Осенний коллоквиум курса «Теория вероятностей»

ФКН НИУ ВШЭ, 2-й курс ОП ПМИ, 2-й модуль, 2016 учебный год

Дата последнего обновления: 11.11.2016

Билет 1

Дискретное вероятностное пространство. Вероятностный алгоритм проверки на простоту. Универсальная хеш-функция.

Дискретное вероятностное пространство

Рассмотрим некоторый эксперимент, все мыслимые исходы которого описываются конечным числом различных исходов ω_1, ..., ω_N. Несущественна природа этих исходов, важно лишь то, что их число N конечно.

Определение 1. Исходы ω_1, ..., ω_N будем называть *элементарными событиями*, а их совокупность

$$\Omega = \{\omega_1, ..., \omega_N\}.$$

(конечным) пространством элементарных событий или *пространством исходов*.

Замечание 1. Можно также называть Ω *множеством элементарных исходов*. Именно так его называют в кратком конспекте лекций.

Определение 2. Всякое подмножество $A \subseteq \Omega$ называется *событием*.

Определение 3. Функцию

$$P\colon 2^\Omega \to [0, 1],$$

удовлетворяющую следующим свойствам:

(a) $P(\Omega) = 1$,

(b) $A \cap B = \varnothing \Rightarrow P(A \cup B) = P(A) + P(B)$ (*правило суммы* или *аддитивность*)

называют *вероятностной мерой*, а значение $P(A)$ *вероятностью события* A.

Замечание 2. Вероятностная мера P полностью определяется значениями $P(\omega_1) = p_1$, ..., $P(\omega_N) = p_N$.

Следствие 1. *Из определения вероятностной меры следует, что*

(a) $p_\omega \geqslant 0$,

(b) $\sum\limits_\omega p_\omega = 1$,

(c) *вероятность произвольного события A вычисляется по формуле*

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

Определение 4. Если все элементарные исходы *равновозможны*, то полагаем, что

$$p_{\omega_1} = ... = p_{\omega_n} = \frac{1}{n}.$$

Замечание 3. В случае, если все элементарные исходы равновозможны, вероятность события A равна отношению количества исходов из A к числу всех исходов в Ω.

Определение 5. *Дискретное вероятностное пространство* — это пара из множества элементарных событий Ω и определенной для него вероятностной мерой.

Вероятностный алгоритм проверки на простоту

Пусть дано некоторое натуральное число $N > 1$. Мы хотим проверить, является ли это число простым. Можно перебирать все простые делители до $\sqrt{N}$, но это очень долго. Хотелось бы иметь быстрый способ проверки.
Если N — простое число, то по малой теореме Ферма для всякого натурального числа b 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.
Пусть основание b 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.

Замечание 4. $\mathbb{Z}_N^*$ - мультипликативная группа обратимых элементов кольца вычетов по модулю N.

Если N проходит тест для основания b, то для основания ab число N уже тест не проходит. В противном случае

$$(ab)^{N-1} \equiv 1 \ (mod\ N),\ (b^{-1})^{N-1} \equiv 1 \ (mod\ N).$$

Следовательно,

$$\begin{cases} a^{N-1} \equiv (b^{-1})^{N-1}(ab)^{N-1} \ (mod\ N), \\ (b^{-1})^{N-1}(ab)^{N-1} \equiv 1 \ (mod\ N), \end{cases} \Rightarrow a^{N-1} \equiv 1 \ (mod\ N),$$

что противоречит предположению. Таким образом, каждому основанию b, для которого N проходит тест, можно сопоставить основание ab, для которого результат теста отрицательный. Значит, оснований, для которых N не проходит тест, не меньше оснований, для которых N проходит тест на простоту. Искомая вероятность не меньше $\frac{1}{2}$. Если независимым образом повторять выбор основания k раз, то вероятность выбрать основание, для которого данное число проходит тест, меньше $\frac{1}{2^k}$.

Замечание 5. Бывают числа, которые проходят тест для всех оснований b. Это числа Кармайкла, например 561.

Замечание 6. Докажем, что $(b^{-1})^{N-1} \equiv 1 \ (mod\ N)$.

Доказательство. Число $b^{N-1} - 1$ делится на N, так N проходит тест для основания b. А значит,

$$b^{N-1} \equiv 1 \ (mod\ N) \mid \cdot (b^{-1})^{N-1},$$
$$1 \equiv (b^{-1})^{N-1} \ (mod\ N).$$

□

Универсальная хеш-функция

Определение 6. Пусть H - конечное множество хеш-функций, которые отображают пространство ключей U ($|U| = n$) в диапазон $\{0, 1, ..., m-1\}$. Такое множество называется *универсальным*, если для каждой пары ключей $k, l \in U, (k \neq l)$, количество хеш-функций $h \in H$, для которых $h(k) = h(l)$ не превышает $\frac{|H|}{m}$.

Иными словами, при случайном выборе хеш-функции из множества H вероятность коллизии между двумя различными ключами k, l не превышает вероятности совпадения двух случайным образом выбранных хеш-значений из множества $\{0, 1, ..., m-1\}$, которая равна $\frac{1}{m}$.
Далее будем считать, что $U = \{0, 1, ..., n-1\}$.

Теорема 1. *Множество хеш-функций* $H_{p,m} = \{h_{a,b} : a \in \mathbb{Z}_p^*, b \in \mathbb{Z}_p\}$, *где*

$$h_{a,b}(k) = ((ak + b) \ mod\ p) \ mod\ m,$$

$\mathbb{Z}_p^* = \{1, 2, ..., p-1\}, \mathbb{Z}_p = \{0, 1, ..., p-1\}$, *p - простое число,* $p > n$, *является универсальным.*

Доказательство. Рассмотрим $k, l \in \mathbb{Z}_p : k \neq l$. Пусть для данной хеш-функции $h_{a,b}$

$$r = (ak + b) \ mod\ p,$$
$$s = (al + b) \ mod\ p.$$

Заметим, что $r \neq s$, так как $r - s \equiv a(k - l) \ (mod\ p)$, а p - простое число, a и $(k - l)$ не равны нулю по модулю p, а значит и разность r и s также отлична от нуля по модулю p. Таким образом, коллизии "по модулю p" отсутствуют. Более того, каждая из $p(p-1)$ возможных пар (a, b) приводит к различным

$(r,s) : r \neq s$. Чтобы доказать это, достаточно рассмотреть возможность однозначного определения a и b по заданным r и s:

$$a = ((r-s)\cdot(k-l)^{-1}) \ mod \ p,$$
$$b = (r - ak) \ mod \ p.$$

(Доказательство приведено ниже, в Замечании под номером 7).
Поскольку имеется только $p(p-1)$ возможных пар (r,s): $r \neq s$, то имеется взаимооднозначное соответствие между парами (a,b) и парами (r,s): $r \neq s$. Таким образом, для любых k,l при равномерном случайном выборе пары (a,b) из $\mathbb{Z}_p^* \times \mathbb{Z}_p$ получаемая в результате пара (r,s) может быть с равной вероятностью любой из пар с отличающимися значениями по модулю p.
Отсюда следует, что вероятность того, что различные ключи k,l приводят к коллизии, равна вероятности того, что $r \equiv s \ (mod \ m)$ при произвольном выборе отличающихся по модулю p значений r и s. Для данного r имеется $p-1$ возможное значение s. При этом число значений $s : s \neq r$ и $s \equiv r \ (mod \ m)$, не превышает

$$\left\lceil \frac{p}{m} \right\rceil - 1 \leqslant \frac{p+m-1}{m} - 1 = \frac{p-1}{m}.$$

Вероятность того, что s приводит к коллизии с r при приведении по модулю m, не превышает $\frac{p-1}{m} \cdot \frac{1}{p-1} = \frac{1}{m}$. Значит, $\forall k \neq l \in \mathbb{Z}_p \ P(h_{a,b}(k) = h_{a,b}(l)) \leqslant \frac{1}{m}$, что означает, что множество хеш-функций $H_{p,m}$ является универсальным. □

Замечание 7. Докажем, что $a = ((r-s)\cdot(k-l)^{-1}) \ mod \ p$.

Доказательство.

$$r = (ak + b) \ mod \ p,$$
$$s = (al + b) \ mod \ p.$$
$$r - s = (a(k-l)) \ mod \ p,$$
$$a \equiv (r-s)\cdot(k-l)^{-1} \ (mod \ p)$$

Так как $a \in \mathbb{Z}_p$, то верно равенство

$$a = ((r-s)\cdot(k-l)^{-1}) \ mod \ p$$

□

Докажем, что $b = (r - ak) \ mod \ p$.

Доказательство. Достаточно вспомнить, что

$$\begin{cases} r = (ak + b) \ mod \ p, \\ b \in \mathbb{Z}_p \end{cases}$$

□

Билет 2

Свойства вероятностной меры. Формула включений и исключений. Парадокс распределения подарков.

Свойства вероятностной меры
Формула включений и исключений

Теорема 1. $\Omega = \{w_1, \ldots, w_n\}$ — *множество всех элементарных исходов. Функция* $P\colon 2^\Omega \to [0,1]$ — *вероятностная мера. Свойства вероятностной меры:*

(1) $P(\varnothing) = 0$.
(2) *Если* $A \subseteq B$, *то* $P(A) \leqslant P(B)$.
(3) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
(4) $P(A_1 \cup \ldots \cup A_k) \leqslant P(A_1) + \ldots + P(A_k)$
(5) *Формула включений и исключений*

$$P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - \ldots - P(A_{n-1} \cap A_n) + \ldots + (-1)^{n-1} P(A_1 \cap \ldots \cap A_n)$$

или

$$P(A_1 \cup \ldots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \sum_{i_1 < \ldots < i_k} P(A_{i_1} \cap \ldots \cap A_{i_k}),$$

что то же самое.

Доказательство.

