

ОСНОВЫ СТАТИСТИКИ

Естественное состояние: Гоббс





Во второй лекции нашего курса мы поговорим о роли статистики в научных исследованиях. Дадим ответы на следующие вопросы: Какую роль могут играть числа и статистические показатели в жизни людей и общества? Дадим определения основным понятиям статистики и изучим достаточно важный материал, это классификация измерительных шкал.

Отвечив на эти вопросы, мы сможем понять, что в общественных науках статистика является необходимым инструментом. Она дает ученому некоторые наиболее полезные методы оценки идей, проверки теоретических предположений и поиска истины. Более подробно рассмотрим взаимосвязи, существующие между теорией, исследованиями и статистикой. Какова же роль статистики в научном исследовании?

На рис.2.1 графически показана роль статистики в исследовательском процессе. Диаграмма классическая и основана на суждениях Уолтера Уоллеса. Она демонстрирует, каким образом база знаний любой научной области растет и развивается.

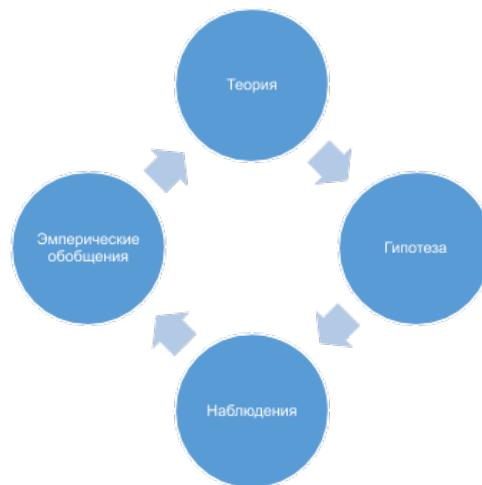


Рисунок 2.1 Роль статистики в исследовательском процессе.

Одна из основных идей диаграммы заключается в том, что научная теория и исследования постоянно влияют друг на друга. Статистика является одним из основных средств коммуникации между исследованиями и теорией. Давайте рассмотрим это колесо более детально.

График является круговым и не имеет ни начала, ни конца, поэтому мы можем начать свое обсуждение с любой точки. Для удобства начнем с вершины и затем будем следовать по стрелкам.

Теория представляет собой объяснение связей между явлениями. Людей вечно интересуют проблемы, существующие в обществе (такие как, например, предубеждения, предрассудки, бедность, жестокое обращение с детьми или серийные убийства), и чтобы попытаться понять эти явления, они дают соответствующие объяснения (недостаток образования приводит к предубеждениям). Безусловно, вам знакомо такого рода неформальное теоретизирование относительно общества. Основное различие между нашими неформальными повседневными объяснениями социальных явлений и научной теорией заключается в том, что последняя подвергается тщательному процессу тестирования. В качестве примера давайте возьмем классический пример социологического исследования, это изучение расовых предубеждений и продемонстрируем, как работает исследовательский процесс.

Гипотеза

Что является причиной расового предубеждения? Возможное объяснение этого явления дается теорией, называющейся гипотезой контакта. Эту теорию более 40 лет назад сформулировал социальный психолог Гордон Аллпорт. И с тех пор она проверялась на множестве реальных проявлений. Теория связывает предубеждение с количеством контактов между членами различных расовых групп. В частности, данная гипотеза утверждает, что контактные ситуации, при которых различные группы имеют одинаковый статус и их поведение основано на принципах сотрудничества, ведут к снижению предубеждения с обеих сторон. Чем выше степень равенства и сотрудничества, тем более вероятно то, что человек будет видеть в другом



личность, а не представителя определенной расовой группы. Например, среди членов спортивной команды со смешанным расовым составом, сотрудничающих друг с другом в целях достижения победы, будет наблюдаться тенденция к снижению уровня предубеждения. С другой стороны, когда различные группы борются между собой за рабочие места, жилье и прочие жизненно важные ресурсы, будет наблюдаться тенденция к росту уровня предубеждения.

