

## ПРОГРАММА КУРСА ПО "МАТМОДЕЛИРОВАНИЕ И".

1. Задачи математической статистики. Параметрические, семипараметрические, непараметрические статистические выводы. Описательная статистика. Теория статистических выводов. Типы данных: количественные и качественные; первичные, вторичные (степень их достоверности). Случайная и систематическая ошибки. Примеры.
2. Выборка. Статистика. Частота. Относительная частота. Связь относительной частоты события и вероятности события.
3. Эмпирическая функция распределения и гистограмма. Их связь с теоретической функцией распределения и плотностью распределения. Уметь их строить.
4. Основные числовые характеристики выборки и общий принцип построения их оценок. Выборочные среднее и дисперсия, абсолютный и центральный моменты.
5. Оценки параметров положения. Медиана, мода, выборочное среднее.
6. Оценки параметров рассеяния. Выборочная дисперсия, стандартное отклонение, абсолютное отклонение, квантили, квартили, персентили, интерквартильный размах, размах.
7. Оценки параметров формы плотности. Асимметрия и эксцесс.
8. Роль нормального распределения в статистике.
9. Теорема об асимптотической нормальности функций от асимптотически нормальных случайных величин. Ее следствия.
10. Многомерное нормальное распределение для произвольной ковариационной матрицы. Ее плотность и график. Болезнь размерности
11. Арифметика нормального распределения. Независимость выборочных средних и дисперсии.
12. Распределения связанные с нормальным: хи-квадрат, Стьюдента, Фишера. Где каждое из них применяется.
13. Теория статистического оценивания. Общая постановка задачи. Определения статистики, статистической оценки, квадратичной функции риска и ее связь с дисперсией оценки. Примеры статистических оценок.
14. Доверительное оценивание (зачем оно нужно?). Доверительный интервал, уровень доверия.
15. Доверительный интервал для математического ожидания нормального распределения (дисперсия известна и неизвестна). Доверительный интервал для дисперсии нормального распределения. Доверительный интервал для биномиального распределения.
16. Построение доверительного интервала на основе асимптотически нормальных статистических оценок. Многомерное нормальное распределение. Доверительный эллипсоид.
17. Метод моментов. Обобщенный метод моментов. Метод наименьших квадратов и его геометрическая интерпретация.
18. Требования к статистическим оценкам. Состоятельность, несмещенность, асимптотическая нормальность.
19. Несмещенные оценки. Смещение оценки. Неравенство Рао -Крамера. Информационное количество Фишера. Эффективность выборочного среднего в нормальной модели. Экспонентные семейства
20. Асимптотическая нормальность. Стандартная ошибка. Асимптотическая эффективность. Нижняя граница в асимптотически эффективном оценивании.
21. Достаточные статистики. Примеры. Характеризация достаточных статистик
22. Теорема Рао-Блекуэлла-Колмогорова
23. Проверка статистических гипотез. Простые и сложные гипотезы. Статистический критерий. Область принятия гипотезы. Критическая область. Тестовая статистика. Вероятности ошибок первого и второго рода. Статистические аспекты выбора гипотезы и альтернативы.
24. Лемма Неймана-Пирсона. Наиболее мощные критерии. Критерий отношения правдоподобия. Замечание к лемме Неймана Пирсона.
25. Равномерно наиболее мощные (РНМ) и равномерно наиболее мощные несмещенные (РНМН) критерии в нормальной модели. Примеры. Содержательный смысл понятия несмещенности.
26. Связь доверительного оценивания и проверки гипотез