(1) Следует из определения.
(2) $B = A \cup (B \setminus A)$. Тогда
$$P(B) = P(A) + \underbrace{P(B \setminus A)}_{\geqslant 0} \geqslant P(A).$$
(3) $P(A \cup B) = \underbrace{P(A \setminus B) + P(A \cap B)}_{P(A)} + \underbrace{P(B \setminus A) + P(A \cap B)}_{P(B)} - P(A \cap B) = P(A) + P(B) - P(A \cap B)$
(4) Докажем по индукции:
База: $n = 1 : P(A_1) \leqslant P(A_1)$ очевидно.
Пусть для n доказано. Докажем для $n+1$:
$$P((A_1 \cup A_2 \cup \ldots \cup A_n) \cup A_{n+1}) \leqslant P(A_1 \cup \ldots \cup A_n) + P(A_{n+1}) \overset{\text{(шаг индукции)}}{\leqslant} P(A_1) + \ldots + P(A_{n+1}).$$
(5) Докажем по индукции:
База: для $n = 1$ и $n = 2$ очевидно (смотри пункт 3).
Пусть для n уже доказано. Докажем для $n+1$:
$$P((A_1 \cup \ldots \cup A_n) \cup A_{n+1}) \overset{(5)}{=} P(A_1 \cup \ldots \cup A_n) + P(A_{n+1}) - P\Big(\underbrace{(A_1 \cup \ldots \cup A_n) \cap A_{n+1}}_{\underbrace{(A_1 \cap A_{n+1}) \cup \ldots \cup (A_n \cap A_{n+1})}_{\text{n штук}}}\Big)$$
$$\overset{\text{(шаг индукции)}}{=} P(A_1) + \ldots + P(A_n) - \sum_{1 \leqslant i < j \leqslant n} P(A_i \cap A_j) + \sum_{1 \leqslant i < j < k \leqslant n} P(A_i \cap A_j \cap A_k) -$$
$$- \ldots + P(A_{n+1}) - \left(P(A_1 \cap A_{n+1}) + \ldots + P(A_n \cap A_{n+1}) - \sum_{1 \leqslant i < j \leqslant n} P(A_i \cap A_j \cap A_{n+1}) + \ldots \right)$$
$$= P(A_1) + \ldots + P(A_{n+1}) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - \ldots - P(A_n \cap A_{n+1}) + \ldots + (-1)^n P(A_1 \cap \ldots \cap A_{n+1})$$

□

Замечание 1. В классическом определении вероятности все элементарные исходы равновероятны. Из ровно одного свойства (свойства (4)) следует определение вероятности произвольного события A, состоящего из k элементов, — $P(A) = \frac{k}{n}$.

Доказательство. Заметим, что из свойства (4) следует аналогичное свойство для k непересекающихся событий $A_1, \ldots, A_k$:

$$P(A_1 \sqcup \ldots \sqcup A_k) = P(A_1) + \ldots P(A_k).$$

Пусть $A = \{w_{i_1}, \ldots, w_{i_k}\}$. Тогда

$$P(A) = P(w_{i_1}) + \ldots + P(w_{i_k}) = \underbrace{\frac{1}{n} + \ldots + \frac{1}{n}}_{k} = \frac{k}{n}.$$

□

Парадокс распределения подарков

Задача 1. N человек принесли подарки друг для друга. Затем эти подарки сложили в мешок и каждый вынул себе из мешка подарок. Какова вероятность того, что конкретный человек вынул подарок, который он принес? Какова вероятность того, что никто не вытащил подарок, который сам принес?

Решение.
Пространство исходов Ω состоит из всех возможных перестановок чисел $1, 2, \ldots, N$, причем все перестановки являются равновозможными. Значит, вероятность конкретной перестановки равна $\frac{1}{N!}$. Событие, состоящее в том, что конкретный человек вытащил подарок, который сам принес, состоит из $(N-1)!$ исходов. Следовательно, вероятность такого события равна $\frac{(N-1)!}{N!} = \frac{1}{N}$. При больших N эта вероятность стремится к нулю.
Можно было бы думать, что вероятность события: ни один человек не вытащил подарок, который сам принес, стремится к единице, но это ошибочное мнение.
Пусть A_k - событие, состоящее в том, что k-й человек вытащил свой подарок. Тогда $A_1 \cup \ldots \cup A_N$ - событие, состоящее в том, что хотя бы один вытащил свой подарок. По формуле включений и исключений

$$P(A_1 \cup \ldots \cup A_N) = \sum_{1 \leqslant i \leqslant N} P(A_i) \; - \sum_{1 \leqslant i < j \leqslant N} P(A_i \cap A_j) + \ldots + (-1)^{N-1} P(A_1 \cap \ldots \cap A_N)$$

$$\begin{cases} P(A_i) = \frac{1}{N}, \\ P(A_i \cap A_j) = \frac{(N-2)!}{N!}, \\ \ldots \\ P(A_{i_1} \cap \ldots A_{i_k}) = \frac{(N-k)!}{N!}, \\ \ldots \\ P(A_1 \cap \ldots \cap A_N) = \frac{1}{N!}, \end{cases} \quad \Rightarrow P(A_1 \cup \ldots \cup A_N) = N \cdot \frac{1}{N} - C_N^2 \cdot \frac{(N-2)!}{N!} + \ldots + (-1)^{N-1} \frac{1}{N!}$$

$$P(A_1 \cup \ldots \cup A_N) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \ldots + \frac{(-1)^{N-1}}{N!}$$

Таким образом, вероятность того, что ни один человек не вытащил подарок, который сам принес, равна

$$1 - P(A_1 \cup \ldots \cup A_N) = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \ldots - \frac{(-1)^{N-1}}{N!}$$

и стремится к $\frac{1}{e}$ при $N \to +\infty$.

Замечание 2. Вспомним разложение e^x в ряд Тейлора:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots$$

Подставляя вместо x число -1, получаем искомую вероятность $1 - P(A_1 \cup \ldots \cup A_N)$.

Билет 3

Условная вероятность. Независимые события. Лемма Ловаса.

Условная вероятность

Пусть $P(B) > 0$.

Определение 1. *Условной вероятностью события A при условии B* называется число

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

Теорема 1. *Если фиксировать событие B, то функция $P(\cdot|B)$ является вероятностной мерой.*

Доказательство. Во-первых, заметим, что

$$P(A \cap B) \leqslant P(B) \Rightarrow P(A|B) \in [0, 1].$$

Проверим теперь выполнимость свойств из определения:

$$(1) P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1,$$

$$(2) P(A \sqcup C|B) = \frac{P((A \sqcup C) \cap B)}{P(B)} =$$

$$= \frac{P((A \cap B) \sqcup (C \cap B))}{P(B)} = \frac{P(A \cap B)}{P(B)} + \frac{P(C \cap B)}{P(B)} =$$

$$= P(A|B) + P(C|B).$$

□

Замечание 1. Равенство из определения часто переписывают в виде

$$P(A \cap B) = P(B)P(A|B)$$

и называют *правилом умножения*.

Независимые события

С точки зрения вычисления вероятностей независимость события A от события B означает, то вероятность A не зависит от того, произошло событие B или нет. Формализовать эту идею помогает условная вероятность. Тогда предположим, что $P(B) > 0$. Событие A не зависит от события B, если

$$P(A|B) = P(A),$$

что по определению условной вероятности можно переписать в следующем виде

$$P(A \cap B) = P(A) \cdot P(B).$$

Это равенство и принимают в качестве определения независимости.

Определение 2. События A и B *независимы*, если

$$P(A \cap B) = P(A) \cdot P(B).$$

Замечание 2. Отметим, что из независимости A и B следует независимость $\overline{A}$ и B.

Определение 3. События $A_1, \ldots, A_n$ называются *независимыми в совокупности*, если

$$P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) = P(A_{i_1}) \cdot P(A_{i_2}) \cdot \ldots \cdot P(A_{i_k})$$

для всякого $2 \leqslant k \leqslant n$ и всяких $1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n$.

Замечание 3. Отметим, что независимость в совокупности не совпадает с попарной независимостью. Это можно проиллюстрировать *парадоксом независимости*.

Задача (парадокс независимости)
Два раза бросаем правильную монету. Событие A - при первом бросании выпал герб. Событие B - при втором бросании выпал герб. Событие C - ровно на одном бросании выпал герб. Эти события попарно независимы, но не являются независимыми в совокупности.
Действительно, любые два однозначно определяют третье, в частности пересечение A и B исключает C, то есть $P(A \cap B \cap C) = 0 \neq P(A) \cdot P(B) \cdot P(C)$.

Лемма Ловаса

Лемма 1. *(Локальная лемма Ловаса в симметричной форме)*
Предположим, что для событий $A_1, \ldots, A_n$ существует натуральное число $d \in \{1, \ldots, n\}$ такое, что

(1) *для всякого A_k найдется набор из не меньше, чем $n - d$, множеств $A_{i_1}, A_{i_2}, \ldots$ с таким свойством:*
A_k независимо с пересечением любого набора из этих множеств.
(2) *$P(A_k) \geqslant 1 - \frac{1}{e(d+1)}$ для каждого k.*

Тогда $P\left(\bigcap\limits_k A_k\right) > 0 \Rightarrow \exists w \in \bigcap\limits_k A_k$.

