Министерство образования и науки Российской Федерации

Федеральное агентство по образованию

ГОУ ВПО

ВСЕРОССИЙСКИЙ ЗАОЧНЫЙ ФИНИНСОВО-ЭКОНОМИЧЕСКИЙ ИНСТИТУТ

Кафедра менеджмента и маркетинга

КОНТРОЛЬНАЯ РАБОТА

*по дисциплине* «Социология»

на тему «Методы анализа эмпирических данных»

Преподаватель: ст. преп. Михальченков Н.В.

Работу выполнил студент: Крюкова О.С.

**Содержание**

Введение

1. Подготовка эмпирических данных к обработке и анализу
2. Сущность и виды группировок. Таблицы и графики: их роль в анализе социологических данных
3. Структура и основные требования отчета о социологическом исследовании

Заключение

Список используемой литературы

**Введение**

Социология не может существовать, не добывая эмпирическую информацию самого разного плана – о мнении избирателей, досуге школьников, рейтинге президента, семейном бюджете, количестве безработных, уровне рождаемости. Первым делом исследователь использует официальную статистику, публикуемую в журналах, бюллетенях, докладах. Недостающую информацию он добирает в социологическом опросе, где выясняются субъективные мнения людей (в анкетировании их называют **респондентами**). Ответы математически усредняются, обобщенные данные представляются в виде статистических таблиц, выводятся и объясняются **закономерности**. Конечный итог – построение **научной теории**, которая позволяет предсказывать будущие явления и разрабатывать практические рекомендации.

Социологическое исследование начинается вовсе не с составления анкеты, как принято думать, а с проработки проблемы, выдвижения целей, гипотез, построения теоретических моделей. Лишь затем социолог переходит к разработке инструментария (чаще всего это действительно анкета), затем сбору первичных данных и их обработке. А на конечной стадии снова теоретический анализ, ибо данные надо правильно, то есть в соответствии с выдвинутой теорией, интерпретировать и объяснить. Только после этого следуют практические рекомендации.

**1. Подготовка эмпирических данных к обработке и анализу**

Анализ получаемых в ходе социологического исследования данных начинается с контроля над качеством заполнения инструментария, исправления ошибок и отбраковки (выбраковки) некачественно заполненных анкет, бланков, карточек и т.п. Категории качества заполнения инструментария разнообразны, подходы здесь неоднозначны. Все зависит от условий работы анкетеров, интервьюеров, места опроса и других факторов.

Вначале проверяется правильность ответа на каждый вопрос и при необходимости исправляется. Например, на вопрос: «Знаете ли вы основные условия, заложенные в трудовом договоре?» чаще всего отвечают: «Да, знаю». Но далее в анкете стоит открытый контрольный вопрос: «Если знаете, назовите их, пожалуйста». Он остается незаполненным. В этом случае альтернативу «да, знаю» надо зачеркнуть и отметить другую, скорее всего «нет ответа», «затрудняюсь ответить» и т.п. Затем подсчитываются неправильные ответы. При исправлении каждого третьего ответа в анкете ее лучше не готовить к машинной обработке. В случае если респондент не ответил на 10-15% основных вопросов, против них контролер ставит отметку «нет ответа», и анкета идет на обработку на ЭВМ.

Более жесткие требования предъявляются к вопросам, касающимся социально-демографических характеристик (пол, возраст, образование), если нет ответов на эти вопросы, то инструментарий изымается из общего массива.

Допущенные к обработке документы нумеруются, в целях контроля за их прохождением. В дальнейшем массив передается кодировщикам. Но перед тем, как передать анкеты в руки операторов, надо закодировать открытые вопросы. На каждый открытый вопрос, как правило составляется не менее пяти шифров-кодов. Выше уже приводился один из открытых вопросов: «Если вы знаете, назовите, пожалуйста». Ответы могут быть самыми разнообразными: от полных, глубоких, обстоятельных, до ответов схем. Шифры-коды должны отражать шкалу интенсивности ответов на открытый вопрос. Обычно кодировка таких ответов проходит в два-три этапа. Сначала варианты ответов выписывают отдельно, подсчитывают число употреблений каждого варианта – частота его повторения. Затем варианты группируются по смысловой близости, совпадению. Таких групп, как показывает практика набирается четыре-пять, и каждой из них присваивается свой шифр или код.