Гипотеза контакта не является исчерпывающим объяснением предубеждения, но она может послужить примером для иллюстрации роли теории. Мы пытаемся объяснить связь между двумя социальными явлениями: (1) предубеждение; (2) контакт, основанный на равенстве членов различных групп и их сотрудничестве между собой. Те люди, которые имеют меньше контактов между собой, будут испытывать больший уровень предубеждения друг к другу, и наоборот. Таким образом, мы начинаем с логически последовательной теории, объясняющей связь между двумя явлениями.

Прежде чем перейти к следующему пункту рисунка 2.1, давайте немного остановимся на детальном рассмотрении теории. Гипотеза контакта, как и большинство теорий, сформулирована в терминах причинно-следственных зависимостей между переменными. Переменная – это свойство, которое может изменять значение от наблюдения к наблюдению. Примерами переменных могут служить пол, возраст, доход или членство в политической партии. В каждой отдельной теории некоторые переменные будут рассматриваться как причины, другие – как следствия или результаты. На научном языке причины называются независимыми переменными, а переменные следствия или результата называются зависимыми переменными. В нашей теории контакт является независимой переменной (или причиной), а предубеждение является зависимой переменной (результатом или следствием). Другими словами, мы утверждаем, что контакт, основанный на равенстве, является причиной предубеждения, или что уровень предубеждения человека зависит от степени его участия в контакте с другими группами, основанном на равенстве и сотрудничестве.

К настоящему моменту у нас есть независимая и зависимая переменные, а также теория предубеждения, объясняющая связь между ними. Чего мы пока не знаем, так это того, истинна ли наша теория. Чтобы выяснить это, нам необходимо сравнить свою теорию с фактами – нужно провести исследование. Последующие этапы процесса будут заключаться в том, чтобы дать более конкретные и точные определения терминов и идей. Одна из проблем, с которой мы часто сталкиваемся при проведении исследований, заключается в том, что научные теории слишком сложны и абстрактны, чтобы их можно было полностью проверить в пределах одного исследовательского проекта. Чтобы провести исследование, из теории следует вывести одну или более гипотез. Гипотеза – это утверждение относительно связи между переменными, которое, хотя и логически выводится из теории, однако является гораздо более конкретным и точным.

Например, если бы мы хотели проверить гипотезу контакта, мы должны были бы точно сказать, что мы подразумеваем под предубеждением, и нам бы пришлось очень подробно объяснить, что такое контакт, основанный на равенстве и сотрудничестве. В отношении вопроса влияния контакта на предубеждение было проведено множество исследований, и мы можем обратиться к исследовательской литературе в целях разработки своих определений для этих понятий. По мере того, как наши определения развиваются, а гипотезы принимают конкретную форму, мы переходим на следующий уровень исследовательского процесса.

Теперь мы рассмотрим этап сбора данных.

Наблюдения

Нам нужно решить, каким образом можно проверить доводы, как их отбирать, как именно будут измеряться переменные и массу подобных вопросов. В конечном счете, эти шаги приведут к фазе наблюдений (нижняя часть колеса науки), где мы уже проводим реальное измерение социальной действительности. Прежде чем это сделать, мы должны иметь очень четкое представление относительно того, что мы ищем, и хорошо продуманную стратегию проведения поиска.

Чтобы проверить, гипотезу контакта, мы можем начать с людей, относящихся к различным социальным группам. Мы могли бы поместить некоторых субъектов в ситуации, которые бы потребовали от них сотрудничества с членами других групп, а других субъектов – в ситуации, в которых происходит конкуренция между группами. Нам следовало бы измерить уровень предубеждения до и после контакта каждого вида. Это можно было бы сделать путем проведения опроса, при котором субъектов бы попросили



согласиться или не согласиться со следующими утверждениями: «Следует прикладывать больше усилий для социальной интеграции в системе средних школ» или «не имеет значения, так как люди остаются людьми независимо от социальной принадлежности». Наша цель заключалась бы в том, чтобы увидеть, действительно ли у людей, оказавшихся в ситуации, включающей контакт с сотрудничеством, произошло снижение уровня предубеждения.

