Математические основы искусственного интеллекта
Комбинаторика

Солодушкин Святослав Игоревич

Кафедра вычислительной математики и компьютерных наук,
УрФУ имени первого Президента России Б.Н. Ельцина

Октябрь 2021

Какие темы изучаем

Первый блок, «Теория вероятностей»

1. Комбинаторика. Теоремы сложения и умножения.
2. Основные понятия теории вероятностей.
3. Условная вероятность. Формула полной вероятности и формула Байеса.
4. Дискретные случайные величины.
5. Непрерывные случайные величины. Функция и плотность распределения.
6. Числовые характеристики случайных величин.
7. Нормальное распределение.
8. Система двух случайных величин.
9. Система двух случайных величин, корреляция.

Какие темы изучаем

Второй блок, «Основные понятия статистики»

1. Генеральная совокупность и выборка. Дизайн исследования.
2. Описательные статистики. Метод моментов. Доверительные интервалы.
3. Метод максимального правдоподобия.
4. Проверка статистических гипотез. Общая теория.
5. Проверка статистических гипотез. Сравнение средних.
6. Проверка статистических гипотез.
7. Проверка статистических гипотез.
8. Анализ статистических связей.
9. Регрессионные модели.
10. Анализ выживаемости.

Основные задачи теории вероятностей

Основной задачей теории вероятностей является изучение различных случайных сущностей

1. случайных событий
2. случайных величин
3. случайных процессов

и их числовая характеристика (например, нахождение вероятностей случайных событий, нахождение средних значений случайных величин).

Традиционно начинают изучение теории вероятностей с рассмотрения случайных событий, а затем переходят к случайным величинам и их совокупностям.

Для того, чтобы выстроить общую теорию, начать надо с рассмотрения простых примеров. На этой лекции изучим комбинаторику.

Теорема 1. Пусть множество $A = \{a_1, \ldots, a_k\}$ состоит из k элементов, а множество $B = \{b_1, \ldots, b_m\}$ — из m элементов. Тогда можно образовать ровно km пар (a_i, b_j), взяв первый элемент из множества A а второй — из множества B.

Пусть ящик содержит n шаров пронумерованных числами 1, 2, 3, …, n. Из ящика достают k шаров; результат этого выбора — набор из k шаров.

Сколькими способами можно выбрать k шаров из n, т. е. сколько различных результатов возможно?

На этот вопрос нельзя дать однозначный ответ. Нужно уточнить:

1. как организован выбор;
2. что понимать под различными результатами выбора.

Выбор с возвращением и без

1. Выбор с возвращением: каждый вынутый шар возвращается в урну, каждый следующий шар выбирается из полной урны. В полученном наборе из k номеров шаров могут встречаться одни и те же номера.
2. Выбор без возвращения: вынутые шары в урну не возвращаются, и в полученном наборе не могут встречаться одни и те же номера.

Важность порядка в результирующей выборке

Условимся, какие результаты выбора (какие наборы номеров шаров) считаются различными. Есть ровно две возможности.

1. Выбор с учетом порядка: два набора номеров шаров считаются различными, если они отличаются составом или порядком номеров. Так, наборы $(1, 5, 2)$, $(2, 5, 1)$ и $(1, 2, 5)$ считаются различными наборами.
2. Выбор без учета порядка: два набора номеров шаров считаются различными, если они отличаются составом. Так, наборы $(1, 5, 2)$ и $(2, 5, 3)$ различны, а наборы $(1, 5, 2)$ и $(2, 5, 1)$ не различаются.

Теорема 2. Общее количество различных наборов при выборе k элементов из n без возвращения и с учетом порядка равняется $A_n^k = n(n-1)\cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$.
Число A_n^k называется числом размещений из n элементов по k элементов, а сами результаты выбора — размещениями.

Теорема 3. Общее количество различных наборов при выборе k элементов из n без возвращения и без учета порядка равняется $C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!\,(n-k)!}$.
Число C_n^k называется числом сочетаний из n элементов по k элементов, а сами результаты выбора — сочетаниями.

Теорема 4. Общее количество различных наборов при выборе k элементов из n с возвращением и с учетом порядка равняется n^k.

Теорема 5. Общее количество различных наборов при выборе k элементов из n с возвращением и без учета порядка равняется

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

Доказательство будем пояснять на примере.
В ящике $n = 4$ вида шаров, извлекают (с возвращением) $k = 5$ шаров.
Примеры наборов: (1, 3, 1, 1, 4) или (3, 2, 1, 2, 4)
Рассмотрим, чем отличаются друг от друга два разных результата такой схемы выбора. Не важен порядок следования номеров, т. е. учитывается только то, сколько раз в наборе из k шаров появился каждый номер. Поэтому результат выбора можно представить набором чисел $k_1, \ldots, k_n$, в котором $k_i \geq 0$ — число появлений шара номер i в наборе, $k_1 + \ldots + k_n = k$.
Вот так: (1, 1, 1, 3, 4) $k_1 = 3, k_2 = 0, k_3 = 1, k_4 = 1$ и
(1, 2, 2, 3, 4) $k_1 = 1, k_2 = 2, k_3 = 1, k_4 = 1$.

Доказательство теоремы 5

Два результата выбора с возвращением и без учета порядка различаются, если соответствующие им упорядоченные наборы $k_1, \ldots, k_n$ не совпадают.
Наборы (1, 3, 1, 1, 4) и (4, 1, 3, 1, 1) совпадают.

Представим себе другой эксперимент, имеющий точно такие же результаты, и посчитаем количество вариантов. В ряд выстроены n ящиков, в которых размещаются k шаров. Важно только число шаров в каждом ящике. Результатом эксперимента снова является набор чисел $k_1, \ldots, k_n$, где $k_i \geq 0$ равно числу шаров в ящике с номером i, $k_1 + \ldots + k_n = k$.

Можно считать, что есть $k + (n - 1)$ позиций, в каждую из которых можно поместить либо одну из $n - 1$ перегородки, либо один из k шаров. Расставляя $n - 1$ перегородку, мы автоматически определяем позиции для шаров.

Согласно теореме 3 существует $C_{n+k-1}^{k} = C_{n+k-1}^{n-1}$ способов выбрать места для $n - 1$ перегородку на $n - 1 + k$ местах.

Задачи простые

1. Имеется 5 видов конвертов без марок и 3 вида марок. Сколькими способами можно выбрать конверт и марку для отправки письма?
2. Из 10 различных книг выбирают 4 для посылки. Сколькими способами это можно сделать?
3. Сколько «слов» (из 10 букв) можно получить, переставляя буквы слова «математика»?
4. Отрезок длиной L делится случайным образом на три части. Какова вероятность, что из полученных частей можно составить треугольник?

Задачи чуть посложнее

1. Сколько чисел среди первых 1000 натуральных чисел не делятся ни на 2, ни на 3, ни на 5?
2. Василий выписал все различные делители числа 2016. Сколько чисел выписал Василий?
3. Пусть $m \leq n$. Найти число возрастающих функций $f : 1, \ldots, m \longrightarrow 1, \ldots, n$, т. е. функций удовлетворяющих условию $i < j \Rightarrow f(i) < f(j)$.
4. Сколько 6-значных чисел можно получить из цифр 1, 2, 3, 4, 5, 6, 7, 8, 9, если число должно состоять из 3 четных и 3 нечетных цифр и никакие две цифры не повторяются?
5. Сколькими способами можно переставить буквы слова «вероятность», чтобы две одинаковые буквы не стояли рядом?