Доказательство. Обозначим через B_k событие противоположное событию A_k. Заметим, что

$$P(A_1 \cap \ldots \cap A_n) = P(A_1|A_2 \cap \ldots \cap A_n) \cdot P(A_2|A_3 \cap \ldots \cap A_n) \cdot \ldots \cdot P(A_{n-2}|A_{n-1} \cap A_n) \cdot P(A_{n-1}|A_n) \cdot P(A_n) =$$

$$= (1 - P(B_1|A_2 \cap \ldots \cap A_n)) \cdot (1 - P(B_2|A_3 \cap \ldots \cap A_n)) \cdot \ldots \cdot (1 - P(B_{n-2}|A_{n-1} \cap A_n)) \cdot (1 - P(B_{n-1}|A_n)) \cdot P(A_n)$$

$P(A_n) > 0$. Это следует из второго условия в формулировке леммы Ловаса.
Если мы докажем, что каждая вероятность $P(B_i|A_{i+1} \cap \ldots \cap A_n)$ строго меньше 1, то тем самым утверждение будет доказано. Будем доказывать более сильное утверждение, что для всякого набора индексов $J \subset \{1, 2, \ldots, n\}$ и всякого $1 \leqslant i \leqslant n$ верна оценка:

$$P\left(B_i | \bigcap_{j \in J} A_j\right) \leqslant \frac{1}{d+1}.$$

Доказательство проведем индукцией по количеству элементов в J.
База: $|J| = 1, j \in J$,

$$P(B_i|A_j) = \frac{P(B_i \cap A_j)}{P(A_j)} \leqslant \frac{P(B_i)}{P(A_j)}$$

$$P(B_i) = 1 - P(A_i) \leqslant \frac{1}{e(d+1)}$$

А $P(A_j) \geqslant 1 - \frac{1}{e(d+1)}$. Тогда

$$P(B_i|A_j) \leqslant \frac{\frac{1}{e(d+1)}}{1 - \frac{1}{e(d+1)}} = \frac{1}{e(d+1) - 1} = \frac{1}{ed + e - 1} \leqslant \frac{1}{d+1}.$$

Шаг индукции:
Нам понадобится следующее утверждение:

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

Доказательство.

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

Делим и числитель, и знаменатель на P(C)

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

□

Пусть Z — пересечение зависимых с B_i множеств A_j, а N — пересечение независимых. Тогда

$$P\left(B_i|\bigcap_{j\in J}A_j\right)=\frac{P(B_i\cap Z|N)}{P(Z|N)}.$$

Числитель не превосходит $P(B_i)$, а $P(B_i)=1-P(A_i)\leqslant\frac{1}{e(d+1)}$.
Оценим знаменатель. Пусть в Z входят множества с индексами $j_1,\ldots,j_s$, то есть $Z=A_{j_1}\cap\ldots\cap A_{j_s}$. Тогда имеем:

$$P(A_{j_1}\cap\ldots\cap A_{j_s}|N)=P(A_{j_1}|A_{j_2}\cap\ldots\cap A_{j_s}\cap N)P(A_{j_2}\cap\ldots\cap A_{j_s}|N)=$$
$$=(1-P(B_{j_1}|A_{j_2}\cap\ldots\cap A_{j_s}\cap N))P(A_{j_2}\cap\ldots\cap A_{j_s}|N).$$

По предложению индукции $P(B_{j_1}|A_{j_2}\cap\ldots\cap A_{j_s}\cap N)\leqslant\frac{1}{d+1}$. Значит, верно следующее неравенство

$$P(A_{j_1}\cap\ldots\cap A_{j_s}|N)\geqslant\left(1-\frac{1}{d+1}\right)\cdot P(A_{j_2}\cap\ldots\cap A_{j_s}|N).$$

Повторяем рассуждения уже относительно $P(A_{j_2}\cap\ldots\cap A_{j_s}|N)$. И так продолжаем, пока не останется $P(A_{j_s}|N)$, так же оцениваем и ее, и, принимая во внимание оценку $s\leqslant d$ (она следует из первого условия формулировки леммы Ловаса), получаем для $P(A_{j_1}\cap\ldots\cap A_{j_s}|N)$ следующую оценку:

$$P(A_{j_1}\cap\ldots\cap A_{j_s}|N)\geqslant\left(1-\frac{1}{d+1}\right)^d.$$

Заметим, что

$$\left(1-\frac{1}{d+1}\right)^d=\left(1+\frac{1}{-d-1}\right)^d=[q=-d-1;-d=q+1]=\left(\left(1+\frac{1}{q}\right)^{q+1}\right)^{-1}$$

Тогда, так как последовательность $\left(\left(1+\frac{1}{q}\right)^{q+1}\right)^{-1}$ сходится к $\frac{1}{e}$ сверху, то

$$P(A_{j_1}\cap\ldots\cap A_{j_s}|N)\geqslant\frac{1}{e}.$$

Возвращаемся к вероятности, которую надо было оценить:

$$P\left(B_i|\bigcap_{j\in J}A_j\right)=\frac{P(B_i\cap Z|N)}{P(Z|N)}\leqslant\frac{1}{e(d+1)}\cdot\frac{1}{\frac{1}{e}}=\frac{1}{d+1}.$$

Шаг индукции доказан. Значит, верно утверждение

$$P\left(B_i|\bigcap_{j\in J}A_j\right)\leqslant\frac{1}{d+1}.$$

□

Пример 1. *(Пример применения леммы Ловаса в задаче)*
Задача.
В научном центре работают специалисты по различным разделам компьютерных наук (количество разделов неизвестно). Известно, что по каждому разделу в центре работает ровно 7 ученых, причем вполне может быть, что один ученый является специалистом сразу по нескольким направлениям, но не более чем по трем. Все ученые должны принять участие в одной (и только одной) из двух конференций, одна из которых проходит в Канаде, а другая — в Австралии. Докажите, что всегда можно распределить ученых по этим конференциям так, что на каждой конференции будут присутствовать специалисты по всем направлениям компьютерных наук.

Доказательство.
Будем для каждого ученого выбирать конференцию простым подбрасыванием правильной монеты. Для каждого направления a рассмотрим событие A, состоящее в том, что ученые направления a поехали на обе конференции. Поскольку каждый ученый является специалистом в не более, чем трех, направлениях, то событие A зависимо самое большее с 14 такими событиями для других направлений (для каждого из семи ученых осталось по 2 направления). Кроме того, $P(A)=\frac{2^7-2}{2^7}=1-2^{-6}$ (нас интересуют все последовательности из единиц и нулей (пусть 0 соответствует Канаде, а 1 — Австралии) длины 7 (так как 7 ученых), кроме двух: из семи нулей (все поехали в Канаду), из семи единиц (все поехали в Австралию)). Заметим, что $2^{-6}\leqslant\frac{1}{15e}$. Утверждение следует из леммы Ловаса с $d=14$. □

Билет 4

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

Формула полной вероятности

Теорема 1. *(Формула полной вероятности)*
Пусть $\Omega = A_1 \cup A_2 \cup \ldots \cup A_n$ *и* $A_i \cap A_j = \varnothing$ *для всех* $i \neq j$*). Тогда для всякого события* B *имеет место равенство*

$$P(B) = \sum_i P(B|A_i)P(A_i).$$

Доказательство.

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_n) =$$
$$\text{Переписываем каждую } P(B \cap A_i) \text{ как } P(B|A_i) \cdot P(A_i)$$
$$= P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + \ldots + P(B|A_n) \cdot P(A_n).$$

□

Формула Байеса

Теорема 2. *(Формула Байеса) Пусть* $P(A) > 0$ *и* $P(B) > 0$*. Тогда имеет место равенство*

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

Доказательство. Достаточно заметить, что

$$P(A|B)P(B) = P(A \cap B) = P(B|A)P(A).$$

□

Задача о сумасшедшей старушке

Задача 1. На посадку в самолет стоят $N \geqslant 2$ пассажиров, среди которых сумасшедшая старушка. Старушка расталкивает всех пассажиров и садится в самолет на произвольное место. Затем пассажиры, когда заходят в самолет, садятся на свое место, если оно свободно, и на произвольное свободное место в противном случае. Какова вероятность того, что последний пассажир сядет на свое место?

Решение.
Пусть эта вероятность равна P_N.
Докажем методом математической индукции, что верно следующее равенство:

$$P_N = \frac{1}{2}$$

База: Если $N = 2$, то $P_N = \frac{1}{2}$.
Шаг индукции:
Предположим, что уже для всех $k \leqslant N$ доказано, что $P_k = \frac{1}{2}$. Докажем равенство $P_{N+1} = \frac{1}{2}$:
Событие B состоит из тех исходов, когда последний пассажир садится на свое место. Событие A_m состоит из тех исходов, когда старушка села на место m-го пассажира. По формуле полной вероятности

$$P_{N+1} = P(B) = \sum_m P(B|A_m)P(A_m). \text{ (*)}$$

Заметим, что $P(A_m) = \frac{1}{N+1}$ и все, кроме двух (когда старушка села на свое место (в этом случае $P(B|A_m) = 1$) или на место последнего пассажира (в этом случае $P(B|A_m) = 0$)), вероятности $P(B|A_m) = \frac{1}{2}$ (когда старушка садится на место m-го пассажира, m-ый пассажир фактически превращается в сумасшедшую старушку и мы получаем задачу для N пассажиров, а по предположению индукции $P_N = \frac{1}{2}$). Таких $P(B|A_m)$, что $P(B|A_m) = \frac{1}{2}$, будет ровно $N - 1$, так как всего слагаемых в (*) $N + 1$. Следовательно, имеем

$$P_{N+1} = \frac{N-1}{2(N+1)} + \frac{1}{N+1} = \frac{1}{2}.$$

Шаг индукции доказан. Значит, верно утверждение

$$P_N = \frac{1}{2}.$$

Билет 5

Схема Бернулли. Теорема Муавра-Лапласа (формулировка, доказательство только локальной теоремы и только для симметричного случая)

Схема Бернулли

Проводится N опытов, в каждом из которых может произойти определенное событие ("успех") с вероятностью p или не произойти ("неуспех") с вероятностью $q = 1 - p$.
Например, рассмотрим следующий эксперимент: N раз бросается монета с вероятностью выпадения орла (успеха) p, причем результат одного бросания не влияет на результат других бросаний.
Нас интересует число успехов, то есть в примере это число выпадения орла.
Можно считать, что множество элементарных исходов состоит из последовательностей длины N из нулей и единиц, где 1 соответствует успеху. Каждому исходу w с k единицами сопоставляем вероятность $p^k q^{N-k}$.

Определение 1. Построенное вероятностное пространство называют *схемой Бернулли*.

