во-первых, при неоднократном применении статистического критерия увеличивается вероятность ошибки, то есть вероятность случайности общего результата исследования. Во-вторых, если между зависимыми переменными

имеется некоторая корреляция, то результат, полученный в отношении каждой из этих переменных в отдельности, не способен отразить этот важный факт.

Отметим, что многомерный дисперсионный анализ, как и одномерный многофакторный, позволяет оценить влияния не только отдельных независимых переменных (главных эффектов), но и их взаимодействий.

Таким образом, в настоящее время разработаны и активно используются исследователями самые разные виды дисперсионного анализа, представляющие собой варианты адаптации метода к тем или иным особенностям эмпирических данных. В целом процедура дисперсионного анализа любой разновидности является довольно громоздкой в математическом плане, но результативной в обработке социологической информации. Появление персонального компьютера третьего поколения и его широкое распространение оказалось тем стимулом, который способствовал популяризации и активному использованию самых сложных методов статистического анализа в социологической практике, в том числе и дисперсионного анализа.

С конца XX века в России активно используется один из самых распространенных программных комплексов для обработки статистической информации - SPSS. В рамках дисперсионного анализа SPSS предлагает множество возможностей и предоставляет возможности пользователю даже без специальной математической подготовки реализовывать дисперсионный анализ разной степени сложности и получать удобные для интерпретации результаты статистического анализа.

Второй раздел «Особенности применения различных типов метода сравнения средних для проверки гомогенности социологических данных» посвящен практической реализации методов сравнения средних на примере данных, полученных в авторском исследовании «Роль социальных сетей в жизни саратовской молодежи».

ЗАКЛЮЧЕНИЕ

Общенаучная ценность группы методов сравнения средних заключается в проверке значимости различий между средними показателями с помощью

сравнения дисперсий. Дисперсия измеряемого признака раскладывается на независимые слагаемые, каждое из которых характеризует влияние того или иного фактора или их взаимодействия. Последующее сравнение таких слагаемых позволяет оценить значимость каждого изучаемого фактора, а также их комбинации.

В настоящее время разработаны и активно используются исследователями самые разные методы сравнения средних, представляющие собой варианты адаптации метода к тем или иным особенностям эмпирических данных. Наиболее простыми и популярными являются методы t-теста и однофакторного дисперсионного анализа, являющегося фактически развернутым вариантом t-теста, когда исследуется влияние лишь одного фактора на дисперсию двух или более выборок.

Усложненной версией однофакторного дисперсионного анализа является многофакторный анализ. Главное отличие последнего заключается в том, что в нем участвуют не одна, а несколько независимых переменных, представленных в виде номинальной шкалы. Многофакторный анализ не меняет общую логику дисперсионного анализа, а лишь усложняет ее, поскольку, кроме учета влияния на зависимую переменную каждого из факторов по отдельности, следует оценивать и их совместное действие. Взаимодействие двух факторов означает, что влияние одного из них проявляется по-разному на разных уровнях другого фактора. При этом решение проблемы взаимодействия факторов принципиально не зависит от их числа, существенно нарастает лишь сложность интерпретации взаимодействий.

Многомерный дисперсионный анализ служит для построения моделей следующего порядка сложности. Особенностью всех типов многомерного анализа, в том числе и дисперсионного, является то, что они обрабатывают все зависимые переменные одновременно. Также многомерный дисперсионный анализ позволяет оценить влияние не только отдельных независимых переменных, но и их взаимодействий.

В целом процедура сравнения средних любой разновидности является

довольно громоздкой в математическом плане, но результативной в обработке социологической информации. Широкое распространение персонального компьютера способствовало популяризации и активному использованию самых сложных методов статистического анализа в социологической практике, в том числе и данного метода.

В рамках сравнения средних программа SPSS предлагает множество возможностей. В зависимости от типа экспериментального плана выделяются несколько разновидностей t-теста (t-тест для двух независимых выборок, t-тест для двух связанных выборок и t-тест для одной выборки), однофакторный, многофакторный, многомерный дисперсионный анализ и некоторые другие. В целом статистический пакет SPSS предоставляет возможности пользователю даже без специальной математической подготовки реализовывать методы сравнения средних разной степени сложности и получать удобные для интерпретации результаты статистического анализа.

