Летний экзамен по алгебре

hse-ami-open-exams

Содержание

1 Дискретное вероятностное пространство. Задача о разделе ставки. Вероятностный алгоритм проверки на простоту. Универсальная хеш-функция.

1.1 Дискретное вероятностное пространство.

Пусть Ω – непустое конечное *множество элементарных исходов.*

Определение 1. *Всякое подмножество $A \subseteq \Omega$ называют* ***событием****.*

Определение 2. *Функцию $P : 2^\Omega \to [0, 1]$, удовлетворяющую следующим свойствам:*

- $P(\Omega) = 1$
- $A \cap B = \varnothing \Rightarrow P(A \cup B) = P(A) + P(B)$

называют ***вероятностной мерой****, а значение $P(A)$* ***вероятностью события*** *A. Вероятностная мера полностью определяется значениями $P(\{\omega\}) = p_\omega$, т.е.*

$$P(A) = \sum_{\omega \in A} p_\omega$$

Если все элементарные исходы равновозможны, то полагаем $p_{\omega_1} = ... = p_{\omega_n} = 1/n$.

1.2 Задача о разделе ставки.

Два человека играют в некоторую игру, причем у обоих шансы победить одинаковые. Они договорились, что тот, кто первым выиграет 6 партий, получит весь приз. Однако игра остановилась раньше, когда первый выиграл пять партий, а второй выиграл три партии. Как справедливо разделить приз?
Предлагается разделить приз в отношении, в котором относятся вероятности выиграть для каждого из игроков в случае продолжения игры. Ясно, что еще надо сыграть не более трех партий. Пространство исходов этих трех партий состоит из восьми элементов, причем только один из этих исходов означает выигрыш второго игрока. Значит приз надо разделить в отношении 7 к 1.

1.3 Вероятностный алгоритм проверки на простоту.

Пусть дано некоторое натуральное число $N > 1$. Если N простое число, то по малой теореме Ферма для всякого натурального числа, такого, что $(b, N) = 1$, число $b^{N-1} - 1$ делится на N. Следовательно, если для некоторого b, удовлетворяющего условию $(b, N) = 1$, число $b^{N-1} - 1$ не делится на N, то N не является простым. Это наблюдение используют для построения простейшего теста на простоту. Если $b^{N-1} - 1$ не делится на N, то говорим, что N не проходит тест для основания b.
Пусть основание мы выбираем случайно из множества $\mathbb{Z}_N^*$. Предположим, что существует такое основание, для которого N не проходит тест. Какова вероятность выбрать такое основание?
Предположим, что для $a \in \mathbb{Z}_N^*$ число N не проходит тест. Если N проходит тест для основания b, то для основания ab число N уже тест не проходит. В противном случае $(ab)^{N-1} \equiv_N 1$ и $(b^{-1})^{N-1} \equiv_N 1$. Следовательно, $a^{N-1} \equiv (b^{-1})^{N-1}(ab)^{N-1} \equiv 1$, что противоречит предположению. Таким образом, каждому основанию b, для которого N проходит тест, можно сопоставить основания ab, для котрого результат теста отрицательный. Значит, оснований, для которых N не проходит тест, не м еньше оснований, для которых N проходит тест на простоту. Искомая вероятность не меньше $1/2$. Если независимым образом повторять набор основания k раз, то вероятность выбрать основание, для которого данное число не проходит тест, меньше $1/2^k$.

1.4 Универсальная хеш-функция.

Пусть $K = \{0, 1, 2, ..., n-1\}$ – множество «ключей». Отображение

$$h : K \to \{0, 1, 2, ..., m-1\}$$

называется хеш-функцией. Предполагается, что $m < n$. Одним из важнейших свойств функции h является равномерность, когда доля ключей k с фиксированным значением $h(k)$ должно быть примерно n/m. Это означает, что вероятность коллизии $h(k_1) = h(k_2)$ при $k_1 \neq k_2$ не больше $1/m$. Ясно, что не всегда можно предполагать, что ключи равномерно распределены по таблице. Предположим, что $h(k) = k \mod m$ и на вход сначала подаются ключи вида $m, 2m, 3m, ...$ Ясно, что всем таким ключам присваивается хеш-код 0 и реально никакого равномерного распределения значений не происходит. Оказывается, с этой проблемой можно справиться, если перед началом хеширования случайным образом выбирать функцию h из некоторого набора таких функций. Зафиксируем простое число $p > n$. Пусть $a \in \{1, 2, ..., o-1\}$ и $b \in \{0, 1, 2, ..., o-1\}$. Положим

