МФТИ ФПМИ

Госэкзамен по ДиСкРеТнОй МаТеМаТиКе

Гагаринов Даниил

Кулапин Артур

Рухадзе Альбина

Хабутдинов Арслан

Декабрь 2021 — Январь 2022

Оглавление

1 Математическая логика и теория алгоритмов **11**

1.1 Понятия множества и подмножества. Операции над множествами, тождества. Отображения и соответствия. Сравнение множеств по мощности. Теорема Кантора-Бернштейна. Счетные множества и их свойства. Теорема кантора. 11

1.1.1 Понятия множества и подмножества. Операции над множествами, тождества. . 11

1.1.1.1 Понятия множества и подмножества. 11

1.1.1.2 Операции над множествами. 12

1.1.1.3 Тождества. 12

1.1.2 Отображения и соответствия. 12

1.1.3 Сравнение множеств по мощности 13

1.1.4 Теорема Кантора-Бернштейна 13

1.1.5 Счетные множества и их свойства. 14

1.1.6 Теорема Кантора 15

1.2 Булевы функции и пропозициональные формулы. Конъюнктивная и дизъюнктивная нормальные формы. Тавтологии. Исчисление высказываний: аксиомы, правила вывода, определение выводимости, примеры выводов. Корректность исчисления высказываний. Лемма о дедукции. Полнота исчисления высказываний: формулировка и идея доказательства. 15

1.2.1 Булевы функции и пропозициональные формулы. 15

1.2.2 Конъюнктивная и дизъюнктивная нормальные формы. 16

1.2.3 Тавтологии. 16

1.2.4 Исчисление высказываний: аксиомы, правила вывода, определение выводимости, примеры выводов. 16

1.2.5 Корректность исчисления высказываний. 18

1.2.6 Лемма о дедукции. 18

1.2.7 Полнота исчисления высказываний: формулировка и идея доказательства. . . . 19

1.3 Языки первого порядка: сигнатуры, термы, правила построения формул. Интерпретации, оценки, определение истинности формулы. Выразимость предикатов: определение, примеры, доказательство невыразимости при помощи автоморфизмов. 20

1.3.1 Языки первого порядка: сигнатуры, термы, правила построения формул. 20

1.3.2 Интерпретации, оценки, определение истинности формулы. 21

1.3.3 Выразимость предикатов: определение, примеры, доказательство невыразимости при помощи автоморфизмов. . . . 21
1.4 Общезначимые формулы первого порядка. Исчисление предикатов: формулы и правила вывода. Примеры выводов: замена порядка кванторов и взаимодействие кванторов и логических операций. Корректность исчисления предикатов: формулировка и идея доказательства. Теорема Гёделя о полноте исчисления предикатов: различные формулировки и общая схема доказательства. . . . 24
1.4.1 Общезначимые формулы первого порядка. . . . 24
1.4.2 Исчисление предикатов: формулы и правила вывода. . . . 26
1.4.2.1 Аксиомы исчисления придикатов . . . 26
1.4.2.2 Правила вывода . . . 26
1.4.3 Примеры выводов: замена порядка кванторов и взаимодействие кванторов и логических операций. . . . 27
1.4.3.1 Замена порядка кванторов . . . 27
1.4.3.2 Взаимодействие кванторов и логических операций . . . 27
1.4.4 Корректность исчисления предикатов: формулировка и идея доказательства. . . . 28
1.4.5 Теорема Гёделя о полноте исчисления предикатов: различные формулировки и общая схема доказательства. . . . 28
1.4.5.1 Дополнительная терминология . . . 28
1.4.5.2 Еще одна формулировка . . . 29
1.4.5.3 Идея доказательства теоремы Геделя о полноте . . . 29
1.5 Машины Тьюринга. Вычислимые функции. Разрешимые и перечислимые множества, их свойства. Неразрешимость проблем самоприменимости и остановки. Теорема Райса–Успенского (б/д). Теорема Клини о неподвижной точке (б/д). Существование программы, печатающей свой собственный текст. . . . 30
1.5.1 Машины Тьюринга. . . . 30
1.5.2 Вычислимые функции . . . 31
1.5.3 Разрешимые и перечислимые множества, их свойства. . . . 31
1.5.4 Неразрешимость проблем самоприменимости и остановки. . . . 32
1.5.4.1 Проблема самоприменимости . . . 32
1.5.4.2 Проблема остановки(останова) . . . 33
1.5.5 Главные универсальные вычислимые функции . . . 33
1.5.6 Теорема Райса–Успенского (б/д). . . . 33
1.5.7 Теорема Клини о неподвижной точке (б/д). Существование программы, печатающей свой собственный текст. . . . 34
1.6 Формальная арифметика. Примеры выводов в аксиоматике Пеано. Моделирование машин Тьюринга в формальной арифметике (б/д). Теорема Гёделя о неполноте: формулировка и идея доказательства. . . . 34
1.6.1 Формальная арифметика. . . . 34

1.6.2 Примеры выводов в аксиоматике Пеано. 35
1.6.3 Моделирование машин Тьюринга в формальной арифметике (б/д). 37
1.6.4 Теорема Гёделя о неполноте: формулировка и идея доказательства. 37
1.7 Лямбда-исчисление. Лямбда-термы и комбинаторы. Преобразования: альфа-конверсия и бета-редукция. Нормальная форма. Теорема Чёрча–Россера (б/д). Нумералы Чёрча. Комбинаторы, представляющие сложение и умножение. Представление логических значений и операций. Представление вычитания (б/д). Комбинатор неподвижной точки, выражение одной из функций: факториал, неполное частное, остаток по модулю или любая другая, где требуется рекурсия. 38
1.7.1 Лямбда-исчисление. Лямбда-термы и комбинаторы. 38
1.7.1.1 Синтаксис 38
1.7.1.2 Построение лямбда термов 39
1.7.1.3 Соглашения о скобках 39
1.7.1.4 Свободные и связанные переменные 39
1.7.2 Преобразования: альфа-конверсия и бета-редукция. 39
1.7.2.1 Альфа-конверсия 39
1.7.2.2 Бета-редукция 40
1.7.3 Нормальная форма. 40
1.7.4 Теорема Чёрча–Россера (б/д). 40
1.7.5 Нумералы Чёрча. Комбинаторы, представляющие сложение и умножение. Представление логических значений и операций. Представление вычитания (б/д). Комбинатор неподвижной точки, выражение одной из функций: факториал, неполное частное, остаток по модулю или любая другая, где требуется рекурсия. 40
1.7.5.1 Нумералы Чёрча. 40
1.7.5.2 Комбинаторы, представляющие сложение и умножение. 41
1.7.5.3 Представление логических значений и операций. 41
1.7.5.4 Представление вычитания (б/д). 42
1.7.5.5 Комбинатор неподвижной точки, выражение одной из функций: факториал, неполное частное, остаток по модулю или любая другая, где требуется рекурсия. 42
1.8 Измерение сложности алгоритма и сложности задачи. Классы P и NP. Сводимость задач по Карпу. NP-полнота. Теорема Кука–Левина: формулировка и идея доказательства. Сводимость общей задачи о выполнимости к задаче о выполнимости 3-КНФ и сводимость задачи о выполнимости 3-КНФ к одной из задач: клика, вершинное покрытие, 3-раскраска, гамильтонов путь, задача о рюкзаке, целочисленное линейное программирование – или другой подобной. 43
1.8.1 Измерение сложности алгоритма и сложности задачи. 43
1.8.2 Классы P и NP. 43

1.8.2.1 Класс P . . . 43
1.8.2.2 Класс NP . . . 44
1.8.3 Сводимость задач по Карпу. NP-полнота. . . . 44
1.8.4 Теорема Кука–Левина: формулировка и идея доказательства. . . . 45
1.8.5 Сводимость общей задачи о выполнимости к задаче о выполнимости 3-КНФ и сводимость задачи о выполнимости 3-КНФ к одной из задач: клика, вершинное покрытие, 3-раскраска, гамильтонов путь, задача о рюкзаке, целочисленное линейное программирование – или другой подобной. . . . 45
1.8.5.1 Сводимость общей задачи о выполнимости к задаче о выполнимости 3-КНФ . . . 45
1.8.5.2 Сводимость 3-КНФ к задаче о клике . . . 46

2 Дискретные структуры **49**

2.1 Основные правила комбинаторики: правило сложения, правило умножения. Принцип Дирихле. Формула включения-исключения: доказательство, применение для вывода формулы для числа беспорядков. Базовые комбинаторные конфигурации: размещения, перестановки и сочетания. Формулы для количеств размещений, перестановок и сочетаний. Формула Стирлинга (б/д). . . . 49
2.1.1 Основные правила комбинаторики: правило сложения, правило умножения. Принцип Дирихле . . . 49
2.1.2 Формула включения-исключения: доказательство, применение для вывода формулы для числа беспорядков . . . 50
2.1.3 Базовые комбинаторные конфигурации: размещения, перестановки и сочетания. Формулы для количеств размещений, перестановок и сочетаний. Формула Стирлинга (б/д) . . . 51
2.2 Формула бинома Ньютона, полиномиальная формула. Свойства биномиальных коэффициентов: симметричность, унимодальность, рекуррентная формула треугольника Паскаля. Знакопеременная сумма биномиальных коэффициентов. Оценки для биномиальных коэффициентов при n: асимптотика C_n^k в случае $k = const \cdot n$ и в случае $k = o(\sqrt{n})$ . . . 52
2.2.1 Бином Ньютона. Свойства биномиальных коэффициентов: симметричность, унимодальность, рекуррентная формула треугольника Паскаля. Знакопеременная сумма биномиальных коэффициентов . . . 52
2.2.2 Полиномиальная формула . . . 53
2.2.3 Оценки для биномиальных коэффициентов при n: асимптотика C_n^k в случае $k = const \cdot n$ и в случае $k = o(\sqrt{n})$ . . . 54

2.3 Формальные степенные ряды: определение, операции над рядами (сумма, разность, произведение, частное, производная). Теорема Коши–Адамара о радиусе сходимости (с доказательством). Производящие функции: определение, примеры производящих функций для последовательности биномиальных коэффициентов и для чисел Фибоначчи. Пример применения производящих функций для доказательства комбинаторных тождеств. . . . 56
2.3.1 Формальные степенные ряды: определение, операции над рядами (сумма, разность, произведение, частное, производная) . . . 56
2.3.2 Теорема Коши–Адамара о радиусе сходимости (с доказательством) . . . 57
2.3.3 Производящие функции: определение, примеры производящих функций для последовательности биномиальных коэффициентов и для чисел Фибоначчи. Пример применения производящих функций для доказательства комбинаторных тождеств . . . 59
2.4 Линейные рекуррентные соотношения с постоянными коэффициентами (л.р.с.п.к.). Пример: числа Фибоначчи. Общий вид решения в произвольном случае (б/д). Доказательство теоремы об общем виде решения у л.р.с.п.к. второго порядка (в том числе при кратных корнях характеристического многочлена). Применение теоремы для нахождения формулы для чисел Фибоначчи . . . 60
2.4.1 Линейные рекуррентные соотношения с постоянными коэффициентами . . . 60
2.4.2 Доказательство теоремы об общем виде решения у л.р.с.п.к. второго порядка для разных корней . . . 61
2.4.3 Пример: Числа Фибоначчи . . . 61
2.4.4 Доказательство теоремы об общем виде решения у л.р.с.п.к. второго порядка для одинаковых корней . . . 62
2.4.5 Общий вид решения в произвольном случае (б/д) . . . 63
2.5 Определение простого графа, орграфа, мультиграфа, псевдографа, гиперграфа. Маршруты в графах, степени вершин. Изоморфизм графов, гомеоморфизм графов. Планарность графов: определение планарного графа, формула Эйлера, верхняя оценка числа рёбер в планарном графе. Критерий Понтрягина–Куратовского (доказательство необходимости; достаточность без доказательства). Эйлеровы и гамильтоновы циклы в графах: критерий эйлеровости, достаточное условие гамильтоновости (теорема Дирака). Признак Эрдеша-Хватала (б/д). . . . 64
2.5.1 Понятия теории графов . . . 64
2.5.1.1 Виды графов . . . 64
2.5.1.2 Структурные определения . . . 64
2.5.1.3 Отношения на графах . . . 65
2.5.2 Планарность графов . . . 65
2.5.3 Эйлеровость . . . 67
2.5.4 Гамильтоновость . . . 68

2.6 Хроматическое число, число независимости, кликовое число. Нижняя оценка хроматического числа через число независимости и через кликовое число; сравнение порядка этих оценок для случайного графа $G(n, 1/2)$ в модели Эрдёша–Реньи (а.п.н. $\alpha(G) \leq 2\log_2(n)$). Теорема Эрдёша о существовании графов с произвольно большим обхватом и хроматическим числом. . . . 69
2.6.1 Характеристики графа . . . 69
2.6.2 Существование графов с произвольно большим обхватом и хроматическим числом 71
2.7 Системы общих представителей (с.о.п.): определение, примеры задач, сводящихся к построению с.о.п.. Тривиальные верхняя и нижняя оценки размера минимальной с.о.п.. Жадный алгоритм построения с.о.п., теорема о верхней оценке размера «жадной с.о.п.». Теорема о неулучшаемости этой оценки в общем случае (б/д). . . . 72
2.7.1 Системы общих представителей . . . 72
2.7.1.1 «Тривиальные» нижние и верхние оценки . . . 72
2.7.2 Жадный алгоритм . . . 73
2.7.3 Неулучшаемость жадной соп . . . 74
2.8 Числа Рамсея: определение, и точные значения $R(s, t)$ при $s \leq 3, t \leq 4$. Верхняя оценка Эрдёша–Секереша, её следствие для диагональных чисел Рамсея; нижняя оценка диагональных чисел с помощью простого вероятностного метода. . . . 74
2.8.1 Числа Рамсея $R(s, t)$ точные значения для $s + t \leq 7$ . . . 74
2.8.2 Рекуррентная верхняя оценка Эрдеша-Секереша . . . 76
2.8.3 Нижняя оценка диагональных чисел с помощью простого вероятностного метода 76
2.9 Гиперграфы. Гиперграфы t-пересечений. Теорема Эрдёша–Ко–Радо (о максимальном числе рёбер в гиперграфе 1-пересечений). Основы линейно-алгебраического метода: теорема Франкла–Уилсона (верхняя оценка на $m(n, r, s)$ через числа сочетаний для случай $r - s = p$, p простое, $r < 2p$), конструктивная нижняя оценка чисел Рамсея (формулировка, определение графа, док-во леммы про число независимости). . . . 77
2.9.1 Гиперграфы, ЭКР . . . 77
2.9.1.1 Гиперграфы t-пересечений . . . 77
2.9.1.2 Теорема Эрдеша–Ко–Радо (о максимальном числе ребер в гиперграфе 1-пересечений) . . . 78
2.9.2 Теорема Франкла-Уилсона . . . 79
2.9.3 Конструктивная нижняя оценка чисел Рамсея . . . 80
2.9.3.1 Формулировка . . . 80
2.9.3.2 Определение графа . . . 80
2.9.3.3 Лемма про число независимости . . . 80

3 Теория вероятностей и математическая статистика 83

3.1 Вероятностное пространство, аксиомы Колмогорова, свойства вероятностной меры (в том числе теорема о непрерывности вероятностной меры с доказательством). Условные вероятности. Формула полной вероятности. Формула Байеса. Независимость. 83
3.1.1 Вероятностное пространство (Ω, $\mathcal{F}$, $\mathcal{P}$). Свойства меры. 83
3.1.1.1 Вероятностное пространство 83
3.1.1.2 Свойства вероятности 84
3.1.1.3 Теорема о непрерывности вероятности в нуле 84
3.1.2 Условные вероятности. Независимость. 86
3.1.2.1 Условные вероятности. 86
3.1.2.2 Формулы полной вероятности и Байеса 86
3.1.2.3 Независимость событий и систем событий 87
3.1.2.4 Пример Бернштейна 87
3.2 Случайные величины и векторы. Характеристики случайной величины и вектора: распределение вероятностей, функция распределения и её свойства, σ-алгебра, порожденная случайной величиной. Примеры конкретных распределений. 87
3.2.1 Случайные величины и векторы, их характеристики 87
3.2.1.1 Случайные величины 87
3.2.1.2 Распределение вероятностей, функция распределения и её свойства . . 88
3.2.1.3 Случайные векторы 89
3.2.1.4 σ-алгебра, порожденная случайной величиной 90
3.2.2 Примеры распределений 91
3.3 Математическое ожидание случайной величины: определение для простых, неотрицательных и произвольных случайных величин. Основные свойства математического ожидания (доказательства только для простых величин). Дисперсия и ковариация, их свойства. 92
3.3.1 Матожидание 92
3.3.1.1 Абсолютно непрерывные величины 92
3.3.1.2 Дискретные величины 93
3.3.1.3 Свойства матожидания 93
3.3.2 Дисперсия и ковариация 93
3.4 Сходимость случайных величин: по вероятности, по распределению, почти наверное, в среднем. Связь между сходимостями (б/д). Лемма Слуцкого (б/д). Теорема о наследовании сходимости. Дельта-метод. 95
3.4.1 Сходимость случайных величин. Связь между сходимостями. 95
3.4.2 Лемма Слуцкого (б/д). Теорема о наследовании сходимости. Дельта-метод. . . . 95
3.4.2.1 Лемма Слуцкого 95
3.4.2.2 Наследование сходимости 96
3.4.2.3 Дельта-метод 97

3.5 Неравенство Маркова, неравенство Чебышёва. Закон больших чисел в форме Чебышёва. Усиленные законы больших чисел (б/д). . . . 97
3.5.1 Неравенства Маркова и Чебышёва . . . 97
3.5.1.1 Неравенство Маркова . . . 97
3.5.1.2 Неравенство Чебышёва . . . 98
3.5.2 ЗБЧ в форме Чебышёва. УЗБЧ . . . 98
3.6 Характеристические функции случайных величин и векторов и их свойства. Теорема непрерывности (б/д). . . . 98
3.6.1 Характеристические функции . . . 98
3.6.1.1 Теорема о единственности (б/д) . . . 99
3.6.1.2 Критерий независимости (б/д) . . . 99
3.6.1.3 Формула обращения (б/д) . . . 99
3.6.1.4 Теорема о непрерывности (б/д) . . . 100
3.6.2 Свойства характеристических функций . . . 100
3.7 Центральная предельная теорема для независимых одинаково распределенных случайных величин. . . . 101
3.7.1 ЦПТ . . . 101
3.8 Выборка, выборочное пространство. Точечные оценки параметров и их основные свойства: несмещенность, состоятельность, асимптотическая нормальность. Выборочные среднее, медиана, дисперсия. Сравнение оценок, функция потерь и функция риска. Подходы к сравнению оценок: равномерный, байесовский, асимптотический. . . . 102
3.8.1 Выборочное пространство (one more iconic trio) . . . 102
3.8.2 Точечные оценки . . . 103
3.8.2.1 Несмещённость оценок . . . 103
3.8.2.2 Состоятельность и сильная состоятельность оценок . . . 103
3.8.2.3 Асимптотическая нормальность оценок . . . 104
3.8.2.4 Выборочные среднее, медиана, дисперсия . . . 104
3.8.3 Сравнение оценок . . . 104
3.8.3.1 Равномерный подход . . . 104
3.8.3.2 Байесовский подход . . . 105
3.8.3.3 Асимптотический подход . . . 105
3.9 Методы построения оценок: метод моментов и метод максимального правдоподобия. Состоятельность оценки метода моментов. Теорема о свойствах оценок максимального правдоподобия (б/д). . . . 106
3.9.1 Метод моментов . . . 106
3.9.2 Свойства оценки по методу моментов . . . 106
3.9.3 Метод максимального правдоподобия . . . 107
3.9.4 Условия регулярности модели . . . 107
3.9.5 Экстремальное свойство функции правдоподобия . . . 107

3.9.6 Состоятельность оценки максимального правдоподобия . . . 108
3.9.7 Асимптотическая нормальность ОМП для одномерного параметра (б/д) . . . 108
3.9.8 Эффективность ОМП . . . 108
3.10 Доверительные интервалы. Метод центральной статистики. Метод построения асимптотических доверительных интервалов. . . . 109
3.10.1 Доверительные интервалы . . . 109
3.10.2 Метод центральной статистики . . . 110
3.10.3 Асимптотические доверительные интервалы . . . 110
3.10.4 Построение асимптотических ДИ с помощью асимптотически нормальных оценок 110
3.11 Статистические гипотезы, ошибки первого и второго рода, уровень значимости критерия. Принципы сравнения критериев, равномерно наиболее мощные критерии. Лемма Неймана–Пирсона. Построение с её помощью наиболее мощных критериев. . . . 111
3.11.1 Проверка статистических гипотез . . . 111
3.11.2 Сравнение критериев . . . 112
3.11.3 Лемма Неймана-Пирсона . . . 112

4 Материалы **115**

Глава 1

Математическая логика и теория алгоритмов

1.1 Понятия множества и подмножества. Операции над множествами, тождества. Отображения и соответствия. Сравнение множеств по мощности. Теорема Кантора-Бернштейна. Счетные множества и их свойства. Теорема кантора.

1.1.1 Понятия множества и подмножества. Операции над множествами, тождества.

1.1.1.1 Понятия множества и подмножества.

Def. *Множество* — набор (совокупность) объектов. Неопределяемое понятие вообще говоря.

Def. A является *подмножеством* B, если

$$A \subset B \Leftrightarrow \forall x \in A \hookrightarrow x \in B$$

Def. $A = B \Leftrightarrow A \subset B \wedge B \subset A$

Properties.

- Рефлективность

$$\forall A \hookrightarrow A \subset A,\ A = A$$

- Антисимметричность

$$A \subset B \wedge B \subset A \Rightarrow A = B$$

- Симметричность равенства
$$A = B \Leftrightarrow B = A$$
- Транзитивность
$$A \subset B \wedge B \subset C \Rightarrow A \subset C$$

1.1.1.2 Операции над множествами.

Def. *Пересечение* $A \cap B = \{x \mid x \in A \wedge x \in B\}$

Def. *Объединение* $A \cup B = \{x \mid x \in A \vee x \in B\}$

Def. *Разность множеств* $A \setminus B = \{x \mid x \in A \wedge x \notin B\}$

Def. *Симметрическая разность* $A \triangle B = \{x \mid (x \in A \wedge x \notin B) \vee (x \notin A \wedge x \in B)\}$

Def. *Дополнение* $\overline{A} = \{x | x \notin A\}$

1.1.1.3 Тождества.

- $\overline{A \cap B} = \overline{A} \cup \overline{B}$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Proof: Доказывается простым разбором случаев

1.1.2 Отображения и соответствия.

Сначала вспомним, что такое декартово произведение и декартова степень

Def. *Упорядоченная пара* $(a, b) = \{\{a, 1\}, \{b, 2\}\} = \{a, \{a, b\}\}$

Def. *Кортеж* задается рекурсивно:

1. кортеж длины 0 — $\varnothing$
2. кортеж длины $k + 1$ — $\{a, \{a, T\}\}$, где T — кортеж длины k

Def. *Декартово произведение* $A \times B \stackrel{def}{=} \{(a, b) \mid a \in B, b \in B\}$

Def. *Декартова степень* $A^n = \{(a_1, \ldots, a_n) \mid a_i \in A\}$

Def. *Соответствие* $F \subset A \times B, y \in F(x)$

Def. *Отображение* — однозначное соответствие

$$\forall x \in A \exists! \, y \in B : (x, y) \in F, \ (y = F(x))$$

Def. *Инъективное соответствие*

$$\forall x_1, x_2 : x_1 \neq x_2 F(x_1) \cap F(x_2) = \varnothing$$

Def. *Инъекция* — инъективное соответствие + отображение

Def. *Сюръективное соответствие*

$$\forall y \in B \exists x \in A : y \in F(x)$$

Def. *Сюръекция* — сюрьективное соответствие + отображение

Def. *Биекция* — сюрьекция + инъекция

Def. Пусть $F : A \to B$ — соответствие и $S \subset A$, тогда *образом* S называется $F(S) = \bigcup\limits_{s \in S} F(s)$

Def. Пусть $F : A \to B$ — соответствие и $T \subset B$, тогда *прообразом* T называется $F^{-1}(T) = \{x \mid F(x) \cap T \neq \varnothing\} = \bigcup\limits_{t \in T} F^{-1}(t) = \bigcup\limits_{t \in T} \{x \mid t \in F(x)\}$

1.1.3 Сравнение множеств по мощности

Def. $\varnothing$ — *0-элементное множество*. Если $a \in A$ и $A \setminus \{a\}$ — *n-элементное множество*, то A — $(n+1)$-элементное

Def. A — *конечное*, если A — n-элементное для некоторого $n \in \mathbb{N}$

Note. Здесь n — количество элементов и обозначается $|A|$

Def. A *равномощно* B ($A \cong B$), если $\exists$ биекция из A в B

Def. A *мощнее* B ($A > B$), если A содержит подмножество, равномощное B, но A не равномощно B

1.1.4 Теорема Кантора-Бернштейна

Th. Кантора-Бернштейна Если $A \lessapprox B$ и $B \lessapprox A$, то $A \cong B$

Proof:

$A \lessapprox B \Rightarrow \exists$ биекция $f : A \to B_1 \subset B$

$B \lessapprox A \Rightarrow \exists$ биекция $g : B \to A_1 \subset A$

Note. $A_0 = A, B_0 = B$

Положим $A_2 = g(B_1)$. Поскольку $B_1 \subset B$, то $A_2 = g(B_1) \subset g(B) = A_1$, при этом $A_2 \cong B_1 \cong A$. Получим утверждение, эквивалентное исходной теореме: если $A_2 \subset A_1 \subset A_0$ и $A_2 \cong A_0$, то $A_1 \cong A_0$. Докажем это утверждение.

Обозначим через h биекцию из A_0 в A_2. $\forall k \in \mathbb{N}, k > 2$ определим $A_k = h(A_{k-2})$. Также определим $C_k \stackrel{def}{=} A_k \setminus A_{k+1}$. Поскольку h — биекция, то $h(C_k) = C_{k+2}$ и h задает биекцию между C_k и C_{k+2}. Однако не обязательно любой элемент элемент A_0 войдет в какой-то слой. Будет еще $C = \bigcap\limits_{k=0}^{\inf} A_k$. Тогда $A_0 = C \cup C_0 \cup C_1 \cup \ldots$ Наконец, построим биекцию α между A_0 и A_1:

$$\alpha(x) = \begin{cases} x, & x \in C \cup C_1 \cup C_3 \cup C_3 \cup \ldots \\ h(x), & x \in C_0 \cup C_2 \cup C_4 \cup \ldots \end{cases}$$

Это действительно биекция: все элементы слоев с нечетными номерами, а также элементы ядра, остаются на месте, а все слои с четными номерами биективно отображаются в слои с номерами на 2 большими. Таким образом, эквивалентное утверждение и исходная теорема доказаны. ■

1.1.5 Счетные множества и их свойства.

Def. Множество A называется *счетным*, если оно равномощно множству натуральных чисел

Statement. Пусть A — счетно, а $B \subset A$. Тогда B либо конечно, либо счетно

Proof: Пронумеруем элементы исходного множества натуральными числами. Будем перебирать все элементы. Если элемент не лежит в B, то пропустим его. Если лежит, то присвоим ему следующий номер. По индукции мы либо дойдем до какого-то $k \in \mathbb{N}$ если подмножество конечное, либо построим биекцию в $\mathbb{N}$ ■

Statement. Пусть A — счетное множество, $b \notin A$. Тогда $A \cup \{b\}$ тоже счетно.

Proof: Пусть f — биекция $\mathbb{N} \to A$. Тогда определим биекцию $g : \mathbb{N} \to A \cup \{b\}$ следующим образом: $g(0) = b, g(n) = f(n-1)$. ■

Statement. Если A счетно, B конечно, то $A \cup B$ также счетно

Proof: Сдвинем все номера на $|B|$

Statement. Если A и B счетны, то $A \cup B$ счетно

$A \cup B = A \cup (B \setminus A)$. Если $B \setminus A$ конечно, то очевидно. Иначе $B \setminus A$ счетно и достаточно доказать для непересекающихся множеств. А это уже очевидно (будем брать по одному элементу из каждого множества по очереди) ■

Statement. Пусть $\forall n \in \mathbb{N} A_n$ — счетно. Тогда множество $\bigcup\limits_{n=0}^{\infty} A_n$ тоже счетно

Proof:

Будем предполагать, что все множества не пересекаются.

Пусть $\alpha_n : \mathbb{N} \to A_n$ это биекция. Построим биекцию $\beta : \mathbb{N} \setminus \{0\} \to \bigcup\limits_{n=0}^{\infty} A_n$ следующим образом: $\beta(2^n(2k+1)) = \alpha_n(k)$. Биективность следует из того, что любое положительное натуральное число единственным образом представляется в виде $2^n(2k+1)$. Дополнить до биекции с $\mathbb{N}$ можно по транзитивности равномощности. ■

1.1.6 Теорема Кантора

Def. Пусть A, B — множества. Тогда множеством B^A называется множество всех отображений из A в B

Def. *Булеаном* множества A называется множество всех подмножеств A. Обозначение: 2^A (или $P(A)$)

Th (Кантора). Любое множество A менее мощно, чем его булеан 2^A

Proof:

Ясно, что $A \lesssim 2^A$, т.к. A равномощно множеству всех одноэлементных подмножеств.

Покажем, что $A \not\cong 2^A$

Предположим обратное, тогда $\exists$ биекция $f : A \to 2^A$

Рассмотрим множество $M = \{x \mid x \notin f(x)\}$

Т.к. f — биекция, то $\exists m \in A : f(m) = M$

Тогда $m \notin f(m) \Leftrightarrow m \in M \Leftrightarrow m \in f(m)$?! ■

1.2 Булевы функции и пропозициональные формулы. Конъюнктивная и дизъюнктивная нормальные формы. Тавтологии. Исчисление высказываний: аксиомы, правила вывода, определение выводимости, примеры выводов. Корректность исчисления высказываний. Лемма о дедукции. Полнота исчисления высказываний: формулировка и идея доказательства.

1.2.1 Булевы функции и пропозициональные формулы.

Def. *Булевой функцией* от n переменных называется любое отображение из $\{0,1\}^n$ в $\{0,1\}$.

Def. Отображение $f : A \to B$ — отношение $f \subset A \times B$, не содержащее пар с одинаковым первым

членом и разными вторыми.

Examples. конъюнкция, дизъюнкция и т.д.

Def. *Пропозициональные формулы* строятся из пропозициональных переменных по следующим правилам:

- всякая пропозициональная переменная есть пропозициональная формула;
- если A — пропозициональная формула, то $\neg A$ — тоже пропозициональная формула;
- если A и B — пропозициональные формулы, то $(A \wedge B), (A \vee B), (A \to B)$ — пропозициональные формулы.

1.2.2 Конъюнктивная и дизъюнктивная нормальные формы.

Def. *Литерал* — переменная либо отрицание переменной.

Def. *Конъюнкт* — конъюнкция литералов.

Def. *Дизъюнкт* — дизъюнкция литералов.

Def. *Конъюнктивная нормальная форма (КНФ)* — конъюнкция дизъюнктов ($n \geq 1$).

Example: При $n > 1$: $(p \vee \neg q \vee r) \wedge (\neg p \vee r) \wedge (p \vee q \vee \neg r)$ — КНФ.

Def. *Дизъюнктивная нормальная форма (ДНФ)* — дизъюнкция конъюнктов ($n \geq 1$).

Example: При $n > 1$: $(p \wedge \neg q \wedge r) \vee (\neg p \wedge r) \vee (p \wedge q \wedge \neg r)$ — ДНФ.

Note. При $n = 1$ один конъюнкт — это и ДНФ как дизъюнкция одного конъюнкта, и КНФ как конъюнкция дизъюнктов, в каждом из которых по одному литералу.

1.2.3 Тавтологии.

Def. *Тавтологии* — формулы, истинные при всех значениях входящих в них переменных.

Example: $((p \wedge q) \to p)$ — тавтология.

1.2.4 Исчисление высказываний: аксиомы, правила вывода, определение выводимости, примеры выводов.

Ниже представлены схемы аксиом, т.е. выражения, которые становятся аксиомами после подстановки в них конкретных формул вместо символов A, B, C.

Аксиомы про импликацию:

1. $A \to (B \to A)$: истина следует из чего угодно.
2. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$: левая дистрибутивность импликации относительно самой себя.

Аксиомы про конъюнкцию:

3. $(A \wedge B) \to A$;
4. $(A \wedge B) \to B$: из конъюнкции двух формул следует каждая из формул;
5. $A \to (B \to (A \wedge B))$: если выполнены обе формулы, то выполнена их конъюнкция.

