



Статистические величины

# СОЦИАЛЬНАЯ СТАТИСТИКА

# Социальная статистика - это

- **отрасль (раздел) статистики, изучающая количественно-качественные характеристики массовых социальных явлений и процессов.**
- **Статистика есть совокупность методов и принципов, согласно которым проводится сбор, анализ, сравнение, представление и интерпретация числовых данных.**

# История

- Исторически возникла под давлением практических потребностей людей задолго до н.э. в различных цивилизациях Древнего мира (Египет, Шумер, сбор налогов).
- Первая опубликованная статистическая информация появляется уже в «Книге чисел» в Ветхом Завете, в которой рассказано о переписи военнообязанных, проведённой под руководством Моисея и Аарона.

# Статистика – это...

В 1954 г. академик АН УССР Б. В. Гнеденко дал следующее определение: «Статистика состоит из трёх разделов:

- сбор статистических сведений, то есть сведений, характеризующих отдельные единицы каких-либо массовых совокупностей;
- статистическое исследование полученных данных, заключающееся в выяснении тех закономерностей, которые могут быть установлены на основе данных массового наблюдения;
- разработка приёмов статистического наблюдения и анализа статистических данных.

# Социальная статистика

- Как наука сформировалась в 19 в. и получила наибольшее распространение в 20 в. Ориентирована на выявление и изучение предпосылок, закономерностей и последствий общественной жизнедеятельности.

# Что изучает?

- Изучает систему показателей, характеризующих образ и уровень жизни людей, социальную структуру и устройство общества, различные аспекты социального взаимодействия, численность и состав населения, трудовые ресурсы, уровень образования и культуры, здравоохранения, социального обеспечения и потребления, условия труда и отдыха.

# Как изучает?

- Широко применяет разнообразные математические и статистические методы, вычислительную технику и т.д. Материалы учета и исследований используются в социальной философии, социологии, политологии, др. науках и дисциплинах, в практической организации общественной жизнедеятельности.

# Шкалы

- Шкала (лат. *scala* — лестница) — сопоставление результатов измерения какой-либо величины и точек числовой прямой.
- Шкала - это множество обозначений, отношения между которыми отражают отношения между объектами эмпирической системы.



# Виды шкал

- Шкалой можно назвать результаты измерения, полученные в исследовании, а также инструмент измерения (т.е. систему вопросов), опросник, тест). Виды шкал:
  - Номинальная
  - Ранговая
  - Порядковая

# Номинальная шкала

- Шкала наименований, строится на отношении тождества. Каждое деление на шкале характеризует критерий, на основании которого производится классификация. С помощью этой шкалы мы приписываем числа, буквы или цифры исходным данным. Примером шкалы такого рода может служить классификация испытуемых на мужчин и женщин, нумерация игроков спортивных команд и др. Номера телефонов, паспортов, штрих-коды товаров, индивидуальные номера налогоплательщиков измерены в шкале наименований.

# Цифра (знак) Наименование (ИМЯ)

- Какая из следующих проблем ощущается Вами наиболее остро?(выберите не более трех самых важных)
- 1Невыплата зарплаты
- 2 Цены на товары и услуги
- 3 Невыплата пенсий
- 4 Работа городского транспорта
- 5 Безработица
- 6 Плата за жилищно-коммунальные услуги
- 7 Низкий уровень заработной платы
- 8Нет никаких проблем.

# Порядковая шкала (или ранговая)

- Строится на отношении тождества и порядка. Субъекты в данной шкале ранжированы. Например, нельзя сказать, что больше круг или треугольник, но можно выделить в этих объектах общее свойство-площадь, и таким образом становится легче установить порядковые отношения.

# Порядковая шкала

- Порядковая шкала предполагает упорядочение объектов относительно какого-либо критерия или свойства. Эта шкала определяется двумя эмпирическими операциями: 1) установлением равенства объектов по отношению какого-либо конкретного значения шкалы и 2) установлением отношения «больше – меньше» между объектами.