$$h_{a,b} = ak + b \mod p \mod m.$$

Пара параметров (a, b) выбирается случайным образом из множества

$$\{1, 2, ..., p-1\} \times \{0, 1, 2, ..., p-1\},$$

причем все элементы этого множества считаем равновероятными. Докажем, что для любых

$$k_1, k_2 \in \{0, 2, ..., n-1\}, \quad k_1 \neq k_2,$$

вероятность коллизии $h_{a,b}(k_1) = h_{a,b}(k_2)$ не превосходит $1/m$.
Заметим, что $ak_1 + b = ak_2 + b \mod p$ тогда и только тогда, когда $k_1 = k2$. Кроме того, для различных k_1 и k_2 по значениям $ak_1 + b \mod p$ и $ak_2 + b \mod p$ однозначно находятся числа a и b. Пусть $k_1 \neq k_2$. Отображение

$$(a, b) \to (ak_1 + b \mod p, ak_2 + b \mod p)$$

является биекцией множества

$$\{1, 2, ..., p-1\} \times \{0, 1, 2, ..., p-1\}$$

на множество

$$(\{0, 1, 2, ..., p-1\} \times \{0, 1, 2, ..., p-1\}) \setminus \{(i, i) \mid 0 \leqslant i \leqslant p-1\}.$$

Остается отметить, что для всякого $t \in \{0, 1, 2, ..., p-1\}$ количетсво чисел $s \in \{0, 1, 2, ..., p-1\}$ таких, что $s \neq t$ и $s = t \mod m$, не превосходит $(p-1)/m$, т.е. вероятность выбора такой пары (t, s) или, что эквивалентно, выбора пары (a, b), у которой $h_{a,b}(k_1) = h_{a,b}(k_2)$, не превосходит $1/m$.

2 Свойства вероятностной меры. Формула включений и исключений. Парадокс распределения подарков. Задача про конференцию.

2.1 Свойства вероятностной меры.

- $P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B)$
- $P(\bigcup\limits_k A_k) \leq \sum\limits_k P(A_k)$
- $P(\bigcup\limits_k A_k) = \sum\limits_{k=1}^{n} (-1)^{k+1} \sum\limits_{i_1 < ... < i_k} P(A_{i_1} \cap ... \cap A_{i_k})$

2.2 Формула включений и исключений.

$$P(\bigcup_k A_k) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{i_1 < ... < i_k} P(A_{i_1} \cap ... \cap A_{i_k})$$

Докажем индуктивно. Для $k = 1$ очевидно, что верно. Для $k = 2$ проверено перывм свойством. Допустим для n верно, докажем для $n+1$: $P(A_1 \cup ... \cup A_n \cup A_{n+1}) = P(A_1 \cup ... \cup A_n) + P(A_{n+1}) - P((A_1 \cap A_{n+1}) \cup ... \cup (A_n \cap A_{n+1})) = P(A_1) + ... + P(A_n) - \sum\limits_{j<k} P(A_j \cap A_k) + P(A_{n+1}) + P(A_1 \cap A_{n+1}) + ... + P(A_n \cap A_{n+1}) - \sum\limits_{j<k} P(A_j \cap A_k \cap A_{n+1})$ и так далее. Таким образом, переместив все, где есть A_{n+1} под знаки суммирования, мы получим исходную формулу.

2.3 Парадокс распределения подарков.

