

Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

кафедра социальной информатики

**ЭФФЕКТИВНОСТЬ ПРИМЕНЕНИЯ ПРОГНОЗНЫХ
КОЭФФИЦИЕНТОВ В СОЦИОЛОГИЧЕСКИХ
ИССЛЕДОВАНИЯХ**

(автореферат бакалаврской работы)

студента 4 курса 451 группы
направления 09.03.03 – Прикладная информатика
профиль подготовки – Прикладная информатика в социологии
социологического факультета
Алёнина Виталия Алексеевича

Научный руководитель
кандидат социологических наук, доцент

_____ С.В. Ситникова
подпись, дата

Зав. кафедрой
кандидат социологических наук, доцент

_____ И.Г.Малинский
подпись, дата

Саратов 2018

Введение. На сегодняшний день в нашей стране все более широкое распространение получают конкретные социологические исследования, которые являются источником разносторонней социальной информации, необходимой для успешного решения важных социально-экономических задач, для научного управления общественными процессами и явлениями.

Так, например, в конце двадцатого столетия в работе Н.Смелзера выделялось пять подходов при изучении социальных явлений: демографический, коллективистический, культурологический, психологический и подход, основанный на выявлении взаимоотношений.

Более традиционно, основываясь на современную научную интерпретацию, можно выделить два основных подхода к изучению процессов, происходящих в обществе через социологическое исследование.

Первый подход, который условно можно назвать *статистическим подходом (или количественным)* основывается на том, что массовые явления имеют статистический характер, то есть, если изучить достаточно большое количество проявлений изучаемого социального явления, то само явление будет познано.

Второй подход – *качественный*. Сторонники данного подхода полагают, что с помощью строго структурированных опросов невозможно исследовать глубину социального явления. Основываясь на качественном подходе социолог зачастую ищет причины и следствия исследуемого явления.

На сегодняшний день в социологических исследованиях активно используются математические методы, и это закономерный этап в развитии прикладной социологии, сочетающий количественный и качественный анализ, что позволяет применять в социологии современные математико-статистические методы так же, как в экономике и иных естественных науках.

Эволюционировали и способы обработки информации. XXI век принято называть веком информационных технологий. Именно стремительная информатизация всех сфер жизни общества дает возможность автоматизировать процесс обработки данных. Использование различных

программных и технических средств в процессе проведения социологического исследования позволяет решить многие проблемы, получить разнообразную информацию о различных областях жизни современного общества. Сегодня на рынке существует большое число специализированных программных продуктов, облегчающих процесс проведения социологических исследований на разных этапах, от составления программы исследования до анализа полученных данных и составления отчёта о проделанной работе.

Говоря о технологиях обработки данных, стоит рассматривать такое важное понятие, как «мера связи». Меры связи дают социологу информацию о тесноте и направлении связей между исследуемыми переменными и зависимостями. С научной точки зрения, меры связи позволяют решать несколько важных задач.

Во-первых, с их помощью можно определить причинно-следственные связи между переменными. Во-вторых, важной областью использования мер связи является предсказание (прогнозирование). То есть если существует связь между двумя переменными, то для некоторого случая возможно предсказать численное значение одной переменной, если известно численное значение другой. Именно в контексте данной научной области вводится понятие значимости прогнозных коэффициентов в социологических исследованиях и степень эффективности их использования. Это обусловило выбор объекта, предмета, цели и задач данного дипломного проекта.

Так как данная тематика до этого не встречалась в научных работах, представленных на русском языке, а процесс качественного улучшения обработки данных исследований постоянно модернизируется, то этим, в том числе, обосновывается **актуальность выбранной темы работы.**

Объектом является прогнозные методики, используемые в анализе результатов социологического исследования.

Предмет – степень эффективности использования прогнозных коэффициентов на примере авторского исследования читательских

предпочтений молодежи Саратовского государственного университета имени Н.Г.Чернышевского.

Цель данной работы – выявить возможности применения прогнозных коэффициентов при обработке результатов исследования и положительные стороны, которые данные коэффициенты могут принести в работе социолога-прикладника в процессе анализа данных.

Достижение цели сопряжено с решением ряда исследовательских **задач**:

1. Представить современные методики обработки социологических данных;
2. Описать теоретические основы и практические возможности использования коэффициентов Лямбда и Гамма в социологическом исследовании при анализе статистической информации;
3. Выявить эффективность прогнозных коэффициентов Лямбда и Гамма на примере авторского исследования;
4. Выявить факторы, влияющие на читательские предпочтения молодежи.