Аксиомы про дизъюнкцию:

6. $A \to (A \vee B)$;
7. $B \to (A \vee B)$: дизъюнкция двух формул следует из каждой из них;
8. $(A \to C) \to ((B \to C) \to ((A \vee B) \to C))$: если формула C следует из каждой из формул A и B, то она следует и из их дизъюнкции. (правило разбора случаев)

Аксиомы про отрицание:

9. $\neg A \to (A \to B)$: из лжи следует всё, что угодно;
10. $(A \to B) \to ((A \to \neg B) \to \neg A)$: правило рассуждения от противного, если из A следует и B, и $\neg B$, то само A обязано быть неверным;
11. $A \vee \neg A$: закон исключенного третьего.

Def. Единственное правило вывода — *modus ponens* (правило отделения, исключение импликации, гипотетический силлогизм): $\dfrac{A \; A \to B}{B}$.

Понимается оно так: если ранее уже выведены формулы A и $A \to B$, то в вывод можно приписать B.

Def. *Вывод* — конечная последовательность формул, каждая из которых либо является аксиомой, либо получается из ранее встретившихся по правилам вывода.

Def. Формула называется *выводимой*, если она встречается в некотором выводе. Утверждение о том, что формула φ выводима в исчислении высказываний (ИВ), записывается так: $\vdash \varphi$.

Example: Вывод формулы вида $A \to A$ в исчислении высказываний.

Statement. Закон тождества выводим, т.е. $A \to A$.

Proof:

1. $A \to ((A \to A) \to A)$ — аксиома 1;
2. $A \to (A \to A)$ — аксиома 1;
3. $(A \to ((A \to A) \to A)) \to ((A \to (A \to A)) \to (A \to A))$ — аксиома 2;
4. $(A \to (A \to A)) \to (A \to A)$ — modus ponens из 1 и 3;
5. $A \to A$ — modus ponens из 2 и 4.

1.2.5 Корректность исчисления высказываний.

Def. Множество *противоречивое*, если из множества формул выводится A и $\neg A$.

Def. Множество формул называется *совместным*, если есть набор, на котором выполнены все формулы.

Th (о корректности). Любая формула исчисления высказываний есть тавтология.

Proof: Все аксиомы — тавтологии. Если A и $A \to B$ это тавтологии, то B — тавтология ■

1.2.6 Лемма о дедукции.

Th (о дедукции). $\Gamma \cup \{A\} \vdash B \iff \Gamma \vdash A \to B$

Note. Лемма о дедукции — это связь между выводимостью и импликацией.

Proof: Справа налево: 1) добавить A во множество посылок 2) получить B применением modus ponens к A и $A \to B$ 3) имея все вышесказанное, продолжить вывод $A \to B$ до вывода B.
Слева направо: по индукции докажем утверждение: если $C_1, ..., C_n$ есть вывод из $\Gamma \cup \{A\}$, то для всех i импликация $A \to C_i$ выводима из Γ. Рассуждение по индукции будет опираться на такой принцип, объединяющийв себе и базу, и переход: если импликации $A \to C_1, ..., A \to C_{i-1}$ выводимы, то и $A \to C_i$ выводима. Разбор случаев:

- C_i - аксиома. Тогда импликация выводится в три шага: C_i; $C_i \to (A \to C_i)$ (аксиома 1); $A \to C_i$ (modus ponens).
- C_i - посылка, т.е. $C_i \in \Gamma \cup \{A\}$
  - $C_i \in \Gamma$. Годится тот же вывод, что и для аксиомы.
  - $C_i = A$. Выводить $A \to A$ умеем.
- C_i выводится по modus ponens из формул $C_i, C_k, (j, k < i)$. Тогда C_k имеет вид $C_j \to C_i$. По предп. инд. $\Gamma \vdash (A \to C_j)$, $\Gamma \vdash (A \to (C_j \to C_i))$. Добавим аксиому 2: $(A \to (C_j \to C_i)) \to ((A \to C_j) \to (A \to C_i))$ и дважды применим modus ponens.

1.2.7 Полнота исчисления высказываний: формулировка и идея доказательства.

Th (о полноте исчисления высказываний). Если формула φ является тавтологией, то тогда φ выводима.

Lemma. (ключевая) Пусть формула $\varphi|_{p_1=a_1,\dots,p_n=a_n} = b$. Тогда из $p_1^{a_1}, \dots, p_n^{a_n} \vdash \varphi^b$, где

$p^a = \begin{cases} p, & a = 1 \\ \neg p, & a = 0 \end{cases}$ (речь идет о синтаксическом равенстве: p^a — обозначение подстроки).

Consequence. Если φ — тавтология, то при любых $a_1, \dots, a_n$ $p_1^{a_1}, \dots, p_n^{a_n} \vdash \varphi$.

Proof: Основная идея доказательства: если φ — это тавтология, то она истинна на любом наборе переменных, то есть из любого набора таких литералов будет выводиться сама формула, а не ее отрицание. Иначе говоря, если φ — тавтология, то:

$$p_1^{a_1}, \dots, p_n^{a_n} \vdash \varphi$$

Далее нужно доказать, что φ выводится и без всяких предположений. К примеру, пусть:

$$p, q, r \vdash \varphi; \quad \neg p, q, r \vdash \varphi$$

Используя правило разбора случаев, а также закон исключенного третьего, можно вывести, что $q, r \vdash \varphi$, то есть:

$$\frac{p, q, r \vdash \varphi \quad \neg p, q, r \vdash \varphi}{q, r \vdash \varphi}$$

Аналогично можно показать, что:

$$\frac{p, \neg q, r \vdash \varphi \quad \neg p, \neg q, r \vdash \varphi}{\neg q, r \vdash \varphi}$$

Теперь применим данные рассуждения еще раз:

$$\frac{q, r \vdash \varphi \quad \neg q, r \vdash \varphi}{r \vdash \varphi} \qquad \frac{q, \neg r \vdash \varphi \quad \neg q, \neg r \vdash \varphi}{\neg r \vdash \varphi}$$

Применив те же рассуждения в последний раз, получим:

$$\frac{r \vdash \varphi \quad \neg r \vdash \varphi}{\vdash \varphi}$$

Такое рассуждение можно провести не только для трех литералов, но и для любого количества. Значит, теорема о полноте доказана. ■

1.3 Языки первого порядка: сигнатуры, термы, правила построения формул. Интерпретации, оценки, определение истинности формулы. Выразимость предикатов: определение, примеры, доказательство невыразимости при помощи автоморфизмов.

1.3.1 Языки первого порядка: сигнатуры, термы, правила построения формул.

Символы алфавита:

- *Индивидные переменные*, обычно буквы x, y, z, t, u, v, w, возможно с индексами. Индивидные переменные обозначают некоторые объекты математического мира: числа, точки, множества, вершины графа и т.д.
- *Символы сигнатуры*, обозначающие те или иные связи между объектами. Сигнатура включает в себя:
  - Функциональные символы, обычно буквы f, g, h, а также стандартные знаки операций $+, -, \cdot, /$ и т.д. С каждым функциональным символом связано натуральное число — его валентность (местность, арность). Иногда валентность пишут как верхний индекс в скобках, например, $f^{(3)}, <^{(2)}$. (функциональные символы валентности ноль часто выделяют отдельно как константные символы. Примерами таких символов являются $0, \pi$ или $\varnothing$).
  - Предикатные символы, обычно буквы P, Q, R, а также стандартные знаки отношений $=, <, \ldots, \subset$ и т.д. С каждым предикатным символом также связана валентность. Предикатные символы нулевой валентности обычно не рассматривают. Валентность также могут писать как верхний индекс в скобках: $P^{(3)}, <^{(2)}, Prime^{(1)}$ (последнее означает простоту числа).
- *Символы логических операций* $\wedge, \vee, \to, \neg$;
- *Кванторы* $\forall, \exists$;
- *Служебные символы:* скобки и запятые.

Def. *Термом* называется строка, рекурсивно построенная по следующим правилам:

- Индивидная переменная есть терм;
- Функциональный символ валентности ноль (т.е. константный символ) есть терм;

- Если $k > 0, f$ — функциональный символ валентности k, а $t_1, \ldots, t_k$ — термы, то $f(t_1, \ldots, t_k)$ также терм.

— это слова, которые обозначают некоторые математические объекты, например: $x, y + 2, \frac{\sqrt{z}+2^7}{\sin(x^2+y^y)}$ и т.д.

Def. *Атомарной формулой* называется выражение вида $P(t_1, \ldots, t_k)$, где $k > 0, t_1, \ldots, t_k$ — термы, а P — предикатный символ валентности k.

Def. *Формулой (первого порядка)* называется строка, рекурсивно построенная по следующим правилам:

- Атомарная формула является формулой;
- Если φ и ψ являются формулами, то строки $(\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \to \psi), \neg\varphi$ также являются формулами;
- Если φ является формулой, а x — индивидная переменная, то $\exists x\varphi$ и $\forall x\varphi$ также являются формулами.

1.3.2 Интерпретации, оценки, определение истинности формулы.

Def. Интерпретацией сигнатуры σ называется набор $\mathcal{I}$ из следующих объектов:

- Непустое множество M, называемое *носителем интерпретации*;
- Для каждого k-местного функционального символа $f \in \sigma$ задана некоторая функция $[F] : M^k \to M$. (Если $k = 0$, рассматриваем просто элемент M — константу. Непустота M в частности требуется для того, чтобы можно было интерпретировать константные символы).
- Для каждого k-местного предикатного символа $P \in \sigma$ задан k-местный предикат, т.е. функция $[P] : M^k \to \{0, 1\}$, также его можно интерпретировать как множество $[P] \subset M^k$.

Def. *Оценкой* переменных называется функция $\pi : Var \to M$, где Var — множество индивидных переменных.

Def. Формула φ называется *общезначимой*, если она истинна при любой интерпретации на любой оценке. Формулы φ и ψ называются *эквивалентными*, если общезначима формула $\varphi \leftrightarrow \psi$.

1.3.3 Выразимость предикатов: определение, примеры, доказательство невыразимости при помощи автоморфизмов.

Зафиксируем некоторую сигнатуру σ и её интерпретацию с носителем M.

Def. Формула с параметрами $x_1, \ldots, x_m$ *выражает* предикат $P : M^m \to \{0, 1\}$, если $\varphi(a_1, \ldots, a_m) = 1 \iff P(a_1, \ldots, a_m) = 1$.

Def. Соответствующие выразимым предикатам подмножества множества M^m (области истинности выразимых предикатов) также называют *выразимыми*.

Example: $x \geq y \iff \exists z : x = y+z$ в $\mathbb{N}$. Предикат $\geq$ выразим в интерпретации $\langle \mathbb{N}, +, = \rangle$ и невыразим в интерпретации $\langle \mathbb{Z}, +, = \rangle$.

Def. Пусть даны две интерпретации одной и той же сигнатуры $\sigma : \mathcal{I}_1$ с носителем M_1 и $\mathcal{I}_\in$ с носителем M_2. Тогда *гомоморфизмом* интерпретаций называется функция $\eta : M_1 \to M_2$, такая что:

- Для любого предикатного символа $P \in \sigma$ валентности k и любых $x_1, \ldots, x_k$ выполнено

$$[P]_1(x_1, \ldots, x_k) \leftrightarrow [P]_2 \left(\eta(x_1), \ldots, \eta(x_k)\right),$$

  где $[P]_j$ — интерпретация P в I_j.

- Для любого функционального символа $f \in \sigma$ валентности k и любых $x_1, \ldots, x_k$ выполнено

$$\eta\left([f]_1(x_1, \ldots, x_k)\right) = [f]_2 \left(\eta(x_1), \ldots, \eta(x_k)\right),$$

  где $[f]_j$ — интерпретация f в $\mathcal{I}_j$.

Неформально говорят, что гомоморфизм «сохраняет» или «уважает» предикаты и функции из сигнатуры. Заметим, что если в сигнатуре есть константные символы, то гомоморфизм должен переводить их интерпретации одну в другую. Действительно, из условия на функции получаем $\eta\left([c]_1\right) = [c]_2$.

Def. Если η является гомоморфизмом и инъекцией, то он называется *мономорфизмом*. Если η является гомоморфизмом и сюръекцией, то он называется *эпиморфизмом*. Если η является гомоморфизмом и биекцией, то он называется *изоморфизмом*.

Заметим, что если в сигнатуре есть равенство, то любой гомоморфизм является мономорфизмом, а любой эпиморфизм — изоморфизмом.

Def. Если $\mathcal{I}_1 = \mathcal{I}_2$, то гомоморфизм называется *эндоморфизмом*, а изоморфизм — *автоморфизмом*.

Доказательства невыразимости строятся на следующем принципе: любой автоморфизм «сохраняет» не только предикаты из сигнатуры, но и любые выразимые предикаты. Значит, если для некоторого автоморфизма предикат не сохраняется, то он не может быть выразимым. Докажем, что этот принцип работает.

Th. Пусть φ — некоторая формула первого порядка с k параметрами в сигнатуре $\sigma, \mathcal{I}$ — некоторая интерпретация этой сигнатуры, а α — автоморфизм этой интерпретации. Тогда $[\varphi](x_1, \ldots, x_k)$ эквивалентно $[\varphi](\alpha(x_1), \ldots, \alpha(x_k))$.

Доказательство. Как обычно, подобные теоремы доказываются индукцией по построению формулы. Начнём с аналогичного утверждения для термов:

Lemma. Пусть t — некоторый терм с k параметрами в сигнатуре $\sigma, \mathcal{I}$ — некоторая интерпретация этой сигнатуры, а α — автоморфизм этой интерпретации. Тогда $\alpha([t](x_1, \ldots, x_k)) = [t](\alpha(x_1), \ldots, \alpha(x_k))$.

Proof: Если терм есть переменная x, то слева и справа будет $\alpha(x)$. Если терм есть константный символ, то слева будет $\alpha(c)$, а справа c, но в силу сохранения констант они должны быть равны. Если же терм составной, то получаем

$$\begin{aligned}\alpha([t](x_1, \ldots, x_k)) &= \alpha(f([t_1](x_1, \ldots, x_k), \ldots, [t_m](x_1, \ldots, x_k))) \\ &= f(\alpha([t_1](x_1, \ldots, x_k)), \ldots, \alpha([t_m](x_1, \ldots, x_k)))) \\ &= f([t_1](\alpha(x_1), \ldots, \alpha(x_k)), \ldots, [t_m](\alpha(x_1), \ldots, \alpha(x_k))) = [t](\alpha(x_1), \ldots, \alpha(x_k)).\end{aligned}$$

■

Теперь перейдём к формулам. Утверждение для атомарных формул получается аналогично утверждению для термов:

$$\begin{aligned}[\varphi](x_1, \ldots, x_k) &\leftrightarrow P([t_1](x_1, \ldots, x_k), \ldots, [t_m](x_1, \ldots, x_k)) \\ &\leftrightarrow P(\alpha([t_1](x_1, \ldots, x_k)), \ldots, \alpha([t_m](x_1, \ldots, x_k)))) \\ &\leftrightarrow P([t_1](\alpha(x_1), \ldots, \alpha(x_k)), \ldots, [t_m](\alpha(x_1), \ldots, \alpha(x_k))) \leftrightarrow [\varphi](\alpha(x_1), \ldots, \alpha(x_k)).\end{aligned}$$

Булевы комбинации получаются непосредственно: например, если $\varphi(x_1, \ldots, x_k)$ эквивалентно $\varphi(\alpha(x_1), \ldots, \alpha(x_k))$ и $\psi(x_1, \ldots, x_k)$ эквивалентно $\psi(\alpha(x_1), \ldots, \alpha(x_k))$, то $(\varphi \wedge \psi)(x_1, \ldots, x_k)$ эквивалентно $(\varphi \wedge \psi)(\alpha(x_1), \ldots, \alpha(x_k))$.

А вот для формул с кванторами потребуется биективность. Рассмотрим для примера формулу $\exists x \varphi(x, x_1, \ldots, x_k)$. По предположению индукции она эквивалентна $\exists x \varphi(\alpha(x), \alpha(x_1), \ldots, \alpha(x_k))$. Из биективности следует, что она эквивалентна $\exists y \varphi(y, \alpha(x_1), \ldots, \alpha(x_k))$. Действительно, для импликации слева направо можно взять $y = \alpha(x)$, а для импликации справа налево — $x = \alpha^{-1}(y)$, что возможно благодаря биективности. Формула с квантором всеобщности разбирается аналогично. ■

Example: $x = \frac{1}{2}$ не выражается в $\langle [0, 1], \leqslant \rangle$, потому что можно взять $\alpha(x) = x^2$

1.4 Общезначимые формулы первого порядка. Исчисление предикатов: формулы и правила вывода. Примеры выводов: замена порядка кванторов и взаимодействие кванторов и логических операций. Корректность исчисления предикатов: формулировка и идея доказательства. Теорема Гёделя о полноте исчисления предикатов: различные формулировки и общая схема доказательства.

1.4.1 Общезначимые формулы первого порядка.

Def. Формула φ называется *общезначимой*, если она истинна при любой интерпретации на любой оценке.

Statement. Все частные случи тавтологий являются общезначимыми формулами

Statement. $\forall\varphi$ общезначима формула $\forall x\varphi \to \exists x\varphi$

Statement. $\forall\varphi$ формулы $\forall x\forall y\varphi \to \forall y\forall x$ и $\exists x\exists y\varphi \to \exists y\exists x\varphi$ общезначимы

Proof: Докажем первую импликацию.

$$[\forall x\forall y\varphi](\pi) = \bigwedge_{m\in M}[\forall y\varphi](\pi_{x=m}) = \bigwedge_{n\in M}\bigwedge_{m\in M}[\varphi](\pi_{x=m,y=n}) = \bigwedge_{m\in M}\bigwedge_{n\in M}[\varphi](\pi_{y=n,x=m}) = \bigwedge_{m\in M}\bigwedge_{n\in M}[\varphi]((\pi_{y=n})_{x=m}) = \bigwedge_{n\in M}[\forall x\varphi](\pi_{y=n}) = [\forall y\forall x\varphi](\pi)$$ ■

Statement. $\forall\varphi$ общнезначима формула $\exists x\forall y\varphi \to \forall y\exists x\varphi$

Statement. $\forall\varphi$ общезначимы эквивалентности $\neg\exists\varphi \leftrightarrow \forall x\neg\varphi$ и $\neg\forall x\varphi \leftrightarrow \exists\neg\varphi$

Statement. $\forall\phi\psi$ общезначимы формулы $\forall x(\varphi\wedge\psi) \leftrightarrow (\forall x\phi\wedge\forall x\psi)$ и $\exists x(\varphi\wedge\psi) \to (\exists x\phi\wedge\exists x\psi)$. Обращение последней импликации не будет общезначимым, но если ψ не зависит от x, то общезначима формула $\exists x(\varphi\wedge\psi) \leftrightarrow (\exists x\phi\wedge\psi)$

Proof: Докажем вторую часть.

Обращение импликации не будет общезначимым, поскольку формулы φ и ψ могуть быть истинными на разных элементах. Например, существуют простые числа и существуют полные квадраты, но не бывает простого числа, являющегося полным квадратом.

С другой стороны, если ψ не зависит от x, то либо ψ ложна при всех x, и тогда любая конъюнкция с ней ложна, так что по обе стороны от эквиваленции стоит ложь, либо ψ истинна при всех x, и тогда её можно вычеркнуть из всех конъюнкций, после чего по обе стороны от эквиваленции останется $\exists x\varphi$

Statement. $\forall\phi\psi$ общезначимы формулы $\exists x(\varphi\vee\psi) \leftrightarrow (\exists x\phi\vee\exists x\psi)$ и $(\forall x\varphi\vee\forall x\psi) \to \forall x(x\phi\vee\psi)$. Об-

ращение последней импликации не будет общезначимым, но если ψ не зависит от x, то общезначима формула $\forall x(\varphi \vee \psi) \leftrightarrow (\forall x\phi \vee \psi)$

Идея доказательства та же самая, поэтому приведём лишь пример, когда обратная импликация необщезначима. Любое натуральное число является чётным или нечётным, но неверно, что любое число чётно или любое число нечётно.

Statement. Если φ зависит только от x, а ψ только от y, то формула $\forall x\exists y(\varphi \wedge \psi) \to \exists y\forall x(\varphi \wedge \psi)$ общезначима. Аналогично для $\vee$

Statement. $\forall\phi\psi$ общезначимы формулы $\forall x(\varphi \to \psi) \to (\forall x\varphi \to \forall x\psi)$, $\forall x(\varphi \to \psi) \to (\exists x\varphi \to \forall x\psi)$, $(\exists x\varphi \to \forall x\psi) \to \forall x(\phi \to \psi)$, $\exists x(\varphi \to \psi) \leftrightarrow (\forall x\varphi \to \exists x\psi)$. Первые три импликации в обратную сторону не общезначимы, но если ψ не зависит от x, то общезначима формула $\forall x(\varphi \to \psi) \leftrightarrow (\exists x\varphi \to \psi)$, а если φ не зависит от x, то общезначима формула $\forall x(\varphi \to \psi) \leftrightarrow (\varphi \to \forall\psi)$

Proof:

Первые две импликации и прямое направление последней по сути обобщают правило modus ponens. В первой говорится, что если везде верно φ и везде верно $\varphi \to \psi$, то везде верно ψ. Во второй: если везде верно $\varphi \to \psi$ и где-то верно φ, то там же верно и ψ. В последней: если везде верно φ и где-то верно $\varphi \to \psi$, то там же верно ψ. Третья импликация обосновывается так: либо φ всюду ложно, и тогда всюду верно $\varphi \to \psi$, либо φ где-то истинно, тогда по предположению везде верно ψ и потому везде верно $\varphi \to \psi$. Альтернативно, можно воспользоваться уже доказанными импликациями:

$$(\exists x\varphi \to \forall x\psi) \to (\neg\exists x\varphi \vee \forall x\psi) \to (\forall x\neg\varphi \vee \forall x\psi) \to \forall x(\neg\varphi \vee \psi) \to \forall x(\varphi \to \psi) \tag{1.4.1}$$

Последнюю импликацию в обратную сторону также можно обосновать двумя способами. Во-первых, можно вновь разобрать случаи: либо ψ где-то истинно, и тогда там же истинно $\varphi \to \psi$, либо ψ всюду ложно, тогда ложно $\exists x\psi$ и для истинности импликации $(\forall x\varphi \to \exists x\psi)$ должно быть ложным $\forall x\psi$. Значит, φ где-то ложно и тогда там же истинно $\varphi \to \psi$. Во-вторых, можно использовать такую цепочку, теперь работающую в обе стороны:

$$(\forall x\varphi \to \exists x\psi) \leftrightarrow (\neg\forall x\varphi \vee \exists x\psi) \leftrightarrow (\exists x\neg\varphi \vee \exists x\psi) \leftrightarrow \exists x(\neg\varphi \vee \psi) \leftrightarrow \exists(\varphi \to \psi) \tag{1.4.2}$$

Обращение первой импликации может быть необщезначимым из-за того, что обе формулы $\forall x\varphi$ и $\forall x\psi$ могут быть неверными, из-за этого для какого-то конкретного x формула φ может быть истинной, а формула ψ — ложной. Например, если бы все натуральные числа были бы простыми, то все они также были бы чётными, но не каждое простое число чётно. Те же рассуждения годятся и для обращения второй импликации. Похожим образом опровергается обращение третьей импликации: $\varphi \to \psi$ может быть истинным и когда обе формулы истинны, и когда обе формулы ложны. Например, любое простое число равно двум или нечётно, и простые числа существуют, но не любое число равно двум или нечётно.

Наконец, если одна из формул не зависит от x, то выводом, аналогичным 1.4.1, можно получить соответствующую эквивалентность. ■

1.4.2 Исчисление предикатов: формулы и правила вывода.

Можно сказать, что исчисление предикатов — это то же, что исчисление высказывание, только в формулах с кванторами.

1.4.2.1 Аксиомы исчисления придикатов

- Аксиомы 1-11: аксиомы исчисления высказываний
- Аксиома 12: $\forall x\varphi \to \varphi(t/x)$, где t/x — это корректная подстановка терм t в φ вместо свободных вхождений x. Здесь могут возникнуть проблемы при неверной замене. Например:

  $$\forall x \exists y x < y \to \exists y y < y$$

  , что неверно из-за некорректной подстановки.

  Корректная подстановка означает, что терм t не содержит переменных, по которым стоят кванторы в φ.
- Аксиома 13: $\varphi(t/x) \to \exists x\varphi$

1.4.2.2 Правила вывода

1. Modus ponens:
   $$\frac{A, A \to B}{B}$$
2. Первое правило Бернайса:
   $$\frac{\varphi \to \psi}{\exists x\varphi \to \psi}$$
3. Второе правило Бернайса:
   $$\frac{\psi \to \varphi}{\psi \to \forall x\varphi}$$

   Для правил Бернайса есть важное ограничение: x не является параметром ψ

Def. *Выводом* называется конечная последовательность формул, каждая из которых либо аксиома, либо получается по одному из правил вывода.

1.4.3 Примеры выводов: замена порядка кванторов и взаимодействие кванторов и логических операций.

1.4.3.1 Замена порядка кванторов

Имеется формула:

$$\exists x \forall y \varphi \to \forall y \exists x \varphi$$

Ее вывод:

1. $\forall y \varphi \to \varphi$ (аксиома 12)
2. $\varphi \to \exists x \varphi$ (аксиома 13)
3. $\forall y \varphi \to \exists x \varphi$ (силлогизм)
4. $\exists x \forall y \varphi \to \exists x \varphi$ (первое правило Бернайса)
5. $\exists x \forall y \varphi \to \forall y \exists x \varphi$ (второе правило Бернайса)

1.4.3.2 Взаимодействие кванторов и логических операций

Имеется формула:

$$\forall x(\varphi \wedge \psi) \leftrightarrow (\forall x \varphi \wedge \forall x \psi)$$

Выведем:

$$\forall x(\varphi \wedge \psi) \to \varphi \wedge \psi; \quad \forall x(\varphi \wedge \psi) \to \varphi$$

Тогда:

$$\forall x(\varphi \to \psi) \to \forall x \varphi$$

Далее запишем:

$$\forall x(\varphi \wedge \psi) \to \forall x \psi$$

В итоге получаем:

$$\forall x(\varphi \wedge \psi) \to (\forall x \varphi \wedge \forall x \psi)$$

,

что и требовалось вывести.

В другую сторону вывод аналогичен:

$$\forall x \varphi \to \varphi; \quad \forall x \psi \to \psi$$

тогда пропозициональная комбинация:

$$(\forall x\varphi \wedge \forall x\psi) \to (\varphi \wedge \psi)$$

.

Итак, получаем искомый результат:

$$(\forall x\varphi \wedge \forall x\psi) \to \forall x(\varphi \wedge \psi)$$

1.4.4 Корректность исчисления предикатов: формулировка и идея доказательства.

Th (О корректности). Если φ — выводима, то она — общезначима

Proof:

Доказательство теоремы о корректности следует из того, что все аксиомы — общезначимы. Правила вывода преобразуют общезначимые в общезначимые.

Рассмотрим на примере аксиомы:

$$\forall x\varphi \to \varphi(t/x)$$

Пусть $\forall x\varphi$ верно в некоторой интерпретации при некоторой оценке π. Тогда φ верна в той же интерпретации при любой оценке π', совпадающей с π всюду, кроме x (так оценивалась истинность формулы с квантором).

В частности, в оценке, при которой:

$$\pi'(x) = [t](\pi)$$

можно записать

$$[\varphi(t/x)]\,(\pi) = [\varphi](\pi')$$

1.4.5 Теорема Гёделя о полноте исчисления предикатов: различные формулировки и общая схема доказательства.

Th (Геделя, о полноте исчисления предикатов). Если φ общезначима, то она выводима в исчислении предикатов.

1.4.5.1 Дополнительная терминология

Def. *Теорией* называется любое множество замкнутых формул (то есть формул, не имеющих параметров)

Def. *Модель теории* — это любая интерпретация, в котрой все формулы из данной теории истинны

Def. *Совместной теорией* называется теория, имеющая модель

Def. *Противоречивой теорией* называется теория, из которой выводится противоречие $(A \wedge \neg A)$

Пользуясь данной теорией, можно сформулировать теорему, из которой будет следовать изначальная теорема

1.4.5.2 Еще одна формулировка

Th. Если теория непротиворечива, то она совместа (имеет модель).

Поясним, почему из этой теоремы следует теорема Геделя о полноте.

Если φ — общезначима, то тогда форма отрицания $\{\neg\varphi\}$ — несовместна. Значит, по теореме система $\{\neg\varphi\}$ — противоречива, следовательно, φ — выводима. Эти рассуждения верны для замкнутых формул, распространим их на незамкнутые.

Если φ — формула со свободной переменной x, и при этом φ — общезначима, то $\forall x\varphi$. Тогда по предыдущему рассуждению получается, что $\forall x\varphi$ выводима, откуда следует, что φ выводима (аксиома 12)

1.4.5.3 Идея доказательства теоремы Геделя о полноте

С полной версией шедевра можно ознакомится тут

Нужно расширить непротиворечивую теорию Γ так, чтобы она стала:

1. Полной, то есть если φ — замнкутая формула, то $\Gamma \vdash \varphi$ или $\Gamma \vdash \neg\varphi$
2. Экзистенциально полной, то есть если $\Gamma \vdash \exists x\varphi$, то

$$\Gamma \vdash \varphi(t/x)$$

, где t — замкнутый терм.

После этого у Γ будет модель из замкнутых термов.

Пополнение Будем считать, что множество переменных, как и сигнатура — счетные. Рассмотрим все замкнутые формулы. Будем строить систему так: $\Gamma_0 = \Gamma_1, \dots,$

$$\Gamma_i = \begin{cases} \Gamma_{i-1} \cup \{\varphi_i\}, & \text{если это непротиворечиво} \\ \Gamma_{i-1} \cup \{\neg\varphi_i\}, & \text{иначе} \end{cases}$$

Тогда пополнение — это:

$$\Delta = \bigcup_{i=0}^{\infty} \Gamma_i$$

Экзистенциальное пополнение Тут нам понадобятся две леммы

Lemma (О свежих константах). Пусть выводима формула φ, в которой переменная x заменена на некоторую константу, то есть $\vdash \varphi(c/x)$, где c — константа, не встречающаяся в формуле φ. Тогда выводима и сама формула φ

Lemma (О добавлении констант). Пусть σ — некоторая сигнутару, а σ' — сигнатура, отличная от σ на некоторое количество константных символов. Тогда если φ — формула в сигнатуре σ — выводима в теории Γ сигнатуры σ при помощи сигнатуры σ', то она выводима в теории Γ сигнатуры σ при помощи сигнатуры σ

Вернемся к доказательству теоремы. Пусть $\Gamma \vdash \exists x\varphi$. Тогда добавим к сигнатуре константу c_φ, а к теории замкнутую формулу $\varphi(c_\varphi/x)$. Поступим аналогично со всеми формулами такого вида. Непротиворечивость сохранится, но автор оставит это упражнение читалям (мы же тут идею приводим).

Возникает проблема: новое Γ может перестать быть полным. Тогда снова используем пополнение.

Новая проблема: экзистенциальная полнота могла быть утрачена. Проведем счетное количество таких операций, тогда полученный результат будет: Непротиворечив, Полон, Экзистенциально полон

Осталось доказать следующую лемму

Lemma. Любая непротиворечивая, полная, экзистенциально полная теория совместа, то есть имеет модель.

С доказательством этого факта можно ознакомиться тут

■

1.5 Машины Тьюринга. Вычислимые функции. Разрешимые и перечислимые множества, их свойства. Неразрешимость проблем самоприменимости и остановки. Теорема Райса–Успенского (б/д). Теорема Клини о неподвижной точке (б/д). Существование программы, печатающей свой собственный текст.

1.5.1 Машины Тьюринга.

Def. Машиной Тьюринга называется следующий набор

$$MT = (\Sigma, \Gamma, Q, q_1, q_0, \delta)$$

, где

- Σ — алфавит (некоторое конечное множество)
- Γ — ленточный алфавит
- Q — множество состояний
- q_0, q_1 — начальное и конечное состояние
- δ — функция перехода

1.5.2 Вычислимые функции

Def. $f : \{0,1\}^* \to \{0,1\}^*$ — вычислимая, если $\exists$ МТ такая, что $\forall x$:

1. f(x) определена, то $M(x) = f(x)$
2. f(x) не определена, то M не останавливается на входе x

Statement. Если f и g вычислимы, то $f \circ g$ вычислима

1.5.3 Разрешимые и перечислимые множества, их свойства.