# Пример порядковой шкалы

Оцените свой жизненный уровень:

- 1 – живу в полном достатке, не испытываю никаких проблем;
- 2 – живу в достатке, но без излишеств;
- 3 – живу средне;
- 4 – живу плохо, собственных доходов не всегда хватает;
- 5 – живу очень плохо, весь в долгах;
- 6 – затрудняюсь ответить.

# Пример порядковой шкалы

- Ваше образование?
- 1 – неполное среднее
- 2 – среднее
- 3 – среднее специальное
- 4 – высшее, незаконченное высшее

# Интервальная шкала

- Построение такой шкалы позволяет большую часть свойств существующих числовых систем приписывать числам, полученным на основе субъективных оценок. Например, шкала возраста. Для данной шкалы допустимым является линейное преобразование. Это позволяет приводить результаты тестирования к общим шкалам и осуществлять, таким образом, сравнение показателей.



# Шкала отношений

- Действует отношение «во столько то раз больше». Это единственная из четырёх шкал имеющая абсолютный ноль. Нулевая точка характеризует отсутствие измеряемого качества. С помощью таких шкал могут быть измерены масса, длина, сила, стоимость (цена).


# Шкала отношений в СОЦИОЛОГИИ

- В социологии шкалы отношений используются в виде естественных (готовых) метрик для таких, например, свойств, как длина (расстояние), уровень доходов, размер жилья и т.п. Однако в чистом виде они используются чрезвычайно редко. Как правило, они переводятся в форму интервальной шкалы.



# Демография

Наука о закономерностях и социальной обусловленности воспроизводства населения, изучает рождаемость, смертность (причины), состояние брака, семьи.





# Объект - население

Тенденции изменения численности населения, проживающего на определенной территории

Возрастной состав

Соотношение мужчин и женщин

Семейное положение





# Структура демографии

Демографическая теория


Сбор первичных данных о населении

Описание демографических процессов

Формальная демография

Демографический анализ





# Основные методы демографии

Статистические методы

Математические методы


Социологические методы

Собственно демографические методы



# Источники данных

Перепись населения (научно организованный процесс сбора и обработки по единой методике демографических и социально-экономических сведений обо всем населении данной страны или территории на определенный момент времени.





# Принципы переписи

Всеобщность

Одномоментность

Самоопределение

Поименность

Конфиденциальность



Системность

Централизация





# Другие источники

Текущий учет демографических событий

Списки и регистры населения

Выборочные и специальные  
демографические исследования

# Относительные величины

- Относительные величины выражают количественные отношения, присущие конкретным общественным явлениям и процессам. Относительные величины рассчитываются путём сравнения (отношения) между собой абсолютных величин и выражаются в коэффициентах, процентах, промилле, децимилле.

# Относительные величины

- Если значение основания или базы сравнения принимается за единицу (приравнивается к единице), то относительная величина (результат сравнения) является коэффициентом и показывает, во сколько раз изучаемая величина больше основания. выражаться и в децимилле, если основание отношения равно 10000 (0/000).

# Виды относительных величин

- Если значение основания или базу сравнения принять за 100 %, результат вычисления относительной величины будет выражаться также в процентах. В тех случаях, когда за базу сравнения принимают 1000, результат сравнения выражаю в промилле (о/оо).  
Относительные величины могут выражаться и в децимилле, если основание отношения равно 10000 (о/ооо).

# Виды относительных величин

- 1. Показатели выполнения договорных соглашений, обязательств, плана.
- 2. Показатели структуры.
- 3. Показатели динамики.
- 5. Показатели сравнения.
- 6. Показатели интенсивности.
- 7. Показатели координации.