И вот наконец-то мы добрались до «статистики».

Эмпирические обобщения

По мере того, как этап наблюдений будет подходить к концу, мы столкнемся с большим количеством численной информации или данных. Если бы наша выборка состояла из 100 человек, мы бы получили 200 заполненных анкет для измерения предубеждения: 100 анкет, заполненных до контактной ситуации, и 100 анкет, заполненных после контактной ситуации. Постарайтесь представить себе работу с 200 заполненными анкетами. Если бы мы задавали каждому респонденту всего пять вопросов с целью измерить его уровень предубеждения, мы бы в итоге получили 1000 отдельных порций информации. Что делать? Нам нужно иметь какой-нибудь способ систематической организации и анализа этой информации, и тогда статистика станет очень ценной. Статистика даст нам множество идей о том, что делать с данными. Мы перейдем к рассмотрению некоторых возможностей в следующей главе. А пока что, позвольте мне высказать две мысли насчет статистики.

Во-первых, статистика является ключевым элементом. Проще говоря, без статистики невозможно провести исследование количественных характеристик. Без проведения исследований количественных характеристик развитие общественных наук понесло бы серьезный урон. Только путем применения статистических методов обычные данные могут помочь нам формировать и совершенствовать свои теории и лучше понимать социальный мир.

Во-вторых, как это ни парадоксально, роль статистики ограничена. Как показывает рисунок 2.1, научное исследование проходит несколько взаимно независимых этапов, и статистика нужна лишь в конце этапа наблюдений. Прежде чем применить какой-либо метод статистического анализа, должны быть успешно завершены предыдущие этапы. Если исследователь задал вопросы, которые были плохо составлены, или допустил серьезные ошибки в разработке плана или в выборе метода исследования, тогда даже самый совершенный статистический анализ будет бесполезен. Ничуть не умаляя ценности статистики, можно сказать, что она все же не может заменить точного определения понятий, подробного и тщательного планирования или творческого применения теории. Статистика не может спасти плохо составленный или проработанный проект исследования. Она не может извлечь пользу из мусора.

С другой стороны, если применение статистики осуществляется некорректно, это может ограничить полезность проекта, который во всем остальном был выполнен очень точно. Проект исследования количественных характеристик может внести свой вклад в понимание вопроса только путем успешного выполнения всех этапов процесса. Правильное понимание полезности и ограниченности статистики является необходимым для образования социолога наряду с изучением теории и методологии.

По мере того, как объект для статистического анализа исчерпывается, мы можем браться за эмпирические обобщения. Хотя мы в первую очередь будем сфокусированы на оценке своей теории, также можно изучать и другие тенденции на основании полученных данных. Предположим, мы обнаружили, что контакт, основанный на равенстве и сотрудничестве, снижает предубеждение в целом. Далее мы можем спросить, относится ли это в равной степени: и к мужчинам, и к женщинам; и к хорошо образованным, и к недостаточно образованным; и к пожилым респондентам, и к молодым. По мере изучения данных мы можем приниматься за рассмотрение некоторых обобщений, основанных на наблюдаемых нами эмпирических моделях. Например, если мы обнаружим, что контакт привел к уменьшению предубеждения молодых респондентов, но не оказал внимания на предубеждение пожилых респондентов? Может ли это означать, что молодые люди, менее устоявшиеся в своих взглядах и что их отношения и чувства более подвержены изменениям? После разработки предварительных объяснений мы можем начать исправлять и развивать свою теорию.

Давайте вкратце повторно рассмотрим наш воображаемый исследовательский проект. Мы начали с идеи или теории о межгрупповом контакте и социальном предубеждении. Мы представили себе некоторые шаги, которые нам придется сделать, чтобы проверить теорию, и вкратце рассмотрели разнообразные этапы исследовательского проекта. Мы вновь оказались на уровне теории, готовые к



тому, чтобы начать новый проект, инициированный измененной теорией. Мы видели, как теория может мотивировать исследовательский проект и как наши наблюдения могут привести нас к пересмотру теории и, таким образом, мотивировать новый исследовательский проект. Колесо науки, изобретенное Уоллесом, иллюстрирует то, как теория стимулирует исследования, и то, как исследование формирует теорию. Это постоянное взаимодействие между теорией и исследованиями представляет для науки источник жизненной силы и является ключом к пониманию социального мира.