Утверждение.
Вероятность того, что в исходе ровно k единиц равна $C_N^k p^k q^{N-k}$.

Доказательство. P(ровно k единиц) равна сумме вероятностей вида $p^k q^{N-k}$ по всем исходам, в которых ровно k единиц, а это значит, что

$$P(\text{ровно } k \text{ единиц}) = C_N^k p^k q^{N-k}.$$

□

Утверждение.

$$\sum_{k=0}^{N} C_N^k p^k q^{N-k} = 1$$

(Если это выполнено, то наше вероятностное пространство корректно определено)

Доказательство.

$$1 = 1^n = (p+q)^n \overset{\text{по биному Ньютона}}{=} \sum_{k=0}^{N} C_N^k p^k q^{N-k}.$$

□

Определение 2. Набор вероятностей $P_{N,0}, P_{N,1}, \ldots, P_{N,N}$, где $P_{N,k} = C_N^k p^k q^{N-k}$, называется *распределением Бернулли*.

Замечание 1. *(Формула Стирлинга)*

$$n! = \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n+\frac{\varepsilon_n}{12n}}, \text{ где } \varepsilon_n \in (0,1).$$

Теорема Муавра-Лапласа
(формулировка, доказательство только локальной теоремы и только для симметричного случая)

Определение 3. Функция

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

называется *функцией Гаусса*.

Теорема 1. *(Теорема Муавра-Лапласа)*
У нас есть схема Бернулли:

$$N \text{ подбрасываний}$$
$$k \text{ - число успехов}$$
$$p \text{ - вероятность успеха}$$
$$q = 1 - p$$

$$X_{N,k} = \frac{k - Np}{\sqrt{Npq}}$$

1 случай. (Локальная теорема Муавра-Лапласа)
Если k выбирается так, что

$$|X_{N,k}| \leqslant C, \text{ где } C \text{ не зависит от } N,$$

то $P_{N,k} = \frac{1}{\sqrt{Npq}}\varphi(X_{N,k}) \cdot (1 + O(\frac{1}{\sqrt{N}}))$.

2 случай. (Интегральная теорема Муавра-Лапласа)
Для любых чисел $a < b$ имеем

$$P\left(a \leqslant \frac{k - Np}{\sqrt{Npq}} \leqslant b\right) \xrightarrow[N\to\infty]{} \int_a^b \varphi(x)dx.$$

Здесь в левой части написана вероятность того, что число единиц k лежит в диапазоне от $Np + a\sqrt{Npq}$ до $Np + b\sqrt{Npq}$.
Отметим, что во втором случае разница между вероятностью и интегралом оценивается через $\frac{p^2+q^2}{\sqrt{Npq}}$ и эта оценка точна. Следовательно, если p близко к нулю или к единице, то вероятность плохо приближается интегралом от φ.

Доказательство. (Доказательство локальной теоремы только для симметричного случая)
Пусть $N = 2n$. Найдем вероятность того, что в последовательности длины N ровно n единиц (ровно половина):

$$P_{2n,n} = C_{2n}^n \cdot \frac{1}{2^{2n}} = \frac{(2n)!}{(n!)^2 \cdot 2^{2n}}$$

Раскрываем по формуле Стирлинга и получаем следующее:

$$\frac{\sqrt{2\pi} \cdot (2n)^{2n+\frac{1}{2}} \cdot e^{-2n+O(\frac{1}{2n})}}{(\sqrt{2\pi})^2 \cdot n^{2n+1} \cdot e^{-2n+O(\frac{1}{n})} \cdot 2^{2n}} = \frac{1}{\sqrt{\pi n}} e^{O(\frac{1}{n})} = \frac{1}{\sqrt{\pi n}}\left(1 + O\left(\frac{1}{n}\right)\right)$$

Таким образом, $P_{2n,n} \sim \frac{1}{\sqrt{\pi n}} = \frac{1}{\sqrt{2n \cdot \frac{1}{4}}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(n-n)^2}{(2n \cdot \frac{1}{4})}}$ (что и требовалось).

Найдем теперь $P_{2n,k}$.
Заметим, что

$$P_{2n,n+a} = P_{2n,n-a} \text{ (следует из равенства } C_N^k = C_N^{N-k}).$$

Тогда будет достаточно найти одну из этих вероятностей. Найдем $P_{2n,n+a}$.

$$\frac{P_{2n,n+a}}{P_{2n,n}} = \frac{C_{2n}^{n+a} \cdot (\frac{1}{2})^{2n}}{C_{2n}^n \cdot (\frac{1}{2})^{2n}} = \frac{C_{2n}^{n+a}}{C_{2n}^n} = \frac{n! \cdot n!}{(n+a)!(n-a)!} =$$
$$= \frac{(n-a+1) \cdot (n-a+2) \cdot \ldots \cdot (n-1) \cdot n}{(n+1) \cdot (n+2) \cdot \ldots \cdot (n+a)} =$$
$$= \frac{(1 - \frac{a-1}{n}) \cdot (1 - \frac{a-2}{n}) \cdot \ldots \cdot (1 - \frac{1}{n}) \cdot 1}{(1 + \frac{1}{n}) \cdot (1 + \frac{2}{n}) \cdot \ldots \cdot (1 + \frac{a}{n})} =$$
$$= e^{ln(1-\frac{a-1}{n}) + ln(1-\frac{a-2}{n}) + \ldots + ln(1-\frac{1}{n}) - ln(1+\frac{1}{n}) - \ldots - ln(1+\frac{a}{n})}$$

Вспомним, что $ln(1 + x) = x + O(x^2)$ при $-1 < x \leqslant 1$. Тогда найдем оценку степени, которую мы нашли:

$$ln\left(1 - \frac{a-1}{n}\right) + ln\left(1 - \frac{a-2}{n}\right) + \ldots + ln\left(1 - \frac{1}{n}\right) - ln\left(1 + \frac{1}{n}\right) - \ldots - ln\left(1 + \frac{a}{n}\right) =$$
$$= \left(-\frac{a-1}{n} - \frac{a-2}{n} - \ldots - \frac{1}{n}\right) + \left(-\frac{1}{n} - \frac{2}{n} - \ldots - \frac{a-1}{n} - \frac{a}{n}\right) + O\left(\frac{a^3}{n^2}\right) =$$
$$= -\frac{2(1 + 2 + \ldots + (a-1)) + a}{n} + O\left(\frac{a^3}{n^2}\right) = -\frac{a(a-1) + a}{n} + O\left(\frac{a^3}{n^2}\right) = -\frac{a^2}{n} + O\left(\frac{a^3}{n^2}\right)$$

Итак,

$$P_{2n,n+a} = P_{2n,n} \cdot e^{-\frac{a^2}{n} + O\left(\frac{a^3}{n^2}\right)} = \frac{1}{\sqrt{\pi n}} \cdot e^{-\frac{a^2}{n} + O\left(\frac{a^3}{n^2}\right)}$$

Пусть $|a| \leqslant c \cdot \sqrt{n}$.

$$P_{2n,n+a} = \frac{1}{\sqrt{\pi n}} e^{\frac{-a^2}{n}}\left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right)$$

□

Билет 6

Теорема Пуассона. Распределение Пуассона. Задача о булочках с изюмом. Пуассоновский процесс.

Теорема Пуассона и распределение Пуассона

Мы рассматриваем серии событий, причем N-я серия состоит из N событий и все N событий независимы в совокупности. Пусть в N-й серии вероятность события равна p_N, причем число $N \cdot p_N = \lambda$ не зависит от N. Нам интересна вероятность $P(A_{k,N})$ наступления ровно k событий в данной серии из N событий. Данная ситуация представляет собой схему Бернулли, следовательно, ее вероятность можно вычислить по формуле $C_N^k p_N^k (1-p_N)^{N-k}$.

Теорема 1. *(Пуассон) Пусть $N \cdot p_N = \lambda$ - не зависит от N. Тогда*

$$P(A_{k,N}) = C_N^k p_N^k (1-p_N)^{N-k} \to \frac{\lambda^k}{k!} e^{-\lambda}, N \to +\infty$$

Доказательство. Учитывая, что $N \cdot p_N = \lambda$, перепишем вероятность $P(A_{k,N})$ в следующем виде:

$$\frac{N(N-1)\dots(N-k+1)}{k!}\left(\frac{\lambda}{N}\right)^k\left(1-\frac{\lambda}{N}\right)^{N-k} = \left[\frac{1\cdot\left(1-\frac{1}{N}\right)\cdot\ldots\cdot\left(1-\frac{k-1}{N}\right)}{\left(1-\frac{\lambda}{N}\right)^k}\right]\cdot\frac{\lambda^k}{k!}\cdot\left(1-\frac{\lambda}{N}\right)^N$$

Так как λ и k не зависят от N, то при $N \to +\infty$ данное выражение стремится к $\frac{\lambda^k}{k!}e^{-\lambda}$. □

Определение 1. При этом набор вероятностей $\left\{\frac{\lambda^k}{k!}e^{-\lambda}\right\}$ называется *распределением Пуассона*.

Задача о булочках с изюмом

Задача 1. Рассмотрим серийное производство булочек с изюмом. Сколько изюма в среднем должны содержать булочки, чтобы вероятность наличия хотя бы одной изюминки в случайно выбранной булочке была не меньше 0,99?

Предположим, что уже изготовлено тесто на некоторое количество булочек. В это тесто добавлено N изюминок так, что отношение числа изюминок к количеству булочек равно λ. Следовательно, количество булочек равно N/λ. Выделим в тесте кусок, из которого будет изготовлена булочка α. Вероятность попадания изюминки β в булочку α равна λ/N. Следовательно, вероятность того, что в булку не попало изюма, равна

$$\left(1-\frac{\lambda}{N}\right)^N$$

А вероятность получить хотя бы одну изюминку в булочке

$$1-\left(1-\frac{\lambda}{N}\right)^N$$

Так как мы рассматриваем серийное производство булочек, то можно предполагать, что $N \to \infty$, т. е. растет объем теста и количество изюма, но не меняется плотность λ. При $N \to +\infty$ получаем $\left(1-\frac{\lambda}{N}\right)^N \to e^{-\lambda}$. Для решения задачи надо найти такое λ, что $e^{-\lambda} < 0,01$. $\lambda = 5$ подходит ($e^{-5} \approx 0,007, e^{-4} \approx 0,02$), то есть ответ 5.