Кодировка выступает связующим звеном между качественной и количественной информацией. На данной основе как раз и осуществляются числовые операции с информацией, введенной в память машины. Если во время кодировки произойдет сбой, замена или потеря кода, то информация окажется неправильной.

Сущность обработки первичной информации заключается в ее обобщении. Результаты обобщения называются *социологической информацией.* Решение о способе обработки инструментария принимается заранее. При опросе 60 – 70 человек обработку вполне можно провести вручную на микрокалькуляторе. Если анкета сравнительно невелика (содержит до 20 вопросов), то ручная обработка возможна при наличии 200-350 анкет. Но если в инструментарии более 20 вопросов, тогда предел для ручной обработки – 100-200 анкет. При обработке на ЭВМ результаты отражаются в виде табуляграмм, поэтому здесь нужна помощь программиста.

Каждый вопрос в анкете или бланке интервью представляет собой в определенной степени шкалу измерений. Единицами измерений выступают соответствующие альтернативы (позиции), варианты ответов. По эти позициям (вариантам ответов) проводится группировка респондентов. Кроме того, определенную шкалу измерений представляют объективные характеристики опрашиваемых, их субъективные оценки, предпочтения и пр.

Измерение производится с помощью различных шкал, которым соответствуют различные правила математического анализа данных. В социологических исследованиях применяются, как правило, шкалы трех основных типов: номинальная, ранговая (порядковая) и интервальная.

Простейшая шкала – *номинальная*. Чаще всего она фиксирует (отражает) дихотомичный ответ: «да» или «нет», «тепло» или «холодно». Применяя *ранговую* шкалу, можно зафиксировать более определенное состояние, определяющее ранжированное распределение типа «холодно» - «тепло» - «горячо». Но цифровые величины пока отсутствуют. Если же мы точку замерзания воды примем за нуль, а точку кипения (пар) за 100 и расстояние между этими точками разделим на 10 равных интервалов, то получим *интервальную* шкалу.

Номинальной шкале в анкете обычно соответствуют вопросы, способствующие выделению мнений, установок, объективных характеристик респондента (пол, возраст, национальность и т.д.). Ранговой шкале (порядковой) соответствует основное число вопросов анкеты или бланков интервью. Варианты ответов в вопросе распределены в строгом порядке убывания или возрастания интенсивности признака. Интервальная шкала более детальна, глубока. Она допускает обстоятельную математическую обработку информации. В социологическом исследовании с ее помощью измеряются те характеристики, которые можно выразить числами: возраст, образование, стаж работы и др. По этой шкале возможно вычисление различных величин.

1. **Сущность и виды группировок. Таблицы и графики: их роль в анализе социологических данных**

Наиболее простой формой обобщения первичной социологической информации является группировка. На этом этапе выделяются существенные признаки или один какой-либо признак (например, пол, возраст, образование), и респондент зачисляется в ту или иную группу в соответствии с выбранным признаком. Когда суммируются ответы респондентов с учетом, например, пола, то осуществляется простая группировка. Точно такую же работу можно проделать, взяв в качестве важнейшего признака уровень образования. Но в данном случае групп будет не две, а как минимум три или четыре.

Выделенные группы можно сопоставить, сравнить, а следовательно, глубже и обстоятельнее проанализировать то или иное социальное явление, мотивы, интересы опрашиваемых. Выбор признака группировки предопределяется задачами социологического исследования, а так же его гипотезами. Ошибка в выборе признака группировки приводит к ошибкам при анализе групп. Исследую один и тот же эмпирический материал, разные исследователи могут сделать диаметрально противоположные выводы.

Социологическая информация может быть сгруппирована по:

- *номинальному* признаку (род занятий, национальность и т.д.);

- признакам, соответствующим *ранговым* шкалам (например, по характеру труда: ручной труд, работа с механизмом, наладка станков, интеллектуальный труд);

- *количественному* признаку (группы характеризуются числовым значением, они качественно сравнимы между собой, например группировка по возрастным интервалам: 18-20 лет, 21-25 лет, 26-30 лет и т.д.).

