

Программа курса по статистическому анализу

У семестр, группы 8382-83 (2020 г).

СТАТИСТИЧЕСКИЙ ЭКСПЕРИМЕНТ

1. Основные понятия математической статистики. Статистический эксперимент. Виды задач математической статистики. Задачи точечного оценивания, доверительного оценивания, проверки статистических гипотез.
2. Понятие статистики. Использование статистик при решении задач точечного, доверительного оценивания и проверки статистических гипотез.
3. Классификация статистических моделей по типу параметра (параметрические, непараметрические, семипараметрические). Способы накопления статистической информации. Понятие выборки, выборочный принцип.
4. Регрессионные модели. Асимптотический подход. Состоятельность.

ТОЧЕЧНОЕ ОЦЕНИВАНИЕ

5. Непараметрическое оценивание. Выборочный метод. Вариационный ряд и порядковые статистики. Ранги. Эмпирическая функция распределения. Теорема Гливленко-Кантелли. Преобразование Смирнова. Теорема Колмогорова. Оценивание теоретической функции распределения эмпирической. Построение доверительной области.
6. Выборочные числовые характеристики. Примеры (моменты, квантили). Состоятельность.
7. Асимптотическая нормальность выборочных квантилей.
8. Оценивание плотности распределения (гистограмма частот, полигон).

ПАРАМЕТРИЧЕСКОЕ ОЦЕНИВАНИЕ

9. Примеры параметрических семейств распределений. Семейства абсолютно непрерывных распределений: нормальное, показательное, гамма, Коши. Семейства дискретных распределений: Бернулли, Биномиальное, Пуассона, Геометрическое.
10. Выборка из нормального распределения, Лемма Фишера. Распределения Хи-квадрат и Стюдента. Следствие и распределение Фишера-Снедекора.
11. Постановка задачи точечного оценивания вещественного параметра. Функция потерь. Риск. Несмещенное оценивание. Асимптотический подход и состоятельное оценивание.
12. Подчиненные и достаточные статистики. Теорема факторизации Неймана-Фишера. Примеры (П).
13. Методы построения статистических оценок. Метод максимального правдоподобия. Примеры вычисления оценок по методу максимального правдоподобия. Метод моментов.
14. Несмещенное оценивание на базе достаточных статистик. Алгоритм построения НРМД оценок. Теорема Рао-Блэкуэлла-Колмогорова, теорема Лемана-Шеффе. Пример(ы) (П).
15. Минимаксный и байесовский подходы. Теорема Лемана.
16. Регулярный эксперимент. Информация Фишера. Свойства. Неравенство Рао-Крамера. Многомерный случай, информационное неравенство.
17. Эффективные по Фишеру оценки (равенство в неравенстве Рао-Крамера). Необходимое условие существования. Эффективность по Фишеру и метод максимального правдоподобия.
18. Многомерный случай. Информационное неравенство.
19. Асимптотические свойства оценок максимального правдоподобия. Асимптотическая нормальность в случае многомерного параметра.

ДОВЕРИТЕЛЬНОЕ ОЦЕНИВАНИЕ

20. Понятие доверительного интервала. Общий метод построения доверительных интервалов.
21. Построение доверительных интервалов для параметров нормального закона (случай одной и двух выборок).
22. Построение доверительных интервалов на базе статистики, распределение которой монотонно зависит от параметра (преобразование Смирнова).
23. Асимптотические доверительные интервалы. Построение асимптотических доверительных интервалов на базе асимптотически нормальной оценки параметра. Построение доверительных интервалов на базе ОМП в регулярном случае. Примеры.

ПРОВЕРКА СТАТИСТИЧЕСКИХ ГИПОТЕЗ

24. Постановка задачи проверки статистических гипотез. Понятие статистической гипотезы, вероятностей ошибок 1-го и 2-го рода, критерия, доверительной, критической области и области сомнений, мощности критерия. Выражение вероятностей ошибок в терминах критерия. Построение нерандомизованного статистического критерия, статистика критерия.