Пусть $S \subset \{0,1\}^*$

Def. χ_S — хар. функция множества S

Def. Множество S называется разрешимым, если χ_S вычислима, т. е. существует алгоритм M, т.ч.:

- $M(x) = 1$, если $x \in S$
- $M(x) = 0$, если $x \notin S$

Statement. S, T — разрешимы, то $S \cup T, S \cap T, \overline{S}$ — разрешимы

Statement. Любое конечное множество разрешимо

Def. Множество S перечислимо, если существует алгоритм, который печатает все элементы множества и только их (в произвольном порядке и с произвольными промежутками времени)

Statement. Множество S разрешимо $\Rightarrow$ оно перечислимо

Proof: Переберем все слова из 0 и 1, проверим лежит ли слово в множестве и напечатаем, если да

Statement. Если S, T — перечислимы, то $S \cup T, S \cap T$ — перечислимы

Note. Дополнение к перечислимому множеству не всегда перечислимо

Def. *Полухарактеристическая функция*

$$\widetilde{\chi_S}(x) = \begin{cases} 1, & x \in S \\ \text{не определено}, & x \notin S \end{cases}$$

Th. S — перечислимо $\Leftrightarrow \widetilde{\chi_S}$ вычислима

Proof:

$\Rightarrow$ Запустим перечислятор, если встретим x, то вернуть 1, иначе алгоритм не остановится (а мы этого и хотим)

$\Leftarrow$ Неформально: запустим параллельно вычисление $\widetilde{\chi_S}$ сразу на всех входах, где ответ 1 там печатаем x

Формально параллельный запуск имеет счетное число стадий. На стадии номер n нужно запустить $\widetilde{\chi_S}(0), \dots \widetilde{\chi_S}(n)$ на n шагов и вывести все x, где вычисление закончилось.

Если $\widetilde{\chi_S}(m)$ вычислится за t шагов, то тогда m будет выведено на стадии $max(m, t)$

Если $\widetilde{\chi_S}(m)$, то m не будет выведено ни на какой стадии ■

Th (Поста). S — разрешимо $\Leftrightarrow S$ и $\overline{S}$ перечислимы

1.5.4 Неразрешимость проблем самоприменимости и остановки.

Def. $U(p, x)$ — результат применения программы p ко входу x. Это *универсальная машина Тьюринга*

Def. $U : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ — универсально вычислимая функция, если:

1. Она сама вычислима
2. Если $f : \mathbb{N} \to \mathbb{N}$ вычислима, то $\exists p : \forall x U(p, x) = f(x)$

1.5.4.1 Проблема самоприменимости

Определена ли $U(p, p)$? (Останавливается ли программа p на своем номере)

Statement. $K = \{p \mid U(p, p) \text{ опр.}\}$ — перечислимо

Th. Проблема самоприменимости не разрешима.

Proof:

Пусть q — программа, разрешающая K (хар функция множества). Тогда существует программа q' такая что:

$$U(q', p) = \begin{cases} \text{не опр.}, & U(p, p) \text{ опр.} \\ 0, & U(p, p) \text{ не опр.} \end{cases}$$

Теперь

$$U(q',q') = \begin{cases} \text{не опр.}, & U(q',q') \text{ опр.} \\ 0, & U(q',q') \text{ не опр.} \end{cases}$$

?! ■

1.5.4.2 Проблема остановки(останова)

Даны программа p и вход x. Вопрос: остановится ли программа p на входе x (определена ли $U(p,x)$)

Statement. $\{(p,x) \mid U(p,x)\text{опр.}\}$ — перечислимо

Proof: Это область определения вычислимой функции U (это еще одно эквивалентное определение перечислимости).

Просто возьмем алгоритм, который для пары (p,x) будет запускать $U(p,x)$. Это будет полухар. функция и она будет вычислима. ■

Th. Проблема остановки не разрешима

Proof: Если проблема остановки разрешима, то и проблема самоприменимости разрешима ?! ■

Consequence. Существует перечислимое, но не разрешимое множество

1.5.5 Главные универсальные вычислимые функции

Def. $U : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ называется ГУВФ, если:

1. U — УВФ
2. Если $V : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ — вычислимая функция, то сущестует вычислимая всюду определенная функция $s : \mathbb{N} \to \mathbb{N} : \forall n \forall x V(n,x) = U(s(n),x)$

Неформально: s это по сути транслятор из языка V в язык U

Th. ГУВФ существует

1.5.6 Теорема Райса–Успенского (б/д).

Def. $A \subset \mathbb{N}$ называется функциональным, если из $\forall x U(n,x) = U(m,x)$ следует, что $n \in A$ и $m \in A$ или $n \notin A$ и $m \notin A$

Th (Райса-Успенского). Если A — функциональное свойство, $A \neq \varnothing, A \neq \mathbb{N} \Rightarrow A$ — неразрешимо

1.5.7 Теорема Клини о неподвижной точке (б/д). Существование программы, печатающей свой собственный текст.

Th (Клини о неподвижной точке). Пусть U — ГУВФ, $h : \mathbb{N} \to \mathbb{N}$ — тотально вычислимая (всюду определена и вычислима). Тогда $\exists n \forall x U(n, x) = U(h(n), x)$

Consequence. $\exists$ программа, печатающая свой текст

Proof: Пусть $h : p \to printp$. По теореме Клини получается что существует программа печатающая свой текст. Вычислимость h это "рукомахание"(c) Мусатов

1.6 Формальная арифметика. Примеры выводов в аксиоматике Пеано. Моделирование машин Тьюринга в формальной арифметике (б/д). Теорема Гёделя о неполноте: формулировка и идея доказательства.

1.6.1 Формальная арифметика.

Def. Формальная теория это список аксиом и все формулы, которые выводятся из них по логическим правилам.

Note. Список аксиом должен быть разрешимым, тогда множество теорем перечислимо.

Def. Вывод это последовательность формул, где каждая либо аксиома теории, либо аксиома логики (исчисление предикатов) либо выводится из предыдущих по логическим правилам.

Note. Множество корректных выводов разрешимо

Символы арифметики: $0, S, +, *, =$, где S — следующий элемент

Аксиомы Пеано:

1. $S(x) = S(y) \Rightarrow x = y$
2. $x = 0 \Leftrightarrow \neg\exists y : x = S(y)$
3. $A(0) \wedge \forall x(A(x) \Rightarrow A(S(x)) \Rightarrow \forall x A(x)$, где A — любая формула с одним параметром
4. $x + 0 = x$
5. $x + S(y) = S(x + y)$
6. $x * 0 = 0$
7. $x * S(y) = x * y + x$

1.6.2 Примеры выводов в аксиоматике Пеано.

Example: Выведем $2+2=4$, что на нашем языке означает $SS0+SS0=SSSS0$

1. $\forall x \forall y x+Sy=S(x+y)$ — аксиома
2. $SS0+SS0=S(SS0+S0)$ — подстановка $x=SS0, y=S0$
3. $SS0+S0=S(SS0+0)$ — подстановка $x=SS0, y=0$
4. $\forall x x+0=x$ — аксиома
5. $SS0+0=SS0$ — подстановка $x=SS0$
6. $\forall x \forall y (x=y \to Sx=Sy$ — аксиома равенства
7. $SS0+0=SS0 \to S(SS0+0)=SSS0$ — подстановка $x=SS0+0, y=SS0$
8. $S(SS0+0)=SSS0$ — modus ponens
9. $SS0+S0=SS0$ — по транзитивности
10. $S(SS0+S0)=SSSS0$ — подстановка $x=S(SS0+0), y=SSS0$
11. $SS0+SS0=SSSS0$ — по транзитивности

Example: Выведем коммутативност $\forall x \forall y x+y=y+x$

Для этого нам понадобятся вспомогательные формулы:

$$\forall x 0+x=x$$

1. $\forall x x+0=x$ — аксиома
2. $0+0=0$ — подстановка $x=0$, база индукции
3. $\forall x \forall y x+Sy=S(x+y)$ — аксиома
4. $0+Sx=S(0+x)$ — подстановка $x=0, y=x$
5. $\forall x \forall y (x=y \to Sx=Sy)$ — аксиома равенства
6. $0+x=x \to S(0+x)=Sx$ — подстановка $x=0+x, y=x$
7. $0+x=x \to 0+Sx=Sx$ — транзитивность + силлогизм
8. $\forall x (0+x=x \to 0+Sx=Sx)$ — правило обобщения

9. $(0+0=0 \wedge \forall x(0+x=x \to 0+Sx=Sx)) \to \forall x 0+x=x$ — принцип индукции
10. $\forall x 0+x=x$ — два раза MP

$$\forall x \forall y Sx+y=S(x+y)$$

1. $\forall x x+0=x$
2. $x+0=x$
3. $Sx+0=Sx$
4. $\forall x \forall y(x=y \to Sx=Sy)$
5. $x+0=x \to S(x+0)=Sx$
6. $S(x+0)=S(x+0)$
7. $Sx+0=S(x+0)$
8. $\forall x \forall y x+Sy=S(x+y)$
9. $SxSy=S(Sx+y)$
10. $x+Sy=S(x+y)$
11. $S(x+Sy)=SS(x+y)$
12. $(Sx+y=S(x+y)) \to (Sx+Sy=S(Sx+y)=SS(x+y)=S(x+Sy))$ т.е. $(Sx+y=S(x+y)) \to (Sx+Sy=S(x+Sy))$
13. $\forall y((Sx+y=S(x+y)) \to (Sx+Sy=S(x+Sy)))$
14. По принципу индукции из $Sx+0=S(x+0)$ и $\forall y((Sx+y=S(x+y)) \to (Sx+Sy=S(x+Sy)))$ получается $\forall y Sx+y=S(x+y)$
15. По правилу обобщения $\forall x \forall y Sx+y=S(x+y)$

И наконец $\forall x \forall y x+y=y+x$

1. $\forall x x+0=x$
2. $\forall x 0+x=x$
3. $x+0=x$
4. $0+x=x$

5. $x + 0 = 0 + x$
6. $\forall x x + 0 = 0 + x$
7. $Sy + x = S(y + x)$ из второй формулы
8. $x + Sy = S(x + y)$
9. $x + y = y + x \to x + Sy = Sy + x$
10. $\forall y(x + y = y + x \to x + Sy = Sy + x)$
11. По принципу индукции из $x + 0 = x$ и $\forall y(x + y = y + x \to x + Sy = Sy + x)$ получаем $\forall y(x + y = y + x)$
12. По правилу обобщения $\forall x \forall y(x + y = y + x)$

1.6.3 Моделирование машин Тьюринга в формальной арифметике (б/д).

Def. Кодирование Смаллиана: взаимно однозначное соответствие между числами и двоичными числами $n \to n + 1 \to bin(n + 1) \to `bin(n + 1)$

$9 \to 10 \to 1010 \to 010$

Дальше: строковые операции и отношения (конкатенация, префикс, подслово и тд) на двоичных словах можно выразить через арифметические операции над соответствующими им числами

Statement. Машину Тьюринга можно смоделировать в формальной арифметике

В чем собственно посыл? Конфигурацию машины можно закодировать как двоичное слово а значит и как число. Один шаг можно закодировать как набор строковых операций над конфигурацией, а значит и как набор арифметических операций над числами.

Далее для любой M можно выразить предикат $Step_M(x, y)$: y есть код конфигурации, следущей за конфигурацией с кодом x.

Можно закодировать и то, что Машина останавливается на входе z с ответом a

$Haly_M(z, a) = \exists(c_1, \ldots, c_n) : c_1$ — начальная конфигурация со входом z, c_n — конечная конфигурация с ответом a, $\forall i Step_M(c_i, c_{i+1})$

Profit

1.6.4 Теорема Гёделя о неполноте: формулировка и идея доказательства.

Def. $Proof(p, x)$ — арифметический предикат, означающий, что цепочка с номером p является доказательством формулы с номером x

Def. $Provable(x) \stackrel{def}{=} \exists p Proof(p, x)$ — предикат доказуемости

Th (Теорема Геделя о неполноте). Какой бы ни была система доказательств в языке арифметики, она либо неадекватна (какое-то неверное утверждение доказуемо), либо неполна (какое-то верное утверждение недоказуемо)

Идея доказательства:

Строится формула G, утверждающая собственную недоказуемость. Если она ложна, то она доказуема, поэтому есть ложная доказуемая формула. Если она истинна, то она и есть истинная недоказуемая формула. Как построить такую формулу?

Строится предикат $Subst(x, y, z)$: замкнутая формула с номером x есть результат подстановки числа z в формулу с одним аргументом номер y

$D(x) = \exists x(Subst(x, z, z) \wedge \neg \exists p Proof(p, x)))$ — результат подстановки числа z в формулу номер z недоказуем

$G = D(D) \Leftrightarrow \exists x(Subst(x, D, D) \wedge \neg \exists p Proof(p, x)) \Leftrightarrow \neg \exists p Proof(p, D(D)) \Leftrightarrow \neg Provable(G)$

1.7 Лямбда-исчисление. Лямбда-термы и комбинаторы. Преобразования: альфа-конверсия и бета-редукция. Нормальная форма. Теорема Чёрча–Россера (б/д). Нумералы Чёрча. Комбинаторы, представляющие сложение и умножение. Представление логических значений и операций. Представление вычитания (б/д). Комбинатор неподвижной точки, выражение одной из функций: факториал, неполное частное, остаток по модулю или любая другая, где требуется рекурсия.

1.7.1 Лямбда-исчисление. Лямбда-термы и комбинаторы.

"Все есть функция"

1.7.1.1 Синтаксис

- Переменные
- Служебные символы ()

- Лямбда-квантор λ.

1.7.1.2 Построение лямбда термов

1. Переменная — λ-терм
2. Если M, N — λ-термы, то (MN) — λ-терм

   Неформальный смысл: В M подставляется N в качестве аргумента
3. Если x — переменная, M — λ-терм, то $(\lambda x.M)$ — λ-терм

Def. $(\lambda x.M)$ — "M, понимаемый как ф-ия от x"

Неформальный пример:

$\lambda x.(x + y)$ — функция от x с параметром y

$\lambda y.(x + y)$ — наоборот

$\lambda x.(\lambda y.(x + y))$ — функция от двух аргументов (принято сокращение $\lambda xy.(x + y)$

1.7.1.3 Соглашения о скобках

1. Если несколько аппликаций подряд без скобок, то приоритет слева направо

   $MNPQ = (((MN)(P)Q)) \neq ((MN)(PQ)) \neq ((M(NP))Q)$
2. Лямбда абстракция имеет низкий приоритет

   $\lambda x.MN = (\lambda x.(MN)) \neq ((\lambda x.M)N)$

1.7.1.4 Свободные и связанные переменные

$\lambda x.(x\lambda z.yz)$ здесь x, z — связанные переменные, а y — свободная

1.7.2 Преобразования: альфа-конверсия и бета-редукция.

1.7.2.1 Альфа-конверсия

Def. Альфа-конверсия — корректное переименование связанной переменной. В общем виде: $\lambda x.M \to \lambda z.M(z/x)$

Note. Корректное = не должно возникать конфликтов имен, и переменные M не должны попадать под действие лямбда-квантора

Example: $\lambda x.xy \to \lambda z.zy$ — корректно

Example: $\lambda x.xy \to \lambda y.yy$ — некорректно

1.7.2.2 Бета-редукция

Def. Бета-редукция — $(\lambda x.M)M \to M(N/x)$. Подстановка тоже должна быть корректной

Example: $(\lambda x.xx)(\lambda z.z) \to (\lambda z.z)(\lambda z.z) \to \lambda z.z$ — корректно

Example: $\lambda xy.(x+y)2z \to (\lambda y.(2+y))z \to 2+z$ — корректно

Example: $(\lambda x.x(\lambda y.xy))y \to y(\lambda y.yy)$ — некорректно

1.7.3 Нормальная форма.

Def. Терм находится в нормальной форме, если к нему нельзя применить β-редукцию, даже после нескольких α-конверсий

Note. Нормальная форма существует не всегда: $(\lambda x.xx)(\lambda x.xx)$ (под действием β-редукции переходит в себя)

Def. Один терм переходит в другой ($M \to\to N$), если есть цепочка α-конверсий и β-редукций, которая переводит M в N.

Note. Равенство термов подразумевается как отношение эквивалентности

Свойства нормальной формы:

1. Если N — нормальная форма M, то не просто $M = N$, а $M \to\to N$
2. Нормальная форма единственная с точностью до α-конверсии

1.7.4 Теорема Чёрча–Россера (б/д).

Th (Черча-Россера (б/д)). Если $M \to\to P$ и $M \to\to Q$, то для некоторого $TP \to\to T$ и $Q \to\to T$

1.7.5 Нумералы Чёрча. Комбинаторы, представляющие сложение и умножение. Представление логических значений и операций. Представление вычитания (б/д). Комбинатор неподвижной точки, выражение одной из функций: факториал, неполное частное, остаток по модулю или любая другая, где требуется рекурсия.

1.7.5.1 Нумералы Чёрча.

Неформально: число k соответствует преобразованию функции f в ее k-ю итерацию

- $\underline{0} = \lambda fx.x$

- $\underline{1} = \lambda fx.fx$
- $\underline{2} = \lambda fx.f(fx)$
- ...
- $\underline{n} = \lambda fx.f(f(f(\ldots(f(fx))\ldots)))$ — n раз f

Def. Комбинатором называется замкнутый λ-терм (без свободных переменных)

1.7.5.2 Комбинаторы, представляющие сложение и умножение.

1. Inc — прибавление 1: $Inc\ \underline{n} = \underline{n+1}$

$$Inc = \lambda nfx.f(nfx)$$

2. Add — сложение: $Add\ \underline{mn} = \underline{m+n}$

$$Add = \lambda mnfx.mf(nf)x$$

3. Mult — умножение

$$Mult = \lambda mnfx.m(nf)x$$

4. Exp — возведение в степень

$$\lambda mnfx.nmfx$$

1.7.5.3 Представление логических значений и операций.

- $False = \lambda xy.y(= \underline{0})$
- $True = \lambda xy.x$

$$TrueMN = M.FalseMN = N$$

- $And = \lambda pq.pqp$
- $Or = \lambda pq.ppq$
- $Not = \lambda pq.pFalseTrue$

1.7.5.4 Представление вычитания (б/д).

Сначала введем функции пары

- $Pair = \lambda xyp.pxy$
- $Left = \lambda p.pTrue$
- $Right = \lambda p.pFalse$

Введем вспомогательную функцию $Decfn = \lambda fp.Pair(f(Leftp))(Leftp)$

Statement. $Dec = \lambda nfx.Right(n(Decfnf)(Pairxx))$

1.7.5.5 Комбинатор неподвижной точки, выражение одной из функций: факториал, неполное частное, остаток по модулю или любая другая, где требуется рекурсия.

Def. Y-комбинатор (комбинатор неподвижной точки): $\forall F \hookrightarrow YF = F(YF)$

Example: $Y = \lambda f.(\lambda x.f(xx))(\lambda x.f(xx))$

Тогда $YF = (\lambda x.F(xx))(\lambda x.F(xx)) = F((\lambda x.F(xx))(\lambda x.F(xx))) = F(YF)$

Тогда YF — неподвижная точка F

Давайте научимся считать факториал. Для этого нам нужен предикат $IsZero$

$IsZero = \lambda n.n(\lambda x.False)True$

Заведем вспомогательную функцию $Factfn = \lambda gn.(IsZeron)\underline{1}(Multn(g(Prevn)))$

Note. Prev получается из Dec, который мы научились делать выше

Тогда $Fact = YFactfn$ (факториал это неподвижная точка $Factfn$)

1.8 Измерение сложности алгоритма и сложности задачи. Классы P и NP. Сводимость задач по Карпу. NP-полнота. Теорема Кука–Левина: формулировка и идея доказательства. Сводимость общей задачи о выполнимости к задаче о выполнимости 3-КНФ и сводимость задачи о выполнимости 3-КНФ к одной из задач: клика, вершинное покрытие, 3-раскраска, гамильтонов путь, задача о рюкзаке, целочисленное линейное программирование – или другой подобной.

1.8.1 Измерение сложности алгоритма и сложности задачи.

M — МТ, x — вход

Def. $time_M(x)$ — число шагов $M(x)$

2 вида агрегирования:

1. Сложность в худшем случае $time_M(n) = \max\limits_{|x|=n}(time_M(x))$
2. Сложность в среднем

Нас интересует как изменяется $time_M(n)$ при $n \to \infty$

Сложность задачи — ?

Def. (Наивное) Сложность задачи это сложность наилучшего алгоритма, который ее решает

Th (Блума). Оптимальный алгоритм не всегда существует

Поэтому на сложность задачи есть только оценки сверху и снизу

1.8.2 Классы P и NP.

1.8.2.1 Класс P

Def. Классом **DTIME** $(T(n))$ называется класс языков, которые распознаются за время $O(T(n))$ детерминированной многоленточной машиной Тьюринга (МТ). Иными словами, время работы машины на любом слове длины n не превосходит некоторой константы, умноженной на $T(n)$.

Def. $\mathbf{P} = \bigcup\limits_{c=1}^{\infty} \mathbf{DTIME}\,(n^c)$

Иными словами классом **P** называется множество языков, распознающихся за полином от длины входа.

Note. **P** замкнут относительно дополнений, объединений и пересечений, звезды Клини.

Note. **P** считается классом эффективно решаемых задач.

1.8.2.2 Класс NP

Def. Класс **NP** состоит в точности из тех языков A, для которых существует некая детерминированная МТ двух аргументов $V(x, s)$, работающая по времени за $poly(|x|)$ такая, что

$$\forall x \in \Sigma^* \; x \in A \iff \exists s : V(x, s) = 1.$$

Тогда V — верификатор, а s — сертификат для данного x.

Def. Классом $\mathbf{NTIME}\,(T(n))$ называется класс языков, которые распознаются за время $O(T(n))$ недетерминированной машиной Тьюринга (МТ). Иными словами, время работы машины на любом слове длины n не превосходит некоторой константы, умноженной на $T(n)$.

Def. $\mathbf{NP} = \bigcup\limits_{c=1}^{\infty} \mathbf{NTIME}\,(n^c)$

Th. Эти определения эквивалентны

Proof:

Пусть A распозанется недетерминированной машиной Тьюринга (МТ), которая работает $O(n^c)$ шагов. Тогда в качестве сертификата можно взять последовательность значений функции перехода, а верификатор будет моделировать работу машины M, используя данные сертификата для выбора одной из ветвей алгоритма. Можно действовать и по-другому: в качестве сертификата взять всё вычисление, а верификатором проверять его соответствие программе.

Пусть для A верно сертификатное определение. Тогда сертификат не может быть длиннее, чем время работы V, т.е. чем некоторый полином $p(|x|)$. Недетерминированная машина будет работать следующим образом: сначала она недетерминированно напишет сертификат s длины не больше $p(|x|)$, а затем запустит $V(x, s)$. ■

Statement. **NP** замкнут относительно объединений, пересечений и звезды Клини, но не дополнений.

1.8.3 Сводимость задач по Карпу. NP-полнота.

Def. Язык A сводится по Карпу к языку B ($A \leqslant_p B$), если существует полиномиально вычислимая f такая, что $x \in A \iff f(x) \in B$.

Def. $A \in$ **NP-hard**, если $\forall B \in \mathbf{NP}\; B \leqslant_p A$.

Note. Если $A \in$ **NP-hard** и $A \leqslant_p B$, то $B \in$ **NP-hard**.

Def. Класс **NPC** (**NP** - полных языков) состоит в точности из тех языков A, которые:

1. $A \in$ **NP**;
2. $A \in$ **NP-hard**.

Note. Если $A \in$ **NPC** и $A \leqslant_p B$, $B \in$ **NP**, то $B \in$ **NPC**.

Def. $\mathsf{SAT} = \{\varphi | \varphi$ — выполнимая, то есть $\exists x : \varphi(x) = 1\}$

Def. $3\mathsf{SAT} = \{\varphi | \varphi$ — записана в 3-КНФ и выполнима$\}$

1.8.4 Теорема Кука–Левина: формулировка и идея доказательства.

Th. (Кука-Левина) Задача SAT является **NP** -полной ($\in$ **NPC**)

Идея доказательства:

1) Очевидно, что $\mathsf{SAT} \in$ **NP**

2) Для произвольной L покажем, что $L \leqslant_p \mathsf{SAT}$. Пусть V, работающий с сертификатами s длины $q(n)$ — верификатор из определения $L \in$ **NP**. Будем так же считать, что V задан МТ с одной, бесконечной вправо лентой и работающей не дольше $p(n)$ шагов. Идея состоит в записи булевой формулы, которая моделирует работу этой машины.

Полное доказательство можно почитать тут или посмотреть тут ■

1.8.5 Сводимость общей задачи о выполнимости к задаче о выполнимости 3-КНФ и сводимость задачи о выполнимости 3-КНФ к одной из задач: клика, вершинное покрытие, 3-раскраска, гамильтонов путь, задача о рюкзаке, целочисленное линейное программирование – или другой подобной.

1.8.5.1 Сводимость общей задачи о выполнимости к задаче о выполнимости 3-КНФ

Statement. $\mathsf{SAT} \leqslant_p 3\mathsf{SAT}$

Пусть φ — пропозициональная формула. Введём переменную для каждой её подформулы (включая исходные переменные и всю формулу). Каждая подформула, составленная из двух, например, задаёт утверждение вида $q = r \wedge s$. Его можно представить как 3-КНФ (в данном случае $(q \vee \neg r \vee \neg s) \wedge (\neg q \vee r \vee \neg s) \wedge (\neg q \vee \neg r \vee s) \wedge (\neg q \vee r \vee s)$). Взяв конъюнкцию всех таких формул, мы получим 3-КНФ, выполнимость которой эквивалентна выполнимости исходной формулы. Действительно, выполняющий набор исходной формулы задаст значения всех подформул, которые будут

выполняющим набором полученной формулы. И наоборот, из выполняющего набора новой формулы можно выделить выпол- няющий набор исходной.

Осталось пояснить, почему этот алгоритм работает полиномиальное время. Ясно, что достаточно построить за полиномиальное время дерево синтаксического разбора формулы, дальнейшие операции занимают константное время для каждой его вершины. Построение дерева также несложно: путём подсчёта скобочного итога можно выделить из формулы две подформулы, из которых она получена, и повторить процедуру с ними рекурсивно. Это займёт линейное время для каждой подформулы, т.е. всего квадратичное время ■

1.8.5.2 Сводимость 3-КНФ к задаче о клике

Def. Пусть дан неориентированный граф G. Кликой называется любой его полный подграф. Независимым множеством называется любое множество вершин, между которыми нет рёбер.

Сформулируем так же задачи распознавания:

- $\mathsf{CLIQUE} = \{(G, k) \mid \text{в графе } G \text{ есть клика из } k \text{ вершин}\}$
- $\mathsf{INDSET} = \{(G, k) \mid \text{в графе } G \text{ есть независимое множество из } k \text{ вершин}\}$

Statement. $\mathsf{CLIQUE} \leqslant_p \mathsf{INDSET}$ и $\mathsf{INDSET} \leqslant_p \mathsf{CLIQUE}$

Proof:

$(G, k) \in \mathsf{INDSET} \Leftrightarrow (\overline{G}, k) \in \mathsf{CLIQUE}$ ■

Теперь покажем, что $\mathsf{INDSET} \in \mathbf{NPC}$

Th. Задача INDSET является **NP** -полной

Proof: Докажем, что $3\mathsf{SAT} \leqslant_p \mathsf{INDSET}$. Построим граф следующим образом: каждому вхождению литерала сопоставим вершину графа. Таким образом, всего будет $3k$ вершин, где k — число дизъюнктов в 3-КНФ. Вершины, соответствующие литералам из одного дизъюнкта, соединим рёбрами и получим треугольник для каждого дизъюнкта. Также соединим все противоположные литералы, например, p и $\neg p$. На этом построение графа заканчивается, размер независимого множества возьмём равным k.

$\Rightarrow$ Если у формулы есть выполняющий набор, то в каждом дизъюнкте есть хотя бы один истинный литерал. Возьмём из каждого треугольника по одной вершине, соответсвтвующей истинному литералу из выполняющего набора. Соответствующие вершины образуют независимое множество, поскольку эти литералы из разных треугольников (дизъюнктов), а литералы вида p и $\neg p$ не могут быть истинны одновременно. Этих вершин как раз k.

$\Leftarrow$ Если, наоборот, есть независимое множество из k вершин (больше не может быть, потому что тогда будет пара вершин из одного треугольника), то эти вершины обязаны соответствовать различным

треугольникам (дизъюнктам). А поскольку среди этих вершин нет одновременно литералов вида p и $\neg p$, то можно взять такие значения переменных, при которых все задействованные литералы истинны. Эти значения и будут выполняющим набором. ■

Note. Остальные сводимости можно найти тут

Глава 2

Дискретные структуры

2.1 Основные правила комбинаторики: правило сложения, правило умножения. Принцип Дирихле. Формула включения-исключения: доказательство, применение для вывода формулы для числа беспорядков. Базовые комбинаторные конфигурации: размещения, перестановки и сочетания. Формулы для количеств размещений, перестановок и сочетаний. Формула Стирлинга (б/д).

2.1.1 Основные правила комбинаторики: правило сложения, правило умножения. Принцип Дирихле

Предположим, что у нас имеются 2 множества (здесь и далее предполагаем, что рассматриваемые множества конечны, если не оговорено обратного) $A = \{a_1, a_2, \ldots, a_n\}, B = \{b_1, b_2, \ldots, b_m\}$.

1. Правило суммы: Количество способов выбрать один объект из A или B (в предположении, что $A \cap B = \varnothing$) равно $n + m$.
2. Правило произведения: Количество способов выбрать один объект из A и к нему в пару один объект из B равно nm.
3. Принцип Дирихле: Предположим, имеется n ящиков и $n + 1$ кролик, которые сидят в этих ящиках. Тогда найдется ящик, в котором сидит $\geqslant 2$ кролика.

Обобщенный приниип Дирихле: Если $nk + 1$ разбит на n множеств, то хотя бы в одном множестве содержится $k + 1$ элемент.

2.1.2 Формула включения-исключения: доказательство, применение для вывода формулы для числа беспорядков

Th. Пусть имеется множество из N объектов и некоторые свойства $\alpha_1, \ldots, \alpha_n$. Обозначим α'_k отрицание свойства α_k. Пусть $N(\alpha_i)$ обозначает количество объектов, удовлетворяющих свойству α_i, $N(\alpha_i, \alpha_j)$ – удовлетворяющих одновременно обоим свойствам $\alpha_i, \alpha_j \ldots N(\alpha_1, \ldots, \alpha_n)$ – удовлетворяющих всем свойствам. Тогда справедлива следующая формула, называемая формулой включений и исключений:

$$N(\alpha'_1, \alpha'_2, \ldots \alpha'_n) = N - N(\alpha_1) - \ldots - N(\alpha_n) + N(\alpha_1, \alpha_2) + \ldots + \\ + N(\alpha_{n-1}, \alpha_n) - \ldots + (-1)^n N(\alpha_1, \ldots, \alpha_n)$$

Proof: Индукция по числу свойств. Для $n = 1 : N(\alpha'_1) = N - N(\alpha_1)$. База доказана. Докажем переход индукции.

Предположим, что для всех $1 \leqslant k \leqslant n$ верно, что для любого N, любого множества из N объектов и любого набора свойств $\alpha_1, \ldots \alpha_k$ выполнена формула из условия теоремы. Рассмотрим только свойства $\alpha_1, \ldots, \alpha_n$. По предположению индукции, для них верна формула включений и исключений. (1)

Рассмотрим теперь только те объекты, которые обладают свойством α_{n+1}.К ним применимо предположение индукции (для $N = N(\alpha_{n+1})$), т.е.

$$N(\alpha'_1, \ldots, \alpha'_n, \alpha_{n+1}) = N(\alpha_{n+1}) - N(\alpha_1, \alpha_{n+1}) - \ldots + (-1)^n N(\alpha_1, \ldots, \alpha_n, \alpha_{n+1}). - (2)$$

Имеем теперь, вычитая (2) из (1)

$$N(\alpha'_1, \ldots, \alpha'_n) - N(\alpha'_1, \ldots, \alpha'_n, \alpha_{n+1}) = N\left(\alpha'_1, \ldots, \alpha'_{n+1}\right) = \\ = N - N(\alpha_1) - \ldots - N(\alpha_n) + N(\alpha_1, \alpha_2) + \ldots + N(\alpha_{n-1}, \alpha_n) - \ldots + (-1)^{n+1} N(\alpha_1, \ldots, \alpha_n, \alpha_{n+1})$$

■

Lemma. Сколько способов рассадить n человек по n местам в аудитории, чтобы никто не сидел на своём месте?

Proof:

Рассмотрим $n!$ перестановок людей. Пусть α_i означает, что i-й человек сидит на своём месте. Тогда: $N(\alpha'_1, \ldots, \alpha'_n) = n! - C_n^1(n-1)! + C_n^2(n-2)! - \ldots + (-1)^n C_n^n 0! = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \ldots + (-1)^n \frac{1}{n!}\right)$ Устремляя $n \to \infty$, получаем ответ $\frac{n!}{e}$.

■

2.1.3 Базовые комбинаторные конфигурации: размещения, перестановки и сочетания. Формулы для количеств размещений, перестановок и сочетаний. Формула Стирлинга (б/д)

Пусть дано множество $A = \{a_1, a_2, \dots, a_n\}$ из n объектов.