Диалог между теорией и исследованиями возникает на многих уровнях и во множестве форм. Статистика – одна из наиболее важных связей между этими двумя сферами.

Статистика дает нам возможность анализировать данные, выявлять и изучать тенденции и зависимости, разрабатывать обобщения, пересматривать и вносить изменения в теории.

Как вы увидите на протяжении этого материала, статистика ограничена во многих отношениях. Однако она является неотъемлемой частью научного или прикладного исследования. Без статистики взаимодействие между теорией и исследованиями стало бы исключительно сложным, и прогресс наших дисциплин существенно затормозился бы.

Перейдем непосредственно к предмету изучения, т.е. к описательной статистике.

О назначении описательной статистики можно судить по ее названию: она имеет дело с числами, характеризующими ту или иную интересующую нас ситуацию. Вот примеры статистической информации:

- уровень преступности в регионе;
- средняя зарплата в различных отраслях региона;
- уровень безработицы;
- число несчастных случаев на шахтах;
- число мобильных телефонов, проданных в текущем месяце;
- таблицы продолжительности жизни;
- уровень заболеваемости;
- уровень достижений учащихся по дисциплинам;
- число граждан СНГ, обучающихся в AlmaU.

Ценность описательной статистики заключается прежде всего в том, что она дает сжатую и концентрированную характеристику изучаемого явления. Рассмотрим следующий пример. Пусть на некотором предприятии работает 1500 человек. Бухгалтерская ведомость на зарплату довольно большая. Информация о том, что средняя месячная зарплата работников этого предприятия составляет 82000 тенге, дает определенное, хотя и неполное представление об уровне заработной платы на этом предприятии.

Предмет исследований во многих сферах отличается исключительной сложностью, изменчивостью, индивидуальным многообразием явлений и процессов. Эти процессы происходят неоднозначно. Поэтому применение одинаковых подходов, средств, технологий дает в каждом конкретном случае различные результаты в зависимости от субъективных факторов, от обстоятельств, которые нельзя контролировать и которые влияют на протекание процесса. Неоднозначность протекания процесса порождается наличием присущего ему случайного. Но это не означает отсутствие общих закономерностей в изучаемых процессах и явлениях. Например, невысокая скорость чтения у отдельного учащегося является случайным событием, но у ученика, любящего читать, она встречается существенно реже, чем у того, кто редко берет книгу в руки. Эта устойчивость появления тех или других случайных событий уже является закономерностью.

Разработкой методов изучения свойств случайных событий и явлений занимается статистика. Статистика имеет различные функции: информационную, прогностическую и аналитическую.

Информационная функция статистики состоит из сбора, обобщения и представления всем заинтересованным лицам достоверной, своевременной информации об исследуемом явлении. В связи с тем, что иногда исследованию подлежат тысячи объектов, необходимым является переход от сплошного изучения к выборочному по многим показателям. Поэтому важное значение приобретают технологии сбора, обработки и анализа данных, которые позволяют использовать информационные возможности частичных первоначальных данных для разработки обобщенной информации о том или ином процессе.

Прогностическая функция статистики состоит в оценивании вероятностей тех или иных случайных событий, которые происходят в изучаемом процессе, показателей тех или иных случайных величин, связанных с этим процессом. Эта функция служит основой для принятия управленческих решений. С



помощью этой функции можно получить сигнал о возможности появления кризисных явлений в изучаемом процессе, если не внести каких-то изменений в управление им.

Аналитическая функция статистики состоит, во-первых, в количественном исследовании тенденций развития процесса; во-вторых, в изучении этого процесса в динамике; в-третьи, в измерении связей между разными факторами, влияющими на процесс, и его результатами.

