

АННОТАЦИЯ К РАБОЧЕЙ ПРОГРАММЕ ДИСЦИПЛИНЫ «Основы статистической обработки экспериментальных данных»

по основной профессиональной образовательной программе по направлению подготовки
09.03.01 «Информатика и вычислительная техника» (уровень бакалавриата)

Направленность (профиль): Программно-техническое обеспечение автоматизированных систем
Общий объем дисциплины – 5 з.е. (180 часов)

Форма промежуточной аттестации – Экзамен.

В результате освоения дисциплины у обучающихся должны быть сформированы компетенции с соответствующими индикаторами их достижения:

- ОПК-1.1: Применяет математический аппарат, методы математического анализа и моделирования для решения задач;
- ОПК-1.3: Участвует в теоретических и экспериментальных исследованиях объектов профессиональной деятельности, в обработке их результатов;
- ОПК-2.1: Выбирает информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности;
- ОПК-2.2: Использует современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности;
- ОПК-9.1: Использует программные средства для решения практических задач на основе существующих методик;

Содержание дисциплины:

Дисциплина «Основы статистической обработки экспериментальных данных» включает в себя следующие разделы:

Форма обучения очная. Семестр 4.

1. Применение математического аппарата теории вероятности и методов математического анализа для решения задач по обработке экспериментальных данных. Понятия классической и геометрической вероятности, правила расчета и примеры задач. Условная вероятность, формулы полной вероятности, Байеса, Бернулли, примеры использования при обработке данных. Случайные величины, функция распределения, плотность распределения. Числовые характеристики случайных величин: математическое ожидание, дисперсия, стандартное отклонение, мода, медиана, асимметрия, эксцесс. Свойства численных характеристик и примеры их вычислений для решения задач обработки экспериментальных данных..

2. Использование современных информационных технологий и программных средств для решения задач статистической обработки экспериментальных данных. Применение Python для статистических вычислений.. Запись на Python разветвляющихся и циклических вычислительных процессов при обработке данных. Функции работы со строками в Python, примеры. Списки и их использование для обработки массивов. Применение библиотеки numpy для обработки одномерных и двумерных числовых массивов. Чтение из файла выборки, вычисление числовых характеристик выборки: среднее, дисперсия, стандартное отклонение, мода, медиана, асимметрия, эксцесс..

3. Критерии согласия Хи-квадрат и Колмогорова-Смирнова. Выборки с заданным законом распределения. Освоение программных средств для решения практических задач.. Выполнение оценки параметров распределений в соответствии с предположением о законе распределения выполнение экспоненциального, нормального, Парето. Применение графической библиотеки Python для построения эмпирических и теоретических функций и плотностей распределения для визуальной оценки соответствия выборки соответствующему распределению. Расчет эмпирического значения Хи-квадрат и эмпирических величин K^+ , K^- для закона Колмогорова-Смирнова. Сравнение с табличными значениями. Вывод о соответствии выборки предполагаемому распределению генеральной совокупности..

4. Описательная статистика и построение доверительных интервалов для нормального закона распределения. Применение существующих методик при разработке программного обеспечения на Python.. Построение статистического отчета описательной статистики программным кодом Python, содержащего частоты, накопленные частоты, относительные частоты

и накопленные относительные частоты. Использование библиотеки `matplotlib` системы программирования Python для построения гистограммы и полигона частот выборки. Использование библиотеки `scipy.stats` для расчета доверительных интервалов для истинного среднего по функции вычисления квантилей распределения Стьюдента и доверительных интервалов истинного значения дисперсии по функции вычисления квантилей распределения Хи-квадрат..

5. Проверка статистических гипотез о равенстве средних и дисперсий нормальных генеральных совокупностей по двум выборкам.. Знакомство с библиотекой Tkinter для построения графических интерфейсов пользователя системы программирования Python. Элементы графического интерфейса: поле вывода Label, поле ввода Entry, кнопка Button, радиокнопка Radiobutton, флажок Checkbutton. Построение меню с помощью библиотеки Tkinter. Примеры. Оценивание средних, дисперсий и стандартных отклонений двух выборок нормальных генеральных совокупностей. Проверка нулевых гипотез о равенстве средних и дисперсий генеральных совокупностей. Вычисление по функциям библиотеки `scipy.stats` теоретических значений распределений Стьюдента и Хи-квадрат для подтверждения или опровержения нулевых гипотез..

