

Курс «Элементы прикладной математической статистики» читается в 4 семестре студентам 2-го года обучения, относится к вариативной части блока математических и естественно-научных дисциплин. Программа курса включает в себя как изучение статистических понятий, моделей и теорем, так и конкретное применение полученных знаний в задачах обработки экспериментальных данных с помощью прикладного программного обеспечения ЭВМ.

Используется статистический пакет, встроенный в Microsoft Excel, а также свободно распространяемый статистический модуль Attestat. Изучение двух программных продуктов позволяет надеяться, что в дальнейшем обучающиеся смогут без особых усилий использовать и другие статистические программные комплексы.

Теоретическая часть состоит из четырех разделов:

- понятие выборки, нахождение распределения выборки: включает в себя основные определения, методы оценки распределения и их состоятельность
- точечные и интервальные оценки параметров распределения выборки: рассматриваются методы построения оценок для определенных моделей, свойства получаемых оценок
- проверка гипотез: определяется понятие статистического критерия, виды критериев, принятие решения, условия применимости, оптимальность критерия; разбирается ряд классических критериев для применения в задачах.
- регрессионный анализ: рассматривается задача регрессии общего вида, методы нахождения параметров

Цели и задачи освоения дисциплины.

Целью освоения дисциплины «Элементы прикладной математической статистики» является формирование у студентов умений и навыков обработки экспериментальных данных с помощью прикладного программного обеспечения ЭВМ.

Требования к результатам освоения содержания дисциплины

В результате освоения дисциплины студент должен **знать**

- основные понятия и теоремы математической статистики
- основные методы обработки данных, возможности и способы их применения

уметь

- формализовать задачу
- выбрать подходящую модель
- применять изученные теоремы на практике
- обрабатывать данные с использованием прикладного программного обеспечения ЭВМ, делать выводы

Структура дисциплины

Общая трудоемкость дисциплины составляет 2 зачетных единицы (72 часа), из них лекции 16 часов, лабораторные работы 16 часов, самостоятельная работа 40 часов.

Вид работы	Всего
Общая трудоёмкость, акад. часов	72
Аудиторная работа:	
Лекции, акад. часов	16
Лабораторные работы, акад. часов	16
Самостоятельная работа, акад. часов	40
Вид итогового контроля (зачёт, зачёт с оценкой, экзамен)	Зачет с оценкой

Лекции

№ раздела	Наименование раздела	Содержание раздела
1	Понятие выборки; распределение выборки.	Основные задачи и методы прикладной математической статистики. Обработка данных, полученных в результате наблюдении; краткий обзор задач, возникающих в практике исследователя-химика. Понятие выборки. Выборочные распределения. Эмпирическая функция распределения. Гистограмма. Эмпирические моменты. Понятие оценки (статистики).
2	Оценки: построение и свойства.	Выборка из нормального распределения. Точные выборочные распределения. Распределения: хи-квадрат, Стьюдента, Фишера. Точечные оценки. Несмещенные и состоятельные оценки. Примеры. Методы получения оценок. Доверительные интервалы. Точные доверительные интервалы для параметров нормального распределения. Приближенный доверительный интервал для вероятности успеха в схеме Бернулли. Приближенный доверительный интервал для математического ожидания в случае выборки из произвольного распределения.
3	Проверка гипотез.	Общие понятия о статистической проверке гипотез. Простые и сложные гипотезы. Ошибки 1-го и 2-го рода. Мощность критерия. Критерий хи-квадрат. Применение критерия хи-квадрат для проверки гипотезы о принадлежности выборки данному распределению (известному полностью или известному с точностью до параметров). Проверка гипотезы о независимости признаков. Примеры. Критерий Колмогорова. Проверка гипотез о равенстве дисперсий и равенстве математических ожиданий по двум выборкам из нормальных распределений. Отбраковка "промахов". Примеры. Оптимальный критерий Неймана-Пирсона для различения двух простых гипотез. Построение оптимального критерия для различения двух простых гипотез о математическом ожидании по выборке из нормального распределения; нахождение объема выборки, который при оптимальном критерии для указанной задачи обеспечивает заданные величины ошибок 1-го и 2-го рода. Примеры.
4	Регрессионный анализ.	Регрессионный анализ. Метод наименьших квадратов и его применение для построения "градуировочных графиков". Примеры.