Работа с номинальными и ранжированными группами ведется с использованием приемов математики, а группы, распределенные по количественному признаку, изучаются с помощью математической статистики.

Если опрашиваемых необходимо сгруппировать по двум и более признакам (например, по полу, возрасту и образованию), тогда речь может идти о *перекрестной*, или *комбинированной* группировке. Она может быть структурной, типологической, аналитической – все зависит от решаемых в ходе исследования задач. Например, надо установить возрастной состав опрашиваемых. В этом случае применяется *структурная* группировка по возрастным интервалам, т.е. респондентов классифицируют по объективному признаку, присущему всей совокупности опрашиваемых. Если же необходимо из респондентов группы по такому, например, признаку, как «отношение к частной собственности», тогда осуществляется *типологическая* группировка (выделяются соответствующие типы респондентов). И, наконец, *аналитическая* группировка производится по двум и более признакам и служит для выявления их взаимосвязи. Если проверить, имеется ли связь между интересом к вопросам нравственного воспитания и чтением литературы (газет, журналов, монографий по данной проблеме), то группировку следует сделать по этим двум признакам.

В социологическом исследовании, как правило, выделяется не одна, а несколько групп респондентов (по возрасту, образованию, месту проживания и т.п.). Каждой группе соответствует некоторое выделенное число (n1, n2, n3, ..., nx), которое характеризует количественный состав группы. Такой ряд чисел, получаемый в результате группировки, социологи называют *рядом распределения*. Существует два вида распределений: вариационный и атрибутивный. *Вариационный* ряд распределения основывается на количественных признаках изучаемых явлений, процессов, *атрибутивный* – отражает результаты группировки опрашиваемых по количественным признакам.

Более глубоко проанализировать социологическую информацию позволяют широко применяемые в эмпирических исследованиях статистические и математические методы анализа получаемой информации. Однако при всем значении получаемых распределений, математических и статистических методов, используемых в исследовании, решающую роль в интерпретации полученных данных играет, прежде всего, сама концепция проводимого исследования, научная эрудиция социолога.

Общая логика интерпретации состоит в превращении статистических данных в показатели, которые выступают уже не в качестве цифровых величин (процент, среднее арифметическое и т.п.), а как социологические данные. Такие показатели – результат интерпретации, несущей определенную смысловую нагрузку.

От возможных крайностей, неумышленных заблуждений исследователя предохраняют ранее выдвинутые гипотезы. Характер проверки гипотез определяется видом исследования.

При *пилотажном* исследовании гипотеза проверяется соотнесением предполагаемого утверждения с выясненной в результате исследования числовой величины. Например, вероятность утверждения о неудовлетворительном состоянии морально-психологического климата в коллективе не вызывает сомнения, если мы располагаем данными полученными в пилотажном исследовании, о том, что 50% опрошенных идут на работу с полным безразличием, а 12% - ожидая неприятности.

Что касается описательного (а тем более аналитического) исследования, то здесь процедура проверки гипотез значительно усложняется. Так, приведенные выше данные (о состоянии морально-психологического климата в коллективе) сами по себе не дают информации о том, кто эти люди, так как в инструментарии отсутствуют конкретные социально-демографические данные. Следовательно, средние величины – лишь первый шаг на пути исследования. Надо сделать второй, и третьи шаги, чтобы подойти как можно ближе к истине. Для этого из всей совокупности следует выделить однородные по социально-демографическим характеристика подгруппы.

Если необходимо превратить в показатель какую-либо среднюю величину, а сравнить ее с другими величинами затруднительно или вообще не представляется возможным (например, из-за новизны), то эталоном оценки выступают знания исследователя или эксперта по данной проблеме. Допустим, коллектив переведен на новые условия оплаты и стимулирования труда. После года его функционирования проведен опрос, задача которого – оценить ответ на главный вопрос инструментария: удовлетворены ли опрашиваемые новой формой оплаты труда. При этом 57% опрошенных ответили положительно. Оценить этот результат (показатель) можно с оптимистической или пессимистической точки зрения. Для получения объективной оценки нужно хорошо знать проблему и конкретные условия, в которых проходит эксперимент. Это делает или сам исследователь, или приглашенный эксперт.