Независимо от того, какой сложности анкетные вопросы или же тестовые методики вы рассматриваете, все их можно разделить на три типа в зависимости от того, к какой измерительной шкале они относятся. Речь в данном случае идет не о специфических методиках построения измерительных инструментов (например, шкала Гутмана или шкала Терстоуна), а о классификации измерительных шкал, предложенной Стэнли Стивенсом в 1946 году. Знание этой классификации имеет решающее значение с точки зрения использования количественного подхода, поскольку применение тех или иных методов математической статистики опирается, в том числе, и на измерительные шкалы, в которой отображены интересующие исследователя переменные.

Поговорим более подробно о понятии «переменная».

Согласно классификации Стивенсона, в самом общем виде, можно выделить три типа шкал:

- номинальную,
- порядковую,
- метрическую.

Номинальная шкала включает в себя класс переменных, значения которых можно разделить на группы, но невозможно проранжировать. Примерами соответствующих переменных являются пол, национальность, религия и т.д. Рассмотрим более подробно такую переменную как национальность. В данном случае респондентов можно разделить на разные группы в зависимости от того, к какой национальности они себя относят. Вместе с тем, на основе этой информации, респондентов невозможно упорядочить в смысле количественной выраженности интересующего нас параметра, ведь национальность не является измеряемым, в традиционном значении этого слова, свойством. Рассмотрим примеры по вышеуказанной классификации.

1. Распределения учащихся по классам, по половому признаку, по месту жительства, по видам спорта, которыми они занимаются, по числу детей в семье являются примерами величин номинальной шкалы. При этом возможно распределение учащихся по двум или более признакам (двумерные или многомерные данные).
2. Перечень фирм, занимающихся производством грузовых и легковых автомобилей, автомобилей специального назначения, автобусов; отличительные признаки автомобилей являются примерами величин номинальной шкалы.

С помощью подсчета можно установить частоту той или иной категории (число мальчиков и девочек в школе; число учащихся, проживающих в каждом микрорайоне; число учащихся в каждом классе; число учащихся, занимающихся тем или иным видом спорта; количество фирм, занимающихся производством автобусов и т. д.). При этом можно определить наиболее часто встречающуюся величину (класс, в котором учится наибольшее число учащихся; вид спорта, пользующийся наибольшей популярностью у учащихся; тип автомобиля, производством которого занимается наибольшее число фирм). Категории данных номинальной шкалы обозначаются, как правило, словесно (вербально).

Порядковая шкала включает в себя класс переменных, значения которых можно не только разделить на группы, но и проранжировать в зависимости от выраженности измеряемого свойства. Например, учащихся можно ранжировать по количеству правильно выполненных тестовых заданий. Пусть учащиеся А, Б, В, Г, Д правильно выполнили соответственно 21, 16, 12, 9 и 3 задания. Графически это можно изобразить так:

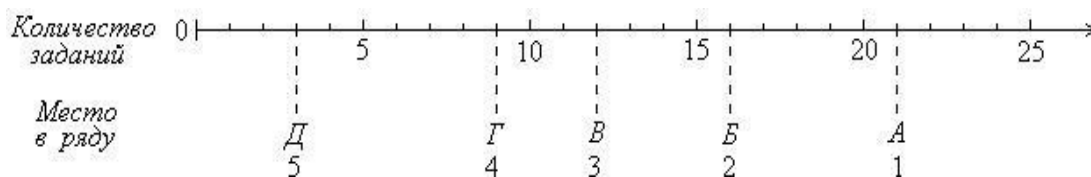


Рисунок 2.2 Порядковая шкала



Эта порядковая шкала имеет величины от 1 до 5, и учащиеся на ней размещены в зависимости от количества правильно выполненных заданий: А - первый, Д - пятый. Из рисунка видно, что интервалы, разделяющие места в ряду, различны по величине. По этой причине нецелесообразно складывать, вычитать, умножать и делить порядковые места.