Def. Произвольный упорядоченный набор (кортеж) из k элементов данного множества, среди которых могут быть повторяющиеся, называется размещением из n элементов по k с повторением. Соответственно, если элементы не могут повторяться, то набор называется размещением без повторений, или просто размещением из n по k. Обозначаются $\overline{A}_n^k$ и A_n^k соответственно.

Def. Перестановка – размещение из n элементов.

Def. Сочетанием из n элементов по k называется набор k элементов этого множества. Наборы, отличающиеся только порядком следования элементов считаются одинаковыми (этим сочетание отличается от размещения). Соответственно, сочетания бывают с повторениями и без, в зависимости от того, разрешаем ли мы элементам набора повторяться или нет. Обозначаются $\overline{C}_n^k$ и C_n^k соответственно.

Lemma.

$$\overline{A}_n^k = n^k$$
$$A_n^k = \frac{n!}{(n-k)!}$$
$$C_n^k = \frac{n!}{k!\,(n-k)!}$$
$$\overline{C}_n^k = C_{n+k-1}^k$$

Proof:

- На каждую из k позиций независимо выбираем n элементов, применяем правило умножения
- Первый элемент выбрать n способов, второй $n-1$ и так далее.
- Так-как нам не важен порядок, нужно разделить на количество перестановок из k элементов
- Фиксируем сочетание с повторениями $(a_1, \dots a_k)$. Построим последовательность из 0 и 1 по следующему правилу: напишем столько 1, сколько раз встречается элемент a_1, потом напишем 0, потом столько 1 сколько встречается элемент a_2 и так далее. В конце ставить 0 не будем. Заметим, что наше сочетане это количество способ выбрать позиции для k единиц в последовательности длины $n+k-1$.

■

Th (б/д). Формула Стирлинга

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

2.2 Формула бинома Ньютона, полиномиальная формула. Свойства биномиальных коэффициентов: симметричность, унимодальность, рекуррентная формула треугольника Паскаля. Знакопеременная сумма биномиальных коэффициентов. Оценки для биномиальных коэффициентов при n: асимптотика C_n^k в случае $k = const \cdot n$ и в случае $k = o(\sqrt{n})$

2.2.1 Бином Ньютона. Свойства биномиальных коэффициентов: симметричность, унимодальность, рекуррентная формула треугольника Паскаля. Знакопеременная сумма биномиальных коэффициентов

Def. *Биномом Ньютона* называется выражение $(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}$.

Proof: n-ю степень суммы можно записать в виде

$$(a+b)^n = \underbrace{(a+b)\cdot(a+b)\cdot\ldots\cdot(a+b)}_{n \text{ раз}}.$$

Чтобы получить слагаемое в сумме, мы должны последовательно выбрать из каждой скобки a или b. Любое слагаемое точно будет иметь вид

$$a^k \cdot b^{n-k}.$$

Более того, чтобы определить слагаемое, нам необходимо и достаточно знать, сколько a мы выбрали. При этом выбирать его можно в любых скобках, а это можно сделать C_n^k способами. Отсюда и формула

$$(a+b)^n = \sum_{k=0}^{n} C_n^k a^k b^{n-k}.$$

■

Th.

1. Симметричность: $C_n^k = C_n^{n-k}$;
2. Унимодальность: Коэффициенты возрастают до $k = \frac{n}{2}$, потом убывают

3. Треугольник Паскаля: $C_n^k = C_{n-1}^k + C_{n-1}^{k-1}$;

4. $C_n^0 - C_n^1 + C_n^2 + \ldots + (-1)^n C_n^n = 0$;

Proof:

1. Выбрать k объектов из n — это то же самое, что оставить $n-k$ объектов из n.

2. $\frac{C_n^k}{C_n^{k-1}} = \frac{n-k+1}{k} = \frac{n+1}{k} - 1$ больше единицы при $k \leq \frac{n}{2}$

3. Количество способов выбрать k-набор из n элементного множества уже известно: C_n^k. Заметим, что каждый набор либо содержит n-й элемент, либо нет. То есть все наборы можно разбить на 2 группы:

   (a) Все наборы, которые содержат n-й элемент. Помимо него в них ещё надо выбрать $k-1$ элемент, а стало быть, их всего C_{n-1}^{k-1} штук.

   (b) Все наборы, которые не содержат n-й элемент. То есть набор выбирается только из первых $n-1$ элементов. Отсюда их C_{n-1}^k штук.

   Так как эти две группы в сумме составляют все возможные наборы, то и очевиден ответ

   $$C_n^k = C_{n-1}^k + C_{n-1}^{k-1}.$$

4. Подставим $a = -1$ и $b = 1$ в определение.

■

2.2.2 Полиномиальная формула

У нас есть n_i объектов a_i для всех $i \in [1, \ldots, t]$. Сколько различных последовательностей можно составить, используя абсолютно все объекты?

Заметим, что у нас всегда одинаковая длина слова, равная n:

$$n = \sum_{i=1}^{t} n_i.$$

Чтобы собрать слово, будем последовательно расставлять все t групп объектов.

- Сколько мест есть для первой группы? Ответ: n.
- Сколько мест есть для второй группы? Ответ: $n - n_1$ (первую поставили первой).
- ...

- Сколько мест есть для последней группы? Ответ: $n - n_1 - \cdots - n_{t-1} = n_t$.

Очевидно, что для i-й группы происходит выбор n_i мест из тех, что остались не занятыми. В итоге имеем

$$P(n_1, n_2, \ldots, n_t) = C_n^{n_1} \cdot C_{n-n_1}^{n_2} \cdot C_{n-n_1-n_2}^{n_3} \cdot \ldots \cdot C_{n-n_1-\ldots-n_{t-1}}^{n_t} =$$
$$= \frac{n!}{n_1!\,(n-n_1)!} \cdot \frac{(n-n_1)!}{n_2!\,(n-n_1-n_2)!} \cdot \frac{(n-n_1-n_2)!}{n_3!\,(n-n_1-n_2-n_3)!} \cdot \ldots \cdot \frac{(n-n_1-\ldots-n_{t-1})!}{n_t!\,(n-n_1-\ldots-n_t)!} =$$
$$= \frac{n!}{n_1! \cdot n_2! \cdot n_3! \cdot \ldots \cdot n_t! \cdot 0!} = \frac{n!}{n_1! \cdot n_2! \cdot n_3! \cdot \ldots \cdot n_t!}.$$

Def. Число $P(n_1, \ldots, n_t)$ называют *полиномиальным коэффициентом.*

Th. *Полиномиальной формулой* называется выражение

$$(x_1 + \ldots + x_t)^n = \sum_{\forall n_1+\ldots+n_t=n} P(n_1, \ldots, n_t) \cdot x_1^{n_1} x_2^{n_2} \cdot \ldots \cdot x_t^{n_t}$$

Proof: Аналогично биному Ньютона распишем скобки:

$$(x_1 + \ldots + x_t)^n = \underbrace{(x_1 + \ldots + x_t) \cdot \ldots \cdot (x_1 + \ldots + x_t)}_{n \text{ раз}}.$$

Чтобы получить полное слагаемое, нам нужно выбрать из каждой скобки по одному объекту. Если его привести, то мы получим выражение вида

$$x_1^{n_1} \cdot \ldots \cdot x_t^{n_t}.$$

При этом верно равенство:

$$n_1 + \ldots + n_t = n,$$

где n_i значит то же, что и до этого: количество раз, сколько мы выбрали (имеем) x_i.

Несложно понять, что такое слагаемое мы могли получить разными способами — в зависимости от того, из каких скобок какие объекты мы брали. Количество различных способов набрать $n_1, n_2, \ldots, n_t$ из скобок определяется полиномиальным коэффициентом. Отсюда

$$(x_1 + \ldots + x_t)^n = \sum_{\forall n_1+\ldots+n_t=n} P(n_1, \ldots, n_t) \cdot x_1^{n_1} x_2^{n_2} \cdot x_t^{n_t}.$$

■

2.2.3 Оценки для биномиальных коэффициентов при n: асимптотика C_n^k в случае $k = const \cdot n$ и в случае $k = o(\sqrt{n})$

Th. Пусть $a \in (0, 1)$, тогда $C_n^{[an]} = \left(\frac{1}{a^a \cdot (1-a)^{1-a}} + o(1)\right)^n = \left(e^{-aln(a)-(1-a)ln(1-a)} + o(1)\right)^n$.

Proof: Будем пользоваться формулой Стирлинга $n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$.

$$C_n^{[an]} = \frac{n!}{[an]!\,(n-[an])!} \sim \frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi[an]}\left(\frac{[an]}{e}\right)^{[an]}\sqrt{2\pi(n-[an])}\left(\frac{n-[an]}{e}\right)^{n-[an]}} =$$

$$= P(n)\cdot\frac{\left(\frac{n}{e}\right)^n}{\left(\frac{[an]}{e}\right)^{[an]}\cdot\left(\frac{n-[an]}{e}\right)^{n-[an]}} = P(n)\cdot\frac{n^n}{[an]^{[an]}\cdot(n-[an])^{n-[an]}}$$

Теперь предположим, что $an \in \mathbb{N}$, тогда формула была бы переписана в виде:

$$C_n^{[an]} \sim P(n)\cdot\frac{n^n}{(an)^{an}\cdot(n-an)^{n-(an)}} = \frac{P(n)}{a^{an}\cdot(1-a)^{n-an}} = P(n)\cdot\left(\frac{1}{a^a\cdot(1-a)^{1-a}}\right)^n =$$

$$= \left(\frac{1}{a^a\cdot(1-a)^{1-a}} + o(1)\right)^n$$

То есть в данном случае теорема доказана. ■

Кокнулись?

Th.

1. Если $k = o(\sqrt{n})$, то $C_n^k \sim \frac{n^k}{k!}$;
2. Если $k^3 = o(n^2)$, то $C_n^k = \frac{n^k}{k!}\cdot e^{-\frac{k^2}{2n}+O\left(\frac{k^3}{n^2}\right)}$.

Proof: Здесь мы не будем использовать формулу Стирлинга.

$$C_n^k = \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{k!} = \frac{n^k}{k!}\cdot\left(1-\frac{1}{n}\right)\cdot\ldots\cdot\left(1-\frac{k-1}{n}\right) = \frac{n^k}{k!}\cdot e^{ln\left(1-\frac{1}{n}\right)+\ldots+ln\left(1-\frac{k-1}{n}\right)} =$$

Далее, используя ряд Тейлора в виде $ln(1-x) = -x + O(x^2)$, продолжим равенство

$$C_n^k = \frac{n^k}{k!}\cdot e^{-\frac{k(k-1)}{2n}+O\left(\frac{1^2+\ldots(k-1)^2}{n^2}\right)} = \frac{n^k}{k!}\cdot e^{-\frac{k(k-1)}{2n}+O\left(\frac{k^3}{n^2}\right)} = \frac{n^k}{k!}\cdot e^{-\frac{k^2}{2n}+O\left(\frac{k^3}{n^2}\right)}$$

Заметим, что последний переход верен при условии, что $k = o(\sqrt{n})$. Это доказывает оба пункта теоремы. ■

2.3 Формальные степенные ряды: определение, операции над рядами (сумма, разность, произведение, частное, производная). Теорема Коши–Адамара о радиусе сходимости (с доказательством). Производящие функции: определение, примеры производящих функций для последовательности биномиальных коэффициентов и для чисел Фибоначчи. Пример применения производящих функций для доказательства комбинаторных тождеств.

2.3.1 Формальные степенные ряды: определение, операции над рядами (сумма, разность, произведение, частное, производная)

Def. Будем рассматривать бесконечную последовательность чисел:

$$(a_0, a_1, \ldots, a_n, \ldots), \quad \forall n \in \mathbb{N}\ a_n \in \mathbb{C}$$

Такую последовательность мы и будем называть *формальным степенным рядом* (ФСР).

Дополнительно введём обозначение A_i для формального степенного ряда A, обозначающее i-ое число в последовательности.

Из определения операций видно, что если положить за единицу ряд

$$1 = (1, 0, \ldots, 0, \ldots)$$

А за некоторое t ряд

$$t = (0, 1, \ldots, 0, \ldots)$$

То тогда t^n будет иметь вид

$$t^n = (\underbrace{0, \ldots, 0}_{n}, 1, 0, \ldots)$$

Из этого следует, что любой ряд $A = (a_1, \ldots, a_n, \ldots)$ можно записать в следующем виде:

$$A = (a_1, \ldots, a_n, \ldots) = a_1 \cdot 1 + a_2 \cdot t + \ldots + a_n \cdot t^n + \ldots$$

То есть формальный степенной ряд — это бесконечный многочлен. Довольно легко проверить, что для формальных степенных рядов, как и для многочленов, верна ассоциативность и дистрибутивность, чем мы и воспользуемся далее.

Теперь можно определить операции над формальными степенными рядами:

- Пусть есть 2 формальных степенных ряда $A = (a_0, \ldots, a_n, \ldots)$ и $B = (b_0, \ldots, b_n, \ldots)$. Тогда назовём их *суммой* формальный степенной ряд $A + B$:

$$A + B = (a_0 + b_0, \ldots, a_n + b_n, \ldots)$$

- Пусть есть 2 формальных степенных ряда A и B. Тогда назовём их *произведением* формальный степенной ряд $A \cdot B$ такой, что

$$(A \cdot B)_n = A_n B_0 + A_{n-1} B_1 + \ldots + A_n B_0 = \sum_{i=0}^{n} A_{n-i} B_i = \sum_{i=0}^{n} A_i B_{n-i}$$

- Пусть есть 2 формальных степенных ряда A и B. Тогда возможно осуществить *деление* $A/B = C$, если существует формальный степенной ряд C такой, что $C \cdot B = A$. Для нахождения чисел C необходимо последовательно решать уравнения из системы:

$$\begin{cases} c_0 b_0 = a_0 \\ c_0 b_1 + c_1 b_0 = a_1 \\ \vdots \\ c_0 b_n + \ldots + c_n b_0 = a_n \\ \vdots \end{cases}$$

  Из записанных уравнений понятно, что необходимым и достаточным условием для деления является $b_0 \neq 0$.

- Взятие производной от ряда. Если ряд A имеет вид:

$$f(t) = a_0 + a_1 t + a_2 t^2 + \ldots + a_n t^n + \ldots$$

  то его *производной* назовётся ряд

$$f'(t) = a_1 + 2a_2 t + \ldots + n a_n t^{n-1} + (n+1) a_{n+1} t^n + \ldots$$

2.3.2 Теорема Коши–Адамара о радиусе сходимости (с доказательством)

Нагло украдено из Матана

Th. Пусть $a_n \geq 0$ при всех $n \in \mathbb{N}$, и $q = \overline{\lim_{n\to\infty}} \sqrt[n]{a_n}$.

- Если $q < 1$, то ряд $\sum_{n=m}^{\infty} a_n$ сходится.
- Если $q > 1$, то $a_n \not\to 0$, и ряд $\sum_{n=m}^{\infty} a_n$ расходится.

Proof:

- Пусть $q < q_0 < 1$. По свойству верхнего предела существует такое $N \in \mathbb{N}$, что $\sqrt[n]{a_n} < q_0$ при всех $n \geq N$. Тогда $a_n < q_0^n-$ и ряд сходится по признаку сравнения с геометрическим рядом, т. к. $\sum_{n=1}^{\infty} q_0^n$ сходится.

- Так как q - частичный предел $\left\{\sqrt[n]{a_n}\right\}$, то неравенство $\sqrt[n]{a_n} > 1$ выполняется для бесконечного множества номеров n. Следовательно, $a_n \nrightarrow 0$, и ряд расходится.

■

Def. Мы можем рассмотреть формальный степенной ряд как функцию $f(x)$:

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

Будем говорить, что $f(x)$ *сходится в точке* x_0, если последовательность частичных сумм этого ряда в этой точке имеет конечный предел. То есть

$$S_n(x_0) = \sum_{k=0}^{n} a_k x_0^k$$
$$\exists \lim_{n\to\infty} S_n = S = f(x_0) \in \mathbb{C}$$

Th. Пусть задан формальный степенной ряд $A = (a_0, \ldots, a_n, \ldots)$, причём $\forall n \in \mathbb{N}\ a_n \in \mathbb{R}$. Обозначим за $\rho = \frac{1}{\overline{\lim\limits_{k\to\infty}} \sqrt[k]{|a_k|}}$. Тогда:

- Если $|x - x_0| < \rho$, то ряд сходится в x.

- Если $|x - x_0| > \rho$, то ряд расходится в x.

- Если же $|x - x_0| = \rho$, то может быть всё, что угодно.

При этом ρ называется радиусом сходимости.

Proof:

Пусть $x \neq x_0$, тогда введём

$$\begin{aligned} q &= \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n (x - x_0)^n|} \\ &= \overline{\lim_{n\to\infty}} |x - x_0| \sqrt[n]{|a_n|} \\ &= |x - x_0| \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|} \end{aligned}$$

Если $|x - x_0| < \rho$, то $q < 1$, и по признаку Коши ряд (1) сходится абсолютно. Если $|x - x_0| > \rho$, то $q > 1$, и по признаку Коши n-й член не стремится к нулю, следовательно, ряд (1) расходится (и абсолютно расходится, т. е. ряд модулей также расходится).

■

2.3.3 Производящие функции: определение, примеры производящих функций для последовательности биномиальных коэффициентов и для чисел Фибоначчи. Пример применения производящих функций для доказательства комбинаторных тождеств

Def. Если нам задана последовательность (ряд) $\{a_k\}_{k=0}^{\infty}$, то её *производящей функцией* называется

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

Lemma. Производящая функция чисел Фибоначчи

Числа Фибоначчи можно выписать в явном виде по формуле из слеующео билета. Если положить $F_0 = 0,\ F_1 = 1$, то формула примет вид:

$$F_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2} \right)^n$$

Посчитаем производяющую функцию для данной последовательности:

$$f(x) = \sum_{k=0}^{\infty} F_k x^k$$

Мы хотим «свернуть» бесконечную сумму в какое-то деление конечных чисел, как для рядов, показанных выше. Посмотрим на следующие ряды:

$$\begin{aligned} xf(x) &= F_0 x + F_1 x^2 + \ldots + F_n x^{n+1} + \ldots \\ x^2 f(x) &= F_0 x^2 + F_1 x^3 + \ldots + F_n x^{n+2} + \ldots \\ xf(x) + x^2 f(x) &= \underbrace{F_0 x}_{0} + \underbrace{(F_0 + F_1)}_{F_2} x^2 + \underbrace{(F_1 + F_2)}_{F_3} x^3 + \ldots = f(x) - x \end{aligned}$$

Получили линейное уравнение относительно $f(x)$. Решив его, получим ответ:

$$f(x) = \frac{x}{1 - x - x^2}$$

Lemma. Производящая функция биномиальных коэффицентов.

$$f(x) = \sum_{k=0}^{n} C_n^k x^k = (1+x)^n$$

Lemma. Почитать $f(x) = \sum_{k=0}^{n} C_n^k k^2$ через производящие функции

Для начала заметим, что $f(x)$ сходится в любой точке, а значит в любой точке $\mathbb{R}$ существует ее производная. Тогда

$$x\left(xf'(x)\right)' = x\left(x'f'(x) + xf''(x)\right) = x\left(f'(x) + xf''(x)\right) = x\sum_{k=1}^{n} k^2 C_n^k x^{k-1} = \sum_{k=0}^{n} k^2 C_n^k x^k$$

С другой стороны

$$x\left(f'(x)+xf''(x)\right)=n^2x(1+x)^{n-1}$$

Подставляя $x=1$ получаем ответ $2^{n-1}n^2$.

2.4 Линейные рекуррентные соотношения с постоянными коэффициентами (л.р.с.п.к.). Пример: числа Фибоначчи. Общий вид решения в произвольном случае (б/д). Доказательство теоремы об общем виде решения у л.р.с.п.к. второго порядка (в том числе при кратных корнях характеристического многочлена). Применение теоремы для нахождения формулы для чисел Фибоначчи

2.4.1 Линейные рекуррентные соотношения с постоянными коэффициентами

Далее: $\mathbb{N}=\{0,1,2\ldots\}$.

Def. Последовательность $\{y_n\}_{n=0}^{\infty}$ удовлетворяет *линейному рекуррентному соотношению (ЛРС) k-го порядка с постоянными коэффициентами*, если для любого $n\in\mathbb{N}$ верно:

$$a_k\cdot y_{n+k}+a_{k-1}\cdot y_{n+k-1}+\ldots+a_1\cdot y_{n+1}+a_0\cdot y_n=0,$$

где $a_0,\ldots,a_k\in\mathbb{C},\ a_k\neq 0,\ a_0\neq 0$.

1. $k=1,\ a_0,a_1\in\mathbb{C}$:

   $$a_1y_{n+1}+a_0y_n=0.$$

   Или же

   $$y_{n+1}=\left(-\frac{a_0}{a_1}\right)y_n.$$

   Несложно понять и доказать общую формулу:

   $$y_n=y_0\cdot\left(-\frac{a_0}{a_1}\right)^n.$$

2. $k=2,\ a_0,a_1,a_2\in\mathbb{C}$:

   $$a_2y_{n+2}+a_1y_{n+1}+a_0y_n=0.$$

Составим *характеристическое уравнение*:

$$a_2x^2 + a_1x + a_0 = 0.$$

Просто заменили все y_i на x^i. Решений у полученного уравнения всегда 2 в поле $\mathbb{C}$. Введём обозначения для этих корней λ_1 и λ_2.

2.4.2 Доказательство теоремы об общем виде решения у л.р.с.п.к. второго порядка для разных корней

Th. Пусть $\lambda_1 \neq \lambda_2$. Тогда

1. $\forall c_1, c_2 \in \mathbb{C}$ последовательность $y_n = c_1\lambda_1^n + c_2\lambda_2^n$ является решением данного ЛРС.
2. Если $\{y_n\}_{n=0}^{\infty}$ удовлетворяет данному ЛРС, то $\exists c_1, c_2 y_n = c_1\lambda_1^n + c_2\lambda_2^n$.

Proof:

1. Подставим выражение для y_n в левую часть ЛРС:

$$a_2(c_1\lambda_1^{n+2} + c_2\lambda_2^{n+2}) + a_1(c_1\lambda_1^{n+1} + c_2\lambda_2^{n+1}) + a_0(c_1\lambda_1^n + c_2\lambda_2^n) =$$
$$= c_1\lambda_1^n \underbrace{(a_2\lambda_1^2 + a_1\lambda_1 + a_0)}_{0} + c_2\lambda_2^n \underbrace{(a_2\lambda_2^2 + a_1\lambda_2 + a_0)}_{0} = 0$$

2. Мы знаем, что $\{y_n\}_{n=0}^{\infty}$ удовлетворет ЛРС. Составим систему

$$\begin{cases} c_1 + c_2 = y_0, \\ c_1\lambda_1 + c_2\lambda_2 = y_1. \end{cases}$$

Определитель её матрицы равен $\Delta = \lambda_2 - \lambda_1 \neq 0$. Значит, решение c_1^*, c_2^* существует и единственно. Рассмотрим последовательность $y_n^* = c_1^*\lambda_1^n + c_2^*\lambda_2^n$. При этом уже по первому пункту y_n^* тоже является решением ЛРС. Более того, $y_0 = y_0^*$ и $y_1 = y_1^*$. Таким образом, последовательности y_n и y_n^* совпадают.

■

2.4.3 Пример: Числа Фибоначчи

Последовательность Фибоначчи задаётся ЛРС второго порядка:

$$F_{n+2} - F_{n+1} - F_n = 0, \quad F_0 = 0, \ F_1 = 1.$$

Характеристическим уравнением будет

$$x^2 - x - 1 = 0.$$

Корни $\lambda_{1,2} = \frac{1 \pm \sqrt{5}}{2}$. Отсюда общая формула

$$F_n = c_1 \left(\frac{1 + \sqrt{5}}{2} \right)^n + c_2 \left(\frac{1 - \sqrt{5}}{2} \right)^n,$$

где $c_{1,2}$ находятся из выбранных начальных значений.

2.4.4 Доказательство теоремы об общем виде решения у л.р.с.п.к. второго порядка для одинаковых корней

Th. Пусть $\lambda := \lambda_1 = \lambda_2$. Тогда

1. $\forall c_1, c_2 \in \mathbb{C}$ последовательность $y_n = (c_1 n + c_2)\lambda^n$ является решением данного ЛРС.
2. Если $\{y_n\}$ является решением, то $\exists c_1, c_2 \in \mathbb{C} y_n = (c_1 n + c_2)\lambda^n$.

Proof:

1. Подставим выражение y_n в левую часть ЛРС:

$$a_2(c_1(n+2) + c_2)\lambda^{n+2} + a_1(c_1(n+1) + c_2)\lambda^{n+1} + a_0(c_1 n + c_2)\lambda^n =$$
$$\lambda^n \left(c_1 n(a_2\lambda^2 + a_1\lambda + a_0) + c_2(a_2\lambda^2 + a_1\lambda + a_0) + a_2 \cdot 2c_1 \cdot \lambda^2 + a_1 \cdot c_1 \cdot \lambda \right) =$$
$$\lambda^n \left(c_1\lambda(2a_2\lambda + a_1) \right)$$

   При этом мы воспользовались тем, что

$$a_2\lambda^2 + a_1\lambda + a_0 = 0$$

   Значит, по теореме Виета:

$$2\lambda = -\frac{a_1}{a_2} \to 2a_2\lambda + a_1 = 0$$

   Отсюда следует, что и оставшееся выражение тоже обнуляется. Следовательно, наша последовательность является решением данного ЛРС.

2. Абсолютно аналогично случаю $\lambda_1 \neq \lambda_2$.

■

2.4.5 Общий вид решения в произвольном случае (б/д)

Общий случай $a_0, \ldots, a_k \in \mathbb{C},\ a_k \neq 0,\ a_0 \neq 0$:

$$a_k y_{n+k} + a_{k-1} y_{n+k-1} + \ldots + a_1 y_{n+1} + a_0 y_n = 0$$

Так же, как и для $k = 2$, составим характеристическое уравнение:

$$a_k x^k + a_{k-1} x^{k-1} + \ldots + a_1 x + a_0 = 0$$

Слева записано не что иное, как многочлен степени k. Если обозначить за $P_m(x)$ - произвольный многочлен степени *не выше* m от x, то выражение перепишется как

$$P_k(x) = 0$$

Согласно основной теореме алгебры, у любого многочлена с комплексными коэффициентами степени k существует ровно k комплексных корней (не обязательно разных), то есть

$$\exists \lambda_1, \ldots, \lambda_k \in \mathbb{C} P_k(x) = a_k(x - \lambda_1) \cdot \ldots \cdot (x - \lambda_k)$$

Разобьём эти k корней на r групп совпадающих. Представителей групп переобозначим как $\mu_1, \ldots, \mu_r$, а размеры групп обозначим за $m_1, \ldots, m_r$.

Th.

1. Для любых многочленов $P_{m_1-1}(n), \ldots, P_{m_r-1}(n)$ последовательность $\{y_n\}_{n=0}^{\infty}$ вида:

   $$y_n = P_{m_1-1}(n) \cdot \mu_1^n + \ldots + P_{m_r-1}(n) \cdot \mu_r^n$$

   удовлетворяет решению данного ЛРС.

2. Если $\{y_n\}_{n=0}^{\infty}$ удовлетворяет решению данного ЛРС, то существуют многочлены $P_{m_1-1}(n), \ldots, P_{m_r-1}(n)$ такие, что можно выписать y_n в общем виде, описанном выше.

2.5 Определение простого графа, орграфа, мультиграфа, псевдографа, гиперграфа. Маршруты в графах, степени вершин. Изоморфизм графов, гомеоморфизм графов. Планарность графов: определение планарного графа, формула Эйлера, верхняя оценка числа рёбер в планарном графе. Критерий Понтрягина–Куратовского (доказательство необходимости; достаточность без доказательства). Эйлеровы и гамильтоновы циклы в графах: критерий эйлеровости, достаточное условие гамильтоновости (теорема Дирака). Признак Эрдеша-Хватала (б/д).

2.5.1 Понятия теории графов

2.5.1.1 Виды графов

Def. *Графом* называется пара множеств $(V, E) = G$, где V — множество каких-то объектов, а E — множество пар объектов из V.

Def. *Ориентированным графом* G называется пара $G = (V, E)$, где V — множество вершин, а $E \subset V \times V$ — множество дуг. *Дуги* в таком графе есть упорядоченные пары вершин v_1, v_2, где v_1 называют началом, а v_2 — концом.

Def. *Граф с петлями* или *псевдограф* — граф, в котором есть ребро, исходящее из вершины и возвращающееся в ту же вершину.

Def. *Граф с кратными ребрами* или *мультиграф* — граф, имеющий ребра, инцидентные одной паре вершин. Если v_1, v_2 — вершины, а $e = (v_1, v_2)$ — соединяющее их ребро, тогда вершина v_1 и ребро e инцидентны, вершина v_2 и ребро e тоже инцидентны.

2.5.1.2 Структурные определения

Def. *Маршрутом* в графе $G = (V, E)$ называется последовательность $v_1 e_1 v_2 \dots e_n v_{n+1}$. ($e_i$ и v_i могут повторяться, ребро e_j связывает вершины v_j и v_{j+1}, между которыми находится).

Def. Если все ребра в маршруте различны, то замкнутый ($v_1 = v_{n+1}$) маршрут называется *циклом*, а незамкнутый — *путем (цепью)*.

Def. Пусть или цикл называются *простыми*, если все вершины в них различны.

Def. Две вершины u и v называются *связанными*, если в графе G существует путь из u в v (обозначение: $u \rightsquigarrow v$).

Def. *Компонентой связности* называется класс эквивалентности относительно связности.

Def. Граф $G = (V, E)$ называется *связным*, если он состоит из одной компоненты связности. В противном случае граф называется *несвязным*.

Def. *Гипеграфом* называется пара $H = (V,\ E)$, где V — множество вершин, а E — произвольное подмножество 2^V (т.е. в отличие от обычного графа, ребро гиперграфа это произвольное неупорядоченное множество вершин).

Def. Гиперграф называется k-*однородным* (для $k \geq 2$), если $\forall a \in E :\ |a| = k$.

2.5.1.3 Отношения на графах

Def. Операция *подразделения* ребра показана на рисунке.

Def. Два графа называются *гомеоморфными*, если от одного можно перейти к другому при помощи операций подразделения ребра и обратных к ним.

Def. Два графа G_1 и G_2 называются *изоморфными*, если существует такая перестановка $\varphi : V(G_1) \to V(G_2)$, что $(u, v) \in E(G_1) \iff (\varphi(u), \varphi(v)) \in E(G_2)$.

2.5.2 Планарность графов

Def. Пусть дан граф $G = (V, E)$. *Укладкой* графа G на плоскости назовем пару отображений (F, H), такую что:

- $F : V \to S$, $S \subset \mathbb{R}^2$, $|S| < \infty$ — биекция
- $H : E \to$ некоторые гладкие кривые, т.ч. $(u, v) \in E \iff H(u, v)$ соединяет $F(u)$ с $F(v)$

Def. *Плоской (планарной)* называют такую укладку, у которой никакая пара кривых, соответствующих ребрам графа G, не пересекается в точках, отличных от образа F, причем если две кривых пересекаются в точке, то эта вершина является концом этих кривых.

Def. Граф называется *планарным*, если существует его плоская укладка на плоскости.

Def. *Грани* планарной укладки — части плоскости, которые образуются в результате «разрезания» плоскости по ребрам планарного графа. Объединение этих граней представляет из себя всю плоскость.

Note. Одна из таких граней будет «бесконечной».

Th (Эйлер). Пусть граф G связен и планарен, $|V|= n$, $|E|= e$. Тогда для любой его планарной укладки с числом граней f верно равенство

$$n - e + f = 2$$

Proof: Индукция по $e - n$.

1. **База.**
   $e - n = -1 \Rightarrow G$ — дерево и $f = 1$.

2. **Переход.**
   Поскольку G не дерево, то в нем имеются циклы. Удалим из G одно ребро (из некоторого цикла), отделяющее две различные грани. Получим граф G', в котором $f' = f - 1,\ e' = e - 1,\ n' = n$. Тогда
   $$2 = n' - e' + f' = n - (e - 1) + (f - 1) = n - e + f$$

■

Consequence. Пусть G — связный планарный граф и есть какая-то его укладка. Тогда $e \leq 3n - 6$.

Proof: Заметим, что каждая грань ограничена хотя бы тремя ребрами, а каждое ребро разделяет две грани, откуда $3f \leq 2e$. Отсюда по формуле Эйлера сразу следует требуемое.

Consequence. Граф K_5 не планарен (в нем 10 ребер и 5 вершин, что не согласуется с неравенством выше).