Пуассоновский процесс

В случайные моменты времени регистрируются некоторые события. Будем отмечать эти моменты времени точками на луче $[0,+\infty)$. Обозначим через $X(t)$ число точек на временном промежутке длины t, а через $P_k(t)$ вероятность того, что $X(t)=k$. Будем предполагать, что:

(a) вероятность попадания k точек в данный промежуток зависит только от длины этого промежутка (не зависит от расположения)
(b) для любой конечной системы промежутков, которые могут попарно пересекаться только по концам, попадания точек в каждый из них являются независимыми в совокупности событиями
(c) вероятность попадания по крайней мере двух точек в интервал длины δ является $\bar{o}(\delta)$

Определение 2. Если эти условия выполняются, то $X(t)$ называют Пуассоновским процессом.

Сначала рассмотрим $P_0(t)$. Разделим промежуток $[0,t]$ на N промежутков. По свойству (a) вероятность отсутствия событий на каждом промежутке разбиения равна $P_0(t/N)$. По свойству (b) регистрация события на каждом промежутке разбиения не зависит от регистрации событий на других промежутках, следовательно, мы находимся в ситуации схемы Бернулли и вероятность отсутствия событий на $[0,t]$ равна $P_0(t)=P_0(t/N)^N$ (*). Положим $P_0(1)=q$. Тогда, подставив в равенство (*) $t=1$, получим $P_0(\frac{1}{N})=q^{\frac{1}{N}}$. Подставив $t=m$ в (*), получим $(P_0(\frac{m}{N}))^N=P_0(m)$, подставив $t=m, N=m$ в (*), получим $P_0(m)=(P_0(1))^m$, следовательно $P_0(\frac{m}{N})=(P_0(1))^{\frac{m}{N}}=q^{\frac{m}{N}}$. Заметим, что $P_0(t+h)=P_0(t)P_0(h)\leqslant P_0(t)$, т. е. $P_0(t)$ не возрастает. Следовательно, для $\frac{m-1}{N}\leqslant t\leqslant\frac{m}{N}$ выполняются неравенства $q^{\frac{m-1}{N}}\geqslant P_0(t)\geqslant q^{\frac{m}{N}}$. Приближая t последовательностью дробей $\{\frac{m}{N}\}$, приходим к равенству $P_0(t)=q^t$. Положив $q=e^{-\lambda}$, получим $P_0(t)=e^{-\lambda t}$.

Теперь вычислим $P_k(t)$ при $k>0$. Разобьем $[0,t]$ на N промежутков. Пусть B - событие, состоящее в том, что хотя бы на одном из промежутков зарегистрированы по крайней мере два события. По свойству (c) $P(B)\leqslant N\cdot\bar{o}(t/N)=\bar{o}(t)\to 0$ при $N\to+\infty$. Теперь рассмотрим событие A_k, состоящее в том, что на промежутке $[0,t]$ зарегистрировано ровно k событий, при условии, что на каждом промежутке разбиения зарегистрировано не более одного события. Вероятность отсутствия события на одном промежутке разбиения равна $P_0(t/N)=e^{-\lambda t/N}$, мы опять находимся в ситуации схемы Бернулли. Следовательно, имеем

$$P(A_k)=C_N^k\left(e^{-\lambda t/N}\right)^{N-k}\left(1-e^{-\lambda t/N}\right)^k=$$

$$=\frac{e^{-\lambda t}}{e^{\lambda tk/N}}\cdot\frac{N\dots(N-k+1)}{k!}\left(1-e^{-\lambda t/N}\right)^k\sim\frac{e^{-\lambda t}}{k!}\cdot\frac{N\dots(N-k+1)}{e^{\lambda tk/N}}\left(\frac{\lambda t}{N}\right)^k\to\frac{(\lambda t)^k}{k!}e^{-\lambda t}\qquad(N\to+\infty)$$

С учетом вышесказанного про $P(B)$ имеем

$$P_k(t)=\frac{(\lambda t)^k}{k!}e^{-\lambda t},$$

т. е. получаем распределение Пуассона. Число λ называется интенсивностью или параметром процесса $X(t)$.

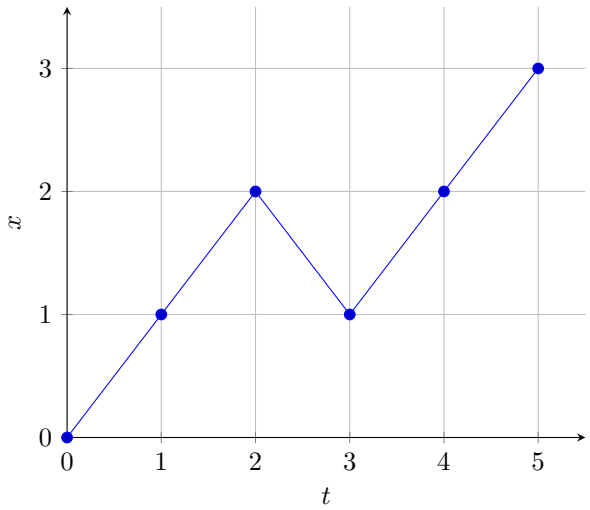
Замечание 1. $P(A_k)$ является условной вероятностью (на промежутке $[0,t]$ зарегистрировано ровно k событий, при условии, что на каждом промежутке разбиения зарегистрировано не более одного события), но $1-P(B)$ является вероятностью этого самого условия, но в силу того, что $P(B)\sim 0$, $P(A_k)=P_k(t)$.

Билет 7

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

Случайное блуждание

Схема Бернулли имеет геометрическую интерпретацию.
По числовой прямой двигается частица, которая каждую секунду перемещается на единицу вправо или на единицу влево, причем выбор обоих направлений равновозможен и не зависит от соответствующего выбора на других шагах. Мы считаем, что в начальный момент времени частица находится в точке $x = 0$. Ясно, что траекторию движения частицы за N перемещений можно закодировать последовательностью из 1 или -1 длины N. Набор таких последовательностей - пространство элементарных исходов. Вероятность каждой траектории равна $\frac{1}{2^N}$. Таким образом, с точностью до обозначений мы получим схему Бернулли, описывающую бросание правильной монеты.
Такая интерпретация называется *случайным блужданием*.
При исследовании случайного блуждания нас будет интересовать вероятность того, что траектория движения частицы обладает некоторым свойством. Для этого будет полезно следующее:
Траектории частицы будем изображать на координатной плоскости переменных (t, x) в виде ломанных, соединяющих точки с целочисленными координатами t и x. Здесь x — положение частицы, а t — время.
Например:



Замечание 1. Количество путей из (t_0, x_0) в (t_1, x_1) вычисляется по формуле

$$C_{t_1-t_0}^{\frac{t_1-t_0+x_1-x_0}{2}}.$$

(Чтобы понять, как возникла такая формула, достаточно представить траекторию именно как последовательность из нулей и единиц длины $t_1 - t_0$.)

Принцип отражения

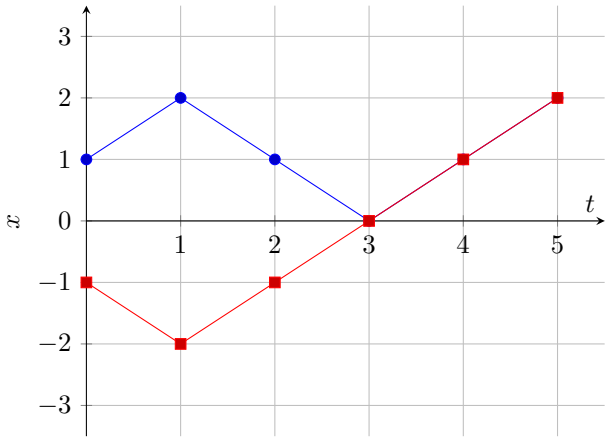
Предложение 1. *(Принцип отражения)*
Пусть $x_0 > 0,\ x_1 > 0$ и $t_0 < t_1$. Число путей из (t_0, x_0) в (t_1, x_1), которые касаются или пересекают ось времени, равно числу путей из $(t_0, -x_0)$ в (t_1, x_1).

Доказательство. Установим биективное соответствие между этими путями. Возьмем путь из (t_0, x_0) в (t_1, x_1), который касается или пересекает ось времени. Пусть t^* — первый момент времени, когда $x = 0$. Отразим часть пути, соответствующую отрезку времени $[t_0, t^*]$, относительно оси времени, а оставшуюся часть оставим без изменений. Получаем путь, соединяющий точки $(t_0, -x_0)$ и (t_1, x_1). □

Замечание 2. Количество таких путей вычисляется по формуле

$$C_{t_1-t_0}^{\frac{t_1-t_0+x_1+x_0}{2}}.$$

Пример:



Задача о баллотировке

Какова вероятность того, что частица, которая вышла из нуля и пришла в точку $k > 0$ за N шагов, все время находилась в точках с положительными координатами? Рассматриваемая задача имеет интересную интерпретацию и называется «теоремой о баллотировке».

Задача 1. Если на выборах один кандидат набрал q голосов, а другой — r голосов и $r > q$, то какова вероятность того, что победивший кандидат все время был впереди? Предполагается, что голосовавшие не имели предпочтений и отдавали свой голос случайно, в подсчет голосов происходил последовательно.