24. Задача проверки простой гипотезы при простой альтернативе. Лемма Неймана-Пирсона.
25. Равномерно наиболее мощный критерий для проверки односторонней гипотезы.
26. Асимптотические критерии. Параметрический критерий отношения правдоподобия.
27. Непараметрические критерии. Концепция построения хи-квадрат критериев. Критерий хи-квадрат для проверки простой гипотезы согласия.
28. Критерий хи-квадрат для проверки сложной параметрической гипотезы согласия. Критерий хи-квадрат для проверки независимости.
29. Критерий согласия Колмогорова и Крамера-Фон Мизеса. Критерий однородности Колмогорова-Смирнова.

РЕГРЕССИОННЫЙ АНАЛИЗ

30. Модель линейной регрессии. Метод наименьших квадратов. Оценка по методу наименьших квадратов. Геометрическая интерпретация.
31. Функции параметра, допускающие несмещенное оценивание. Теорема Гаусса-Маркова.
32. Примеры регрессионных моделей, простая и полиномиальная регрессии, модель однофакторного дисперсионного анализа.
32. Несмещенная оценка дисперсии, дисперсии и ковариационные характеристики НРМД-оценок допускающих несмещенное оценивание функций параметра.
33. Доверительное оценивание параметров модели в предположении нормальности. Построение доверительных эллипсоидов и доверительных интервалов.
34. Постановка задач проверки статистических гипотез, F -критерий. Два способа вычисления статистики F -критерия.
35. Дисперсионный анализ. Однофакторный дисперсионный анализ (простая группировка). Гипотеза отсутствия влияния фактора на результат, F -критерий. Параметризация модели с использованием сравнений (главных эффектов). Запись гипотезы в терминах главных эффектов.
36. Двухфакторный дисперсионный анализ. Главные эффекты и взаимодействия. Основные гипотезы и их интерпретация. Многофакторные модели.
37. Обобщенные линейные модели (мотивация, определение). Экспоненциальные семейства распределений, свойства.
38. Построение оценок максимального правдоподобия параметров регрессии в обобщенной линейной модели. Каноническая функция связи, уравнения для построения оценок максимального правдоподобия.
39. Анализ категориальных данных с использованием обобщенных линейных моделей. Модель логистической регрессии. Пуассоновская модель.
40. Использование пуассоновской модели при анализе таблиц сопряженности трех признаков. Гипотезы однородности зависимости, условной независимости двух признаков при известном значении третьего, независимости трех признаков в терминах вероятностей и в терминах параметров регрессии.

Список рекомендованной литературы:

1. Ивченко Г. И., Медведев Ю. И. Математическая статистика. М.: Высш. шк., 1984.
2. Леман Э. Теория точечного оценивания. М.: Наука, 1991.
3. Леман Э. Проверка статистических гипотез. М.: Наука, 1964.
4. Малов С.В. Регрессионный анализ: теоретические основы и практические рекомендации. – СПб.: Изд-во СПбГУ, 2013. – 276 стр.
5. Шеффе, Г. Дисперсионный анализ. М.: Наука, 1980.

Учебные пособия:

- Коробейников А.И., Малов С.В., Матвеева И.В. Основные алгоритмы численного анализа: использование пакета R (S-plus) для анализа статистических данных. – СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2011, 40 с.
- Лившиц А.Н., Малов С.В. Математическая статистика. СПбГЭТУ «ЛЭТИ», СПб., 1999
- Егоров В.А., Ингстер Ю.И., Лившиц А.Н., Малова И.Ю., Малов С.В. Анализ однородных статистических данных. СПбГЭТУ «ЛЭТИ», СПб., 2005.

Список дополнительной литературы:

1. Боровков А. А. Математическая статистика. М.: Наука, 1984.
2. Крамер Г. Математические методы статистики. 2-е изд. М.: Мир, 1975.
3. Кокс Д., Хинкли Д. Теоретическая статистика. М. Мир, 1978.
4. Рао, С.Р. Линейные статистические методы и их применения. М.: Наука, 1968.