# Показатели структуры

- Показатели структуры характеризуют состав той или иной совокупности явления, процесса. Показатель структуры - это относительная доля или удельный вес части в целом, выражающаяся в коэффициентах или процентах:

# Показатели динамики

- Показатели динамики характеризуют изменение изучаемого явления во времени, выявляют направления развития, измеряют интенсивность развития. Эти показатели исчисляются путём деления величины признака текущего периода на величину одного из предшествующих (базисных) периодов.

# Показатели сравнения

- Если при расчёте показателей за базу взята какая-либо часть совокупности, и другие части исчисляются относительно её, тогда в результате такого расчёта получают показатели сравнения. Показатели сравнения могут применяться и для сравнения одноимённых показателей, относящихся к разным совокупностям. Относительная величина сравнения показывает во сколько раз одна из сравниваемых величин больше или меньше другой.



# Показатели интенсивности

- ...строятся путём сравнения разноимённых показателей, но находящихся в определённой связи друг с другом. К ним относятся показатели плотности населения и др. ...относительная величина интенсивности характеризует степень насыщенности изучаемым явлением определенной среды, и исчисляется как отношение величины изучаемого явления к объёму той среды, в которой происходит развитие явления. Чаще всего показатели интенсивности измеряются в промилле и децимилле.

# Промилле и децимилле

- нареч. [латин. promille] (спец.), на тысячу, децимилле на 10 000.
  - 0,5 промилле означает, что на 1 литр, содержащейся в организме жидкости приходится 0,5 г алкоголя.
- Обеспеченность врачами на 10 тыс. населения, показатель составляет 32,2 децимилле (2005г.)

# Показатели координации

- Показатели координации показывают соотношение частей целого между собой. За базу сравнения, как правило, принимают одну из составных частей целого, а затем находят отношение к ней всех остальных частей. Результат показывает, во сколько раз сравниваемая часть больше или меньше части, принятой за базу или сколь единиц данной части приходится на 1, 10, 100, 1000 единиц части, принимаемой за базу сравнения

# Средние величины

- Средняя показывает некоторое центральное значение, около которого группируются реально существующие варианты. В статистических исследованиях в качестве вспомогательных описательных статистических характеристик распределения варьирующего признака широко применяются мода и медиана.

# Рабочий пример (меры главной или основной тенденции)

В качестве рабочего примера возьмем уровень личного богатства (включая все виды благ, такие, как дома, автомобили, банковские счета и капиталовложения) тринадцати индивидов. Предположим, что эти тринадцать человек владеют следующим объемом благ:

- 1.  $J_0$
- 2.  $J_{5000}$
- 3.  $J_{10000}$
- 4.  $J_{20000}$
- 5.  $J_{40000}$
- 6.  $J_{40000}$
- 7.  $J_{40000}$
- 8.  $J_{80000}$
- 9.  $J_{100000}$
- 10.  $J_{150000}$
- 11.  $J_{200000}$
- 12.  $J_{400000}$
- 13.  $J_{10000000}$

# Средние величины: эффективность?

- Среднее здесь соответствует *усреднению в его* обычном понимании, и получается сложением вместе личного богатства всех тринадцати человек и делением результата на их общее число, то есть на 13. Итоговая сумма будет  $11085000$ , разделив это на тринадцать, получаем значение, равное  $852692$ . Среднее часто полезно потому, что оно основывается на использовании всего объема имеющихся данных. Тем не менее, эта операция может ввести в заблуждение там, где один или небольшая часть случаев очень сильно отличаются от большинства. В приведенном примере среднее значение фактически не будет мерой *главной тенденции*, поскольку присутствие одной очень большой величины  $10000000$  искажает все остальное. Может сложиться впечатление, что большинство этих людей владеет гораздо большим объемом благ, чем на самом деле.