Решение.
Заметим, что первым шагом частица обязана подняться наверх. Начиная со второго шага, необходимая траектория частицы исходит из точки $(1, 1)$ и идет в точку (N, k), не касаясь и не пересекая ось времени. По принципу отражения мы умеем считать число остальных траекторий, соединяющих $(1, 1)$ и (N, k). Таких траекторий ровно столько же, сколько всего траекторий из $(1, -1)$ в (N, k), а их количество равно $C_{N-1}^{\frac{N+k}{2}}$. Всего траекторий из $(1, 1)$ в (N, k) равно $C_{N-1}^{\frac{N+k}{2}-1}$. Следовательно, число нужных траекторий частицы равно

$$C_{N-1}^{\frac{N+k}{2}-1} - C_{N-1}^{\frac{N+k}{2}} = \frac{k}{N} C_N^{\frac{N+k}{2}}.$$

Здесь $C_N^{\frac{N+k}{2}}$ — количество путей, соединяющих начало координат и точку (N, k). Значит, искомая вероятность равна $\frac{k}{N}$. В условиях задачи о баллотировке соответствующая вероятность равна $\frac{r-q}{r+q}$.

Задача о возвращении в начало координат

Пусть частица вышла из начала координат. Обозначим через u_{2n} вероятность того, что в момент времени $t = 2n$ частица вернулась в точку $x = 0$, а через f_{2n} вероятность того, что это произошло в первый раз.
Ясно, что $u_{2n} = C_{2n}^n 2^{-2n}$ (в таких траекториях частица n раз поднимается и столько же спускается, тогда достаточно найти количество исходов, в которых ровно n раз частица выбирает подъем, и разделить на количество всех возможных исходов).
Найдем теперь f_{2n}. Частица приходит в точку x_0 в момент времени $2n$ из точек $x = 1$ или $x = -1$ (в момент времени $t = 2n - 1$). Число путей в точку $(2n - 1, 1)$ из начала координат таких, что все координаты точек, через которые проходит путь, положительные, равно $\frac{1}{2n-1} C_{2n-1}^n$ (смотри задачу о баллотировке). Столько же путей в точку $(2n-1, -1)$ из начала координат таких, что все координаты точек, через которые проходит путь, отрицательные (так как этот случай фактически симметричен предыдущему). Следовательно, всего нужных нам путей $\frac{2}{2n-1} C_{2n-1}^n$ и

$$f_{2n} = \frac{2}{2n-1} C_{2n-1}^n \cdot 2^{-2n} = \frac{1}{2n} u_{2n-2}.$$

Формула Стирлинга позволяет найти асимптотику таких вероятностей:

$$f_{2n} = \frac{1}{2n} \cdot u_{2n-2} \sim \frac{1}{n^{\frac{3}{2}}} \cdot \frac{1}{2\sqrt{\pi}}.$$

Билет 8

Геометрическая вероятность. Бесконечное бросание правильной монеты. Игра «Penney Ante». Парадокс Бертрана.

Геометрическая вероятность

Чтобы преодолеть недостаток классического определения вероятности, состоящий в том, что оно неприменимо к бесконечным множествам элементарных исходов, вводят *геометрические вероятности* — вероятности попадания точки в область (отрезок, часть плоскости и т.д.).
Пусть отрезок l составляет часть отрезка L. На отрезок L наудачу поставлена точка. Это означает выполнение следующих предположений:

- поставленная точка может оказаться в любой точке отрезка L,
- вероятность попадания точки на отрезок l пропорциональна длине этого отрезка и не зависит от его расположения относительно отрезка L.

В этих предположениях вероятность попадания точки на отрезок l определяется равенством

$$P = \frac{\text{Длина } l}{\text{Длина } L}.$$

Пусть плоская фигура g составляет часть плоской фигуры G. На фигуру G наудачу брошена точка. Это означает выполнение следующих предположений:

- брошенная точка может оказаться в любой точке фигуры G,
- вероятность попадания брошенной точки на фигуру g пропорциональна площади этой фигуры и не зависит ни от ее расположения относительно G, ни от формы g.

В этих предположениях вероятность попадания точки на фигуру g определяется равенством

$$P = \frac{\text{Площадь } g}{\text{Площадь } G}.$$

Замечание 1. Приведенные определения являются частными случаями общего определения геометрической вероятности. Если обозначить меру (длину, площадь, объем) области через mes, то вероятность попадания точки, брошенной наудачу (в указанном выше смысле) в область g — часть области G, равна

$$P = \frac{mes\ g}{mes\ G}.$$

Замечание 2. В случае классического определения вероятность достоверного (невозможного) события равна единице (нулю); справедливы и обратные утверждения (например, если вероятность события равна нулю, то событие невозможно). В случае геометрического определения вероятности обратные утверждения не имеют места. Например, вероятность попадания брошенной точки в одну определенную точку области G равна нулю, однако это событие может произойти, и, следовательно, не является невозможным.

Бесконечные подбрасывания монеты

Построим вероятностное пространство, моделирующее бесконечное подбрасывание монеты. Множеством элементарных исходов Ω в данном случае будет множество бесконечных двоичных последовательностей, и таким образом, Ω будет континуальным множеством. Каждая такая последовательность задает некоторую точку из отрезка $[0,1]$ при помощи двоичной записи. То есть определена функция $f : \Omega \to [0,1]$, сопоставляющая каждой последовательности $\omega = \{\omega_n\}$ число $f(\omega) = \sum\limits_{n=1}^{\infty} \frac{\omega_n}{2^n}$. Вероятности для такого вероятностного пространства не могут быть определены для каждой последовательности, и поэтому определены только для событий $A_n = \{\omega : \omega_1 = \varepsilon_1, \ldots, \omega_n = \varepsilon_n\}$ и $P(A_n) = 2^{-n}$, где $\varepsilon_i \in \{0,1\}$. Такое определение вероятностного пространства позволяет нам рассматривать события, касающиеся только первых n бросков, при этом вероятности таких событий вычисляются исходя из схемы Бернулли для n бросаний и от дальнейших бросков эти вероятности не зависят.

Игра «Penney Ante»

Задача 1. Алиса и Боб играют в следующую игру:
Правильная монета бросается до тех пор, пока не встретится комбинация 110 или 100 (где, например, 1 соответствует решке, а 0 — орлу).
Алиса выигрывает, если первой появилась комбинация 110, а Боб в случае, когда первой появилась комбинация 100. Кто будет выигрывать чаще?

Решение.
Будем считать, что монету продолжают подбрасывать и в случае, когда одна из комбинаций 110 или 100 уже выпала, полагая, что результат дальнейших подбрасываний не имеет значения. Таким образом, эксперимент выглядит так: случайно выбираем бесконечную последовательность из нулей и единиц, а затем смотрим, какая из последовательностей 110 или 100 появилась раньше.
Обозначим за R_N событие, которое заключается в том, что на N-ом броске выигрывает Алиса (то есть на N-ом броске в первый раз встретилась комбинация 110 и ни разу до этого не встретилась комбинация 100). Ясно, что $P(R_1) = P(R_2) = 0$, а $P(R_3) = \frac{1}{8}$.
Обозначим за Q_N вероятность того, что за первые N бросков Алиса не выигрывает (то есть комбинация 110 так и не появилась после первых N бросков). Тогда

$$P(Q_N) = 1 - P(R_1) - P(R_2) - \ldots - P(R_N)$$

и

$$P(R_{N+3}) = P(R_{N+3} \cap Q_N) = P(R_{N+3}|Q_N) \cdot P(Q_N) = \frac{1}{8}P(Q_N).$$

Тогда

$$P(R_{N+3}) = \frac{1}{8}(1 - P(R_1) - P(R_2) - \ldots - P(R_N))$$

Мы получили рекуррентную формулу для вычисления $P(R_{N+3})$, и теперь можем вычислить, например, $P(R_4) = \frac{1}{8} = P(R_5)$ и $P(R_6) = \frac{7}{64}$.
Для того чтобы вычислить то же самое для выигрыша Боба можно провести аналогичные рассуждения и получить похожую формулу. И может показаться, что обе комбинации, 110 и 100, равноправны, однако эта игра все же не является справедливой и одна из комбинаций выигрывает у другой. Докажем это:
Обозначим через A_N событие, представляющее собой выигрыш Алисы на N-ом шаге, через B_N — событие, представляющее собой выигрыш Боба на N-ом шаге, а C_N — событие, при котором ни Алиса, ни Боб не выиграли на N-ом шаге. Тогда:

$$P(A_{N+3}) = P(A_{N+3} \cap C_N) = P(A_{N+3}|C_N) \cdot P(C_N) = \frac{1}{8}P(C_N).$$

Пусть D_N — событие, состоящее в том, что после N-го бросания никто не выиграл, а потом при следующих трех бросаниях выпадает 100. Тогда

$$P(D_N) = \frac{1}{8}P(C_N).$$

С другой стороны, событие D_N заключается в том, что либо на $N+2$ шаге выиграла Алиса, либо на $N+3$ шаге выиграл Боб, то есть

$$P(D_N) = \frac{1}{2}P(A_{N+2}) + P(B_{N+3}).$$

Вспомним, что

$$P(A_{N+3}) = \frac{1}{8}P(C_N).$$

Тогда имеем:

$$\begin{cases} P(A_{N+3}) = \frac{1}{8}P(C_N), \\ P(D_N) = \frac{1}{8}P(C_N), \\ P(D_N) = \frac{1}{2}P(A_{N+2}) + P(B_{N+3}) \end{cases} \Rightarrow P(A_{N+3}) = \frac{1}{2}P(A_{N+2}) + P(B_{N+3}).$$

Пусть A — событие, состоящее в том, что Алиса выиграла. Тогда $A = \bigsqcup\limits_{N=1}^{\infty} A_N$. И пусть B — событие, состоящее в том, что Боб выиграл. Тогда $B = \bigsqcup\limits_{N=1}^{\infty} B_N$.

$$P(A) = \sum_{N=1}^{\infty} P(A_N) = \sum_{N=1}^{\infty} P(A_{N+3}) + P(A_3) = \sum_{N=1}^{\infty} \left(\frac{1}{2}P(A_{N+2}) + P(B_{N+3})\right) + \frac{1}{8} =$$

$$= \frac{1}{2}\sum_{N=1}^{\infty} P(A_{N+2}) + \sum_{N=1}^{\infty} P(B_{N+3}) + \frac{1}{8} = \frac{1}{2}\sum_{N=1}^{\infty} P(A_{N+2}) + \sum_{N=1}^{\infty} P(B_N) - \frac{1}{8} + \frac{1}{8} =$$

$$= \frac{1}{2}P(A) + P(B).$$

Тогда

$$P(A) = 2P(B).$$

Уже, не вычисляя сами вероятности, мы можем сделать вывод, что Алиса будет выигрывать вдвое чаще Боба.