Эмпирической базой исследования являются данные проведенного анкетирования по квотной выборке объемом в 83 респондента (рассчитано с помощью таблицы уровня точности выборки и данных представленных на сайте <http://www.sgu.ru/>). Источниковедческой базой дипломного проекта стали научные статьи, публикации, пособия и справочники по теме исследования.

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

Метод исследования: системный анализ, применением социологического опроса на начальном этапе исследования.

Структура выпускной бакалаврской работы представлена введением, тремя разделами, заключением, списком использованных источников и приложением.

Основное содержание работы. Для анализа количественных данных существует большое количество различных специализированных программ – статистических пакетов, каждый из которых обладает рядом достоинств и

недостатков. Для реализации различных задач можно использовать разные статистические пакеты, в которые включены разнообразные методы статистики от простой описательной статистики до анализа главных компонент.

В первом разделе работы были описаны наиболее интересные и распространенные из программ, учитывая разные этапы проведения социологического исследования.

Основными стадиями проведения социологического исследования являются три основных этапа: подготовительный этап, сбор данных и анализ данных. На каждом из них, так или иначе, используются различные информационные технологии. В работе рассмотрены возможности использования информационных технологий на каждом из этих этапов.

Были выделены следующие три группы программ:

1. интегрированные методоориентированные пакеты общего назначения (SPSS, STATA, STATISTIKA) – достоинством данных программ можно назвать их универсальность, так как они не ориентированы на какую-то конкретную предметную область. Они имеют относительно простой интерфейс, что позволяет работать с пакетом не только профессиональным исследователям, но и обычным пользователям;
2. специализированные методоориентированные пакеты (STADIA, Олимп, SAS, BMDP) – данные программы ориентированы в первую очередь на корреляционно-регрессионный анализ, анализ временных рядов, кластерный или факторный анализ;
3. предметно- (или проблемно-) ориентированные пакеты (BioStat, MESOSAUR, DATASCOPE) – предназначены для решения вопросов, связанных с конкретной предметной областью и ориентированы на высококвалифицированных специалистов.

Подробно изложены возможности программы SPSS для обработки первичной социологической информации. В пакете реализованы следующие методы статистической обработки информации: расчёт суммарных статистик для отдельных переменных; расчёт частот, суммарных статистик и графиков

для произвольного числа переменных; построение таблиц сопряженности; расчёт средних, стандартных отклонений; корреляционный анализ; дискриминантный анализ; однофакторный и многофакторный дисперсионный анализ; факторный анализ; кластерный анализ; непараметрические тесты; множественная регрессия и другие. В рамках исследуемой темы для нас особый интерес вызывает возможность использования SPSS для анализа связи между переменными и возможность использования встроенных компонент для вывода прогнозных коэффициентов Лямбда и Гамма, прибегая к процедуре Crosstabs (таблицы сопряженности). Даны теоретические основы использования SPSS для нахождения данных мер связи.

Подводя итог вышесказанному, можно утверждать, что с приходом информационных технологий в социологию возможности проведения исследований серьёзно расширились. Во многом усовершенствовался процесс подготовки исследования. Глобальная компьютеризация позволила уменьшить объёмы бумажной работы, что в свою очередь значительно сократило временные затраты на подготовку исследования.

Претерпел значительные изменения характер проведения исследования и обработки результатов. С внедрением информационных технологий в социологию усовершенствовались уже существующие методы проведения исследования и появились новые более эффективные. Также появилось много возможностей для обработки данных, полученных в ходе исследования.

Во **втором разделе** работы описаны возможности прогнозирования в социологических исследованиях с использованием коэффициентов Лямбда и Гамма. Данные коэффициенты служат дополнением к прежним мерам связи, основанным на хи-квадрат, и расширяют возможности более точной обработки данных исследования.

Если между переменными существует статическая связь, то дополнительная информация, полученная в ходе расчета прогнозных коэффициентов, снижает количество ошибок прогнозирования. Чем сильнее связь между переменными, тем более значительным будет сокращение

количества ошибок. Благодаря данному подходу мы имеем возможность прогнозирования социального явления на всю генеральную совокупность на основе проведенного исследования по определенной выборке. Для практического применения данного подхода в социологии введены статистические показатели – прогнозные коэффициенты. Если мы рассматриваем номинальные переменные, то статистикой является прогнозный коэффициент Лямбда, для порядковых переменных – это коэффициент Гамма.

В данном разделе подробно описан расчет коэффициентов вручную, используя данные и формулы. На сегодняшний день развитие информационных технологий дает возможность обрабатывать данные и получать необходимые результаты исследований, используя автоматизированные пакеты. В ходе выполнения выпускной квалификационной работы использовался пакет SPSS Statistics. В разделе проведены способы обработки описательной статистики SPSS.