# Средняя взвешенная

- Предположим, что 270 человек владеют следующим объемом благ:
- 0 рублей(10 чел)
- 5000 рублей(100 чел)
- 10000 рублей(50 чел)
- 20000 рублей (20 чел)
- 40000 рублей (50 чел)
- 80000 рублей(20 чел)
- 100000 рублей(10 чел)
- 150000 рублей(5 чел)
- 200000 рублей (2 чел)
- 400000 рублей (2 чел)
- 10000000 рублей (1 чел)
- Средняя – 64 814 рублей 81 копейка

# Мода

- **Модой** в статистике называется величины признака (варианта), которая чаще всего встречается в данной совокупности.
- В примере, приведенном здесь, это 140000 (или 5000 рублей). Проблема с модой заключается в том, что этот метод не учитывает общее *распределение* данных, т. е. весь диапазон величин. Наиболее часто встречающийся случай не обязательно будет представительным для распределения в целом, и поэтому в качестве “средней величины” не очень полезен. В нашем случае 140000 не дают точного представления об основной тенденции, поскольку эта сумма располагается слишком близко к нижнему уровню приведенных значений.



# Медиана

Третьей мерой является **медиана** — значение, находящееся *в середине* набора. В примере, приведенном здесь, это седьмое значение — 140000 (в нашем примере — 10 000 рублей). В первом примере дано нечетное количество значений. Если бы оно было четным, например, двенадцать вместо тринадцати то медиана исчислялась бы средним двух чисел, находящихся в середине - шестого и седьмого. Как и мода, медиана не дает представления о реальном диапазоне полученных данных.

# Интервальный ряд

- Рассмотрим, как определяется мода и медиана для интервального ряда.
- Прежде закрывают открытые интервалы (первый и последний) и определяют интервалы, в которых находятся мода и медиана. Их называют соответственно модальным и медианным интервалом.

# Модальный интервал

- Модальный интервал – интервал с наибольшей частотой.
- Медианный интервал (содержащий частоту, который делит ряд пополам) определяется по накопленным частотам. Это будет интервал, накопленная частота которой равна или превышает половину суммы частот.

# Квантили

- Квантили (процентили, децили, квартили)
- Квантиль – это такое значение признака  $q$ , которое делит диапазон его изменения на две части так, чтобы отношение числа элементов выборки, имеющих значение признака, меньшее  $q$ , к числу элементов, имеющих значение признака, большее  $q$ , было равно заранее заданной величине.

# Семейства

- Квантили одного семейства делят диапазон изменения признака на заданное число равнонаполненных частей. Семейство определяется тем, сколько частей получается. Наиболее популярными квантилями являются **квартили**, разбивающие диапазон изменения признака на 4 равнонаполненные части; **децили** - на 10 равнонаполненных частей; **проценти**ли – на 100 частей.

# Использование

- Разного рода квантилями пользуются очень часто. Так, в газетах пишут о том, что, например, 10% наиболее богатых "россиян" имеют месячный доход свыше 100 тысяч рублей, а 10% наиболее бедных – ниже 300 рублей. Ясно, что 100 тысяч рублей – это девятый дециль  $D_9$ , а 300 рублей – это первый дециль  $D_1$ .

# Коэффициенты корреляций

Как связаны друг с другом две (или больше) переменных? Если две переменные полностью коррелируют, мы можем говорить о полной положительной корреляции, выражаемой коэффициентом 1. Там, где связи между двумя переменными не обнаружено (они могут быть вовсе не связаны), коэффициент будет нулевым. Абсолютная отрицательная корреляция, выражаемая как -1, существует там, где две переменные находятся в точном *обратном* отношении друг к другу. В общественных науках абсолютные корреляции никогда не обнаруживаются. Корреляции порядка 0,6 и более, будь то положительные или отрицательные, обычно являются индикатором сильной связи между любыми анализируемыми переменными. Положительные корреляции такого уровня можно, например, обнаружить между классовым происхождением и поведением на выборах. Чем выше англичанин располагается по социально-экономической шкале, тем вероятнее он предпочтет консерваторов лейбористам.