Consequence. Пусть G — связный планарный граф и есть какая-то его укладка. Пусть t — длина наименьшего цикла в G. Тогда

$$e \geq \frac{t}{2} f$$

Proof: Пусть e_i — число ребер, отделяющих i грань от других, $i = 1, \ldots, f$. Тогда

$$2e = \sum_{i=1}^{f} e_i \geq ft.$$

Consequence. Пусть G — связный планарный граф и есть какая-то его укладка. Пусть t — длина наименьшего цикла в G. Тогда

$$e \leq \frac{t}{t-2}(n - 2).$$

Proof: По теореме Эйлера

$$2 = n - e + f \leq n - e + \frac{2}{t} e = \frac{2 - t}{t} e + n.$$

Statement. Графы K_5 и $K_{3,3}$ являются не планарными

Proof: Применяем следствие. Для K_5 : $n = 5, e = 10, t \geq 3$; для $K_{3,3} : n = 6, e = 9, t \geq 4$.

Th (Критерий Понтрягина–Куратовского) (б/д). Граф планарен тогда и только тогда, когда он не содержит подграфа, гомеоморфного K_5 или $K_{3,3}$.

Proof: Ясно, что гомеоморфные графы являются или не являются планарными одновременно, из утверждения выше следует необходимость.

■

2.5.3 Эйлеровость

Def. Граф называется *эйлеровым*, если он является циклом (т.е. существует замкнутый маршрут, проходящий по всем ребрам этого графа ровно один раз).

Def. *Ориентированным графом* G называется пара $G = (V, E)$, где V — множество вершин, а $E \subset V \times V$ — множество дуг. *Дуги* в таком графе есть упорядоченные пары вершин v_1, v_2, где v_1 называют началом, а v_2 — концом.

Def. Неориентированный граф *связен*, если для произвольной вершины все остальные вершины достижимы из нее.

Th. *(Критерий эйлеровости графа)* Для связного графа следующие утверждения эквивалентны:

1. граф эйлеров,
2. степень каждой вершины графа четная,
3. множество ребер графа распадается в объединение непересекающихся по ребрам простых циклов.

Proof:

- 1) $\Rightarrow$ 2) :
  Эйлеров граф есть цикл, а цикл есть замкнутый маршрут. Вершина, встречающаяся в этом маршруте, как составляющая, задается конструкцией $e_i v_i e_{i+1}$, т.е. каждое ее вхождение в маршрут увеличивает степень на 2 $\Rightarrow$ степень вершины четная.

  Вхождение в маршрут первой и последней вершины увеличили степень только на 1, но т.к. $v_1 = v_{n+1}$, суммарно степень увеличилась на 2.

- 2) $\Rightarrow$ 3) :
  Зафиксируем вершину x_1 ненулевой степени (актуально для повторных итераций). Выберем любую ее соседнюю вершину x_2. Так как $deg(x_2) > 0 \wedge deg(x_2) \vdots 2$, то $\exists x_3 \neq x_1 \in V$, связанная ребром с x_2. Будем идти далее по произвольному ребру из только что выбранной вершины x_k, пока не вернемся в одну из уже выбранных вершин. Тогда мы найдем некоторый простой цикл

Z_1. Удалим все его ребра из G и получим новый граф, возможно с несколькими компонентами связности. Проделаем аналогичную операцию в каждой из новых компонент связности и заметим при этом, что величина $|E|$ уменьшается. Проделав данную операцию индуктивно для каждой компоненты, мы разобьем множество E на требуемое объединение.

- 3) $\Rightarrow$ 1) :
  Доказательство по индукции. Для одного простого цикла утверждение очевидно. Предположим, что в G больше простых циклов. Удалим один простой цикл C. Полученный граф G' распадется на некоторые компоненты связности, каждая из которых распадается на простые циклы. Начнем обходить граф G по вершинам цикла C, причем если мы попали в вершину $v \in V$, лежащую в одной из компонент связности G', то обойдем ее по предположению индукции и вернемся в v. Продолжим идти по циклу C, обходя еще не посещенные компоненты связности G'. Таким образом, мы обойдем весь граф G.

Th. *(Критерий эйлеровости ориентированного графа)* В ориентированном графе $G = (V, E)$ существует эйлеров цикл тогда и только тогда, когда:

1. граф эйлеров,
2. входная степень любой вершины равна ее выходной степени,
3. множество ребер графа распадается в объединение непересекающихся по ребрам простых циклов.

Proof: Аналогично случаю неориентированного графа.

2.5.4 Гамильтоновость

Def. *Гамильтонов путь* — простой путь, проходящий через все вершины графа (причем по одному разу).

Def. *Гамильтонов цикл* — замкнутый гамильтонов путь (или простой цикл, проходящий через все вершины графа)

Def. Граф называется *гамильтоновым*, если в нем существует гамильтонов цикл

Th (признак Дирака). Если в связном графе n вершин степень любой вершины $\geq \frac{n}{2}$, то этот связный граф — гамильтонов.

Proof: Пусть $P = v_1 v_2 \dots v_k$ — самый длинный путь в графе G. Если v_1 смежна с некоторой вершиной $x \notin P$, то существует путь длиннее P — противоречие. Аналогичное рассуждение с $v_k \Rightarrow v_1$ и v_k смежны **только** с вершинами из P. Поскольку $\deg(v_1) \geq \frac{n}{2}$ и в графе нет петель, то $k \geq \frac{n}{2} + 1$.

Statement. Существует $1 \leq j \leq k$, такое что v_j инцидентна с v_k, а v_{j+1} с v_1.

Proof: Предположим, что такой ситуации не оказалось. Тогда в P есть как минимум $\deg(v_1)$ вершин, несвязанных с v_k (предыдущие в пути от соседей v_1). Поскольку все вершины, связанные с v_k находятся в пути и v_k не инцидентная сама с собой, то в P хотя бы $\deg(v_1) + \deg(v_k) + 1 = n + 1$ вершин. Противоречие.

Из утверждения следует, что в G существует простой цикл $C = v_{j+1} \dots v_k v_j v_{j-1} \dots v_1 v_{j+1}$. Покажем, что этот цикл — гамильтонов. Предположим, что существует $v \in V \setminus C$. Поскольку граф связен, v должна быть связана каким-то путем с некоторой $v_i \in C$. Но тогда существует путь $P' = v$ — путь от v до C — круг по C, длиннее чем P, что противоречит выбору P.

■

Def. *Вершинной связностью* графа G называется минимальное количество вершин, в результате удаления которых граф перестает быть связным. Обозначается $\kappa(G)$.

Def. *Числом независимости* графа G называется максимальное количество вершин в свободном от ребер подмножестве, называемым назависимым подмножеством:

$$\alpha(G) = max\{|W|\colon \forall x, y \in W \hookrightarrow \{x, y\} \neq E\}$$

Th (Эрдёш, Хватал) (б/д). Пусть количество вершин в графе $G(V, E)$ $|v| \geqslant 3$ и $\alpha(G) \leqslant \kappa(G)$. Тогда граф гамильтонов.

2.6 Хроматическое число, число независимости, кликовое число. Нижняя оценка хроматического числа через число независимости и через кликовое число; сравнение порядка этих оценок для случайного графа $G(n, 1/2)$ в модели Эрдёша–Реньи (а.п.н. $\alpha(G) \leq 2\log_2(n)$). Теорема Эрдёша о существовании графов с произвольно большим обхватом и хроматическим числом.

2.6.1 Характеристики графа

Def. *Хроматическим числом* графа G называется величина

$$\chi(G) = min\{k \in \mathbb{N} : V = V_1 \sqcup \dots \sqcup V_k \ \forall i \, \forall x, \ y \in V_i (x, y) \notin E\}.$$

— минимальное число цветов, в которые можно раскрасить вершины графа G так, чтобы концы любого ребра имели разные цвета.

Def. Пусть дан граф $G = (V,\ E)$. Тогда его *числом независимости* называется число

$$\alpha(G) = \max\{k \in \mathbb{N} :\ \exists W \subseteq V :\ |W| = k \wedge \forall x,\ y \in W\ (x,\ y) \notin E\}.$$

— размер максимальной антиклики в графе G.
Множество вершин W, между любыми двумя из которых нет ребра, называется *независимым* множеством вершин.

Statement. $\chi(G) \geq \frac{|V|}{\alpha(G)}$.

Proof: Действительно, $|V| = \sum\limits_{i=1}^{\chi(G)} |V_i| \leq \chi(G) \cdot \max|V_i| \leq \chi(G) \cdot \alpha(G)$.

Def. *Кликовым числом* графа G называется величина

$$\omega(G) = \max\{k \in \mathbb{N} :\ \exists W \subseteq V :\ |W| = k\ \wedge \forall x,\ y \in W\ (x,\ y) \in E\}.$$

Заметим следующие очевидные соотношения между числом независимости, кликовым числом и хроматическим числом графа:

1. $\alpha(G) = \omega(\overline{G})$
2. $\chi(G) \geq \max\{\omega(G),\ \frac{|V|}{\alpha(G)}\}$

Для удобства будем считать, что множеством вершин случайного графа является $V = V_n = \{1, \ldots, n\}$.

Def. Зафиксируем число вершин в графе n. Рассмотрим полный граф K_n и занумеруем все его ребра в некотором порядке: $e_1, e_2, \ldots, e_N$, где $N = C_n^2$. Пусть вероятность «появления» ребра в графе равна p, все ребра появляются с равной вероятностью и независимо друг от друга. Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, \mathsf{P})$, где $\Omega = \{$множество последовательностей из 0 и 1 длины $C_n^2\}$, $\mathcal{F} = 2^\Omega$, $\mathsf{P} = p^{|E|}(1-p)^{C_n^2 - |E|}$ и $|E|$ — это число единиц в $w \in \Omega$. Тогда граф $G = G(n,p) = (\Omega_n, \mathcal{F}_n, \mathsf{P}_{n,p})$ является *моделью Эрдеша-Реньи случайного графа*. При этом 0 на i-том месте означает отсутствие ребра e_i в графе, а 1 — присутствие.

Th. Вероятность того, что граф $G(n, 0.5)$ имеет $\omega(G) \leq 2\log_2 n$ стремится к единице при $n \to \infty$.

Note. Так как вероятность проведения ребра равна 0.5, получим, что абсолютно так же, что $\alpha(G) < 2\log_2 n$ при тех же условиях.

Proof: Пусть $k = [2log_2 n]$, $X_k(G)$ — число k-клик в G. Если в графе есть $(k+1)$-клика, то есть $k+1$ k-клик, то есть они могут иметь общие вершины.

Тогда $P(\omega(G) < k) = P(X_k = 0) = 1 - P(X_k \geq 1)$.

Воспользуемся неравенством Маркова: $P(|\xi| < a) \geq \frac{\mathbb{E}\xi}{a}$, тогда введем величину $X_{k,i}(G)$:

$$X_{k,i}(G) = \begin{cases} 1 & \text{, если } i\text{-й набор вершин образует } k\text{-клику} \\ 0 & \text{, иначе} \end{cases}$$

Тогда верно, что

$$P(\omega(G) < k) = P(X_k = 0) = 1 - P(X_k \geq 1) \geq 1 - \mathbb{E}X_k = 1 - \mathbb{E}\left(\sum_{i=1}^{C_n^k} X_{k,i}\right) = 1 - C_n^k \cdot \left(\frac{1}{2}\right)^{\frac{k(k-1)}{2}}$$

$$C_n^k \cdot 2^{-\frac{k(k-1)}{2}} \leq \frac{n^k}{k!} \cdot 2^{-\frac{k^2}{2}+\frac{k}{2}} = \frac{2^{k\log_2 n}}{k!} \cdot 2^{-\frac{k^2}{2}+\frac{k}{2}} \leq \frac{2^{2\log_2^2 n}}{k!} \cdot 2^{\left(-\frac{(2log_2 n-1)^2}{2}+\log_2 n\right)} = \frac{2^{\left(2\log_2 n-\frac{1}{2}\right)}}{k!} \to 0$$

Значит $P(\omega(G) < k) \to 1$, что и требовалось доказать.

■

2.6.2 Существование графов с произвольно большим обхватом и хроматическим числом

Def. *Обхватом графа* называют длину наименьшего цикла в нем. Если граф ацикличен, то обхват равен бесконечности. Обозначается как $g(G)$

Th (Эрдёш). $\forall l, k \in \mathbb{N}\ \exists G:\ \chi(G) > k, g(G) > l.$

(Если сложно читать док-во, можно посмотреть на харизматичного лектора по ссылке)

Proof: Пусть $\theta = \frac{1}{2l}$, тогда положим, что у нас модель $G(n, n^{\theta-1})$. Рассмотрим такой случайный граф. Пусть случайная величина $X(G)$ — количество циклов в G, имеющих длину не более l, то есть число «плохих» циклов. тогда

$$\mathbb{E}X = \sum_{i=1}^{l} C_n^i \frac{(i-1)!}{2} p^i < \sum_{i=1}^{l} (np)^i = \sum_{i=1}^{l} n^{\theta i} < l n^{\theta l} = l\sqrt{n}$$

Откуда по неравенству Маркова

$$P\left(X > \frac{n}{2}\right) \leq \frac{\mathbb{E}X}{n/2} \to 0 \Longrightarrow P\left(X \leq \frac{n}{2}\right) \to 1$$

То есть $\exists n_0:\ \forall n > n_0\ P\left(X \leq \frac{n}{2}\right) > \frac{1}{2}$.

Пусть теперь Y_m — число независимых множеств размера m в графе.

$$\mathbb{E}Y_m = C_n^m (1-p)^{C_m^2} \leq n^m e^{-pC_m^2} = \exp\left(m\log n - \frac{pm^2}{2}(1-o(1))\right)$$

Заметим, что если $m \sim \frac{3\log n}{p}$, откуда $\mathbb{E}Y_m \to 0$ с ростом n, откуда существует n_1, начиная с которого $\mathbb{E}Y_m < \frac{1}{2}$. По неравенству Маркова:

$$P(Y_m = 0) = 1 - P(Y_m \geq 1) \geq 1 - \frac{\mathbb{E}Y_m}{1} = 1 - \mathbb{E}Y_m > \frac{1}{2}$$

То есть у нас есть два события с вероятностью большей $\frac{1}{2}$, а значит они пересекаются! То есть $\exists n > \max\{n_0, n_1\}$ такое, что существует G с следующими свойствами:

$$X(G) \leq \frac{n}{2} \quad Y_m(G) = 0$$

Заметим, что $\alpha(G) < m$, так как $Y_m(G) = 0$, то есть независимых множеств размера m просто нет.

Выкинем из каждого цикла длины не большей l по одной вершине, то есть не более половины вершин, получим G'.

$$|V(G')| \geq \frac{n}{2},\ g(G') > l,\ \alpha(G') \leq \alpha(G) < m$$

откуда

$$\chi(G') \geq \frac{|V(G)|}{\alpha(G)} \geq \frac{n}{2m} \sim \frac{np}{6\log n} = \frac{n^\theta}{6\log n} \to \infty$$

То есть $\exists n_2$, начиная с которого $\chi(G') > k$. А значит для $n > \max\{n_0, n_1, n_2\}$ верно, что есть G' такой, что

$$\chi(G') > k, g(G') > l$$

■

2.7 Системы общих представителей (с.о.п.): определение, примеры задач, сводящихся к построению с.о.п.. Тривиальные верхняя и нижняя оценки размера минимальной с.о.п.. Жадный алгоритм построения с.о.п., теорема о верхней оценке размера «жадной с.о.п.». Теорема о неулучшаемости этой оценки в общем случае (б/д).

2.7.1 Системы общих представителей

Определим гиперграф (множество вершин) $R_n := \{1, \dots, n\}$ и совокупность его подмножеств, которые в терминах гиперграфа можно считать ребрами: $\mathcal{M} := \{M_1, \dots, M_s \mid \forall i\, M_i \subseteq R_n \ \wedge\ M_i \neq M_j\}$.

Def. *Системой общих представителей* (далее — соп) для совокупности множеств $\mathcal{M}$ назовем любое $S \subseteq R_n$, т.ч. $\forall i\ M_i \cap S \neq \varnothing$. (т.е. S содержит хотя бы по одной вершине из каждого ребра)

$$\tau(\mathcal{M}) := min\{\tau \in \mathbb{N} \mid \exists S \subseteq R_n, |S| = \tau \ \wedge\ S \text{ — соп для } \mathcal{M}\}.$$

Note. Для гиперграфа $H = (R_n, \mathcal{M})$ соп системы $\mathcal{M}$ — это вершинное покрытие H.

2.7.1.1 «Тривиальные» нижние и верхние оценки

Пусть $\forall i\, |M_i| = k$, $|\mathcal{M}| = s$ и $M_i \subseteq R_n$. При фиксированных n, s, k обозначим $\mathcal{M}$ с такими параметрами за $\mathcal{M}(n, k, s)$. Заметим, что количество $\mathcal{M}$ с такими параметрами равно $C^s_{C^k_n}$.

Th.

$$\forall \mathcal{M}: \ \tau(\mathcal{M}) \leq \min\{s, n-k+1\}$$

Proof: От $n-k+2$ до n ровно $k-1$ число, а значит, взяв все числа от 1 до $n-k+1$, мы получим соп.

Th.

$$\exists \mathcal{M}: \ \tau(\mathcal{M}) \geq \min\left\{\left[\frac{n}{k}\right], s\right\}$$

Proof: Возможно два случая:

1. $s \leq \left[\frac{n}{k}\right]$. Тогда $\mathcal{M} = \{\{1, 2, \ldots, k\}, \{k+1, \ldots, 2k\}, \ldots, \{(s-1)k+1, \ldots, sk\}\}$
2. $s > \left[\frac{n}{k}\right]$. В таком случае систему $\mathcal{M}$ так же, как и в первый раз, добирая новые множества пока можем, а после добавляем произвольные множества, пока $|\mathcal{M}|$ не равна s.

2.7.2 Жадный алгоритм

Пусть $\forall i \, |M_i| = k$, $|\mathcal{M}| = s$ и $M_i \subseteq R_n$. При фиксированных n, s, k обозначим $\mathcal{M}$ с такими параметрами за $\mathcal{M}(n, k, s)$. Заметим, что количество $\mathcal{M}$ с такими параметрами равно $C^s_{C^k_n}$.

Th.

$$\forall n, k, s \ \ \forall \mathcal{M} = \mathcal{M}(n, k, s) \ \ \tau(\mathcal{M}) \leq \max\{\frac{n}{k}, \frac{n}{k} \ln \frac{sk}{n}\} + \frac{n}{k} + 1$$

Proof: Зафиксируем $\mathcal{M}$. Возможны следующие случаи:

1. $s \leq \frac{n}{k} \Rightarrow \tau(\mathcal{M}) \leq s \leq \frac{n}{k}$
2. $\frac{n}{k} \ln \frac{sk}{n} \geq n \Rightarrow \tau(\mathcal{M}) \leq n \leq \frac{n}{k} \ln \frac{sk}{n}$
3. $s > \frac{n}{k}$, $\frac{n}{k} \ln \frac{sk}{n} < n$.

Для доказательства последнего случая воспользуемся *жадным алгоритмом* построения соп.

Возьмем любой элемент $\nu_1 \in R_n$, который принадлежит наибольшему числу множеств в $\mathcal{M}$. Пусть их ρ_1 штук. Тогда $\rho_1 \geq \frac{sk}{n}$, поскольку $sk = \sum\limits_{i=1}^{n} \sum\limits_{M \in \mathcal{M}} I_{\{i \in M\}} \leq \rho_1 n$. Выкинем из $\mathcal{M}$ все множества, содержавшие ν_1. Осталась совокупность $\mathcal{M}_1$, $|\mathcal{M}_1| = s - \rho_1 = s_1$. Снова сделаем шаг жадного алгоритма.

Всего сделаем $N = \left[\frac{n}{k} \ln \frac{sk}{n}\right] + 1$ шагов ж.а, причем $\rho_i \geq \frac{s_{i-1}k}{n}$. После этого имеем построенное ж.а. множество $S = \{\nu_1, \ldots, \nu_N\}$ и совокупность $\mathcal{M}_N$ т.ч.

$$|\mathcal{M}_N| = s_N = s_{N-1} - \rho_N \leq s_{N-1} - \frac{s_{N-1}k}{n} \leq \ldots \leq s(1 - \frac{k}{n})^N = se^{N \ln(1-\frac{k}{n})} \leq se^{-\frac{kN}{n}} \leq se^{-\frac{k}{n} \cdot \frac{n}{k} \ln \frac{sk}{n}} = \frac{n}{k}$$

Итого $\tau(\mathcal{M}) \leq N + \frac{n}{k} \leq \frac{n}{k} \ln \frac{sk}{n} + 1 + \frac{n}{k}$

■

2.7.3 Неулучшаемость жадной соп

Th (б/д). Пусть $n \to \infty$, $k = k(n) \to \infty$, $\frac{sk}{n} \to \infty$.

Пусть также $\ln\ln k = o\left(\ln \frac{sk}{n}\right)$, $k^2 = o(n)$, $\ln^2 \frac{sk}{n} = o(k)$.

Тогда

$$\exists n_0 : \forall n \geq n_0 \, \exists \mathcal{M}(n,k,s) : \tau(\mathcal{M}) \geq \frac{n}{k}\ln\frac{sk}{n} - \frac{n}{k}\ln\ln\frac{sk}{n} - \frac{n}{k}\ln\ln k - \frac{3n}{k}$$

Note. То есть по сути у нас асимптотическое равенство верхней и нижней оценок. Более того, из теоремы следует асимптотическая неулучшаемость оценки жадного алгоритма.

2.8 Числа Рамсея: определение, и точные значения $R(s,t)$ при $s \leq 3, t \leq 4$. Верхняя оценка Эрдёша–Секереша, её следствие для диагональных чисел Рамсея; нижняя оценка диагональных чисел с помощью простого вероятностного метода.

2.8.1 Числа Рамсея $R(s,t)$ точные значения для $s+t \leq 7$

Def. Пусть $s,t \in \mathbb{N}$. *Число Рамсея* $R(s,t) := \min\{n \in \mathbb{N} :$ при любой раскраске ребер K_n в красный и синий цвета либо найдется K_s, все ребра у которого красные, либо K_t, все ребра которого синие$\}$.

Эквивалентное определение $R(s,t) := \min\{n \in \mathbb{N} : \forall\ G = (V,\ E), |V| = n$ и либо $\omega(G) \geq s$, либо $\alpha(G) \geq t\}$. ($\omega(G)$ — кликовое число, $\alpha(G)$ — число независимости)

Note.

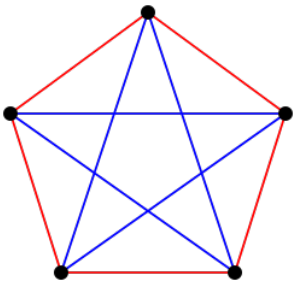
0. Число Рамсея симметрично по своим аргументам.
1. $R(1,t) = 1$
2. $R(2,t) = t$
3. $R(3,t)$ никто не знает.

Th. $R(3,3) = 6$

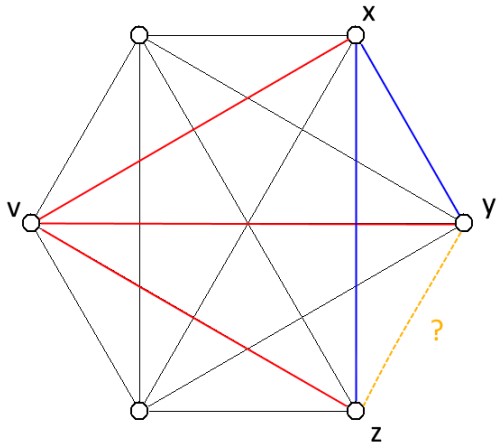
Proof:

1. Покажем, что для K_5 найдется такая раскраска, при которой в нем не будет ни синего, ни красного K_3. Так, $R(3,3) > 5$.

n, m	1	2	3	4	5	6
1	1	1	1	1	1	1
2	1	2	3	4	5	
3	1	3	6	9		
4	1	4	9			
5	1	5				
6	1					



2. Покажем, что $R(3,3) \leq 6$. Рассмотрим некоторую вершину v и исходящие из нее 5 ребер. Хотя бы 3 из них будут одного цвета, б.о.о. — красного $((v,x),(v,y),(v,z))$. Теперь рассмотрим ребра в треугольнике, построенном на вершинах x, y, z. Если хотя бы одно из них красное, мы получили красный K_3. Так, все эти ребра должны быть синими, но тогда мы получили синий K_3.

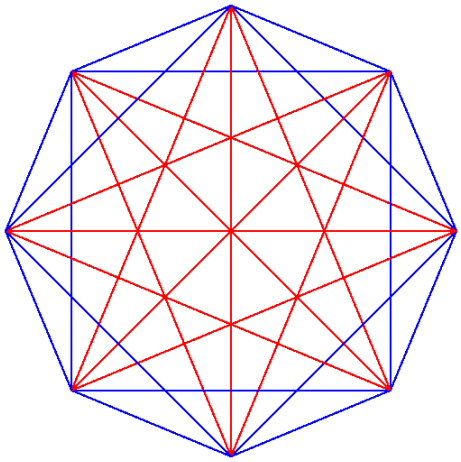


Th. $R(3,4) = 9$

Proof:

1. Покажем, что для K_8 найдется такая раскраска, при которой в нем не будет красных K_3 и синих K_4 (пикча ниже).

2. Воспользуемся рекуррентной верхней оценкой. $R(3,4) \leq R(2,4) + R(3,3) - 1 = 9$. (следствие из т. Эрдеша-Секереша)



2.8.2 Рекуррентная верхняя оценка Эрдеша-Секереша

Th (Эрдеш, Секереш, 1935).

$$R(s,t) \leq R(s-1,t) + R(s,t-1)$$

Proof:

$n := R(s-1,t) + R(s,t-1)$. Зафиксируем произвольную раскраску K_n в 2 цвета и вершину $v \in V$. Из нее выходит $n-1$ ребро. От противного получаем, что из нее выходит либо $\geq R(s-1,t)$ красных ребер, либо $\geq R(s,t-1)$ синих. Без ограничения общности считаем, что $\geq R(s-1,t)$ красных. $V_1 := \{u \in V,\ (v,u) - \text{красное}\}$. Поскольку $|V_1| \geq R(s-1,t)$, то в V_1 есть либо синий K_t, либо красный K_{s-1}, который вместе с вершиной v дает искомый красный K_s.

Consequence. $R(s,t) \leq C_{s+t-2}^{t-1}$ — индукция по s и t.

Consequence. $R(s,s) \leq C_{2s-2}^{s-1} = \dfrac{4^{s-1}}{\sqrt{\pi s}}(1+o(1))$

Consequence. $R(3,3) \leq C_4^2 = 6$, при этом $R(3,3) > 5$ (цикл на 5 вершинах).

Consequence. (скатано с вики) Если $R(s-1,t)$ и $R(s,t-1)$ - четные числа, можно усилить оценку: $R(s,t) \leq R(s-1,t) + R(s,t-1) - 1$

2.8.3 Нижняя оценка диагональных чисел с помощью простого вероятностного метода

Th. Нижняя оценка $R(s,s)$

$$R(s,s) \geq (1+o(1))\frac{s}{e\sqrt{2}} \cdot 2^{s/2}$$

Proof: $n := (1+o(1))\frac{1}{e\sqrt{2}} \cdot s2^{s/2}$. Покажем, что существует раскраска K_n, в которой нет одноцветного K_s. Рассмотрим случайную раскраску ребер K_n в два цвета, где $\mathcal{P}(e - \text{красное }) = \mathcal{P}(e - \text{синее}) = \frac{1}{2}$.

Пусть $S \subset V$, $|S|= s$ и $A_S = \{K_s,$ порожденный S, — одноцветный $\}$. Тогда $\mathcal{P}(A_S) = 2 \cdot 2^{-C_s^2} = 2^{1-C_s^2}$.

$$\mathcal{P}(\bigcup_{\substack{S\subset V\\|S|=s}} A_S) \le \sum_{\substack{S\subset V\\|S|=s}} \mathcal{P}(A_S) = C_n^s 2^{1-C_s^2} \le \frac{n^s}{s!} 2^{1-C_s^2} = \frac{n^s}{s!} 2^{1-s^2/2+s/2} =$$

$$= \frac{1}{s!}(1+o(1))^s \cdot \frac{1}{e^s 2^{s/2}} \cdot s^s 2^{s^2/2} 2^{1-s^2/2+s/2} = \frac{(1+o(1))^s}{\sqrt{2\pi s}\frac{s^s}{e^s}(1+o(1))}\frac{s^s}{e^s} \cdot 2 = \frac{2}{\sqrt{2\pi s}} \cdot \frac{(1+o(1))^s}{(1+o(1))}$$

Подбором $o(1)$ в числителе и знаменателе можно сделать так, что полученное число < 1 при всех s.

■

2.9 Гиперграфы. Гиперграфы t-пересечений. Теорема Эрдёша–Ко–Радо (о максимальном числе рёбер в гиперграфе 1-пересечений). Основы линейно-алгебраического метода: теорема Франкла–Уилсона (верхняя оценка на $m(n, r, s)$ через числа сочетаний для случай $r - s = p$, p простое, $r < 2p$), конструктивная нижняя оценка чисел Рамсея (формулировка, определение графа, док-во леммы про число независимости).

2.9.1 Гиперграфы, ЭКР

2.9.1.1 Гиперграфы t-пересечений

Def. *Гипеграфом* называется пара $H = (V,\ E)$, где V — множество вершин, а E — произвольное подмножество 2^V (т.е. в отличие от обычного графа, ребро гиперграфа это произвольное неупорядоченное множество вершин).

Def. Гиперграф называется *k-однородным* (для $k \ge 2$), если $\forall a \in E :\ |a|= k$.

Def. $f(n,\ k,\ t) = \max\{f \in \mathbb{N} :\ \exists\ k$-однородный гиперграф $H = (V,\ E)$, $|V|= n$, $|E|= f$, $\forall A,\ B \in E :\ |A \cap B| \ge t\}$

Def. $m(n,\ k,\ t) = \max\{f \in \mathbb{N} :\ \exists\ k$-однородный гиперграф $H = (V,\ E)$, $|V|= n$, $|E|= f$, $\forall A,\ B \in E :\ |A \cap B| \ne t\}$

Для примера рассмотрим граф $G(n,\ r,\ s)$ (надпись для тех, кто не помнит что это). Интерпретируем вершины этого графа как ребра некоторого r-однородного гиперграфа, а пару вершин, пересекающихся по s элементам — ребром. Легко понять, что $\alpha(G(n,\ 3,\ 1)) = m(n,\ 3,\ 1)$, и вообще

$$\alpha(G(n,\ k,\ t)) = m(n,\ k,\ t)$$

2.9.1.2 Теорема Эрдеша–Ко–Радо (о максимальном числе ребер в гиперграфе 1-пересечений)

Th. (Эрдеш-Ко-Радо)

$$f(n,\ k,\ 1) = \begin{cases} C_n^k & 2k > n \\ C_{n-1}^{k-1} & 2k \le n \end{cases}$$

Proof: Первый случай очевиден. Верхняя оценка $f(n,k,1) \ge C_{n-1}^{k-1}$ в случае $2k \le n$ тоже проста: достаточно рассмотреть совокупнсоть $\mathscr{M} = \{M \subset [n],\ |M| = k \wedge 1 \in M\}$.

Покажем теперь, что $f(n,k,1) \le C_{n-1}^{k-1}$. Рассмотрим совокупность $\mathcal{F} = H_E = \{F_1,\ \ldots,\ F_s\}$, $\forall i\ |F_i| = k$, $\forall i,\ j:\ |F_i \cap F_j| \ge 1$. Наша цель показать, что $s \le C_{n-1}^{k-1}$.

Рассмотрим семество множеств $\mathscr{A} = \{A_1, \ldots, A_n\}$, где $A_1 = \{1, 2, \ldots, k\}$, $A_2 = \{2, \ldots, k+1\}, \ldots, A_n = \{n,\ 1,\ 2,\ \ldots,\ k-1\}$. Докажем сначала следующую лемму:

Lemma. $|\mathcal{F} \cap \mathscr{A}| \le k$ *(круговой метод Катона)*

Proof: Если $\mathcal{F} \cap \mathscr{A} = \varnothing$, то все очевидно. Иначе, без ограничения общности, считаем, что $A_1 \in \mathcal{F}$. Все остальные $A_i \in \mathcal{F} \cap \mathscr{A}$ должны пересекать A_1 и пересекаться между собой. Разобем их на пары следующим образом: $(A_i,\ A_{n-k+i})$ для $i \ge 2$ (например, пара $(A_2,\ A_{n-k+2})$ — A_2 начинается с 2, а A_{n-k+2} кончается в 1). Тогда $A_i \cap A_{n-k+i} = \varnothing$.

Рассмотрим следующие пары: $(A_2,\ A_{n-k+2}),\ \ldots,\ (A_k,\ A_n)$. В этих парах все множества пересекают A_1, но при этом два множетсва из одной пары не пересекаются. Это означает, что в $\mathscr{A} \cap \mathcal{F}$ не более одного множества из каждой пары, откуда следует

$$|\mathscr{A} \cap \mathcal{F}| \le 1\ (A_1) + (\text{количество пар}) = 1 + k - 1 = k$$

и лемма доказана.

