

Applied Methods of Statistical Analysis

Annotation

Учебная дисциплина предусматривает изучение методов статистического моделирования, используемых при решении прикладных задач. Программа рассчитана как на студентов, владеющих базовыми знаниями теории вероятностей и математической статистики, так и на студентов без этих знаний, но имеющих достаточный уровень математической подготовки. Для последних в режиме смешанного обучения предусмотрено расширенное рассмотрение базовых тем, адаптирующее студентов к прохождению основного материала. В качестве информационной и методической поддержки при изучении курса используются материалы записанного автором онлайн курса "Прикладной статистический анализ" на Национальной платформе открытого образования.

1. Basic concepts and tasks of statistics. Types and sources of statistical data

Основные определения:

- Генеральная совокупность – сведения о всех анализируемых объектах;
- Выборка – множество результатов, отобранных из генеральной совокупности (репрезентативность);
- Объем совокупности – число единиц, образующих совокупность;
- Неопределенность и вариация (основные характеристики статистики) – при многократных измерениях происходят изменения;
- Признак – характеристика единицы совокупности;
- Показатель (индикатор) – количественная характеристика явления;
- Параметр – относительно постоянная величина, характеризующая генеральную совокупность;
- Выборочная характеристика (статистика) – эмпирический аналог параметра;
- Статистические выводы – заключения, формируемые анализом эмпирических данных.

Виды и источники статистических данных

Статистические данные разделяются на:

- Пространственные: сведения об объектах наблюдения с различным порядком;
- Временные: хронологический порядок (моментные – сумма только сумма значений и интервальные – суммирование дает общую характеристику и может быть проинтерпретирована);
- Пространственно-временные: набор объектов в хронологическом порядке со сведениями об объектах.

Статистические данные могут быть одномерными и многомерными, количественные и категориальные, первичные (регистрируемые для одного конкретного объекта) и агрегированные (объект – совокупность других объектов).

Шкалы измерения данных: качественные данные – номинальная (профессия, пол, город), порядковая (место в рейтинге), количественные (непрерывные и дискретные) – интервальная (температура воздуха), относительная (количество наличных денег, времени, объектов).

Источники статистических данных: непосредственные измерения, мнения экспертов, документированные значения.

- Статистическое наблюдение – планомерный и систематический сбор данных об исследуемых явлениях и процессах, бывает сплошным (на генеральной совокупности) и несплошным (на выборке).

Задачи статистики

В узком смысле: сжатие информации и наглядное представление результатов.

В широком смысле: обобщение результатов выборочного исследования на генеральную совокупность.

Первичная обработка: пример данных качественного характера

- таблица частот;
- столбиковая диаграмма;
- круговая диаграмма.

Первичная обработка: количественного характера.

- гистограмма

Этапы статистического моделирования:

Определение цели и задач моделирования

1. Формализация – преобразование объектов и отношений в математическую абстрактную модель;
2. Сбор и квантификация данных – предусматривает отражение данных в шкалах, их предварительная обработка – избавление от ошибок;
3. Спецификация модели – представление в виде формул;
4. Идентификация модели и ее анализ – оценка параметров модели, ее характеристик;

Верификация модели.