27. Байесовские критерии проверки сложных гипотез. Критерий максимума апостериорной вероятности и максимума правдоподобия для проверки сложных гипотез
  28. Локальная асимптотическая нормальность. Связь нормальной модели статистических выводов и моделей асимптотической статистики. В каждом вопросе по асимптотической статистике предварительно объяснить как он выглядит для нормальной модели.
  29. Асимптотическая эффективность по Питману
  30. Связь асимптотически эффективных критериев и асимптотически эффективных оценок.
  31. Байесовский подход к теории оценивания. Априорная и апостериорная плотности распределения. Апостериорный байесовский риск.
  32. 3 интерпретации байесовского подхода. Примеры байесовских оценок. Оценки максимума апостериорной вероятности.
  30. Минимаксный подход к задачам статистического оценивания. Связь минимаксного и байесовского рисков. Минимаксность выборочного среднего.
  31. Функция правдоподобия. Оценки максимума правдоподобия (ОМП). Содержательный смысл ОМП. ОМП как оценки минимума относительной энтропии.
  32. Состоятельность омп. Закон больших в пространстве  $C$ .
  33. Асимптотическая нормальность и эффективность ОМП.
  34. Построение по любой состоятельной оценке асимптотически эффективной и асимптотически нормальной оценки.
  35. М-оценки Их асимптотическая нормальность. Связь с робастностью.
  36. Критерии максимума правдоподобия, Вальда и Рао.
  37. Корреляционный и регрессионный анализ. Выборочная ковариация и корреляция. Уметь считать.
  38. Функция регрессии. Многомерная линейная регрессия. Метод наименьших квадратов. Матричная запись метода наименьших квадратов. Геометрическая интерпретация метода наименьших квадратов.
  39. Многомерные гауссовские наблюдения и модель линейной регрессии.
  40. Простая линейная регрессия. Свойства оценок МНК простой линейной регрессии.
  41. Коэффициент детерминации, его содержательный смысл и связь с выборочным коэффициентом корреляции. Основное тождество " $SST = SSE + SSR$ " его содержательный смысл.
  42. Гребневая регрессия. Решение линейной некорректной задачи методом регуляризации
  43. МНК как ОМП. Метод регуляризации как байесовская оценка. LASSO.
  44. Нелинейная регрессия. МНК. Его геометрическая интерпретация. Асимптотическая нормальность оценок МНК (зачем она нужна). Оценки МНК для регрессии как ОМП.
  45. Логистическая регрессия. Оценка максимума правдоподобия (уметь ее написать и составить уравнение ОМП. ). Где она применяется.
  46. Критерии согласия Критерии Колмогорова и омега-квадрат. Преобразование Смирнова. Критерий хи-квадрат.
  47. Джекнайф. Оценка дисперсии и устранение смещения
  48. Бутстрап. Параметрический и непараметрический. Оценка дисперсии, оценка квантилей, устранение смещения. Доверительный интервал.
  49. Классификация с обучением и без. Функция риска. Экссесс риск. Избыточный риск. Переобучение и недообучение.
  50. Задача байесовской классификации. Байесовский классификатор. Связь с задачей оценивания регрессии. Плагин метод.
  51. Плагин метод. Оценка его функции риска
  52. Задача минимизации эмпирического риска. Функции потерь для SVM, ADABoost, логистическая, нейронных сетей.
  53. Оценка избыточного риска через равномерные законы больших чисел. 2 их примера
- Очень сложные формулы, распределения, вид доверительных интервалов, константы разрешается точно не помнить - однако это на мое усмотрение На каждый вопрос программы иметь краткий

ответ без шпаргалки -чтобы ответить перед моими глазами Обязательно прочтите в конце подготовки программу и поймите, что вы можете ответить, а что нет

При решении задачи, если не уверены в правильности перевода, пишите его и подходите ко мне. Когда задача решена открываете тетрадку проверяете по основным ошибкам упомянутым к ней. При решении задачи. на листке обязательно указывается фамилия, постановка задачи (например вид гипотезы и альтернативы), а также основная формула, по которой решается задача. На решение задачи дается 35 минут. На вопрос 45 минут. Если вы подошли ко мне и задача неправильна минус бал, а для тех у кого не было допуска минус 2 балла. Если второй раз Вы подошли ко мне (на исправление дается 10 минут) и задача решена неправильно – уходите домой.

Последние 5 вопросов, если попались в билете отвечаете. Когда отвечаете мне как довопросы их можно отвечать на только на пять.

Кто не получит допуск -5 допзадач на экзамене + за каждое неверное решение еще две

Не пуха , ни пера -однако ошипанными не приходить

М.Ермаков