

Статистика — это искусство и наука сбора и анализа данных. Статистические методы следует рассматривать как важную часть процесса принятия решений, позволяющую вырабатывать обоснованные стратегические решения, сочетающие интуицию специалиста с тщательным анализом имеющейся информации.

Использование статистики становится все более важным преимуществом в конкуренции, в том числе территориальной.

Предметом государственной и муниципальной статистики является количественная и качественная оценка социально-экономических явлений и процессов общественной жизни, входящих в компетенцию органов государственного и муниципального управления и обеспечивающих выполнение его функций на разных уровнях.

На федеральном уровне статистикой оцениваются результаты деятельности отраслей народного хозяйства и секторов экономики, объемы производства промышленности, строительства, сельского хозяйства; дается характеристика различных форм собственности, численности населения страны, структуры произведенного продукта, характеристика инфляции и развития социальной сферы, показатели доходов общества и т.п.

На региональном уровне дается статистическая оценка уровня развития регионов, объемы производства в отраслевом и секторальном разрезах; оценка регионального бюджета, характеристика социальной сферы, доходы населения и др.

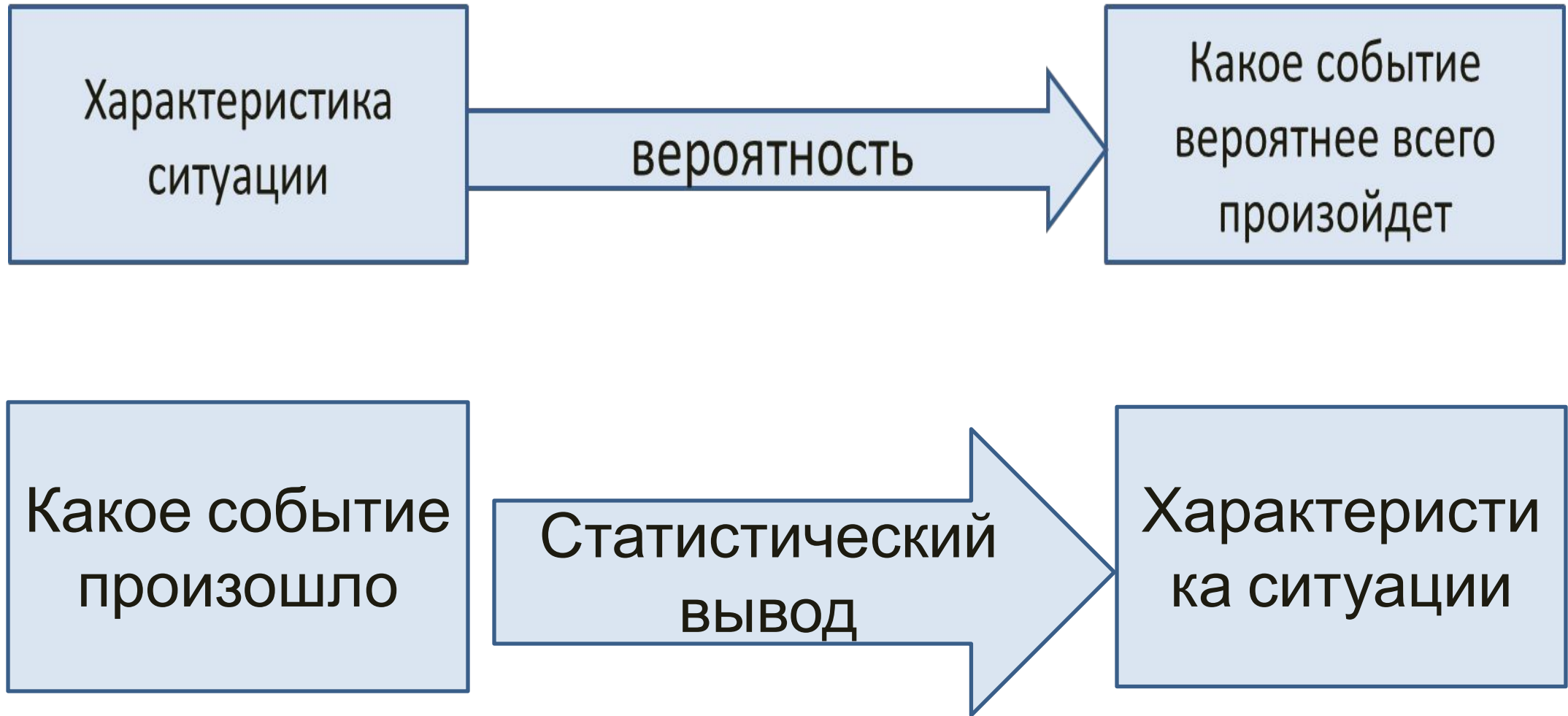
На муниципальном уровне предоставляется информация о муниципальном образовании и его социально-экономическом положении.