Изначально $V = \{1,\ 2,\ \ldots,\ n\}$. Рассмотрим любую перестановку $\sigma \in S_n$. Определим множества $V_\sigma = \{\sigma(1),\ \ldots,\ \sigma(n)\}$ и $\mathscr{A}_\sigma = \{\sigma(A_1),\ \ldots,\ \sigma(A_n)\}$, где $\sigma(A_i)$ означает множетсво $\{\sigma(i),\ \sigma(i+1),\ \ldots\}$. Например, для $n = 7$ и σ такой, что $V_\sigma = \{2,\ 5,\ 1,\ 3,\ 4,\ 6,\ 7\}$ совокупность $\mathscr{A}_\sigma$ это множество $\{\{2,\ 5,\ 1\},\ \{5,\ 1,\ 3\},\ \ldots,\ \{6,\ 7,\ 2\},\ \{7,\ 2,\ 5\}\}$

Lemma. $|\mathcal{F} \cap \mathscr{A}_\sigma| \le k$ — доказательство аналогично предыдущей лемме.

Определим индикаторы $I(\sigma,\ F_i) = \begin{cases} 1 & F_i \in \mathscr{A}_\sigma, \\ 0 & \text{иначе} \end{cases}$ и посмотрим на следующую величину:

$$\sum_\sigma \sum_{i=1}^{s} I(\sigma,\ F_i) = \sum_{i=1}^{s} \sum_\sigma I(\sigma,\ F_i)$$

При фиксированной перестановке сумма $\sum\limits_{i=1}^{s} I(\sigma,\ F_i) = |\mathscr{A}_\sigma \cap \mathcal{F}| \le k$, а значит сумма слева не превосходит $n!\,k$. С другой стороны, при фиксированном i, F_i можт оказаться на одном из n мест в множестве

$\mathscr{A}_\sigma$, и перестановок, в которых возникает F_i, ровно $k!\,(n-k)!$, а значит $\sum\limits_\sigma I(\sigma,\ F_i) = nk!\,(n-k)!$ и вся сумма справа $= snk!\,(n-k)!$. Окончательно получаем

$$snk!\,(n-k)! = \sum_{i=1}^{s}\sum_{\sigma} I(\sigma,\ F_i) = \sum_{\sigma}\sum_{i=1}^{s} I(\sigma,\ F_i) \le kn! \Rightarrow s \le C_{n-1}^{k-1}$$

■

2.9.2 Теорема Франкла-Уилсона

Th. Пусть $r - s = p$, p простое, $r < 2p$, тогда

$$m(n, r, r-p) \le \sum_{k=0}^{p-1} C_n^k$$

Proof: Рассмотрим произвольное множество:

$$W = \{A_1, \ldots, A_t \,|\, |A_i| = r,\ |A_i \bigcap A_j \ne s\} \Longleftrightarrow W = \{\overline{x_1}, \ldots, \overline{x_t}\,|\overline{x_i} \in \{0,1\}^n,\ x_i^1 + \ldots + x_i^n = r,\ (\overline{x_i}, \overline{x_j}) \ne s\}$$

Теперь введем многочлены для: $\overline{x_i}$, тогда $P_{\overline{x_i}} \in \mathbb{Z}_p[y_1, \ldots, y_n]$,

$$P_{\overline{x_i}}(\overline{y}) = \prod_{j=0,\ j \not\equiv r(p)}^{p-1} (j - (\overline{x_i}, \overline{y})) \to P'_{\overline{x_i}}(\overline{y})$$

Где $P'_{\overline{x_i}}(\overline{y})$ получится путем того, что все степени больше единицы у y_i станут равными единице, что не меняет значения многочлена, так как $x_i^k, y_i \in \{0,1\}$.

Тогда покажем, что многочлены вида P' задают базис, то есть линейно независимы, а всего их $\sum\limits_{k=0}^{p-1} C_n^k$.

$$c_1 P'_{\overline{x_1}} + \ldots + c_t P'_{\overline{x_t}} = 0 \Longrightarrow \forall \overline{y} \in \{0,1\}^n\ c_1 P'_{\overline{x_1}}(\overline{y}) + \ldots + c_t P'_{\overline{x_t}}(\overline{y}) \equiv 0\ (mod\ p) \Longrightarrow$$

$$c_1 P_{\overline{x_1}} + \ldots + c_t P_{\overline{x_t}} = 0 \Longrightarrow \forall \overline{y} \in \{0,1\}^n\ c_1 P_{\overline{x_1}}(\overline{y}) + \ldots + c_t P_{\overline{x_t}}(\overline{y}) \equiv 0\ (mod\ p)$$

Теперь рассмотрим

$$P_{\overline{x_i}}(\overline{x_i}) = \prod_{j=0,\ j \not\equiv r(p)}^{p-1} (j - (\overline{x_i}, \overline{x_i})) \not\equiv 0\ (mod\ p),\ P_{\overline{x_i}}(\overline{x_j}) = \prod_{j=0,\ j \not\equiv r(p)}^{p-1} (j - (\overline{x_i}, \overline{x_j})) \equiv 0\ (mod\ p)$$

Последнее сравнение по модулю верно, так как $(\overline{x_i}, \overline{x_j}) \ne s, r, s-p$ (для разности верно, так как $r < 2p$), то есть найдется скобка в произведении в определении многочлена, для которой $j \equiv (\overline{x_i}, \overline{x_j})\ (mod\ p)$. Тогда

$$\forall i\ P_{\overline{x_1}}(\overline{x_i}) + \ldots + c_t P_{\overline{x_t}}(\overline{x_i}) \equiv 2c_i\ (mod\ p) \Longrightarrow 2c_i \equiv 0\ (mod\ p) \Longrightarrow c_i \equiv 0\ (mod\ p)$$

Все переходы верны в силу простоты p, а значит было получено, что эти одночлены $P'_{\overline{x_i}}$ действительно линейно независимы, а значит задают базис. Таким образом, теорема доказана.

Note. Порождающие мономы: $1, y_1, \ldots, y_{p-1}, y_1y_2, \ldots y_{p-2}y_{p-1}, \ldots, y_1 \cdot \ldots \cdot y_{p-1}$

■

2.9.3 Конструктивная нижняя оценка чисел Рамсея

2.9.3.1 Формулировка

Th (Франкл-Уилсон).

$$R(s,s) \geq \left(\sqrt[4]{e} + o(1)\right)^{\frac{\log^2 s}{\log\log s}}$$

2.9.3.2 Определение графа

Рассмотрим простое число p. Пусть $m = p^3$, $k = p^2$. Множество вершин:

$$V = \{\overline{x} = (x_1, \ldots, x_m) \mid x_i \in \{0,1\},\ x_1 + \ldots + x_m = k\}$$

А множество ребер:

$$E = \{(\overline{x}, \overline{y}) \in V \times V \mid \langle \overline{x}, \overline{y} \rangle \equiv 0 \pmod{p}\}$$

2.9.3.3 Лемма про число независимости

Lemma. Для графа выше $\alpha(G) \leq \sum\limits_{i=0}^{p-1} C_m^i$.

Proof: пусть $W = \{\overline{x_1}, \ldots, \overline{x_t}\}$ — независимое множество графа G. Сопоставим каждому $\overline{x_i}$ многочлен:

$$F_{\overline{x_i}}(\overline{y}) = \prod_{j=1}^{p-1} (j - \langle \overline{x_i}, \overline{y} \rangle)$$

Сразу отметим, что $deg F_{\overline{x_i}}(\overline{y}) \leq p-1$.

$$F_{\overline{x_i}}(\overline{x_i}) = \prod_{j=1}^{p-1} (j - \langle \overline{x_i}, \overline{x_i} \rangle) = \prod_{j=1}^{p-1} (j-k) = \prod_{j=1}^{p-1} (j-p^2) \not\equiv 0 \pmod{p}$$

С другой стороны, $F_{\overline{x_j}}(\overline{x_i}) \equiv 0 \pmod{p}$ так как они из независимого множества, то есть $\langle \overline{x_i}, \overline{x_j} \rangle$ не сравнимо с нулем, а значит имеет какой-то ненулевой остаток, который j пробегает.

Наша великая цель: доказать, что $F_{\overline{x_i}}(\overline{y})$ линейно-независимы над $\mathbb{Z}_p$. Допустим, что это не так, то есть найдутся $c_1, \ldots, c_t$ такие, что для любого y

$$c_1 F_{\overline{x_i}}(\overline{y}) + \ldots + c_t F_{\overline{x_t}}(\overline{y}) \equiv 0 \pmod{p}$$

Пусть $y = \overline{x_1}$, тогда все члены, кроме первого, зануляются, а значит

$$c_1 F_{\overline{x_i}}(\overline{x_1}) \equiv 0 \pmod{p} \Longrightarrow c_1 \equiv 0 \pmod{p}$$

Аналогично получаем, что $c_i \equiv 0 \pmod{p}$, то есть показали линейную независимость. Разберемся с размерностью.

Рассмотрим многочлен $\widetilde{F_{\overline{x_i}}}(\overline{y})$, полученный по следующему правилу:

$$F_{\overline{x_i}}(\overline{y}) = \sum c y_{i_1}^{\alpha_{i_1}} \ldots y_{i_q}^{\alpha_{i_q}} \to \widetilde{F_{\overline{x_i}}}(\overline{y}) = \sum c y_{i_1} \ldots y_{i_q}$$

Заметим, что $\forall i, j$ верно, что $F_{\overline{x_i}}(\overline{x_j}) = \widetilde{F_{\overline{x_i}}}(\overline{x_j})$, откуда следует и линейная независимость многочленов с волной, а их число как раз не превышает $\sum\limits_{i=0}^{p-1} C_m^i$, что и требовалось показать.

■

Глава 3

Теория вероятностей и математическая статистика

3.1 Вероятностное пространство, аксиомы Колмогорова, свойства вероятностной меры (в том числе теорема о непрерывности вероятностной меры с доказательством). Условные вероятности. Формула полной вероятности. Формула Байеса. Независимость.

3.1.1 Вероятностное пространство $(\Omega, \mathcal{F}, \mathcal{P})$. Свойства меры.

3.1.1.1 Вероятностное пространство

Def. Ω — (произвольное) множество элементарных событий (исходов)

Def. ω — элементарный исход, если $\omega \in \Omega$

Def. Семейство σ подмножеств X — сигма-алгебра, если:

1. $X \in \sigma$
2. Если $E \subset X$ и $E \in \sigma$, то и $X \setminus E \in \sigma$
3. σ замкнуто относительно счетного пересечения (или счетного объединения), эти формулировки эквивалентны

Def. $\mathcal{F}$ — множество событий, при этом $\mathcal{F} \subset 2^{\Omega}$ и $\mathcal{F}$ — сигма-алгебра

Def. A — событие, если $A \in \mathcal{F}$

Def. Функция $\mu : \mathcal{F} \to [0, \infty]$ — счетно-аддитивная мера на $(\Omega\,, \mathcal{F})$, если:

1. $\mu(\varnothing) = 0$
2. Пусть $\forall i \in \mathbb{N}\ E_i \in \mathcal{F}$ и $\forall i, j,\ i \neq j\ E_i \cap E_j = \varnothing$, тогда $\mu\left(\bigsqcup\limits_{i=1}^{\infty} E_i\right) = \sum\limits_{i=1}^{\infty} \mu(E_i)$.

Def. P — счетно-аддитивная мера на $(\Omega\,, \mathcal{F})$ такая, что $P(\Omega) = 1$

Тогда тройка $(\Omega,\ \mathcal{F},\ \mathcal{P})$ — вероятностное пространство

3.1.1.2 Свойства вероятности

Во всех свойствах подразумевается, что $A, B \in \mathcal{F}$.

1. $P(\varnothing) = 0$ — из определения P
2. $P(A) = 1 - P(\Omega \setminus A)$ — из аддитивности P
3. Если $A \subset B$, то $P(A) \leq P(B)$ — из аддитивности P и неотрицательности
4. Формула включений и исключений, то есть $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ для двух событий, для большего числа по аналогии — из аддитивности P

3.1.1.3 Теорема о непрерывности вероятности в нуле

Th. Пусть даны $(\Omega,\ \mathcal{F})$ и P — конечно-аддитивная мера на $(\Omega,\ \mathcal{F})$ такая, что $P(\Omega) = 1$, тогда следующие утверждения эквивалентны:

1. P — счетно-аддитивная мера
2. P непрерывна сверху, то есть

$$\forall n \in \mathbb{N}\ A_n \in \mathcal{F}\ \left(A_n \uparrow A \iff A_1 \subset A_2 \subset \ldots,\ \bigcup_{n=1}^{\infty} A_n = A\right) \Longrightarrow P(A) = \lim_{n\to\infty} P(A_n)$$

3. P непрерывна снизу, то есть

$$\forall n \in \mathbb{N}\ A_n \in \mathcal{F}\ \left(A_n \downarrow A \iff A_1 \supset A_2 \supset \ldots,\ \bigcap_{n=1}^{\infty} A_n = A\right) \Longrightarrow P(A) = \lim_{n\to\infty} P(A_n)$$

4. P непрерывна в нуле, то есть

$$\forall n \in \mathbb{N}\ A_n \in \mathcal{F},\ A_n \downarrow \varnothing \Longrightarrow P(A) = \lim_{n\to\infty} P(A_n)$$

Proof:

(1) $\Longrightarrow$ (2):

Рассмотрим A_i, удовлетворяющие (2):

$$A = \bigcup_{n=1}^{\infty} A_i = A_1 \cup (A_2 \setminus A_1) \cup \ldots \cup (A_{k+1} \setminus A_k) \cup \ldots$$

Тогда для них верно, согласно (1), что:

$$P(A) = P(A_1) + \sum_{n=1}^{\infty} P(A_{n+1} \setminus A_n) = P(A_1) + \sum_{n=1}^{\infty} (P(A_{n+1}) - P(A_n)) = \lim_{n\to\infty} P(A_n)$$

(2) $\Longrightarrow$ (3):

$$P(A_n) = P(A_1 \setminus (A_1 \setminus A_n)) = P(A_1) - P(A_1 \setminus A_n)$$

Заметим, что последовательность $B_n = A_1 \setminus A_n$ не убывает и

$$\bigcup_{n=1}^{\infty} B_n = A_1 \setminus \bigcap_{n=1}^{\infty} A_n$$

Тогда, согласно (2), верно, что:

$$\lim_{n\to\infty} P(A_1 \setminus A_n) = \lim_{n\to\infty} P(B_n) = P\left(\bigcup_{n=1}^{\infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} (A_1 \setminus A_n)\right) = P\left(A_1 \setminus \bigcap_{n=1}^{\infty} (A_n)\right)$$

Откуда следует, что:

$$\lim_{n\to\infty} P(A_n) = P(A_1) - \lim_{n\to\infty} P(A_1 \setminus A_n) = P(A_1) - P\left(\bigcup_{n=1}^{\infty} (A_1 \setminus A_n)\right) = P(A_1) - P\left(A_1 \setminus \bigcap_{n=1}^{\infty} (A_n)\right) =$$

$$= P(A_1) - P(A_1) + P\left(\bigcap_{n=1}^{\infty} A_n\right) = P\left(\bigcap_{n=1}^{\infty} A_n\right)$$

(3) $\Longrightarrow$ (4):

Так как (4) является частным случаем (3), из (3) немедленно следует (4).

(4) $\Longrightarrow$ (1):

Пусть $\forall n \in \mathbb{N}\ A_n \in \mathcal{F}$ и A_i попарно не пересекаются. Тогда

$$P\left(\bigsqcup_{i=1}^{\infty} A_i\right) = P\left(\bigsqcup_{i=1}^{n} A_i\right) + P\left(\bigsqcup_{i=n+1}^{\infty} A_i\right)$$

Тогда заметим, что $\bigsqcup\limits_{i=n+1}^{\infty} A_i \downarrow \varnothing$. Теперь рассмотрим

$$\sum_{i=1}^{\infty} P(A_i) = \lim_{n\to\infty} \left(\sum_{i=1}^{n} P(A_i)\right) = \lim_{n\to\infty} \left(P\left(\bigsqcup_{i=1}^{n} A_i\right)\right) = \lim_{n\to\infty} \left(P\left(\bigsqcup_{i=1}^{\infty} A_i\right) - P\left(\bigsqcup_{i=n+1}^{\infty} A_i\right)\right) =$$

$$= P\left(\bigsqcup_{i=1}^{\infty} A_i\right) - \lim_{n\to\infty}\left(P\left(\bigsqcup_{i=n+1}^{\infty} A_i\right)\right) = P\left(\bigsqcup_{i=1}^{\infty} A_i\right)$$

Откуда следует счетная аддитивность P, а значит и (1).

■

3.1.2 Условные вероятности. Независимость.

3.1.2.1 Условные вероятности.

Пусть нам дано некоторое знание о множестве событий (или о множестве элементарных исходов), например, произошло событие B. Тогда у нас будет меняться вероятностная мера событий, если рассматривать их при условии, что произошло событие B.

Def. Вероятность события A при условии того, что произошло B определяется так:

$$P(A|B) = \begin{cases} \frac{P(A\cap B)}{P(B)}, & P(B) > 0, \\ 0 & P(B) = 0 \end{cases}$$

Note. Данное доопределение нулем корректно, так как если $P(B) = 0$, то и $P(A \cap B) = 0$, а тогда можно определять так, как нам удобно.

3.1.2.2 Формулы полной вероятности и Байеса

Note. Обе эти теоремы верны как для конечного, так и для бесконечного разбиения Ω

Th (Формула полной вероятности). Если $\Omega = B_1 \sqcup B_2 \sqcup \ldots$, то

$$P(A) = \sum_{i=1}^{\infty} (P(A|B_i) \cdot P(B_i))$$

Proof:

$$A = \bigsqcup_{i=1}^{\infty}(B_i \cap A) \Longrightarrow P(A) = \sum_{i=1}^{\infty} P(B_i \cap A) = \sum_{i=1}^{\infty} (P(A|B_i) \cdot P(B_i))$$

Th (Байес). Если $\Omega = B_1 \sqcup B_2 \sqcup \ldots$, то

$$P(B_n|A) = \frac{P(A|B_n) \cdot P(B_n)}{\sum\limits_{i=1}^{\infty} P(A|B_i) \cdot P(B_i)}$$

Proof:

$$P(B_n|A) = \frac{P(B_n \cap A)}{P(A)} = \frac{P(A|B_n) \cdot P(B_n)}{\sum\limits_{i=1}^{\infty} P(A|B_i) \cdot P(B_i)}$$

3.1.2.3 Независимость событий и систем событий

Def. События A и B независимы, то есть $A \perp\!\!\!\perp B$, если $P(A \cap B) = P(A) \cdot P(B)$

Def. Система событий $A_1, \ldots, A_n$ независима *(в совокупности)*, если $\forall n_1, \ldots, n_k$ $(2 \le k \le n), \forall i, j \in \overline{1,k} : n_i \ne n_j$ верно, что $P(A_{n_1} \cap \ldots \cap A_{n_k}) = \prod\limits_{i=1}^{k} P(A_{n_i})$

Note. Для бесконечного набора верно то, что любой конечный поднабор независим в совокупности

Def. Системы событий $\mathcal{A}_1, \ldots, \mathcal{A}_n \in \mathcal{F}$ независимы, если $\forall A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n$ события $A_1, \ldots, A_n$ независимы.

Def. Система событий $A_1, \ldots, A_n$ независима *попарно*, если $\forall i, j (i \ne j)$ верно, что $A_i \perp\!\!\!\perp A_j$

3.1.2.4 Пример Бернштейна

Рассмотрим тетраэдр с покрашенными в три цвета гранями (всего граней четыре) так, что по одной грани каждого цвета и четвертая грань покрашена сразу во все три цвета. Событиями назовем то, что выбранная грань покрашена в данный цвет, то есть три события, вероятность каждого равна 0.5, при этом они попарно независимы, так как грань с «миксом» из двух цветов только одна, а всего их четыре. При этом события не независимы в совокупности.

3.2 Случайные величины и векторы. Характеристики случайной величины и вектора: распределение вероятностей, функция распределения и её свойства, σ-алгебра, порожденная случайной величиной. Примеры конкретных распределений.

3.2.1 Случайные величины и векторы, их характеристики

3.2.1.1 Случайные величины

Note. В данном билете почти всюду все случайные величины в рамках одного утверждения определены на одном и том же вероятностном пространстве

Def. Измеримая $\xi : \Omega \to \mathbb{R}$ — случайная величина

Note. ξ измерима, если $\forall B \in \beta(\mathbb{R})\ \{\omega|\xi(\omega) \in B\} \in \mathcal{F}$

Def. Измеримая $\xi : \Omega \to \mathbb{R}^n$ — случайный вектор

Note. ξ измерима, если $\forall B \in \beta(\mathbb{R}^n)\ \{\omega|\xi(\omega) \in B\} \in \mathcal{F}$

3.2.1.2 Распределение вероятностей, функция распределения и её свойства

Def. Распределение — вероятность на $(\mathbb{R}, \beta(\mathbb{R}))$.

Def. Функцией распределения, соответствующей распределению P называют функцию $F : \mathbb{R} \to [0, 1]$ такая, что $F(x) = P((-\infty, x])$

Свойства функций распределения:

1. F не убывает
2. F непрерывна справа
3. $\lim\limits_{x\to-\infty} F(x) = 0,\ \lim\limits_{x\to\infty} F(x) = 1$

Proof:

1. Пусть $x < y$, тогда $F(y) - F(x) = P((-\infty, y]) - P((-\infty, x])) = P((x, y]) \geq 0$
2. Рассмотрим $x_n \downarrow x$. Докажем, что $\lim\limits_{n\to\infty} F(x_n) = F(x)$. Этого хватит для того, чтобы доказать второе свойство, так как F не убывает из свойства выше.

$$\lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} P((-\infty, x_n]) = P\left(\bigcap_{n=1}^{\infty}(-\infty, x_n]\right) = P((-\infty, x]) = F(x)$$

   Последний переход верен по теореме о непрерывности вероятностной меры (билет 1).

3. Абсолютно аналогично свойству выше.

Th (О построении вероятности по ф.р.) (б/д). Если F обладает свойствами выше, то найдется такое распределение, что она является его функцией.

Def. Если $F(x) = \int\limits_{-\infty}^{x} p(t)dt$ для некоторой функции $p : \mathbb{R} \to \mathbb{R}_+$, то F абсолютно непрерывна и распределение абсолютно непрерывно.

Def. p из определения выше — плотность распределения.

Note. Если F дифференцируема, то $p(t) = F'(t)$, то есть в качестве плотности можно взять производную функции распределения, то есть равенство не тождественное, ибо вообще там класс плотностей с отношением эквивалентности «почти всюду»

Statement. $\forall B \in \beta(\mathbb{R})\ P(B) = \int\limits_B p(t)d\mu(t)$

Proof: Из теоремы о единственности продолжения меры и определения $F(x)$ ■

3.2.1.3 Случайные векторы

Def. Вероятностная мера P на $(\mathbb{R}^n, \beta(\mathbb{R}^n))$ называется n-мерным распределением.

Def. Функцией распределения, соответствующей распределению P называют функцию $F : \mathbb{R}^n \to [0,1]$ такая, что $F(x) = P((-\infty, x_1] \times \ldots \times (-\infty, x_n])$

Def. $\Delta^i_{(a,b)}\ f(x_1, \ldots, x_n) = f(x_1, \ldots, x_{i-1}, b, x_{i+1}, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n)$

Свойства функций распределения:

1. F не убывает, то есть $\forall a_1 \le b_1, \ldots, a_n \le b_n$ верно, что $\Delta^1_{(a_1,b_1)} \ldots \Delta^n_{(a_n,b_n)} F \ge 0$
2. F непрерывна справа, то есть $\lim\limits_{x_n \downarrow x} F(x_n) = F(x)$, где $x_m \downarrow x \iff \forall i < n\ x^1_m \downarrow x^1, \ldots, x^n_m \downarrow x^n$
3. $\exists k \le n \lim\limits_{x_k \to -\infty} F(x) = 0$, $\forall k \le n \lim\limits_{x_1 \to \infty, \ldots, x_n \to \infty} F(x) = 1$

Proof:

1. Докажем по индукции, что

   $$\forall k \le n\ \Delta^1_{(a_1,b_1)} \ldots \Delta^k_{(a_k,b_k)} F(x_1, \ldots, x_n) = P((a_1, b_1] \times \ldots \times (a_k, b_k] \times (-\infty, x_{k+1}] \times \ldots \times (-\infty, x_n])$$

   База ($k = 1$) очевидна (из линейности Δ). Шаг: пусть $\forall i < k$ верно утверждение, докажем его для k:

   $$\Delta^1_{(a_1,b_1)} \ldots \Delta^k_{(a_k,b_k)} F(x_1, \ldots, x_n) =$$
   $$= \Delta^1_{(a_1,b_1)} \ldots \Delta^{k-1}_{(a_{k-1},b_{k-1})} F(x_1, \ldots, x_{k-1}, b_k, \ldots x_n) - \Delta^1_{(a_1,b_1)} \ldots \Delta^{k-1}_{(a_{k-1},b_{k-1})} F(x_1, \ldots, x_{k-1}, a_k, \ldots x_n) =$$
   $$= P((a_1, b_1] \times \ldots \times (a_{k-1}, b_{k-1}] \times (-\infty, b_k] \times \ldots \times (-\infty, x_n]) -$$
   $$- P((a_1, b_1] \times \ldots \times (a_{k-1}, b_{k-1}] \times (-\infty, a_k] \times \ldots \times (-\infty, x_n]) =$$
   $$= P((a_1, b_1] \times \ldots \times (a_{k-1}, b_{k-1}] \times (a_k, b_k] \times \ldots \times (-\infty, x_n])$$

   Тогда из неотрицательности меры получаем, что

   $$\Delta^1_{(a_1,b_1)} \ldots \Delta^n_{(a_n,b_n)} F(x_1, \ldots, x_n) = P((a_1, b_1] \times \ldots \times (a_n, b_n]) \ge 0$$

   Note. По сути применение n раз оператора высекает n-мерный прямоугольный гиперпараллелипипед.

2. Доказательство аналогично функции одномерного распределения, только теперь у нас сходятся вектора, а их сходимость по сути покоординатная.

3. Абсолютно аналогично свойству выше.

Th (О построении вероятности по ф.р. (многомерный случай)) (б/д). Если F обладает свойствами выше, то найдется такое многомерное распределение, что она является его функцией

Def. Пусть $X \subset \mathbb{R}^n$, X не более чем счетно, тогда если $P(X) = 1$, то P — дискретное распределение.

Def. Если $F(x_1, \dots, x_n) = \int\limits_{-\infty}^{x_1} \dots \int\limits_{-\infty}^{x_n} p(t_1, \dots, t_n) dt_1 \dots dt_n$ для некоторой функции $p : \mathbb{R}^n \to \mathbb{R}_+$, то F абсолютно непрерывна и распределение абсолютно непрерывно.

Def. p из определения выше — плотность распределения.

Statement. $\forall B \in \beta(\mathbb{R}^n)\ P(B) = \int\limits_B p(t_1, \dots, t_n) d\mu(t_1) \dots d\mu(t_n)$

Statement. Если F дифференцируема по всем аргументам, то $\frac{\partial F}{\partial x_1 \dots \partial x_n} = p(x_1, \dots, x_n)$

Statement. Если $p : \mathbb{R}^n \to \mathbb{R}_+$ и $\int\limits_{\mathbb{R}^n} p(t_1, \dots, t_n) d\mu(t_1) \dots d\mu(t_n) = 1$, то p — плотность некоторого абсолютно непрерывного распределения.

Proof: Найдем $F(x)$ и проверим ее на то, что она удовлетворяет трем свойствам функции распределения. Неубывание из того, что p действует в $\mathbb{R}_+$, непрерывность из свойств интеграла с переменным верхним пределом, стремление к 0 и 1 на бесконечностях из того, что интеграл по $\mathbb{R}^n$ конечен и равен единице.

3.2.1.4 σ-алгебра, порожденная случайной величиной

Def. $\mathcal{F}_\xi = \{\xi^{-1}(B) | B \in \beta(\mathbb{R})(\beta(\mathbb{R}^n))\}$ — σ-алгебра, порожденная случайной величиной (вектором) ξ

Statement. $\mathcal{F}_\xi$ действительно является σ-алгеброй.

Proof: Непосредственно проверяем все свойства σ-алгебры. ■

Statement. Пусть ξ, η — случайные вектора, тогда $\mathcal{F}_\xi \subset \mathcal{F}_\eta \iff \exists g : g(\eta) = \xi$ борелевская.

Proof: В левую сторону тривиально, в правую б/д (на лекции так было). ■

Note. ξ называется тогда η-измеримой.

Statement. Пусть ξ — n-мерный случайный вектор и $g : \mathbb{R}^n \to \mathbb{R}^k$ борелевская, то $g(\xi)$ — случайный k-мерный вектор.

Proof: Утверждение выше все доказывает. ■

Consequence. Если $\xi = (\xi_1, \dots, \xi_n)$ — случайный вектор, то ξ_1, …, ξ_n — случайные величины.

Th (критерий измеримости). Пусть $M \in \mathcal{E}$ и $\sigma(M) = \mathcal{E}$, $\xi : \Omega \to E$ тогда ξ — случайный элемент $\iff \forall B \in M\ \xi^{-1}(B) \in \mathcal{F}$.

Proof: В курсе теории меры. ■

Statement. Если ξ_1, …, ξ_n — случайные величины, то $\xi = (\xi_1, \dots, \xi_n)$ — случайный вектор.

Proof: Пусть $M = \{B_1 \times \dots \times B_n | \forall i \le n B_i \in \beta(\mathbb{R})\}$, тогда $\sigma(M) = \beta(\mathbb{R}^n)$.

Тогда $\xi^{-1}(B_1 \times \ldots \times B_n) = \xi_1^{-1}(B_1) \cap \ldots \cap \xi_n^{-1}(B_n) \in \mathcal{F}$. Тогда по критерию измеримости утверждение доказано. ■

Consequence. Пусть ξ, η — случайные величины, тогда $\xi + \eta, \xi - \eta, \xi \cdot \eta, \frac{\xi}{\eta}$ — случайные величины.

3.2.2 Примеры распределений

Examples.

1. Распределение Бернулли $\mathrm{Bern}(p)$

$$P(\xi = 1) = p, \quad P(\xi = 0) = 1 - p$$

2. Биномиальное с параметрами, то есть $\mathrm{Bin}(n, p)$

$$P(\xi = k) = C_n^k p^k (1-p)^{n-k}$$

Идея — вероятность получить ровно k исходов в серии из n независимых испытания с вероятностью успеха p.

3. Пуассоновское распределение $\mathrm{Poiss}(\lambda)$:

$$P(\xi = k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad \xi \in \mathbb{N} \cup \{0\}$$

4. Геометрическое распределение с параметром p:

$$P(\xi = k) = (1-p)^{k-1} p$$

То есть вероятность получения первого успеха в схеме Бернулли на k-м шаге.

5. $U[a, b]$ — равномерное распределение:

$$p(t) = \frac{I(t \in [a,b])}{b-a} \Longrightarrow F(x) = \begin{cases} 0, & x < a, \\ \frac{x-a}{b-a}, & x \in [a,b] \\ 1, & x > b \end{cases}$$

6. Нормальное распределение с параметрами или $\mathcal{N}(a, \sigma^2)$:

$$p(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-a)^2}{2\sigma^2}} \Longrightarrow F(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-a)^2}{2\sigma^2}} d\mu(t)$$

Note. Стандартное нормальное распределение — распределение $\mathcal{N}(0, 1)$

7. Экспоненциальное с параметром λ или $Exp(\lambda)$:

$$p(t) = \lambda e^{-\lambda t} I(t \geq 0) \Longrightarrow F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

8. Бета-распределение $\mathrm{Beta}(\alpha, \beta)$, $\alpha > 0$, $\beta > 0$:

$$p(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathsf{B}(\alpha,\beta)} I(0 < x < 1)$$

9. Гамма-распределение $\mathrm{Gamma}(\alpha, \beta)$, $\alpha > 0$, $\beta > 0$:

$$p(x) = \frac{\alpha^\beta x^{\beta-1} e^{-\alpha x}}{\Gamma(\beta)} I(x > 0)$$

3.3 Математическое ожидание случайной величины: определение для простых, неотрицательных и произвольных случайных величин. Основные свойства математического ожидания (доказательства только для простых величин). Дисперсия и ковариация, их свойства.

3.3.1 Матожидание

Def. Математическим ожиданием случайной величины ξ называют величину

$$\mathbb{E}\xi = \int\limits_\Omega \xi(\omega) dP$$

3.3.1.1 Абсолютно непрерывные величины

Если ξ абсолютно непрерывна, то ее математическое ожидание можно вычислить по формуле

$$\mathbb{E}\xi = \int\limits_{\mathbb{R}} x \cdot p_\xi(x) dx$$

Examples.