Третий раздел данной бакалаврской работы посвящен прогнозным оценкам по выявлению факторов, влияющих на читательские предпочтения среди молодежи на примере студентов Саратовского национального исследовательского государственного университета имени Н.Г.Чернышевского на основе авторского исследования.

По данным современных исследований наблюдается трансформация культурных ценностей среди молодого поколения. Стремительное развитие информационных технологий, максимальная доступность информации посредством интернета влекут за собой изменения в образовательной и досуговой деятельности молодежи. Несомненно, на характере чтения студенческой молодежи сказывается специфика ее образа жизни и социализации.

Проведенное исследование позволило подтвердить основную гипотезу и сделать следующие выводы. Выявлены читательские предпочтения современной молодежи, а так же определены основные факторы, влияющие на читательские предпочтения молодежи – это возраст и курс обучения, выбор

места для чтения, наличие домашней библиотеки. К факторам, которые не влияют на читательские предпочтения, относится выбор носителя чтения.

В данном разделе проведены анализ прогнозируемости данных факторов через коэффициент Лямбда в части зависимости читательских предпочтений от половой принадлежности респондентов, влияние возраста на чтение книг и влияние наличия домашней библиотеки на чтение. Рассчитана связь между типом читаемой литературы и полом респондентов через коэффициент Гамма. Представлены данные рассчитанные в ручную и с использованием автоматизированных средств (пакет программы SPSS Statistics)/

Рассчитав коэффициент Лямбда для ряда критериев мы подтвердили выводы по частотным характеристикам, что основным фактором, влияющим на читательские предпочтения молодежи, является гендерная принадлежность. Рассчитав коэффициент Гамма, установили зависимость чтения определенной литературы от гендерной принадлежности респондентов.

Заключение. На практике социологу приходится работать с разнообразными областями исследования социальных явлений, преследуя при этом различные задачи. Особую степень интереса имеют те исследования, которые дают вклад в прогнозирование исследуемого явления. Само по себе прогнозирование в социологии – это способ научного предвидения с той или иной степенью вероятности итога, направленности или характера протекания социальных процессов в течение некоторого промежутка времени.

Главная особенность прогнозирования – нацеленность на будущее и попытка преодолеть *неопределенность*, обусловленную отсутствием знаний о точном значении статистических параметров характеристик социальных процессов. В связи с тем, что речь идет об определении статистических величин, любой прогноз в прикладной социологии носит вероятностный характер, т.е. предсказание является достоверным с той или иной степенью вероятности. Отсюда следует, что для любой обоснованной прогностической модели необходимы два критерия:

1) форма числовой оценки характеристик социальных явлений, определяемая видом шкалы измерения, на основе которой строится прогностическая модель;

2) степень погрешности оценки, которая зависит от числа факторов, воздействующих с той или иной степенью интенсивности на вариацию характеристик социального процесса в течение временного периода.

Для оценки степени погрешности возможно использования специальных прогностических коэффициентов.

Целью дипломной работы было изучение данных коэффициентов, изучение их характеристик, оценка эффективности применения в социологическом исследовании.

В данной работе представлено описание двух статистик – меры пропорционального уменьшения ошибки для номинальной шкалы Лямбда (L) и меры пропорционального уменьшения ошибки для порядковых шкал Гамма (G). Даны подробные теоретические характеристики и подходы к их расчету.

Проведено авторское исследование по выявлению факторов, влияющих на читательские предпочтения молодежи на примере студентов Саратовского национального исследовательского государственного университета имени Н.Г.Чернышевского. Обработав результаты исследования, были получены выходные данные, которые подтвердили гипотезы по влиянию на читательских предпочтений следующих основных факторов – это возраст и курс обучения, выбор места для чтения, наличие домашней библиотеки. Проводя исследование в рамках определенной выборки социолог-прикладник должен быть уверен в применении полученных результатов на всю генеральную совокупность, именно с этой позиции эффективным методом прогнозирования является использования статистик Лямбда и Гамма, что подтверждается расчетами проведенными с результатами авторского исследования.

Основной гипотезой авторского исследования служило утверждение, что основные факторы, влияющие на читательские предпочтения молодежи – это

возраст и курс обучения, выбор места для чтения, наличие домашней библиотеки.

Проверив полученные в ходе исследования факторы, влияющие на читательские предпочтения по статистике Лямбда, сделали следующий вывод, что знание о половой принадлежности случайного респондента повышает наши возможности прогнозирования читательских предпочтений на 32%. Остальные факторы имеют малый коэффициент прогнозирования и требуют уточняющих характеристик, либо рассмотрение иной выборочной совокупности.