Заметим, что $P(C) = 0$, где C — событие, состоящее в том, что никто не выиграл. Почему так? В это множество входит множество всех таких последовательностей, что если их разбить на подряд идущие тройки цифр, то среди этих троек нет 100. Вероятность того, что если первые $3N$ цифр разбить на N троек, то среди нет 100, равна $\left(\frac{7}{8}\right)^N$ и стремится к нулю при $N \to \infty$. Итак,

$$P(A) + P(B) = 1 \Rightarrow P(A) = \frac{2}{3},\ P(B) = \frac{1}{3}.$$

Парадокс Бертрана

Задача 2. В круге единичного радиуса проводят случайным образом хорду. Какова вероятность того, что эта хорда длиннее стороны правильного вписанного треугольника?

Решение.

Задавая различными способами вероятностное пространство в этой задаче можно получить совершенно различные, но абсолютно верные ответы.

Первый способ:

Фиксируем одну точку на окружности, а вторую выбираем наудачу и проводим через них хорду. Чтобы посчитать искомую вероятность, представим, что треугольник повернут так, что одна из его вершин совпадает с фиксированным концом хорды. Что в данном случае означает, что хорда длиннее стороны правильного вписанного треугольника? Это значит, что вторая точка лежит на дуге между двумя другими вершинами треугольника. Длина такой дуги равна одной трети длины окружности, а значит искомая вероятность равна $\frac{1}{3}$.

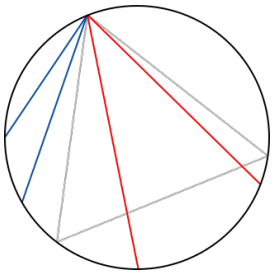


Рис. 1. Первый способ

Красным цветом на данном рисунке изображены те хорды, длина которых больше длины правильного вписанного треугольника, а синим - те, длина которых короче.

Замечание 3. *(Комментарий к первому способу)*

Почему длина дуги, стягиваемой двумя вершинами правильного вписанного треугольника, равна одной трети длины окружности?

Вспомним, что нам дан правильный треугольник. У него все углы равны, а значит и равны дуги, на которые эти углы опираются. Но треугольник делит окружность только на три части, а значит одна такая дуга имеет длину, равную одной трети длины окружности.

Замечание 4. Определим вероятностное пространство:

$$\Omega = [0, 2\pi)$$

$$\mathfrak{A} = \mathfrak{B}([0, 2\pi))$$

Разъединим окружность в фиксированной точке и представим ее как отрезок от 0 до 2π, не включая 2π. Тогда пусть L — длина отрезка, в который должна попасть выбираемая случайным образом точка.

$$P = \frac{\mathrm{L}}{2\pi}$$

Второй способ:
Фиксируем радиус окружности и наудачу выбираем точку на этом радиусе. Построим хорду, перпендикулярную зафиксированному радиусу, проходящую через выбранную точку. Для нахождения искомой вероятности, представим, что треугольник повернут так, что одна из его сторон перпендикулярна зафиксированному радиусу. Хорда длиннее стороны треугольника, если ее центр ближе к центру окружности, чем точка пересечения стороны треугольника с зафиксированным радиусом. Сторона треугольника делит радиус пополам, а значит вероятность выбрать хорду, длиннее стороны треугольника, равна $\frac{1}{2}$.

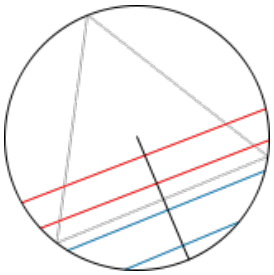


Рис. 2. Второй способ
Красным цветом на данном рисунке изображены те хорды, длина которых больше длины правильного вписанного треугольника, а синим - те, длина которых короче.

Замечание 5. *(Комментарий ко второму способу)*
Почему сторона правильного вписанного треугольника, перпендикулярная зафиксированному радиусу, делит этот радиус пополам?
В этом треугольнике перпендикуляр h (он же биссектриса и медиана, так как треугольник правильный), проведенный к стороне, которая перпендикулярна фиксированному радиусу, делится центром описанной окружности в соотношении $2:1$, считая от вершины, из которой он проведен. Получается, что радиус описанной окружности равен $\frac{2}{3}\cdot h$, а $h - R = R\cdot\frac{1}{2}$, из чего и следует утверждение из вопроса.

Замечание 6. Определим вероятностное пространство:

$$\Omega = [0,1]$$

(длина радиуса равна единице)

$$\mathfrak{A} = \mathfrak{B}([0,1])$$

Пусть L — длина отрезка на радиусе, в который должна попасть середина хорды.

$$P = L$$

(определяется как длина подоотрезка отрезка длины единица)

Третий способ:
Выбираем наудачу произвольную точку внутри круга и строим хорду с центром в выбранной точке. Хорда длиннее стороны равностороннего вписанного треугольника, если выбранная точка находится внутри круга, вписанного в этот треугольник. Площадь вписанного круга есть $\frac{1}{4}$ от площади описанного, а значит искомая вероятность равна $\frac{1}{4}$.

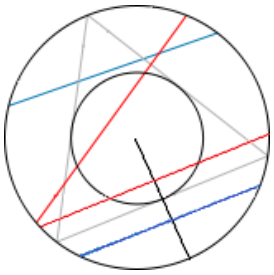


Рис. 3. Третий способ
Красным цветом на данном рисунке изображены те хорды, длина которых больше длины правильного вписанного треугольника, а синим - те, длина которых короче.

Замечание 7. *(Комментарий к третьему способу)*
1) Почему площадь вписанного круга в четыре раза меньше площади описанного?
Достаточно вспомнить, что площадь круга равна πR^2 и то, что радиус вписанной в правильный треугольник окружности вдвое меньше радиуса описанной.
2) Почему хорда будет больше стороны треугольника, только если ее центр попал внутрь вписанной окружности?
Вспомним, что радиус вписанной окружности равен половине радиуса описанной окружности и что точка касания вписанной окружности и стороны правильного треугольника делит эту сторону пополам. Зафиксируем нашу случайно проведенную хорду и повернем треугольник так, чтобы одна из его сторон была параллельна нашей зафиксированной хорде и при этом его середина лежала на том же радиусе описанной окружности, что и середина хорды. Тогда заметим, что теперь ситуация аналогична ситуации во втором способе решения данной задачи: сторона треугольника делит радиус описанной окружности пополам, при этом отрезок от центра описанной окружности до самой стороны является радиусом вписанной окружности, и хорда будет длиннее, если ее середина находится на радиусе вписанной окружности.

Замечание 8. Определим вероятностное пространство:

$$\Omega - \text{множество из всех точек круга}$$
$$\mathfrak{A} = \mathfrak{B}(\Omega)$$
$$P = \frac{\text{площадь области, в которую должна попасть случайно выбраннная точка}}{\pi\ (\text{площадь круга})}$$

Но парадокс Бертрана — это парадокс в кавычках. Никакого парадокса нет, просто это разные задачи. Они словесно звучат одинаково: "случайно проводится хорда". Но слово "случайно" можно расшифровать разными способами. Поэтому нет ничего удивительного в том, что мы получили разные ответы.

Билет 9

Вероятностное пространство: сигма алгебра событий и вероятностная мера. Борелевская сигма алгебра. Мера Лебега.

Вероятностное пространство: сигма алгебра событий и вероятностная мера

Определение 1. Пусть Ω — некоторое множество точек w. Система $\mathfrak{A}$ подмножеств Ω называется *алгеброй событий*, если

- $\Omega \in \mathfrak{A}$,
- $A, B \in \mathfrak{A} \Rightarrow A \cup B \in \mathfrak{A}, A \cap B \in \mathfrak{A}$,
- $A \in \mathfrak{A} \Rightarrow \overline{A} \in \mathfrak{A}$.

Пример 1.

$$\{\varnothing, \Omega, B, \Omega \setminus B\},\ B \subset \Omega.$$

А бывают "не алгебры"? Да, бывают. Например:

$$\{\Omega\}.$$

Определение 2. Если дополнительно верно, что

$$\{A_n\}_{n=1}^{\infty} \in \mathfrak{A} \Rightarrow \bigcup_n A_n, \bigcap_n A_n \in \mathfrak{A},$$

то $\mathfrak{A}$ называется *σ-алгеброй событий ("сигма-алгеброй событий")*.

Пример 2.
Пример алгебры, которая не является σ-алгеброй:
Пусть $\Omega = (0, 1]$, $\mathfrak{A}$ — множество, состоящее из конечных объединений попарно непересекающихся полуинтервалов $(a, b]$, $0 \leqslant a \leqslant b \leqslant 1$. Тогда $\mathfrak{A}$ является алгеброй, но не является σ-алгеброй. Проверим это:

(1) $\Omega \in \mathfrak{A}$
(2) Если $A, B \in \mathfrak{A}$, то и $\Omega \setminus A \in \mathfrak{A}$, и $A \cap B \in \mathfrak{A}$. Для объединения уже доказано из-за дополнений. Пусть

$A = A_1 \cup \ldots \cup A_n$ (то есть множество из конечного объединения попарно непересекающихся множеств),

$B = B_1 \cup \ldots \cup B_m$ (то есть множество из конечного объединения попарно непересекающихся множеств).