Метрическая шкала включает в себя класс переменных, значения которых можно как разделить на группы и проранжировать, так и определить их величину в точных терминах (те самые «на сколько?» и «во сколько?»). Типичными примерами соответствующих переменных являются возраст, заработная плата, количество детей и т.д. Измерение каждой из них можно осуществить максимально точно: возраст в годах, заработную плату в тенге, количество детей в штуках). Естественно, если переменная может быть потенциально выражена в метрической шкале, то эту же переменную можно выразить и в порядковой. Например, возраст можно выразить в возрастных группах (молодежь, средний возраст, пожилой возраст), которые дают лишь приблизительную информацию о респонденте, несмотря на возможность их ранжирования. Принадлежность переменной к метрической шкале открывает возможность использования любых статистических методов. В свою очередь принадлежность к порядковой или номинальной ограничивает выбор математических инструментов (в случае порядковой шкалы в меньшей мере, а в случае номинальной - в большей).

Для того, чтобы сделать различия между номинальной, порядковой и метрической шкалами еще более очевидными, приведу дополнительный пример:

В эмпирическом исследовании могут встречаться, к примеру, следующие переменные (указано их наиболее вероятное кодирование):

Таблица 2.1 Эмпирическое исследование

Пол	1 = мужской
	2 = женский
Семейное положение	1 = холост/не замужем
	2 = женат/замужем
	3 = вдовец/вдова
	4 = разведен(а)
Курение	1 = некурящий
	2 = изредка курящий
	3 = интенсивно курящий
	4 = очень интенсивно курящий
Месячный доход	1 = до 150000 тенге
	2 = 150001 - 200000 тенге
	3 = более 200000 тенге
Коэффициент интеллекта (I.Q.)	
Возраст (лет)	

Рассмотрим сначала графу «Пол». Мы видим, что назначение соответствия цифр 1 и 2 обоим полам абсолютно произвольно, их можно было поменять местами или обозначить другими цифрами. Мы, конечно, не имеем в виду, что женщины стоят на ступеньку ниже мужчин, или мужчины значат меньше, чем женщины. Следовательно, отдельным числам не соответствует никакое эмпирическое значения. В этом случае говорят о переменных, относящихся к номинальной шкале. В нашем примере рассматривается переменная с номинальной шкалой, имеющая две категории. Такая переменная имеет еще одно название – дихотомическая.

Такая же ситуация и с переменной «Семейное положение». Здесь также соответствие - между числами и категориями семейного положения не имеет никакого эмпирического значения. Но в отличие от Пола, эта переменная не является дихотомической – у нее четыре категории вместо двух. Возможности обработки переменных, относящихся к номинальной шкале очень ограничены. Собственно говоря, можно провести только частотный анализ таких переменных. К примеру, расчет среднего значения для переменной



Семейное положение, совершенно бессмысленно. Переменные, относящиеся к номинальной шкале часто используются для группировки, с помощью которых совокупная выборка разбивается по категориям этих переменных. В частичных выборках проводятся одинаковые статистические тесты, результаты которых затем сравниваются друг с другом.

В качестве следующего примера рассмотрим переменную «Курение». Здесь кодовым цифрам присваивается эмпирическое значение в том порядке, в котором они расположены в списке. Переменная Курение, в итоге, сортирована в порядке значимости снизу вверх: умеренный курильщик курит больше, нежели некурящий, а сильно курящий – больше, чем умеренный курильщик и т.д. Такие переменные, для которых используются численные значения, соответствующие постепенному изменению эмпирической значимости, относятся к порядковой шкале.

Однако эмпирическая значимость этих переменных не зависит от разницы между соседними численными значениями. Так, несмотря на то, что разница между значениями кодовых чисел для некурящего и изредка курящего и изредка курящего и интенсивно курящего в обоих случаях равна единице, нельзя утверждать, что фактическое различие между некурящим и изредка курящим и между изредка курящим и интенсивно курящим одинаково. Для этого данные понятия слишком расплывчаты.