Другой способ превращения описательного исследования в показатель – сравнение рядов распределения по относительно однородным группам из обследуемой совокупности с помощью внутреннего и внешнего соотнесения. *Внутреннее соотнесение* – это сравнение между собой элементов числового ряда, *внешнее* – сравнение двух или более рядов распределения, построенным по двум или более признакам, из которых один – общий для соотносимых рядов. Например, распределение двух разных групп – работающих в новых условиях и по прежним формам оплаты – можно сравнить по уже упомянутому признаку: с каким чувством они ежедневно идут на работу.

Внутреннее соотнесение позволит однозначно оценить результаты группировки в тех случаях, когда в числовом ряде четко видна наибольшая (модальная) величина. Соотнесение элементов числового ряда в подобном положении заключается в их ранжировании. При ответе на вопрос: «Как вы относитесь к своей работе?» 58% респондентов выбрали альтернативу (ответ): «стремлюсь отдать работе все силы, знания», 37% высказались: «выполняю все, что от меня требуется, но не более», 5% ответили: «как правило, работаю без желания, по необходимости». Из этих ответов видно, как выстроится ранг опрошенных.

Если внутреннее сравнение затруднено, то применяют внешнее сравнение числового ряда.

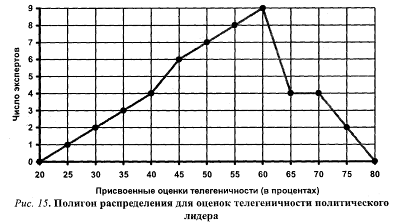
Для большей наглядности и читабельности данных прибегают к табличному или графическому отображению данных. Стремление к компактности и «читабельности» данных не должно вести к крайностям. Руководствуясь соображениями здравого смысла, исследователь должен избегать ситуаций, когда перегруппировка ведет к тому, что полученная переменная оказывается слишком грубым средством классификации наблюдений, не позволяющим выявить существенные для анализа различия. Важно также следить за тем, чтобы объединение категорий или числовых градаций переменной-признака не привело к искусственному созданию отношений и взаимосвязей, которые в действительности отсутствуют в данных.

Помимо табличного представления частотных распределений обычно используют и различные методы *графического представления.* Самый распространенный метод графического представления одномерных распределений - это *гистограмма,* или столбиковая диаграмма. Каждый столбик соответствует интервалу значений переменной, причем его середина совмещается с серединой данного интервала. Высота столбика отражает частоту (абсолютную или относительную) попадания наблюдавшихся значений переменной в определенный интервал. При построении гистограмм часто приходится использовать некоторые конвенции, основанные на сугубо практических соображениях. Так, используя при группировке значений переменной неравные интервалы либо оставляя крайние градации открытыми («старше 65 лет», «свыше 24000 рублей» и т. д.), мы все же отображаем эти интервалы на гистограмме с помощью столбиков, имеющих одинаковую ширину. Другое практическое правило позволяет сделать гистограмму визуально уравновешенной, т. е. более привлекательной: масштаб шкалы обычно выбирают так, чтобы общая высота гистограммы составляла приблизительно 40-60% ее ширины. Пример гистограммы приведен на рис. 1.

социологический исследование группировка



Если просто соединить между собой точки, соответствующие абсолютным или относительным частотам (ось ординат) для середин интервалов, мы получим так называемый *полигон распределения.* Эта операция, разумеется, будет иметь какой-то смысл лишь для количественных переменных, которые мы в принципе можем представить себе как непрерывные. На *рисунке 2* изображен полигон распределения для экспертных оценок телегеничности политического лидера (50 экспертов оценивали политика в процентах по отношению к некоторому абсолютному эталону телегеничности).



Еще один популярный способ графического представления, обычно используемый для качественных данных (т. е. для номинальных или ординальных измерений),- это *круговая диаграмма.* Каждый сектор круговой диаграммы представляет дискретную категорию переменной. Величина сектора пропорциональна частоте категории для данной выборки. На *рисунке 3* приведена круговая диаграмма, иллюстрирующая распределение подростков, страдающих вялотекущей формой шизофрении, по возрасту на момент начала («дебюта») заболевания.