Основные этапы статистического анализа

1. Планирование исследования включает составление подробного плана сбора данных, возможно, с использованием случайной выборки из генеральной совокупности.
2. Предварительное исследование данных включает рассмотрение набора данных с разных точек зрения, описание и обобщение данных. Выполнение этого этапа помогает убедиться, что запланированный анализ адекватен данным, а при необходимости позволяет внести в процесс анализа определенные коррективы.
3. Оценивание неизвестной величины дает наиболее обоснованное возможное предположение о значении, основанное на исходных данных. Кроме того, есть возможность вычислить величину ошибки, которая возникает при использовании оценки вместо фактического, но неизвестного значения.
4. Проверка статистических гипотез заключается в использовании данных для выбора одной из двух (или больше) различных возможностей при решении вопроса в неопределенной ситуации. Такая проверка позволяет убедиться, действительно ли данные обладают определенным интересным свойством, или мы имеем дело с «чистой случайностью», которая не представляет интереса.

Вероятность исходя из предположений об изучаемой ситуации показывает возможность или шанс наступления в будущем каждого из нескольких потенциальных событий.

Вероятность — это понятие, в некотором смысле обратное статистике: вероятность показывает, какие данные вы скорее всего получите, если известна характеристика ситуации, а статистика помогает охарактеризовать ситуацию в результате анализа и обобщения данных.



Статистика лучше всего работает в сочетании с вашими собственными экспертными заключениями и здравым смыслом. Если результаты статистического анализа расходятся с вашей интуицией, необходимо разобраться, чтобы установить причину. Статистический анализ может оказаться некорректным, если в его основу положены неверные допущения, или ваша интуиция может вас подвести, если она не основана на фактах.

Набор данных содержит одно или несколько значений для каждого из отдельных объектов, называемых элементарными единицами. В качестве таких объектов могут выступать люди, домохозяйства, города, телевизионные приемники или что угодно, что представляет интерес для изучения. Для каждого из объектов регистрируют один и тот же признак (или признаки). Признак, который регистрируют для каждого из объектов (например, стоимость), называется **переменной**.

Существуют **три основных способа классификации наборов данных**:

- по количеству переменных (одномерный, двумерный и многомерный);
- по типу представленной каждой из переменных информации (числа или категории);
- в зависимости от того, является ли набор данных временным рядом, или это данные об одном временном срезе.

Одномерные наборы данных (одна переменная) содержат информацию только об одном признаке, зарегистрированную для каждого объекта. Одномерный набор данных позволяет определить типичное значение и характеристику изменчивости данных, а также выделить специфические особенности или проблемы в данных.

Двумерные наборы данных (две переменные) содержат два признака, значения которых регистрируются для каждого объекта. Двумерные данные в дополнение к информации о каждой переменной как наборе одномерных данных позволяют изучить связь между двумя переменными и предсказать значение одной переменной на основе значения другой.

Многомерные наборы данных (много переменных) содержат три или больше признаков, значения которых регистрируются для каждого объекта. Многомерные данные в дополнение к информации о каждой переменной как наборе одномерных данных дают возможность изучить связь между переменными и предсказать значение одной переменной на основе значения других. Значения переменных, которые регистрируются как числа, имеющие содержательный смысл, называют **количественными данными**. **Дискретная количественная переменная** может принимать значения только из некоторого списка конкретных чисел (таких как, например, 0 или 1, или перечень чисел 0, 1, 2, 3, . . .). Любую количественную переменную, которая не является дискретной, будем называть **непрерывной**. Значения непрерывной переменной не ограничены простым перечнем возможных значений. Если переменная содержит информацию о том, какой из нескольких нечисловых категорий принадлежит объект, то она называется **качественной переменной**. Если категории можно естественным образом и содержательно осмысленно упорядочить, то речь идет о **порядковой качественной переменной**. Если такой порядок отсутствует, то речь идет о **номинальной качественной переменной**. Несмотря на то что часто значения качественной переменной можно записать с помощью чисел, такая переменная все равно остается качественной, а не количественной, поскольку эти числа не имеют какой-либо интерпретации, содержательно присущей этой переменной.

К количественным данным можно применять те же операции, что и к обычным числам: подсчет частоты, ранжирование, арифметические действия. С порядковыми данными можно выполнять только подсчет частоты и ранжирование, с номинальными данными — только подсчет частоты.

Если последовательность записи данных имеет содержательный смысл, то соответствующий набор данных представляет собой **временной ряд**. Если последовательность записи данных не важна, то соответствующий набор содержит данные об **одном временном срезе**. Анализ временных рядов сложнее анализа данных об одном временном срезе.

Если вы самостоятельно планируете сбор данных (даже если собственно сбор данных делают другие), то получаете **первичные данные**. Если вы используете данные, предварительно собранные другими людьми и для других целей, то вы имеете дело с **вторичными данными**. Получение первичных данных часто обходится дорого и занимает много времени, но вы получаете то, что вам необходимо. Вторичные данные можно получить дешевле (или даже бесплатно), но вы можете найти, а можете и не найти то, что вам необходимо.