К классическими примерами переменных с порядковой шкалой относятся также переменные, полученные в результате объединения величин в классы, как «Месячный доход» в нашем примере.

Кроме частотного анализа, переменные с порядковой шкалой допускают также вычисление определенных статистических характеристик, таких как медианы. В некоторых случаях возможно вычисление среднего значения. Если должна быть установлена связь (корреляция) с другими переменными такого рода, для этой цели можно использовать коэффициент ранговой корреляции.

Для сравнения различных выборок переменных, относящихся к порядковой шкале, могут применяться непараметрические тесты, формулы которых оперируют рангами.

Рассмотрим теперь «Коэффициент интеллекта (IQ)». Не только его абсолютные значения отображают порядковое отношение между респондентами, но и разница между двумя значениями также имеет эмпирическую значимость. Например, если у Армана IQ равен 80, у Александра – 120 и у Куаныша – 160, можно сказать, что Александр в сравнении с Арманом настолько же интеллектуальнее насколько Куаныш в сравнении с Александром (а именно – на 40 единиц IQ). Однако, основываясь только на том, что значение IQ у Армана в два раза меньше, чем у Куаныша, исходя из определения IQ нельзя сделать вывод, что Куаныш вдвое умнее Александра.

Такие переменные, у которых разность (интервал) между двумя значениями имеет эмпирическую значимость, относятся к интервальной шкале. Они могут обрабатываться любыми статистическим методами без ограничений. Так, к примеру, среднее значение является полноценным статистическим показателем для характеристики таких переменных.

Наконец, мы достигли наивысшей статистической шкалы, на которой эмпирическую значимость приобретает и отношение двух значений. Примером переменной, относящейся к такой шкале является «Возраст»: если Макс 30 лет, а Тимуру 60, можно сказать, что Тимур вдвое старше Макса. Шкала, к которой относятся данные - называется шкалой отношений. К этой шкале относятся все интервальные переменные, которые имеют абсолютную нулевую точку. Поэтому переменные, относящиеся к интервальной шкале, как правило, имеют и шкалу отношений.

Подводя итоги, можно сказать, что существует четыре вида статистических шкал, на которых могут сравниваться численные значения:

Таблица 2.2 Виды статистических шкал

Статистическая шкала	Эмпирическая значимость	Примеры
Номинальная	Нет	Пол, семейное положение
Порядковая	Порядок чисел	Курение, месячный доход
Интервальная	Разность чисел	Коэффициент интеллекта (I.Q.)
Шкала отношений	Отношение чисел	Возраст (лет)

На практике, в том числе в SPSS, различие между переменными, относящимися к интервальной шкале и шкале отношений обычно несущественно. То есть в дальнейшем практически всегда речь будет идти о переменных, относящихся к интервальной шкале.



Пользователь SPSS должен четко разбираться в видах статистических шкал и при выборе метода обращать внимание на то, чтобы были определены надлежащие виды шкал.

Мы уже указывали, что переменные, относящиеся к номинальной шкале допускают весьма ограниченные возможности для проведения анализа. Исключение в некоторых ситуациях составляют дихотомические переменные. Для них можно, по крайней мере, определять ранговую корреляцию. Если, например, обнаруживается корреляция коэффициента интеллекта с полом, то положительный коэффициент корреляции означает, что женщины интеллектуальнее, чем мужчины. Однако если переменные, относящиеся к номинальной шкале не являются дихотомическими, вычисление коэффициентов ранговой корреляции не имеет смысла.

Итак, подведем итоги по всему сказанному:

Во-первых, цель статистики состоит в том, чтобы организовать, управлять и проанализировать данные так, чтобы исследователи могли проверить теории и ответить на вопросы. Статистические данные - основной инструмент социологии.

Во-вторых, статистика имеет различные функции: информационную, прогностическую и аналитическую.

В-третьих, переменные могут быть измерены на любом из трех разных уровней. На номинальном уровне мы можем сравнить размеры категории. На порядковом уровне очки могут быть оценены от высоко до низко. На уровне отношения интервала разрешены все математические операции.