1. **Структура и основные требования отчета о социологическом исследовании**

Итоги анализа полученной информации отражаются, как правило, в отчете о проведенном социологическом исследовании, который содержит в себе информацию, интересующую заказчика (исследователя), научные выводы, рекомендации. Структура отчета по итогам исследования чаще всего соответствует логике операционализации основных понятий, но социолог, готовя этот документ, идет путем индукции, постепенно сводя социологические данные в показатели. Число разделов в отчете обычно соответствует числу гипотез, сформулированных в программе исследования. Первоначально дается ответ на главную гипотезу.

Как правило п е р в ы й раздел отчета содержит краткое обоснование актуальности изучаемой социальной проблемы, характеристику параметров исследования (выборка, методы сбора информации, количество участников исследования, сроки проведения работы и т.п.). Во в т о р о м разделе дается характеристика объекта исследования по социально-демографическим признакам (пол, возраст, образование и др.) Последующие разделы включают поиск ответов на выдвинутые в программе гипотезы.

Разделы (или главы) отчета при необходимости могут быть разбиты на параграфы. Каждый раздел или даже параграф целесообразно завершать выводами. Заключение отчета лучше давать в виде практических рекомендаций, базирующихся на общих выводах. Отчет может быть изложен на двух-трех сотнях страниц. Это зависит от объема материала, целей и задач исследования.

Приложение к отчету содержит все методологические и методические документы исследования: программу, план, инструментарий, инструкции и т.п. Кроме того, в приложение чаще всего вносят таблицы, графики, индивидуальные мнения, ответы на открытые вопросы, которые не вошли в отчет. Это необходимо делать потому, что данные документы, ответы могут быть использованы при подготовке программы нового исследования.

**Заключение**

Социологию называют наукой о социально типичных явлениях. Это единственная наука, которая точно знает о том, что думает и чего хочет среднестатистический человек. Действительно, при помощи количественных распределенных ответов в анкете и последующего анализа, социология выявляет типичное мнение большой группы людей. В данной контрольной работе были рассмотрены понятие и основы анализа эмпирических данных.

Методы, применяемые социологами для анализа данных, многообразны. Выбор конкретного метода зависит, в первую очередь, от характера исследовательских гипотез, т. е. от того, на какие вопросы мы хотим получить ответ. Если целью является описание одной характеристики выборки в определенный момент времени, разумно ограничиться *одномерным анализом,* т. е. описанием распределения наблюдений («случаев»). Разнообразные техники *многомерного анализа* позволяют одновременно исследовать взаимоотношения двух и более переменных и в той или иной форме проверять гипотезы о причинных связях между ними. В реальном исследовании каждое уточнение исходных гипотез или выдвижение новой гипотезы в ходе анализа результатов приводит к необходимости выбора новой техники анализа данных. Так, если изначальная модель взаимоотношения двух переменных (скажем, профессии и дохода) не позволяет выявить определенную закономерность в собранных данных, исследователь выбирает одну из статистических техник, позволяющих контролировать влияние какой-то третьей переменной, например пола, на интересующее его отношение.

Помимо характера исследовательских гипотез на выбор методов статистического анализа влияет и природа полученных социологом данных. Для того чтобы охарактеризовать распределение в выборке такого номинального признака, как «пол», мы не можем воспользоваться его среднеарифметическим значением и, следовательно, нам потребуются какие-то другие приемы компактного и точного представления полученной информации.

**Список используемой литературы**

1. Кравченко А.И., социология: учеб. – М.: ТК Велби, Изд-во Проспект, 2005.

2. Девятко И.Ф. Методы социологического исследования.- Екатеринбург: Изд-во Урал, ун-та, 1998.

3. Социология: Учебник для вузов/ под. Ред. Проф. Лавриненко В.Н. – М.: Культура и спорт, ЮНИТИ, 1998.

4. Фролов С.С. Социология: Учебник для вузов. – М., 2007.