1. $\xi \sim Exp(\lambda)$, тогда $\mathbb{E}\xi = \int\limits_0^\infty \lambda x e^{-\lambda x} dx = -\int\limits_0^\infty x d\left(e^{-\lambda x}\right) = \int\limits_0^\infty e^{-\lambda x} dx = \lambda^{-1}$

2. $\xi \sim \mathbb{N}(a, \sigma^2)$, тогда

$$\mathbb{E}\xi = \int\limits_{\mathbb{R}} \frac{x}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} dx = \int\limits_{\mathbb{R}} \frac{x-a+a}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-a)^2}{2\sigma^2}} = \sqrt{\frac{\sigma^2}{2\pi}} \int\limits_{\mathbb{R}} e^{-\frac{(x-a)^2}{2\sigma^2}} d\left(\frac{(x-a)^2}{2\sigma^2}\right) + a \int\limits_{\mathbb{R}} p_\xi(x) dx = a$$

3.3.1.2 Дискретные величины

Если ξ дискретна, то ее математическое ожидание можно вычислить по формуле

$$\mathbb{E}\xi = \sum_x (x \cdot P(\xi = x))$$

Examples.

1. $\xi \sim Bern(p)$, тогда $\mathbb{E}\xi = p$
2. $\xi \sim Bin(n, p)$, тогда $\mathbb{E}\xi = \mathbb{E}\left(\sum\limits_{i=1}^{n} \xi_i\right) = \sum\limits_{i=1}^{n} \mathbb{E}\xi_i = np$
3. $\xi \sim U\{1, \dots, n\}$, тогда $\mathbb{E}\xi = \frac{1}{n}\sum\limits_{i=1}^{n} i = \frac{n+1}{2}$
4. $\xi \sim Pois(\lambda)$, тогда $\mathbb{E}\xi = \sum\limits_{k=0}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \lambda e^{-\lambda} \sum\limits_{k=0}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda$

3.3.1.3 Свойства матожидания

1. Пусть c — константа, тогда $\mathbb{E}c = c$.
2. Если $\xi \leq \eta$ почти наверное и если $\exists\ \mathbb{E}\xi, \mathbb{E}\eta$, то $\mathbb{E}\xi \leq \mathbb{E}\eta$ почти наверное
3. $\mathbb{E}(c\xi) = c \cdot \mathbb{E}\xi$, если $\exists |\mathbb{E}\xi| < \infty$
4. $\mathbb{E}(\xi + \eta) = \mathbb{E}\xi + \mathbb{E}\eta$, если $\exists \mathbb{E}\xi, \mathbb{E}\eta, |\mathbb{E}\xi|, |\mathbb{E}\eta| < \infty$
5. $|\mathbb{E}\xi| \leq \mathbb{E}|\xi|$, если $\exists\ \mathbb{E}\xi$
6. Если $\forall A \in \mathcal{F}\ \mathbb{E}\xi \cdot I_A \leq \mathbb{E}\eta \cdot I_A$, то $\xi \leq \eta$ почти наверное

   Proof: Пусть $A = \{\xi > \eta\} \in \mathcal{F} \Longrightarrow \mathbb{E}\xi I_A \leq \mathbb{E}\eta I_A \Longrightarrow \mathbb{E}(\xi - \eta) I_A \leq 0$. При этом $(\xi - \eta) I_A > 0 \Longrightarrow \mathbb{E}(\xi - \eta) I_A \geq 0 \Longrightarrow \mathbb{E}(\xi - \eta) I_A = 0 \Longrightarrow P((\xi - \eta) I_A) = 0 \Longrightarrow P(A) = 0$ ■
7. $\xi \perp\!\!\!\perp \eta \Longrightarrow \mathbb{E}\xi\eta = \mathbb{E}\xi\mathbb{E}\eta$, если $|\mathbb{E}\xi|, |\mathbb{E}\eta|, \mathbb{E}|\xi\eta| < \infty$ (по теореме Фубини)
8. $\xi_n \to \xi$ почти наверное и $\forall n\ |\xi_n| \leq \eta, \mathbb{E}\eta < \infty$, тогда $\mathbb{E}\xi_n \to \mathbb{E}\xi < \infty$ и $\mathbb{E}|\xi_n - \xi| \to 0$

3.3.2 Дисперсия и ковариация

Def. Дисперсия случайной величины: $\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2$

Statement. $\mathbb{D}\xi = \mathbb{E}\xi^2 - \mathbb{E}^2\xi$

Proof: $\mathbb{D}\xi = \mathbb{E}(\xi - \mathbb{E}\xi)^2 = \mathbb{E}(\xi^2 - 2\xi\mathbb{E}\xi + \mathbb{E}^2\xi) = \mathbb{E}\xi^2 - 2\mathbb{E}\xi\mathbb{E}\xi + \mathbb{E}^2\xi = \mathbb{E}\xi^2 - \mathbb{E}^2\xi$ ■

Def. Ковариация двух независимых случайных величин: $cov(\xi, \eta) = \mathbb{E}((\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta))$

Note. $\mathbb{D}\xi = cov(\xi, \xi)$

Statement. $cov(\xi, \eta) = \mathbb{E}\xi\eta - \mathbb{E}\xi\mathbb{E}\eta$

Proof:

$cov(\xi, \eta) = \mathbb{E}((\xi - \mathbb{E}\xi)(\eta - \mathbb{E}\eta)) = \mathbb{E}(\xi\eta - \eta\mathbb{E}\xi - \xi\mathbb{E}\eta + \mathbb{E}\xi\mathbb{E}\eta) = \mathbb{E}\xi\eta - 2\mathbb{E}\xi\mathbb{E}\eta + \mathbb{E}\xi\eta = \mathbb{E}\xi\eta - \mathbb{E}\xi\mathbb{E}\eta$

Statement. Ковариация билинейна.

Proof:

$cov(a_1\xi_1 + a_2\xi_2, \eta) = \mathbb{E}(a_1\xi_1 + a_2\xi_2)\eta - \mathbb{E}(a_1\xi_1 + a_2\xi_2)\mathbb{E}\eta = a_1\mathbb{E}\xi_1\eta + a_2\mathbb{E}\xi_2\eta - a_1\mathbb{E}\xi_1\mathbb{E}\eta - a_2\mathbb{E}\xi_2\mathbb{E}\eta = a_1\mathbb{E}\xi_1\eta - a_1\mathbb{E}\xi_1\mathbb{E}\eta + a_2\mathbb{E}\xi_2\eta - a_2\mathbb{E}\xi_2\mathbb{E}\eta = a_1 cov(\xi_1, \eta) + a_2 cov(\xi_2, \eta)$ ■

Examples.

1. $\xi \sim Bern(p) \Longrightarrow \xi = \xi^2 \Longrightarrow \mathbb{D}\xi = \mathbb{E}\xi - \mathbb{E}^2\xi = p(1-p)$
2. $\xi \sim Bin(n, p)$ Представим как сумму независимых бернуллиевских, тогда $\mathbb{D}\xi = np(1-p)$
3. $\xi \sim Poiss(\lambda)$, тогда

$$\mathbb{D}\xi = \mathbb{E}\xi^2 - \mathbb{E}^2\xi = \mathbb{E}\xi^2 - \lambda^2 = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k e^{-\lambda}}{k!} - \lambda^2 = \sum_{k=1}^{\infty} k \frac{\lambda^k e^{-\lambda}}{(k-1)!} - \lambda^2 =$$
$$= \sum_{k=1}^{\infty} (k-1)\frac{\lambda^k e^{-\lambda}}{(k-1)!} + \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-1)!} - \lambda^2 = \lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2} e^{-\lambda}}{(k-2)!} + \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1} e^{-\lambda}}{(k-1)!} - \lambda^2 =$$
$$= \lambda^2 \sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} + \lambda \sum_{k=0}^{\infty} (k-1)\frac{\lambda^k e^{-\lambda}}{k!} - \lambda^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$$

4. $\xi \sim U[a, b]$, тогда

$$\mathbb{D}\xi = \mathbb{E}\xi^2 - \mathbb{E}^2\xi = \mathbb{E}\xi^2 - \left(\frac{b-a}{2}\right)^2 = \frac{1}{b-a}\int_a^b x^2 dx - \left(\frac{b-a}{2}\right)^2 = \frac{b^3 - a^3}{3(b-a)} - \frac{b^2 - 2ab + a^2}{4} = \frac{(b-a)^2}{12}$$

5. $\xi \sim \mathbb{N}(a, \sigma^2)$, тогда с учетом того, что $F_{c\xi+d}(x) = F_\xi\left(\frac{x-d}{c}\right)$

$$p_{c\xi+d}(x) = \frac{1}{c} p_\xi\left(\frac{x-d}{c}\right) = \frac{1}{\sqrt{2\pi\sigma^2 c^2}} e^{-\frac{(x-d-ac)^2}{2\sigma^2 c^2}} \Longrightarrow \xi \sim \mathbb{N}(ac + d, c^2\sigma^2)$$

Пусть $\eta \sim \mathbb{N}(0, 1), \mathbb{D}\xi = \sigma^2\mathbb{D}\eta$

$$\mathbb{D}\eta = \mathbb{E}\eta^2 = \int_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = -\int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} d\left(e^{-\frac{x^2}{2}}\right) = \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 1 \Longrightarrow D\xi = \sigma^2$$

Th (б/д). Пусть $\mathbb{E}\xi^2, \mathbb{E}\eta^2 < \infty$, тогда $(\mathbb{E}|\xi\eta|)^2 \leq \mathbb{E}\xi^2\mathbb{E}\eta^2$

Statement. $\xi \perp\!\!\!\perp \eta \Longrightarrow cov(\xi, \eta) = 0$

Note. Обратное утверждение неверно.

Statement. $\mathbb{D}(\xi_1 + \ldots + \xi_n) = \sum\limits_{i=1}^{n} \mathbb{D}\xi_i + 2\sum\limits_{i \neq j} cov(\xi_i, \xi_j)$

Consequence. Если $\xi_1, \ldots, \xi_n$ независимы (достаточно попарной независимости), то $\mathbb{D}\left(\sum\limits_{i=1}^{n} \xi_i\right) = \sum\limits_{i=1}^{n} \mathbb{D}\xi_i$

3.4 Сходимость случайных величин: по вероятности, по распределению, почти наверное, в среднем. Связь между сходимостями (б/д). Лемма Слуцкого (б/д). Теорема о наследовании сходимости. Дельта-метод.

3.4.1 Сходимость случайных величин. Связь между сходимостями.

Пусть $\xi_1, \xi_2, \ldots$ — случайные величины. Тогда виды сходимостей к ξ:

1. Сходимость почти наверное или с вероятностью 1, т.е. $P(\xi_n \to \xi) = 1$
2. Сходимость по вероятности $\forall \varepsilon > 0 \lim\limits_{n \to \infty} P(|\xi_n - \xi| \geq \varepsilon) = 0$
3. Сходимость в L_p, если $\mathbb{E}|\xi|^p, \mathbb{E}|\xi_i|^p < \infty$, то $\lim\limits_{n \to \infty} \mathbb{E}|\xi_n - \xi|^p = 0$
4. Сходимость по распределению (слабая)

$$\forall f : \mathbb{R} \to \mathbb{R},\ f \in C(\mathbb{R})\ \exists C > 0 : \forall x \in \mathbb{R}\ |f(x)| < C \hookrightarrow \mathbb{E}f(\xi_n) \to \mathbb{E}f(\xi)$$

Th (о взаимосвязи видов сходимости) (б/д).

1. Из сходимости почти наверное следует сходимость по вероятности
2. Из сходимости по вероятности следует сходимость по распределению
3. Из сходимости в L_p следует сходимость по вероятности
4. Никаких других следствий нет.

3.4.2 Лемма Слуцкого (б/д). Теорема о наследовании сходимости. Дельта-метод.

3.4.2.1 Лемма Слуцкого

Th (Лемма Слуцкого) (б/д). Пусть ξ_n, η_n — последовательности с.в. такие, что $\xi_n \xrightarrow{d} \xi$ и $\eta_n \xrightarrow{d} C$, тогда

1. $\xi_n + \eta_n \xrightarrow{d} \xi + C$.

2. $\xi_n \cdot \eta_n \xrightarrow{d} C \cdot \xi$.

3.4.2.2 Наследование сходимости

Th.

1. Пусть $X_k \xrightarrow{\text{п.н.}} X$ и борелевская $f : \mathbb{R}^n \to \mathbb{R}^m$ такова, что $P_X(f \text{ непрерывна}) = 1$, тогда $f(X_k) \xrightarrow{\text{п.н.}} f(X)$

2. Пусть $X_k \xrightarrow{\text{P}} X$ по вероятности и борелевская $f : \mathbb{R}^n \to \mathbb{R}^m$ такова, что $P_X(f \text{ непрерывна}) = 1$, тогда $f(X_k) \xrightarrow{\text{P}} f(X)$

3. Пусть $X_k \xrightarrow{d} X$ и $f : \mathbb{R}^n \to \mathbb{R}^m$ непрерывна, тогда $f(X_k) \xrightarrow{d} f(X)$

Proof:

1. Пусть $B \in \beta(\mathbb{R}^n)$ такое, что $B = \{x : f \text{ непрерывна в } x\}$, событие $A = \{X_k \to X\}$, $\widetilde{A} = \{f(X_k) \to f(X)\}$.

$$P_X(B) = P(\{\omega : \ X(\omega) \in B\}) = P(\{\omega : f \text{ непрерывна в } X(\omega)\})$$

   Заметим, что $\widetilde{A} \supset A \bigcap \{\omega : f \text{ непрерывна в } X(\omega)\}$, при этом $P(\{\omega : f \text{ непрерывна в } X(\omega)\}) = 1$ и $P(A) = 1$, а значит и $P(\widetilde{A}) = 1$

2. Предположим противное, то есть $f(X_k) \not\to f(X)$ по вероятности, а значит $\exists k_m \ \exists \varepsilon > 0 \ \exists \delta > 0$ $\forall m \ P(||f(X_{k_m} - f(X)||_2 > \varepsilon) > \delta$. При этом так как $X_{k_m} \to X$ по вероятности, а из нее можно выбрать подпоследовательность $X_{k_{m_j}}$, сходящуюся почти наверное. То есть $f(X_{k_{m_j}}) \to f(X)$ почти наверное, но для сходимости почти наверное сходимость наследуется, то есть $f(X_{k_{m_j}}) \to f(X)$ почти наверное, а значит и по вероятности, противоречие.

3. $\forall g : \mathbb{R}^m \to \mathbb{R}$ непрерывной и ограниченной. Заметим, что $\mathbb{E}(g(f(K_k))) = \mathbb{E}(g \circ f(X_k))$, при этом композиция $g \circ f$ непрерывна и ограничена. Откуда следует требуемое по определению, так как $X_k \to X$ сходится по распределению.

■

3.4.2.3 Дельта-метод

Th (Дельта-метод). Пусть ξ_n — последовательность с.в. и $\xi_n \xrightarrow{d} \xi$, $h(x)$ дифф. в точке a, $b_n \to 0$ и $b_n \neq 0\ \forall n$, тогда

$$\frac{h(a+\xi_n b_n) - h(a)}{b_n} \xrightarrow{d} \xi \cdot h'(a)$$

Proof: Рассмотрим

$$H(x) = \begin{cases} \frac{h(x+a)-h(a)}{x}, & x \neq 0 \\ h'(a), & x = 0 \end{cases}$$

Так как h — дифф. в a, H — непрерывна в 0. По лемме Слуцкого $\xi_n b_n \xrightarrow{d} \xi \cdot 0 = 0$, а, значит, к H можно применить теорему о наследовании сходимости (H непрерывна на множестве значений $\xi \cdot 0$, что есть на $\{0\}$). Тогда

$$\frac{h(a+\xi_n b_n) - h(a)}{\xi_n b_n} = H(\xi_n b_n) \xrightarrow{d} H(0) = h'(a)$$

Домножим на ξ_n и применим лемму Слуцкого:

$$\frac{h(a+\xi_n b_n) - h(a)}{b_n} = H(\xi_n b_n) \cdot \xi_n \xrightarrow{d} H(0) \cdot \xi = h'(a) \cdot \xi$$

■

Th (Многомерный дельта-метод) (б/д). Пусть ξ_n — последовательность с.в. из $\mathbb{R}^k$ и $\xi_n \xrightarrow{d} \xi$, $H : \mathbb{R}^k \to \mathbb{R}^l$ имеет в точке a матрицу частных производных (матрицу Якоби), $b_n \to 0$ и $b_n \neq 0\ \forall n$, тогда

$$\frac{H(a+\xi_n b_n) - H(a)}{b_n} \xrightarrow{d} \langle \nabla H|_a, \xi \rangle$$

3.5 Неравенство Маркова, неравенство Чебышёва. Закон больших чисел в форме Чебышёва. Усиленные законы больших чисел (б/д).

3.5.1 Неравенства Маркова и Чебышёва

3.5.1.1 Неравенство Маркова

Th (Марков). Если $\xi \geq 0$, и существует $\mathbb{E}\xi$, то $P(\xi \geq a) \leq \frac{\mathbb{E}\xi}{a}$

Proof:

$$\xi \geq a \cdot I(\xi \geq a) \Longrightarrow \mathbb{E}\xi \geq a \cdot P(\xi \geq a)$$

■

3.5.1.2 Неравенство Чебышёва

Th (Чебышёв). Если существует $\mathbb{D}\xi$, то $P\{|\xi - \mathbb{E}\xi| \geq \varepsilon\} \leq \frac{\mathbb{D}\xi}{\varepsilon^2}$.

Proof: Применяем неравенство Маркова к $\eta = (\xi - \mathbb{E}\xi)^2$.

■

3.5.2 ЗБЧ в форме Чебышёва. УЗБЧ

Th (ЗБЧ в форме Чебышёва). Пусть $\xi_1, \xi_2, \ldots$ — независимые случайные величины, при этом $\forall i\ \exists C > 0:\ \mathbb{E}\xi_i^2 < \infty, \mathbb{D}\xi_i \leq C$, тогда верно, что

$$\forall\ \delta > 0 \forall \varepsilon > 0 \quad \lim_{n\to\infty} P\left(\left|\frac{\sum\limits_{i=1}^{n}\xi_i - \mathbb{E}\left(\sum\limits_{i=1}^{n}\xi_i\right)}{n^{\frac{1}{2}+\delta}}\right| > \varepsilon\right) = 0$$

Proof: Применим неравенство Чебышёва:

$$P\left(\left|\sum_{i=1}^{n}\xi_i - \mathbb{E}\left(\sum_{i=1}^{n}\xi_i\right)\right| > \varepsilon \cdot n^{\frac{1}{2}+\delta}\right) \leq \frac{\mathbb{D}\left(\xi_1 + \ldots + \xi_n\right)}{\varepsilon^2 n^{1+2\delta}} = \frac{\mathbb{D}\xi_1 + \ldots + \mathbb{D}\xi_n}{\varepsilon^2 n^{1+2\delta}} \leq \frac{nC}{\varepsilon^2 n^{1+2\delta}} = \frac{C}{\varepsilon^2 n^{2\delta}} \to 0$$

■

Consequence. Если в знаменателе поставить не $n^{\frac{1}{2}+\delta}$, а $\sqrt{n}\cdot f(n)$, где $\lim\limits_{n\to\infty} f(n) = \infty$, то теорема остается верной.

Note. н.о.р.с.в. — независимые одинаково распределенные случайные величины.

Th (УЗБЧ) (б/д). Пусть $\xi_1, \xi_2, \ldots$ — последовательность н.о.р.с.в. с $\mathbb{E}|\xi_1| < \infty$, тогда

$$\frac{S_n - \mathbb{E}S_n}{n} \xrightarrow{\text{п.н.}} 0,\ n \to \infty$$

3.6 Характеристические функции случайных величин и векторов и их свойства. Теорема непрерывности (б/д).

3.6.1 Характеристические функции

Def. Пусть ξ, η — случайные величины, тогда $\mathbb{E}(\xi + i\eta) = \mathbb{E}\xi + i\mathbb{E}\eta$

Def. Пусть ξ — случайная величина, тогда $\varphi_\xi(x) = \mathbb{E}e^{ix\xi}$, $\varphi_\xi(x) : \mathbb{R} \to \mathbb{C}$ — характеристическая функция ξ

Def. Пусть ξ — случайный вектор, тогда $\varphi_\xi(x) = \mathbb{E}e^{i(\xi,x)}$, $\varphi_\xi(x) : \mathbb{R}^n \to \mathbb{C}$ — характеристическая функция ξ, где (x, y) — скалярное произведение

Examples.

1. $\xi \sim Bern(p)$, тогда $\varphi_\xi(x) = pe^{ix} + (1-p)$
2. $\xi \sim U[a,b]$, тогда $\varphi_\xi(x) = \mathbb{E}e^{ix\xi} = \frac{1}{b-a}\int\limits_a^b e^{ixt}dt = \frac{e^{itb}-e^{ita}}{it(b-a)}$
3. $\xi \sim \mathcal{N}(0,1)$, тогда $\varphi_\xi(x) = \int\limits_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{itx-\frac{t^2}{2}}dt = \int\limits_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(t-ix)^2}{2}-\frac{x^2}{2}}dt = e^{-\frac{x^2}{2}}$
4. $\xi \sim \mathcal{N}(a,\sigma^2) \Longrightarrow \frac{\xi-a}{\sigma} \sim \mathcal{N}(0,1)$

$$\varphi_{\frac{\xi-a}{\sigma}}(x) = \mathbb{E}e^{i\frac{\xi-a}{\sigma}x} = e^{-\frac{x^2}{2}} \Longrightarrow \varphi_{\frac{\xi}{\sigma}}(x) = e^{ia\frac{x}{\sigma}-\frac{x^2}{2}} \Longrightarrow \varphi_\xi(x) = \varphi_{\frac{\xi}{\sigma}}(\sigma x) = e^{iax-\frac{\sigma^2x^2}{2}}$$

3.6.1.1 Теорема о единственности (б/д)

Th (б/д). Если $\varphi_\xi(x), \varphi_\eta(x)$ — характеристические функции случайных векторов одинаковой размерности и $\forall x \in \mathbb{R}^n\ \varphi_\xi(x) = \varphi_\eta(x)$, то $\xi = \eta$ по распределению.

3.6.1.2 Критерий независимости (б/д)

Th (б/д). $\xi_1, \ldots, \xi_n$ независимы $\Longleftrightarrow \varphi_{(\xi_1,\ldots,\xi_n)}(x_1,\ldots,x_n) = \prod\limits_{j=1}^n \varphi_{\xi_j}(x_j)$

3.6.1.3 Формула обращения (б/д)

Th (б/д). Пусть $\varphi_\xi(x)$ — характеристическая функция ξ.

- Для любых двух точек непрерывности $a < b$ функции распределения F_ξ верно, что

$$F_\xi(b) - F_\xi(a) = \frac{1}{2\pi}\lim_{c\to\infty}\int\limits_{-c}^{c}\frac{e^{-ita}-e^{-itb}}{it}\varphi_\xi(t)\,dt$$

- Если $\int\limits_{\mathbb{R}}|\varphi_\xi(t)|dt < \infty$, то $p_\xi(x) = \frac{1}{2\pi}\int\limits_{\mathbb{R}} e^{-itx}\varphi_\xi(t)\,dt$

Note. Обратим внимание, что формулы выше по сути являются преобразованием Фурье (либо прямым, либо обратным).

3.6.1.4 Теорема о непрерывности (б/д)

Th (о непрерывности) (б/д).

- Если $\xi_n \to \xi$ по распределению, то $\forall t \in \mathbb{R} \lim\limits_{n\to\infty} \varphi_{\xi_n}(t) \to \varphi_\xi(t)$
- Пусть $\xi_1, \xi_2, \ldots$ — случайные величины и $\forall t \in \mathbb{R} \lim\limits_{n\to\infty} \varphi_{\xi_n}(t) \to \varphi(t)$, причем $\varphi(t)$ непрерывна в нуле. Тогда $\varphi(t)$ — характеристическая функция некоторой случайной величины ξ и $\xi_n \to \xi$ по распределению.

3.6.2 Свойства характеристических функций

1. $|\varphi_\xi(t)| \leq \varphi_\xi(0) = 1$

   Proof: $|\varphi_\xi(t)| = |\mathbb{E}e^{it\xi}| \leq \mathbb{E}|e^{it\xi}| = \mathbb{E}(1) = 1 = \varphi_\xi(0)$

2. $\varphi_\xi(t)$ равномерно непрерывна на $\mathbb{R}$

   Proof: Пусть $\varepsilon > 0$. Так как $\sin(t\xi) \to 0$ почти наверное при $t \to 0$, по теореме Лебега о мажорируемой сходимости $\mathbb{E}|\sin(t\xi)| \to 0$ при $t \to 0$

   $$\exists \delta > 0 : \forall t \in (-\delta, \delta)\ \mathbb{E}|\sin(t\xi)| < \frac{\varepsilon}{4}$$

   Пусть $|s - t| < \delta$, тогда

   $$|\varphi_\xi(t) - \varphi_\xi(s)| = |\mathbb{E}e^{is\xi}\left(e^{i(t-s)\xi} - 1\right)| \leq \mathbb{E}|e^{i(t-s)\xi} - 1| = \mathbb{E}\sqrt{(\cos((t-s)\xi) - 1)^2 + \sin^2((t-s)\xi)} =$$
   $$= \mathbb{E}\sqrt{2 - 2cos((t-s)\xi)} = 2\mathbb{E}\left|\sin\left(\frac{t-s}{2}\xi\right)\right| < \varepsilon$$

3. $\forall t \in \mathbb{R}\ \varphi_\xi(t) \in \mathbb{R} \iff \xi = -\xi$ по распределению.

   Proof:

   $\Longrightarrow$

   $\forall t \in \mathbb{R}\ \varphi_\xi(t) \in \mathbb{R}$, тогда

   $$\varphi_\xi(t) = \mathbb{E}e^{it\xi} = \mathbb{E}(\cos(t\xi) + i\sin(t\xi)) = \mathbb{E}(\cos(t\xi)) = \mathbb{E}(\cos(t\xi) - i\sin(t\xi)) = \varphi_\xi(-t)$$

   $\Longleftarrow$

   $\xi = -\xi$ по распределению, тогда $\mathbb{E}\sin(t\xi) = \mathbb{E}\sin(-t\xi) = -\mathbb{E}\sin(t\xi) = 0$

4. $\forall t \in \mathbb{R}\ \varphi_\xi(-t) = \overline{\varphi_\xi(t)}$

   Proof: $\varphi_\xi(-t) = \mathbb{E}(\cos(-t\xi) + i\sin(-t\xi)) = \mathbb{E}(\cos(t\xi) - i\sin(t\xi)) = \mathbb{E}\overline{e^{it\xi}} = \overline{\varphi_\xi(t)}$

5. Если $\mathbb{E}|\xi|^n < \infty$, тогда $\forall r \le n\ \varphi_\xi^{(r)}(t) = \int\limits_{\mathbb{R}} (ix)^r e^{itx} dF_\xi(x)$, $\mathbb{E}\xi^r = \frac{\varphi_\xi^{(r)}(0)}{i^r}$

$$\varphi_\xi(t) = \sum_{r=0}^{n} \frac{(it)^r}{r!}\mathbb{E}\xi^r + \frac{(it)^n}{n!}\varepsilon_n(t)$$

Где $|\varepsilon_n(t)| \le e\mathbb{E}|\xi|^n$ и $\lim\limits_{t\to 0}\varepsilon_n(t) = 0$.

6. Если $|\varphi^{(2n)}(0)| < \infty$, то $\mathbb{E}\xi^{2n} < \infty$

Example: $\varphi(t) = e^{-t^4}$. Может ли она быть характеристической?

Solution. $\varphi'(0) = 0$, $\varphi''(0) = 0$, тогда по свойству 6 существуют $\mathbb{E}\xi^2$ и $\mathbb{E}\xi$. Тогда по свойству 5: $\mathbb{E}\xi^2 = \mathbb{E}\xi = 0 \Longrightarrow \xi = 0$ почти наверное, но при этом $\varphi_\xi(0) = 1$. Значит не может быть.

3.7 Центральная предельная теорема для независимых одинаково распределенных случайных величин.

3.7.1 ЦПТ

Th. Пусть $\xi_1, \xi_2, \ldots$ — н.о.р.с.в. такие, что $\mathbb{E}\xi^2 < \infty$, $S_n = \xi_1 + \ldots + \xi_n$. Тогда имеется сходимость по распределению

$$\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} \xrightarrow{d} \eta \sim \mathcal{N}(0,1)$$

Proof: Пусть $\mathbb{E}\xi = a$, $\mathbb{D}\xi = \sigma^2$. Пусть $\widetilde{\xi}_j = \frac{\xi_j - a}{\sqrt{\sigma^2}}$.

Тогда заметим, что $\mathbb{E}\widetilde{\xi} = 0$, $\mathbb{D}\widetilde{\xi} = 1$, $\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}} = \frac{\widetilde{S}_n}{\sqrt{n}}$

$$\varphi_{\frac{S_n - \mathbb{E}S_n}{\sqrt{\mathbb{D}S_n}}}(x) = \mathbb{E}e^{i\frac{\widetilde{S}_n}{\sqrt{n}}x} = \prod_{j=1}^{n} \varphi_{\widetilde{\xi}_j}\left(\frac{x}{\sqrt{n}}\right) = \left(\varphi_{\widetilde{\xi}_1}\left(\frac{x}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{1}{2}\left(\frac{x}{\sqrt{n}}\right)^2 + o\left(\frac{1}{n}\right)\right)^n =$$

$$= e^{n\left(-\left(\frac{x}{\sqrt{2n}}\right)^2\right)+o(1)} = e^{-\frac{x^2}{2}+o(1)} \to e^{-\frac{-x^2}{2}}$$

Тогда по теореме о непрерывности и о единственности получаем требуемое, так как предел характеристической функций совпадет с характеристической функцией стандартного нормального распределения. ■

3.8 Выборка, выборочное пространство. Точечные оценки параметров и их основные свойства: несмещенность, состоятельность, асимптотическая нормальность. Выборочные среднее, медиана, дисперсия. Сравнение оценок, функция потерь и функция риска. Подходы к сравнению оценок: равномерный, байесовский, асимптотический.

3.8.1 Выборочное пространство (one more iconic trio)

Def. Все возможные значения одного эксперимента образуют выборочное пространство $\mathcal{X}$.

Def. Вероятностно-статистическая модель — кортеж $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mathcal{P})$, где $\mathcal{X}$ — выборочное пространство (обычно $\mathbb{R}$ или $\mathbb{R}^n$), $\mathcal{B}_\mathcal{X}$ — борелевская сигма-алгебра на $\mathcal{X}$ и $\mathcal{P}$ — семейство вероятностных мер на $(\mathcal{X}, \mathcal{B}_\mathcal{X})$.

Def. Пусть $\forall x \in \mathcal{X}\ X(x) = x$, тогда, с одной стороны, X — результат эксперимента, а с другой: $\forall P \in \mathcal{P}\ X$ — случайная величина (вектор) на $(\mathcal{X}, \mathcal{B}_\mathcal{X}, P)$ со значениями в $\mathcal{X}$ и распределением $P_X = P$. Тогда такое отображение X называют наблюдением.

Построим вероятностное пространства, в котором будет n независимых результатов случайного эксперимента. Для этого возьмём прямое произведение $(\mathcal{X}, \mathcal{B}_\mathcal{X}, P)$ вероятностных пространств, получим $(\mathcal{X}^n, \mathcal{B}_\mathcal{X}^n, P^n)$:

1. $\mathcal{X}^n = \mathcal{X} \times \ldots \times \mathcal{X}$, $\overline{x} = (x_1, \ldots, x_n) \in \mathcal{X}^n$
2. $\mathcal{B}_\mathcal{X}^n = \sigma\left(B_1 \times \ldots \times B_n\right)$, где $B_i \in \mathcal{X}$
3. $P^n = P \otimes \ldots \otimes P$ — продолжение прямого произведения мер с алгебры прямоугольников, то есть $P^n(B_1 \times \ldots \times B_n) = \prod\limits_{i=1}^{n} P(B_i)$. Такое продолжение существует и единственно по теореме Каратеодори.

Из такой конструкции $X_i(x_1, \ldots, x_n) = x_i$ — случайная величина с распределением P, а $x_1, \ldots, x_n$ — независимые в совокупности о.р.с.в. (по построению).

Def. Н.о.р.с.в. $X_1, \ldots, X_n$ — выборка размера n из неизвестного распределения P_X.

Def. Реализация выборки $(X_1(\omega), \ldots, X_n(\omega))$ — числовой кортеж, полученный подстановкой события в выборку.

Note. Можно построить бесконечную выборку. В ней $\mathcal{X}^\infty$ будет пространством последовательностей, $\mathcal{B}_\mathcal{X}^\infty$ как минимальная сигма-алгебра на всех конечных декартовых произведениях, а мера P^∞ будет определена для любых конечномерных гиперпараллелипипедов.