Несколько человек решили сделать друг другу подарки следующим образом. Каждый приносит подарок. Подарки складываются вместе, перемешиваются и случайно распределяются среди участников. Этот справедливый способ раздачи подарков применяется часто, так как считают, что для больших групп людей вероятность совпадения, т. е. получения кем-то собственного подарка, очень мала. Парадоксально, но вероятность по крайней мере одного совпадения намного больше вероятности того, что совпадений нет (кроме случая, когда группа состоит из двух человек, тогда вероятность отсутствия совпадений равна 1/2).

Доказательство. Рассмотрим компанию из n человек, тогда число подарков также равно n. Подарки могут быть распределены $n!$ различными способами. (Это общее число исходов.) Число исходов, в которых никто не получит свой собственный подарок, равно

$$\binom{n}{0} n! - \binom{n}{1} (n-1)! + ... + (-1)^n 0!,$$

так что отношение числа благоприятных исходов к общему числу исходов вычисляется по формуле

$$P = \frac{1}{2!} - \frac{1}{3!} + ... + \frac{(-1)^n}{n!}$$

и P действительно меньше 1/2 при $n > 2$. Таким образом, при $n = 6$ имеем

$$\frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \frac{1}{5!} + \frac{1}{6!} = \frac{53}{144} \approx 0.36$$

□

2.4 Задача про конференцию.

В научном центре работают специалисты по 60 различным разделам компьютерных наук. Известно, что по каждому разделу в центре работает ровно 7 ученых, причем вполне может быть, что один ученый является специалистом сразу по нескольким направлениям. Все ученые должны принять участие в одной (и только одной) из двух конференций, одна из которых проходит в Канаде, а другая в Австралии. Оказывается, что всегда можно так распределить ученых по этим конференциям, что на каждой конференции будут присутствовать специалисты по всем 60 направлениям компьютерных наук.

Доказательство. Будем для каждого ученого выбирать конференцию простым подбрасыванием правильной монеты. Для каждого направления компьютерной мысли рассмотрим событие, состоящее в том, что среди ученых этого направления окажутся и те, которые поехали в Канаду, и те, которые поехали в Австралию. Вероятность этого события равна $1 - 2^{-6}$ (нас устроят все исходы кроме двух, когда все отправились на конференцию в одну страну). Остается заметить, что число событий равно 60 и вероятность каждого события больше $1 - 60^{-1}$. □

3 Условная вероятность. Независимые события. Отличие попарной независимости и независимости в совокупности.

3.1 Условная вероятность.

Условной вероятностью A при событии B называется число

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Если зафиксировать B, то $P'(x) = P(x|B)$ является вероятностной мерой. Равенство часто записывают как $P(A \cap B) = P(B) \cdot P(A|B)$ и называют правилом произведения.

3.2 Независимые события.

События A и B называются независимыми, если $P(A \cap B) = P(A) \cdot P(B)$.

3.3 Отличие попарной независимости и независимости в совокупности.

4 Формула полной вероятности. Формула Байеса. Задача о сумасшедшей старушке. Парадокс Байеса.

4.1 Формула полной вероятности.

4.2 Формула Байеса.

4.3 Задача о сумасшедшей старушке.

4.4 Парадокс Байеса.

5 Схема Бернулли. Моделирования бросания правильной монеты. Теорема Муавра-Лапласа. Закон больших чисел для схемы Бернулли.

5.1 Схема Бернулли.

5.2 Моделирования бросания правильной монеты.

5.3 Теорема Муавра-Лапласа.

5.4 Закон больших чисел для схемы Бернулли.

6 Случайное блуждание: принцип отражения, задача о баллотировке и задача о возвращении в начало координат. Броуновское движение.

6.1 Случайное блуждание: принцип отражения, задача о баллотировке и задача о возвращении в начало координат.

6.2 Броуновское движение.

7 Теорема Пуассона. Распределение Пуассона. Задача про изюм. Пуассоновский процесс.

7.1 Теорема Пуассона.

7.2 Распределение Пуассона.

7.3 Задача про изюм.

7.4 Пуассоновский процесс.

8 Марковские цепи. Существование стационарного распределения и сходимость к стационарному распределению.

8.1 Марковские цепи.

8.2 Существование стационарного распределения и сходимость к стационарному распределению.