

Тема 1.

Введение

Прикладной статистический анализ

Для успешного бизнеса необходимо постоянно собирать данные, отражающие текущее состояние дел. Чтобы принять обоснованное решение, эти данные следует преобразовывать в информацию. В настоящее время существует много способов **извлечь информацию из собранных фактов с помощью методов статистики** — отрасли математики, изучающей методы обработки и анализа данных.

Генеральная совокупность и выборка

Генеральная совокупность (population) – вся интересующая исследователя совокупность изучаемых объектов.

Выборка (sample) – некоторая, обычно небольшая, часть генеральной совокупности, отбираемая специальным образом и исследуемая с целью получения выводов о свойствах генеральной совокупности.

Репрезентативная выборка *хорошо* представляет генеральную совокупность. Это означает, что каждое свойство (или комбинация свойств) наблюдается в выборке с той же частотой, что и в генеральной совокупности.

Параметры — это суммарные количественные показатели, характеризующий всю генеральную совокупность.

Статистики - суммарные количественные показатели, вычисленные по выборке и позволяющие оценить параметры генеральной совокупности. .

Мы будем использовать статистики для оценки тех параметров генеральной совокупности, которым они соответствуют.

Генеральная совокупность и выборка (2)

Что рассматривается?	Генеральная совокупность	Выборка
Как называется числовая характеристика?	Параметры	Статистики
Примеры	Все жители города (150 тыс.) Доля мужчин в городе Средний возраст жителей	Участники опроса (250 чел.) Доля мужчин в выборке Средний возраст в выборке

Переменные, признаки

Переменная, признак – это некоторая общая для всех изучаемых объектов характеристика или свойство, конкретные проявления которого могут меняться от объекта к объекту. Проявления признака называют значениями, альтернативами, градациями.

Переменная Возможные значения

«Пол» два значения: «мужчина» и «женщина»

«Профессия» большое число значений, например, «менеджер»

«Рост» может меняться от «очень низкий» до «очень высокий»
или от 150 см до 210 см

Шкалы

Типы признаков:

Качественные (номинальные, категориальные) - названия, имена и т.п. Объекты не могут быть упорядочены по качественному признаку и с ними не могут производиться арифметические операции.

Порядковые признаки измеряются в условной шкале и объекты можно упорядочить по степени проявления признака.

Количественные признаки позволяют не только упорядочить объекты, но и производить арифметические операции между значениями признаков.

Примеры признаков

Номинальные Хобби студента, пол, профессия

Порядковые Место на соревнованиях, рейтинг региона,
баллы, полученные в семестре

Количественные Температура воздуха, ВРП,
средняя заработная плата

Пример. Какой тип признака?

Температура воздуха в лекционной аудитории

Возраст

Пол студента

Семейное положение

Место жительства

Религиозные предпочтения

Время на подготовку домашнего задания

Трудолюбие

Данные

Данные представляют собой значения признаков, полученные в результате наблюдений, испытаний, накапливаемые с целью последующего изучения и анализа.

Респондент	Возраст	Пол	Образование	Семейное положение
1	29	0	12	2
2	23	1	14	1
3	37	1	16	2
4	46	0	10	4
5	34	1	14	1

Дискретные данные представляют собой отдельные значения признака, общее число которых конечно.

Непрерывные данные могут принимать любое значение в некотором интервале.