Def. Если $\mathcal{P}$ имеет естественную параметризацию, то есть $\mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$, $\Theta \subset \mathbb{R}$ или $\mathbb{R}^n$, то необходимо будет лишь оценить θ. Тогда говорят, что статистическая модель параметрическая.

3.8.2 Точечные оценки

Def. Пусть $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mathcal{P})$ — вероятностно статистическая модель, X — наблюдение. Пусть $(E, \mathcal{E})$ — измеримое пространство, а $S : \mathcal{X} \to E$ — измеримое отображение, тогда $S(X)$ — статистика от наблюдения X.

Def. Пусть $(\mathcal{X}, \mathcal{B}_\mathcal{X}, \mathcal{P} = \{P_\theta \mid \theta \in \Theta\})$ — параметрическая модель, а X — наблюдение на ней и $S(X)$ — статистика. Тогда если $S : \mathcal{X} \to \Theta$, то S — оценка параметра θ.

Note. Можно оценивать не только параметр θ, но и функции $\tau(\theta)$. В этом случае $S : \mathcal{X} \to \tau(\Theta)$.

Def. Выборочное среднее $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$.

Note. В термине выборочного среднего предельные теоремы:

УЗБЧ: $\overline{X} - \mathbb{E}\overline{X} \xrightarrow{\text{п.н.}} 0$.

Многомерная ЦПТ: $\sqrt{n}\left(\overline{X} - a\right) \xrightarrow{d} \eta \sim \mathcal{N}(0, \mathrm{Var}X_1)$.

3.8.2.1 Несмещённость оценок

Пусть $X = (X_1, \ldots, X_n)$ — наблюдение из распределения $P \in \mathcal{P} = \{P_\theta \mid \theta \in \Theta\}$.

Def. Оценка $\theta^*(X)$ параметра θ — несмещённая, если $\forall \theta \in \Theta\ \mathbb{E}_\theta \theta^*(X) = \theta$.

Example: Пусть $X_1, \ldots, X_n$ — выборка из $\mathcal{U}[0, \theta]$. Тогда оценка $2\overline{X}$ — несмещённая.

Solution.

$$\mathbb{E}_\theta 2\overline{X} = \frac{2}{n}\sum_{i=1}^{n} \mathbb{E}_\theta X_i = 2\mathbb{E}_\theta X_1 = 2 \cdot \frac{\theta}{2} = \theta$$

Note. В общем случае нужно искать распределение $\theta^*(X)$, а потом считать матожидание.

3.8.2.2 Состоятельность и сильная состоятельность оценок

Def. Посл. оценок $\{\theta_n^*(X)\}_{n\in\mathbb{N}}$ параметра θ — состоятельная, если $\forall \theta \in \Theta\ \theta_n^* \xrightarrow{P_\theta} \theta$.

Def. Посл. оценок $\{\theta_n^*(X)\}_{n\in\mathbb{N}}$ параметра θ — сильно состоятельная, если $\forall \theta \in \Theta\ \theta_n^* \xrightarrow{P_\theta\text{-п.н.}} \theta$.

Note. Легко заметить, что из сильной состоятельности следует обычная состоятельность.

Example: Пусть $X_1, \ldots, X_n$ — выборка из $\mathcal{U}[0, \theta]$. Покажите, что оценка $2\overline{X}$ — сильно состоятельная.

Solution. По УЗБЧ: $\overline{X} \xrightarrow{P_\theta\text{-п.н.}} \mathbb{E}_\theta X_1 = \frac{\theta}{2}$, откуда легко получить необходимое.

3.8.2.3 Асимптотическая нормальность оценок

Def. Посл. оценок $\{\theta_n^*(X)\}_{n\in\mathbb{N}}$ параметра θ — асимптотически нормальная, если $\forall\theta \in \Theta \Rightarrow \sqrt{n}\,(\theta_n^* - \theta) \xrightarrow{d_\theta} \mathcal{N}(0, \sigma^2(\theta))$, где $\sigma^2(\theta)$ называется асимптотической дисперсией.

Note. Из асимптотической нормальности следует состоятельность. (Доказывается с помощью леммы Слуцкого и утверждения о том, что $\xi_n \xrightarrow{d} \mathrm{C} \Longrightarrow \xi_n \xrightarrow{\mathrm{P}} C$).

Example: $2\overline{X}$ является асимптотически нормальной оценкой θ из распределения $\mathcal{U}[0,\theta]$.

Solution. Можно применить многомерную ЦПТ: $\sqrt{n}\left(2\overline{X} - \theta\right) \xrightarrow{d} \eta \sim \mathcal{N}(0, \frac{\theta^2}{3})$, тогда $\sigma^2(\theta) = \frac{\theta^2}{3}$.

3.8.2.4 Выборочные среднее, медиана, дисперсия

Def. Выборочным средним называют статистику $\overline{X} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$.

Def. Пусть $F(x)$ — ф.р. на $\mathbb{R}$. Пусть $p \in (0,1)$. Тогда p-квантиль $z_p = \inf\{x \mid F(x) \geqslant p\}$.

Def. Пусть $X_1, \ldots, X_n$ — выборка из распределения P. Тогда выборочной p-квантилью называют статистику

$$\hat{z}_{n,p} = \begin{cases} X_{([np]+1)}, & np \notin \mathbb{Z} \\ X_{(np)}, & np \in \mathbb{Z} \end{cases}$$

Def. Выборочной медианой называют $\frac{1}{2}$-квантиль.

Def. Выборочной дисперсией называют статистику $S^2 = \overline{X^2} - \overline{X}^2$.

3.8.3 Сравнение оценок

Def. Борелевская функция $g(x,y) \geqslant 0$ такая, что $g(x,x) = 0$, называется функцией потерь.

Def. Величиной потери оценки θ^* параметра θ называют $g(\theta^*, \theta)$.

Def. Пусть g — функция потерь, тогда функцией риска θ^* называется $R(\theta^*, \theta) = \mathbb{E}_\theta g(\theta^*, \theta)$.

3.8.3.1 Равномерный подход

Def. Оценка θ^* лучше оценки $\hat{\theta}$ в равномерном подходе, если $\forall\theta \in \Theta\ R(\theta^*, \theta) \leqslant R(\hat{\theta}, \theta)$ и на каком-то θ неравенство строгое.

Def. Оценка θ^* — наилучшая в некотором классе оценок $\mathcal{K}$ в равномерном подходе, если она лучше всех оценок данного класса.

Note. Наилучшая оценка не всегда существует. Например, $g(x,y) = (x-y)^2$, $\mathcal{K}$ — всевозможные оценки. Тогда тривиальная оценка $\hat{\theta} = \theta_0 \in \Theta$ удовлетворяет условию $R(\hat{\theta}, \theta_0) = 0$.

Значит, если существует наилучшая оценка θ^*, то

$$\forall \theta_0 \in \Theta \ 0 \leqslant R(\theta^*, \theta_0) \leqslant R(\hat{\theta}, \theta_0) = 0 \Longrightarrow\Longrightarrow \theta^* \equiv \theta_0 \ P_\theta\text{-п.н.}$$

, что невозможно, так как тогда θ^* равна сразу всем тривиальным оценкам.

Def. Если $g(x, y) = (x - y)^2$ и $\mathcal{K}$ — класс несмещённых оценок, то равномерный подход называют среднеквадратичным.

Def. Оценка θ^* — допустимая оценка параметра θ, если нет оценки $\hat{\theta}$ — лучшей в равномерном подходе. (Оценки могли быть получены другими подходами)

3.8.3.2 Байесовский подход

В этом подходе предполагается, что на Θ задано некое распределение вероятности Q,
а θ случайно выбран из Θ по закону Q.

Def. Пусть θ^* — оценка параметра θ, а $R(\theta^*, \theta)$ — функция риска, тогда функция сравнения в байесовском подходе

$$\widetilde{R}(\theta^*) = \mathbb{E}_Q R(\theta^*, \theta) = \int\limits_\Theta R(\theta^*, t) Q(dt)$$

Def. Оценка θ^* — наилучшая в некотором классе оценок $\mathcal{K}$ в байесовском подходе, если на ней достигается минимум $\widetilde{R}$.

3.8.3.3 Асимптотический подход

Def. Пусть θ_1^* и θ_2^* — две асимптотически нормальные оценки параметра θ с асимптотическими дисперсиями $\sigma_1^2(\theta)$ и $\sigma_2^2(\theta)$, соответственно. Тогда если $\forall \theta \in \Theta \ \sigma_1^2(\theta) \leqslant \sigma_2^2(\theta)$ и на каком-то θ неравенство строгое, то θ_1^* лучше θ_2^*.

Def. Оценка θ^* — наилучшая в некотором классе оценок $\mathcal{K}$ в асимптотическом подходе, если она лучше всех оценок данного класса.

Example: Пусть $X_1, \ldots, X_n$ — выборка из $\mathcal{N}(\theta, 1)$. Сравните $\overline{X}$ и $\hat{\mu}$ в асимптотическом подходе.

Solution. По ЦПТ для $\overline{X}$: $\sigma_1^2 = 1$, и по теореме о выборочной квантили для $\hat{\mu}$: $\sigma_2^2 = \pi/2$.

Тогда $\overline{X}$ лучше $\hat{\mu}$ в асимптотическом подходе.

3.9 Методы построения оценок: метод моментов и метод максимального правдоподобия. Состоятельность оценки метода моментов. Теорема о свойствах оценок максимального правдоподобия (б/д).

3.9.1 Метод моментов

Пусть $X_1, \dots, X_n$ — выборка из распределения $P \in \{P_\theta \mid \theta \in \Theta\}$, $\Theta \subset \mathbb{R}^k$. Пусть $g_1(x), \dots, g_k(x)$ — такие борелевские функции, что набор $m(\theta) = (m_1(\theta), \dots, m_k(\theta)) = (\mathbb{E}_\theta g_1(X_1), \dots, \mathbb{E}_\theta g_k(X_1))$ однозначно определяет θ, то есть $m(\theta)$ — биекция.

Def. Оценкой параметра θ по методу моментов с пробными функциями $g_1(x), \dots, g_k(x)$ называется решение следующей системы относительно θ^*:

$$\begin{cases} m_1(\theta^*) = \overline{g_1(X)} \\ \dots \\ m_k(\theta^*) = \overline{g_k(X)} \end{cases}$$

Откуда $\theta^*(X) = m^{-1}\left(\overline{g_1(X)}, \dots, \overline{g_k(X)}\right)$.

Note. На практике обычно используют стандартные пробные функции: $g_i(x) = x^i$.

Example: Пусть $X_1, \dots, X_n$ — выборка из $\mathcal{N}(a, \sigma^2)$. Найти оценку $\theta = (a, \sigma^2)$ по методу моментов.

Solution. Воспользуемся стандартными пробными функциями, откуда получим, что

$$\begin{cases} a^* = \overline{X} \\ (a^*)^2 + (\sigma^2)^* = \overline{X^2} \end{cases} \Longrightarrow \begin{cases} a^* = \overline{X} \\ (\sigma^2)^* = \overline{X^2} - (\overline{X})^2 = s^2 \end{cases}$$

То есть $\theta^*(X) = \left(\overline{X}, s^2\right)$ — оценка параметра $\theta = (a, \sigma^2)$ по методу моментов.

3.9.2 Свойства оценки по методу моментов

Th. Пусть $\theta^*(X) = m^{-1}\left(\overline{g_1(X)}, \dots, \overline{g_k(X)}\right)$ — оценка параметра θ по методу моментов, тогда если $m^{-1}(\cdot) \in C(\Theta)$ и $\forall i\ \mathbb{E}_\theta|g_i(X_1)| < \infty$, то θ^* — сильно состоятельная оценка параметра θ.

Proof: По УЗБЧ: $\forall i\ \overline{g_i(X)} = \frac{1}{n}\sum\limits_{j=1}^{n} g_i(X_j) \xrightarrow{P_\theta\text{-п.н.}} \mathbb{E}_\theta g_i(X_1) = m_i(\theta)$.

Так как $m^{-1}(\cdot) \in C(\Theta)$, применим её к $\overline{g_i(X)}$ и воспользуемся теоремой о наследовании сходимости:

$$m^{-1}\left(\overline{g_1(X)}, \dots, \overline{g_k(X)}\right) \xrightarrow{P_\theta\text{-п.н.}} m^{-1}(m_1(\theta), \dots, m_k(\theta)) = m^{-1}(m(\theta)) = \theta$$

■

3.9.3 Метод максимального правдоподобия

Def. Пусть X — наблюдение из распределения $P \in \{P_\theta \mid \theta \in \Theta\}$ — семейство распределений, доминируемое относительно меры μ. Функцией правдоподобия называют с.в. $f_\theta(X) = p_\theta(X)$, где $p_\theta(X)$ — плотность P_θ по мере μ. И если $X = (X_1, \dots, X_n)$ — выборка с плотностью $p_\theta(x)$, то

$$f_\theta(X) = \prod_{i=1}^{n} p_\theta(X_i)$$

Def. Оценкой θ максимального правдоподобия (ОМП) называют $\hat{\theta}(X) = \operatorname{argmax} f_\theta(X)$.

Def. Логарифмическая функция правдоподобия $L_\theta(X) = \ln f_\theta(X)$

3.9.4 Условия регулярности модели

0. $\{P_\theta \mid \theta \in \Theta\}$ – параметрическое семейство распределений, доминируемое относительно меры μ. При этом $P_{\theta_1} \not\equiv P_{\theta_2}$ при $\theta_1 \neq \theta_2$, и $\forall \theta \in \Theta$ определена $p_\theta(x)$ — плотность P_θ по μ.

1. Носитель плотности $\mathcal{A}$ не зависит от θ.

2. Наблюдение $X = (X_1, \dots, X_n)$ — выборка из неизвестного распределения $P \in \{P_\theta \mid \theta \in \Theta\}$.

3. $\Theta \subset \mathbb{R}$ — открытый интервал.

4. Плотность $p_\theta(x)$ непрерывно дифференцируема по θ при всех $x \in \mathcal{A}$.

5. Плотность $p_\theta(x)$ трижды непрерывно дифференцируема по θ при всех $x \in \mathcal{A}$.

6. Интеграл $\int\limits_{\mathcal{A}} p_\theta(x)\mu(dx)$ можно трижды дифференцировать по θ под знаком интеграла.

7. Информация Фишера элемента выборки $i(\theta)$ существует, положительна и конечна.

8. $\forall \theta_0 \in \Theta\ \exists c > 0\ \exists H(x) : \forall x \in \mathcal{A}\ \forall \theta \in (\theta_0 - c, \theta_0 + c) \Longrightarrow \left|\frac{\partial^3}{\partial \theta^3} \ln p_\theta(x)\right| \leqslant H(x)$ и $\mathbb{E}_\theta H(X_1) < \infty$.

3.9.5 Экстремальное свойство функции правдоподобия

Th (б/д). $\forall \theta, \theta_0 \in \Theta$, $\theta \neq \theta_0$ в условиях 0-2 выполнено, что

$$P_{\theta_0}\left(f_{\theta_0}(X_1, \dots, X_n) > f_\theta(X_1, \dots, X_n)\right) \to 1 \tag{3.9.1}$$

Note. То есть предполагается, что выборка $X = (X_1, \dots, X_n)$ из распределения P_{θ_0}.

3.9.6 Состоятельность оценки максимального правдоподобия

Consequence. Если Θ — конечно, то ОМП $\exists$, и она состоятельная оценка θ.

Proof: Из теоремы выше пересечём все события типа 3.9.1.

Th (состоятельность ОМП) (б/д). В условиях регулярности 0-4 уравнение правдоподобия $\frac{\partial}{\partial\theta} f_\theta(X) = 0$ с вероятностью, стремящейся к 1, имеет решение $\widetilde{\theta}_n(X)$ — состоятельная оценка параметра θ.

3.9.7 Асимптотическая нормальность ОМП для одномерного параметра (б/д)

Def. Пусть X — наблюдение из неизвестного параметрического распределения $P \in \{P_\theta \mid \theta \in \Theta\}$ и $\forall\theta \in \Theta$ распределение P_θ имеет плотность $p_\theta(x)$ по одной и той же мере μ (либо считающей, либо Лебега). Иными словами, все распределения либо дискретные, либо абсолютно непрерывные. Тогда семейство $\{P_\theta \mid \theta \in \Theta\}$ называется доминируемым относительно меры μ.

Def. $p_\theta(X) = p_\theta(X_1, \dots, X_n) = \prod\limits_{i=1}^{n} p_\theta(X_i)$ — функция правдоподобия.

Def. Случайная величина $\mathrm{U}_\theta(X) = \frac{\partial}{\partial\theta} \ln p_\theta(X)$ называется вкладом выборки.

Def. Информация Фишера $I_X(\theta) = \mathbb{E}_\theta \left(\mathrm{U}_\theta(X)\right)^2 = \mathbb{E}_\theta \left(\frac{\partial}{\partial\theta} \ln p_\theta(X)\right)^2$ — количество информации о параметре θ, содержащейся в наблюдении X.

Statement. (Аддитивность информации) Если $X = (X_1, \dots, X_n)$ — выборка, то $I_X(\theta) = n \cdot i(\theta)$, где $i(\theta) = \mathbb{E}_\theta \left(\frac{\partial}{\partial\theta} \ln p_\theta(X_1)\right)^2$.

Th. В условиях регулярности 0-8 любая состоятельная последовательность $\widetilde{\theta}_n(X)$ решений уравнения правдоподобия удовлетворяет соотношению

$$\forall\theta \in \Theta \ \sqrt{n}\left(\widetilde{\theta}_n - \theta\right) \xrightarrow{d_\theta} \mathcal{N}\left(0, \frac{1}{i(\theta)}\right) \tag{3.9.2}$$

Consequence. (Асимптотическая нормальность ОМП) В условиях предыдущей теоремы, если $\forall n \ \forall X_1, \dots, X_n \ \exists!$ решение уравнения правдоподобия, то ОМП является асимптотически нормальной оценкой параметра θ с асимптотической дисперсией $\sigma^2(\theta) = \frac{1}{i(\theta)}$.

3.9.8 Эффективность ОМП

Note. Работаем в среднеквадратичном подходе (равномерный подход в классе несмещённых оценок с функцией потерь $g(x, y) = (x - y)^2$). Как тогда искать наилучшую оценку?

Условия регулярности Крамера-Рао:

1. $\Theta \subseteq \mathbb{R}$ — открытый интервал.

2. Носитель плотности $\mathcal{A} = \{x \in \mathcal{X} \mid p_\theta(x) > 0\}$ не зависит от θ.

3. Для любой статистики $S(X)$ такой, что $\mathbb{E}_\theta(S(X))^2 < \infty$, $\forall \theta \in \Theta$ выполнено

$$\frac{\partial}{\partial\theta}\mathbb{E}_\theta S(X) = \mathbb{E}_\theta\left(S(X)\cdot \mathrm{U}_\theta(X)\right) = \mathbb{E}_\theta\left(S(X)\cdot\frac{\partial}{\partial\theta}\ln p_\theta(X)\right)$$

   Пояснение: это требование о возможности занесения производной под знак интеграла.
   И здесь же используется условие регулярности 2: множество интегрирования не зависит от θ.

4. Информация Фишера положительна и конечна, то есть $0 < I_X(\theta) < \infty$ $\forall\theta \in \Theta$.

Th (Неравенство Рао-Крамéра) (б/д). Пусть выполнены условия регулярности 1-4, а $\hat\theta(X)$ — несмещённая оценка $\tau(\theta)$ и $\forall\theta\in\Theta$ $\mathbb{E}_\theta(\hat\theta(X))^2 < \infty$. Тогда верно, что

$$\mathrm{Var}_\theta\hat\theta(X) \geqslant \frac{(\tau'(\theta))^2}{I_X(\theta)}$$

Это значит, что нельзя уточнять оценку бесконечное число раз: есть предел, указанный выше.

Def. Оценка $\hat\theta$ — эффективная, если на ней достигается равенство в неравенстве Рао-Крамера.

Note. Эффективная оценка является наилучшей в среднеквадратическом подходе в классе несмещённых оценок (обратное неверно).

Th (б/д). Пусть выполнены условия регулярности 1-4 для неравенства Рао-Крамера и $\hat\theta(X)$ — эффективная оценка параметра θ. Тогда $\hat\theta(X)$ — ОМП.

3.10 Доверительные интервалы. Метод центральной статистики. Метод построения асимптотических доверительных интервалов.

3.10.1 Доверительные интервалы

Пусть X — наблюдение из распределения P_θ, $\theta \in \mathbb{R}$.

Def. Пара статистик $(T_1(X), T_2(X))$ называется доверительным интервалом (ДИ) уровня доверия γ для параметра θ, если

$$\forall\theta\in\Theta\ P_\theta(T_1(X) < \theta < T_2(X)) \geqslant \gamma \tag{3.10.1}$$

Def. Если в 3.10.1 стоит знак равенства, то такой ДИ называется точным.

Def. ДИ, где одна из статистик заменена на $\pm\infty$ называется односторонним.

Def. Область $S(X) \subset \Theta \subset \mathbb{R}^k$ называется доверительной областью уровня доверия γ для параметра $\theta = (\theta_1, \ldots, \theta_k)$, если

$$\forall\theta\in\Theta\ P_\theta(\theta \in S(X)) \geqslant \gamma$$

3.10.2 Метод центральной статистики

Def. Пусть известна одномерная функция $G(X, \theta)$ такая, что её распределение не зависит от параметра θ. Тогда $G(X, \theta)$ называется центральной статистикой.

Example: Пусть $X_1, \dots, X_n \sim \mathcal{N}(\theta, 1)$, тогда $\sum\limits_{i=1}^{n} X_i \sim \mathcal{N}(n\theta, n)$, значит, $\sum\limits_{i=1}^{n}(X_i - \theta) \sim \mathcal{N}(0, n)$ — не зависит от θ. Таким образом, построили центральную статистику $G(X, \theta) \sim \mathcal{N}(0, n)$.

Пусть $\gamma_1, \gamma_2 \in (0, 1) : \ \gamma_2 - \gamma_1 = \gamma$ и пусть $\exists g_1, g_2$, которые являются γ_i-квантилями распределения статистики $G(X, \theta)$, соответственно. Тогда

$$\forall \theta \in \Theta \ P_\theta(g_1 \leqslant G(X, \theta) \leqslant g_2) \geqslant F_{G(X,\theta)}(g_2) - F_{G(X,\theta)}(g_1) \geqslant \gamma_2 - \gamma_1 = \gamma \tag{3.10.2}$$

Note. Если $S(X) = \{\theta \in \Theta \mid g_1 \leqslant G(X, \theta) \leqslant g_2\}$, то получим доверительную область.
Для получения ДИ или доверительной области необходимо решить два неравенства из левой части 3.10.2 относительно θ.

Note. Центральная статистика может быть далеко не единственной.

3.10.3 Асимптотические доверительные интервалы

Def. Пусть $X = (X_1, \dots, X_n)$ — н.о.р.с.в. из неизвестного распределения $P \in \{P_\theta \mid \theta \in \Theta\}$.
Тогда последовательность пар статистик $(T_{n,1}(X), T_{n,2}(X))$ называется асимптотическим доверительным интервалом, если

$$\forall \theta \in \Theta \ \underline{\lim_{n\to\infty}} P_\theta(T_{n,1}(X) < \theta < T_{n,2}(X)) \geqslant \gamma \tag{3.10.3}$$

Def. Если в 3.10.3 достигается равенство, то этот асимптотический ДИ называется точным.

3.10.4 Построение асимптотических ДИ с помощью асимптотически нормальных оценок

Пусть $\hat{\theta}$ — асимптотически нормальная оценка θ. Тогда, по определению:

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{d_\theta} \mathcal{N}(0, \sigma^2(\theta)) \Longrightarrow \sqrt{n} \cdot \frac{\hat{\theta}_n - \theta}{\sigma(\theta)} \xrightarrow{d_\theta} \mathcal{N}(0, 1) \tag{3.10.4}$$

Реальность не так проста, ибо далеко не всегда из-за выражения $\sigma(\theta)$ удастся явно разрешить 3.10.4.

Поэтому положим, что $\sigma(\theta)$ непрерывна. Но $\hat{\theta}_n \xrightarrow{P_\theta} \theta$ (ас. норм. $\Rightarrow$ сост.), тогда по теореме о наследовании сходимости $\sigma(\hat{\theta}_n) \xrightarrow{P_\theta} \sigma(\theta)$. И, таким образом, по лемме Слуцкого

$$\sqrt{n} \cdot \frac{\hat{\theta}_n - \theta}{\sigma(\theta)} \cdot \frac{\sigma(\theta)}{\sigma(\hat{\theta}_n)} = \sqrt{n} \cdot \frac{\hat{\theta}_n - \theta}{\sigma(\hat{\theta}_n)} \xrightarrow{d_\theta} \mathcal{N}(0, 1) \tag{3.10.5}$$

Тогда пусть $u_{\frac{1+\gamma}{2}}$ — $\frac{1+\gamma}{2}$-квантиль $\mathcal{N}(0,1)$, тогда с учетом того, что будет $u_{\frac{1-\gamma}{2}} = -u_{\frac{1+\gamma}{2}}$, получаем

$$P_\theta\left(-u_{\frac{1+\gamma}{2}} \leqslant \sqrt{n}\cdot\frac{\hat\theta_n-\theta}{\sigma(\hat\theta_n)} \leqslant u_{\frac{1+\gamma}{2}}\right)\to\gamma \iff P_\theta\left(\hat\theta_n - u_{\frac{1+\gamma}{2}}\cdot\frac{\sigma(\hat\theta_n)}{\sqrt{n}} \leqslant \theta \leqslant \hat\theta_n + u_{\frac{1+\gamma}{2}}\cdot\frac{\sigma(\hat\theta_n)}{\sqrt{n}}\right)\to\gamma$$

Значит, асимптотический ДИ уровня γ таков:

$$\left(\hat\theta_n - u_{\frac{1+\gamma}{2}}\cdot\frac{\sigma(\hat\theta_n)}{\sqrt{n}}, \hat\theta_n + u_{\frac{1+\gamma}{2}}\cdot\frac{\sigma(\hat\theta_n)}{\sqrt{n}}\right)$$

3.11 Статистические гипотезы, ошибки первого и второго рода, уровень значимости критерия. Принципы сравнения критериев, равномерно наиболее мощные критерии. Лемма Неймана–Пирсона. Построение с её помощью наиболее мощных критериев.

Пусть наблюдение $X \sim P \in \mathcal{P}$, где $\mathcal{P}$ — некое произвольное семейство распределений.

3.11.1 Проверка статистических гипотез

Def. Статистическая гипотеза — предположение вида $P \in \mathcal{P}_0 \subset \mathcal{P}$ — подсемейство распределений.

Основная задача проверки гипотез: либо принять гипотезу $H_0 : P \in \mathcal{P}_0$, либо отвергнуть.

Если H_0 отвергается, то переходим к рассмотрению другой гипотезы $H_1 : P \in \mathcal{P}_1 \subset (\mathcal{P} \setminus \mathcal{P}_0)$.

Note. Если $\mathcal{P}$ — параметризовано, то основная гипотеза выглядит так: $H_0 : \theta \in \Theta_0 \subset \Theta$.

Def. Пусть $X \in \mathcal{X}$, а $S \subset \mathcal{X}$ — некое подмножество. Если правило проверки H_0 выглядит так:

$$\begin{cases} X \notin S & \Rightarrow \quad H_0 \text{ не отвергается} \\ X \in S & \Rightarrow \quad H_0 \text{ отвергается} \end{cases}$$

тогда S называется критическим множеством для проверки H_0.

Результаты проверки:

1. $X \in S \Rightarrow H_0$ отвергается; результат статистически значим.
2. $X \notin S \Rightarrow H_0$ не отвергается; результат статистически не значим.

Note. Никогда не говорим «H_0 принимается».

Def. Правило из определения выше называется критерием проверки H_0 (против H_1).

Def. Ошибка I рода: отвергаем верную гипотезу.

Def. Ошибка II рода: не отвергаем неверную гипотезу.

Note. Ошибка I рода опаснее, так как после отвержения гипотезы, к ней большей не возвращаемся. Поэтому критерий нужно выбирать так, чтобы вероятность ошибки I рода была меньше заранее выбранного α, а вероятность ошибки II рода была как можно меньше.

Def. Пусть S — критическое множество для проверки $H_0 : P \in \mathcal{P}_0$.

Функция $\beta(Q, S) = Q(X \in S)$, где $Q \in \mathcal{P}$, называется функцией мощности критерия.

Def. Если для критерия с критическим множеством S выполнено $\beta(Q, S) \leqslant \alpha \ \forall Q \in \mathcal{P}_0$
(то есть вероятность ошибки I рода). Тогда говорят, что критерий имеет уровень значимости α.

Def. Минимальный уровень значимости $\varepsilon(S) = \sup\limits_{Q \in \mathcal{P}_0} \beta(Q, S)$ называется размером критерия.

3.11.2 Сравнение критериев

Def. Критерий с критическим множеством S уровня значимости α называется более мощным, чем критерий с критическим множеством R уровня значимости α, если $\forall Q \in \mathcal{P}_1 \ \beta(Q, S) \geqslant \beta(Q, R)$
(то есть вероятность ошибки II рода у первого критерия равномерно меньше, чем у второго).

Note. Задавая уровень значимости, регулируем вероятность ошибки I рода, а далее сравниваем по ошибке II рода (у более мощного критерия она меньше).

Def. Критерий с критическим множеством S проверки гипотезы H_0 называется равномерно наиболее мощным критерием (р.м.н.к.) уровня значимости α, если он мощнее любого другого критерия проверки той же гипотезы H_0 того же уровня значимости α.

3.11.3 Лемма Неймана-Пирсона

Def. Гипотеза простая, если она имеет вид $H_0 : P = P_0$ — конкретное известное распределение.

Note. Пусть имеются две простые гипотезы $H_0 : P = P_0$ и $H_1 : P = P_1$, причём P_i имеют плотность $p_i(x)$ по одной и той же мере μ. Рассмотрим критерий с критическим множеством

$$S_\lambda = \{x \in \mathcal{X} \mid p_1(x) - \lambda p_0(x) \geqslant 0\}$$

Lemma. Пусть есть критерий с крит. множеством R таким, что $P_0(X \in S_\lambda) \geqslant P_0(X \in R)$. Тогда

1. $P_1(X \in R) \leqslant P_1(X \in S_\lambda)$ (то есть S_λ — р.н.м.к.)
2. $P_0(X \in S_\lambda) \leqslant P_1(X \in S_\lambda)$ (то есть S_λ — несмещённый)

Proof:

1. Для любого $x \in \mathcal{X}$ выполнено

$$I_R(x) \cdot (p_1(x) - \lambda p_0(x)) \leqslant I_R(x) \cdot (p_1(x) - \lambda p_0(x)) \cdot I_{S_\lambda}(x) \leqslant (p_1(x) - \lambda p_0(x)) \cdot I_{S_\lambda}(x)$$

Отсюда

$$P_1(X \in R) - \lambda P_0(X \in R) = \mathbb{E}_1 I_R(X) - \lambda \mathbb{E}_0 I_R(X) = \int\limits_{\mathcal{X}} I_R(x) \, (p_1(x) - \lambda p_0(x)) \, d\mu \leqslant$$

$$\leqslant \int\limits_{\mathcal{X}} I_{S_\lambda}(x) \, (p_1(x) - \lambda p_0(x)) \, d\mu = P_1(X \in S_\lambda) - \lambda P_0(X \in S_\lambda) \Longleftrightarrow$$

$$\Longleftrightarrow P_1(X \in S_\lambda) - P_1(X \in R) \geqslant \lambda \, (P_0(X \in S_\lambda) - P_0(X \in R)) \geqslant 0 \Longleftrightarrow P_1(X \in S_\lambda) \geqslant P_1(X \in R)$$

2. Пусть $\lambda \geqslant 1$, тогда $\forall x \in S_\lambda \; p_0(x) \leqslant p_1(x)$ и

$$P_0(X \in S_\lambda) = \int\limits_{\mathcal{X}} I_{S_\lambda}(x) p_0(x) d\mu \leqslant \int\limits_{\mathcal{X}} I_{S_\lambda}(x) p_1(x) d\mu = P_1(X \in S_\lambda)$$

Теперь $\lambda < 1$, тогда

$$\forall x \in \overline{S_\lambda} \; p_0(x) \geqslant p_1(x) \Longrightarrow 1 - P_0(X \in S_\lambda) = P_0(X \in \overline{S_\lambda}) \geqslant P_1(X \in \overline{S_\lambda}) = 1 - P_1(X \in S_\lambda)$$

■

Consequence. Если λ удовлетворяет соотношению $P_0(X \in S_\lambda) = \alpha$, то критерий с критическим множеством S_λ — р.н.м.к. уровня значимости α. То есть мы должны решить относительно λ уравнение

$$\int\limits_{\{x : p_1(x) - \lambda p_0(x) \geqslant 0\}} p_0(x) d\mu = \alpha$$

Глава 4

Материалы

1. Вопросы ГОСа 2021
2. Текущий статус
3. Ссылка на этот тех