6. Математический аппарат корреляционного анализа. Применение современных информационных технологий и программных средств для решения задач корреляционного анализа при обработке данных эксперимента. Объектно-ориентированное программирование на Python. Классы и объекты. Атрибуты и методы. Разграничение доступа, пометка `private` полей и методов через двойное подчеркивание. Конструктор, его использование. Класс с атрибутами вспомогательного класса. Коэффициент парной корреляции Пирсона. Вычисление коэффициента парной корреляции и построение диаграммы рассеяния функцией `scatter` библиотеки `matplotlib`. проверка значимости коэффициента парной корреляции по распределению Стьюдента. Корреляционное отношение как обобщение парной корреляции. Коэффициенты частной и множественной корреляции для нескольких выборок. Коэффициент ранговой корреляции Спирмена..

7. Регрессионный анализ. Решение задач профессиональной деятельности при обработке данных эксперимента.. Производные классы в объектно-ориентированном программировании на Python. Особенность конструктора производного класса с параметрами. Линейная регрессия двух переменных. Построение прямой регрессии методом наименьших квадратов. Коэффициенты детерминации и множественной корреляции, их вычисление на Python. Варианты регрессионных зависимостей, сводящихся к линейной регрессии: гиперболическая, экспоненциальная, логарифмическая, степенная. Полиномиальная регрессия. Исследование значимости уравнения регрессии в целом по распределению Фишера и значимости отдельных коэффициентов по распределению Стьюдента в предположении нормальности выборок. Построение доверительного коридора прогнозных значений..

8. Математический аппарат факторного анализа. Применение современных информационных технологий и программных средств для решения задач факторного анализа при обработке данных эксперимента. Построение главных компонент корреляционной матрицы. Определение числа факторов по количеству собственных чисел больших единицы корреляционной матрицы и по величине в процентах накопленной дисперсии. Построение матрицы факторных нагрузок варимаксным методом через собственные вектора корреляционной матрицы. Интерпретация факторов. Вычисление количества каждого фактора у наблюдаемых переменных..

9. Дисперсионный анализ. Решение задач профессиональной деятельности при обработке данных эксперимента.. Таблица значений измеряемой величины при наборах уровней одного или двух неизмеримых факторов. Вычисление межгрупповой и внутригрупповой дисперсии. Проверка значимости влияния факторов и их взаимодействия на измеряемую величину по критерию Фишера..

10. Дискриминантный анализ.. Построение канонических линейных дискриминантных функций для двух и более классов. Вычисление коэффициентов дискриминантных функций методом наименьших квадратов. Статистический анализ классификации выборки на основе статистики Уилкса и частного F-отношения..

11. Кластерный анализ. Решение задач профессиональной деятельности при обработке данных эксперимента.. Методы кластерного анализа выборки K внутригрупповых средних (K -mean) и максиминное расстояние. Особенности их использования. Библиотека ImageDraw системы программирования Python для отображения графических образов и их обработки. Применение библиотеки ImageDraw для графического отображения построенных кластеров методами кластерного анализа..

12. Метод Arima для статистического анализа данных, заданных в виде временного ряда.. Модели авторегрессии, проинтегрированного скользящего среднего для анализа выборки в виде временного ряда. Смешанные модели. Автокорреляционная функция и частная автокорреляционная функция, их оценки библиотеками Python. Линейный метод наименьших квадратов для оценки параметров модели авторегрессии и метод Марквардта для оценки параметров скользящего среднего. Сезонные и несезонные разности модели. Построение прогнозов по модели с заданными вероятностными пределами..

Разработал:
доцент
кафедры ИВТиИБ

А.Н. Тушев

Проверил:
Декан ФИТ

А.С. Авдеев