В рамках данной квалификационной работы с целью выявления познавательного потенциала методов сравнения средних было проведено исследование роли социальных сетей в разных группах городской молодежи. При помощи программы SPSS были проведены t_тест для двух независимых выборок, а также однофакторный, многофакторный и многомерный дисперсионный анализ, позволившие продемонстрировать особенности технологии применения и социологической интерпретации результатов данных методов анализа.

Применение серии t-тестов и однофакторного дисперсионного анализа позволило выявить ряд закономерностей в пользовании социальными сетями молодыми саратовцами, а именно то, что холостые несколько меньше, по сравнению семейными респондентами, уделяют времени общению в социальных сетях. У респондентов со средним профессиональным образованием меньше друзей в соцсетях, но они чаще в них заходят, чем респонденты с высшим образованием, у которых друзей больше. Также было выявлено, что респонденты с меньшим опытом пребывания в соцсетях реже

общаются в них и имеют меньшее число друзей в соцсетях, с которыми общаются и в реальной жизни, тогда как респонденты с большим опытом – наоборот. Молодежь, ориентированная на общение, имеет меньшее количество друзей, чем респонденты, интересующиеся всевозможной информацией: фото, видео, музыкой. Меньше всего времени в социальных сетях проводят респонденты, зарегистрированные в Одноклассниках и ВКонтакте. Кроме того, респонденты, готовые отказаться от посещения соцсетей, заходят в них реже, чем те, кто не готов к этому. Тот же принцип сохраняется и в случае со временем, проводимом в соцсетях. Наконец, наименьшее количество друзей в соцсетях у тех, кто легко откажется от присутствия виртуального общения, больше всего у тех, кто не сможет отказаться.

Проведение многофакторного дисперсионного анализа позволило несколько обобщить уже имеющуюся информацию и выявить сложные нелинейные связи между целым комплексом переменных. Так, респонденты, состоящие в браке, зарегистрированные в соцсетях менее пяти лет и предпочитающие пользоваться сетью ВКонтакте, больше времени проводят в соцсетях, чем те же респонденты, но выбирающие другие соцсети. В свою очередь, семейные респонденты, кто в соцсетях состоит уже более пяти лет и выбирает другие сети, больше общаются в соцсетях, чем те же семейные респонденты с большим опытом пользования соцсетями, но предпочитающие Одноклассников и ВКонтакте.

Что касается холостых респондентов, то для них оказываются характерны совсем другие тенденции. Больше всего времени в социальных сетях проводят респонденты как с большим опытом пользования соцсетями, так и с маленьким, но только те, кто предпочитают сеть ВКонтакте. На несколько меньшее время заходят в сеть респонденты с опытом менее пяти лет и предпочтением других соцсетей, совсем мало проводят времени в сети респонденты с опытом менее пяти лет и предпочтением Одноклассников. Что касается холостых респондентов с большим опытом пользования соцсетями, то здесь среднюю позицию по времени, проводимом в соцсетях, занимают любители

Одноклассников, чуть меньше времени в соцсетях проводят предпочитающие другие сети.

С другой стороны, чаще всего посещают социальные сети респонденты со средним профессиональным образованием, которые отмечают, что могут легко отказаться от пользования соцсетями. Те же респонденты со средним профессиональным образованием и которые могут, но не хотели бы отказываться, посещают соцсети значительно реже, еще чуть реже это делают респонденты с тем же уровнем образования, но не желающие отказываться от пользования соцсетями. Что касается респондентов с высшим образованием, то их уровень пользования соцсетями намного стабильнее и в зависимости от степени готовности отказаться от социальных сетей практически не меняется. Лишь чуть чаще пользуются соцсетями те, кто утверждает, что легко откажется от соцсетей в случае необходимости.

Согласно многомерному дисперсионному анализу, больше всего социальных контактов в виртуальном и реальном мире поддерживают респонденты с большим опытом пользования соцсетями и предпочитающие другие социальные сети, тогда как у респондентов с небольшим опытом число друзей много меньше. Респонденты, пользующиеся сетью ВКонтакте, имеют практически равное число друзей вне зависимости от опыта. То же можно сказать и о поклонниках Одноклассников, с одной пометкой относительно того, что у них число друзей чуть ниже, чем у пользователей сети ВКонтакте.

Таким образом, применение методов сравнения средних в социологии на сегодняшний день является не только технологией обработки статистических баз данных, но и способом усовершенствования алгоритмов аналитической работы, отвечающей вызовам методологии современных научных исследований.