Лабораторные работы

№ раздела	№ ЛР	Наименование лабораторных работ
1,2	1	Построение и расчет оценок и доверительных интервалов для параметров распределений. Построение гистограмм и нормальной бумаги. Проверка предположений о распределении выборки с помощью критериев.
3	2	Исследование независимости 2х выборок.
	3	Сравнение дисперсий и средних 2х выборок.
	4	Дисперсионный анализ. Выявление зависимости ошибок измерения от измеряемых величин.

4	5	Линейная одномерная и многомерная регрессии. Построение градуировочных графиков.
	6	Нелинейная регрессия и дискриминантный анализ

Вопросы к зачету

1. Понятие случайной выборки. Вариационный ряд. Размах выборки. Эмпирическая функция распределения. Гистограмма. Эмпирические моменты.
2. Понятие оценки (статистики). Точечные оценки. Несмещенные, состоятельные оценки параметров распределения. Достаточные условия состоятельности несмещенной оценки.
3. Эмпирическая функция распределения как несмещенная состоятельная оценка теоретической функции распределения.
4. Эмпирические моменты как оценка теоретических моментов. Несмещенные состоятельные оценки для математического ожидания и дисперсии.
5. Выборка из нормального распределения. Точное выборочное распределение хи-квадрат, Стьюдента, Фишера. Таблицы.
6. Теорема Фишера о распределении среднего арифметического и выборочной дисперсии для нормальных выборок.
7. Функции от среднего арифметического и выборочной дисперсии нормальных выборок, имеющие распределение Стьюдента и Фишера.
8. Интервальные оценки. Понятие доверительного интервала для неизвестного параметра распределения. Доверительная вероятность. Вид оценок, с помощью которых возможно построить доверительные интервалы.
9. Построение доверительного интервала для математического ожидания (при известной и неизвестной дисперсии) и дисперсии (при неизвестном и известном математическом ожидании) в случае нормальных выборок.
10. Приближенные доверительные интервалы для математического ожидания в случае выборок из произвольного распределения. Приближенные доверительные интервалы для вероятности успеха в схеме Бернулли.
11. Проверка гипотезы о равенстве дисперсий двух независимых нормальных выборок. Пример.
12. Проверка гипотезы о равенстве математических ожиданий двух независимых нормальных выборок при равенстве их (неизвестных) дисперсий. Отбрасывание “промахов”.
13. Критерий хи-квадрат Пирсона для проверки гипотезы о распределении выборки (при полностью известной и известной с точностью до параметра гипотетической функции распределения). Критерий Колмогорова.
14. Применение критерия хи-квадрат для проверки гипотезы о независимости признаков. Пример.
15. Общие понятия проверки статистических гипотез. Простые и сложные гипотезы. Ошибки 1го и 2го рода. Мощность критерия. Критерий Неймана-Пирсона для различия двух простых гипотез.
16. Градуировочные графики. Линейная регрессионная модель. Метод наименьших квадратов. Пример.
17. Оптимальный критерий Неймана-Пирсона.

Примеры расчетно-графических заданий.

1. Заданы 3 набора чисел. Требуется
 - построить диаграммы рассеяния; визуально выдвинуть предположения о зависимости или независимости соответствующих выборок
 - рассчитать оценки коэффициентов корреляции; выдвинуть предположения о зависимости или независимости соответствующих выборок на основе полученных значений
 - подтвердить или опровергнуть сделанные предположения на основе применения критериев, приняв величину в 1% за уровень значимости
 2. Заданы 2 набора чисел.
 - Провести линейную регрессию и построить градуировочный график. Оценить качество регрессии
 - Сделать вывод о наличии/отсутствии постоянной систематической ошибки и линейной систематической ошибки
- Заданы результаты измерения в зависимости от 3х факторов.
- Провести многомерную линейную регрессию. Оценить качество регрессии.
 - Сделать вывод о зависимости измеряемой величины от каждого из факторов.
 - Выписать уравнение зависимости измеряемой величины от факторов.

Учебные пособия

1. В. П. Чистяков, Курс теории вероятностей, 7-е изд., Дрофа, Москва, 2007, 256 с.
2. Гнеденко Б.В. Курс теории вероятностей. 8-е изд., испр. и доп.—М.: Едиториал УРСС, 2005.— 448 с.

Интернет-ресурсы

1. А.Д.Манита. Теория вероятностей и математическая статистика. <http://teorver-online.narod.ru/>

Программное обеспечение современных информационных компьютерных технологий

1. Microsoft Excel, Microsoft Office
2. Attestat – бесплатно распространяемый статистический модуль для Microsoft Excel