Тогда

$$A \cap B = (A_1 \cup \ldots \cup A_n) \cap (B_1 \cup \ldots \cup B_m) =$$

$$= (A_1 \cap (B_1 \cup \ldots \cup B_m)) \cup \ldots \text{ (объединение попарно непересекающихся множеств)} =$$

$$= (A_1 \cap B_1) \cup (A_1 \cup B_2) \cup \ldots \text{ (объединение попарно непересекающихся множеств)}$$

То есть $\mathfrak{A}$ действительно является алгеброй, но не является σ-алгеброй, так как, например, $\bigcap\limits_n \left(\frac{1}{2} - \frac{1}{n}, 1\right] = \left[\frac{1}{2}, 1\right] \notin \mathfrak{A}$.

Важнейшим примером σ-алгебры является борелевская σ-алгебра, которая рассмотрена в следующем пункте.

Определение 3. *Вероятностная мера P на σ-алгебре $\mathfrak{A}$ подмножества Ω* — это такая функция $P\colon \mathfrak{A} \to [0, 1]$, удовлетворяющая следующим свойствам:

(1) $P(\Omega) = 1$,
(2) $P\left(\bigcup\limits_n A_n\right) = \sum\limits_n P(A_n)$ при условии, что $\forall i, j\ A_i \cap A_j = \varnothing$.

Определение 4. Тройка $(\Omega, \mathfrak{A}, P)$ называется *вероятностным пространством*.

Борелевская сигма алгебра

Определение 5. Пусть S — набор подмножеств Ω. $\sigma(S)$ — *сигма алгебра, порожденная* S, то есть $\sigma(S)$ — наименьшая по вложению σ-алгебра, содержащая S.

Определение 6. *Борелевская σ-алгебра* $\mathfrak{B}(\mathbb{R})$ — σ-алгебра, порожденная всеми возможными полуинтервалами $\{(a,b]|a \in \mathbb{R}, b \in \mathbb{R}, a < b\}$.

Утверждение.
$\mathfrak{B}(\mathbb{R})$ — σ-алгебра, порожденная

(a) интервалами (a,b)
(b) лучами $(-\infty, c]$
...

Доказательство. Докажем только пункт a:

$$(a,b] = \bigcup_n \left(a, b + \frac{1}{n}\right)$$

$$(a,b) = \bigcup_n \left(a, b - \frac{1}{n}\right].$$

□

Мера Лебега

Определение 7. *Мерой Лебега* называется обычная длина, то есть такая $\lambda : \mathfrak{B}(\mathbb{R}) \to \mathbb{R}$, что

$$\lambda([a,b]) = b - a.$$

Как вычислить вероятность попадания случайной точки в некоторый подотрезок отрезка $[0,1]$ на прямой? Мы не можем приписать положительную вероятность каждому такому подотрезку, так как если мы каждой точке (подотрезку $[x_0, x_0]$) присвоим положительную вероятность, то, так как отрезок $[0,1]$ содержит бесконечное число различных точек, какую бы маленькую вероятность мы ни присвоили каждой точке, первое свойство вероятностной меры ($P(\Omega) = 1$) выполняться не будет.
Для таких событий (попадание точки x_0, случайно выбранной из отрезка $[0,1]$, в некоторый подотрезок $[a,b]$ отрезка $[0,1]$) определить вероятностную меру можно при помощи меры Лебега — каждому событию вида $\{x_0 \in [a,b], a \leqslant b\}$ сопоставляется вероятность $\lambda([a,b])$.
Зададим вероятностную меру для каждого события из $\mathfrak{B}([0,1])$ (можно обобщить на $\mathfrak{B}([a,b])$, где $a < b$) и проверим выполнимость свойств заданной нами вероятностной меры. Длина отрезка $[0,1]$ равна единице, поэтому первое свойство вероятностной меры выполнено. Вероятности для каждого подотрезка определены корректно (то есть не больше единицы и не меньше нуля). В случае объединения конечного числа непересекающихся подотрезков определим вероятностную меру как сумму вероятностей для каждого из подотрезков, входящих в сумму (нетрудно убедиться, что свойства по-прежнему будут выполняться). Аналогично определим вероятностную меру для объединения счетного числа непересекающихся отрезков. Утверждение, что при таком определении вероятностной меры свойства по-прежнему будут выполняться, оставим без доказательства.

Билет 10

Случайная величина и ее распределение. Функция распределения вероятностной меры и функция распределения случайной величины.

Случайная величина и ее распределение

Случайной называют величину, которая в результате испытания принимает одно и только одно возможное значение, наперед неизвестное и зависящее от случайных причин, которые заранее не могут быть учтены. Дадим формальное определение:

Определение 1. Функция $\xi\colon \Omega \to \mathbb{R}$ называется *случайной величиной*, если

$$\{w\colon \xi(w) \in \langle a, b\rangle\} \in \mathfrak{A}.$$

Замечание 1. $\langle a, b\rangle$ — промежуток от a до b, это может быть отрезок, интервал или полуинтервал.

Замечание 2. Напомним, что если есть некоторая функция $f : X \to Y$, то прообразом множества $S \subseteq Y$ называется множество $f^{-1}(S) = \{x \in X\colon f(x) \in S\}$.
Тогда определение случайной величины можно переписать следующим образом:
Функция $\xi\colon \Omega \to \mathbb{R}$ называется *случайной величиной*, если

$$\xi^{-1}(\langle a, b\rangle) \in \mathfrak{A}.$$

Замечание 3. *Дискретной (прерывной)* называют случайную величину, которая принимает отдельные, изолированные возможные значения с определенными вероятностями. При этом число возможных значений дискретной случайной величины может быть конечным или бесконечным.
Законом распределения дискретной случайной величины называют соответствие между возможными значениями и их вероятностями. Его можно задавать таблично, аналитически (в виде формулы) и графически. При табличном задании закона распределения дискретной случайной величины первая строка таблицы содержит возможные значения, а вторая — их вероятности. При этом важно понимать, что так как случайная величина в одном испытании может принимать одно и только одно возможное значение, то сумма вероятностей во второй строке таблицы должна быть равна единице.

Пример 1. (Пример дискретной случайной величины)
Бросание монетки:

$$\xi(w) = \begin{cases} 1, \text{ если } w \text{ — Орел} \\ 0, \text{ если } w \text{ — Решка} \end{cases}$$

При этом закон распределения этой дискретной случайной величины можно задать следующим образом (аналитически):

$$P(\xi = w) = \begin{cases} p, \text{ если } w \text{ — Орел} \\ 1-p, \text{ если } w \text{ — Решка} \end{cases}$$

где p — вероятность выпадания Орла.

Утверждение.
Если ξ — случайная величина, то $\forall B \in \mathfrak{B}(\mathbb{R})\ \xi^{-1}(B) \in \mathfrak{A}$ или (что есть то же самое) $\xi^{-1}(B)$ является событием.

Доказательство. Определим σ как $\{C\colon \xi^{-1}(C) \in \mathfrak{A}\}$ и докажем, что σ является σ-алгеброй.
$\xi^{-1}(\mathbb{R}) = \Omega \in \mathfrak{A} \Rightarrow \mathbb{R} \in \sigma$ и $\xi^{-1}(\varnothing) = \varnothing \in \mathfrak{A} \Rightarrow \varnothing \in \sigma$.
Пусть теперь $C_1, C_2 \in \sigma$. Докажем, что $C_1 \cup C_2,\ C_1 \cap C_2,\ C_1 \setminus C_2 \in \sigma$:

$$\xi^{-1}(C_1 \cup C_2) = \xi^{-1}(C_1) \cup \xi^{-1}(C_2)$$

Но $\xi^{-1}(C_1) \in \mathfrak{A}$ и $\xi^{-1}(C_2) \in \mathfrak{A}$, а значит и $\xi^{-1}(C_1) \cup \xi^{-1}(C_2) \in \mathfrak{A}$, то есть $C_1 \cup C_2 \in \sigma$.
Аналогично для объединения и разности событий C_1, C_2. Более того, это выполняется и для счетного объединения (пересечения). Значит, σ является σ-алгеброй.
По определению случайной величины ξ эта σ-алгебра содержит промежутки. А минимальная σ-алгебра, содержащая все промежутки — $\mathfrak{B}(\mathbb{R})$. Значит, $\mathfrak{B}(\mathbb{R}) \subseteq \sigma$. А что есть σ? Это $\{C\colon \xi^{-1}(C) \in \mathfrak{A}\}$ (множества, чьи прообразы лежат в $\mathfrak{A}$). Следовательно, $\forall B \in \mathfrak{B}(\mathbb{R})\ \xi^{-1}(B) \in \mathfrak{A}$ (прообраз всякого B лежит в $\mathfrak{A}$). □

На $\mathfrak{B}(\mathbb{R})$ определена вероятностная мера

$$\mu_\xi(B) = P\left(\xi^{-1}(B)\right).$$

Определение 2. Вероятностная мера μ_ξ называется *распределением* ξ.

Функция распределения вероятностной меры и функция распределения случайной величины

Вспомним, что дискретная случайная величина может быть задана перечнем всех ее возможных значений и их вероятностей. Такой способ задания не является общим. Для того, чтобы дать общий способ задания любых типов случайных величин, вводят функции распределения вероятностей случайной величины.
Пусть x — действительное число. Вероятность события, состоящего в том, что ξ примет значение, меньшее или равное x, то есть вероятность события $\xi \leqslant x$, обозначим через $F_\xi(x)$. Разумеется, если x изменяется, то, вообще говоря, изменяется и $F_\xi(x)$, то есть $F_\xi(x)$ — функция от x.

Определение 3. *Функцией распределения случайной величины* ξ называют функцию $F_\xi(x)$, определяющую вероятность того, что случайная величина ξ в результате испытания примет значение, меньшее или равное x, то есть

$$F_\xi(x) = P(\xi \leqslant x).$$

Замечание 4. Геометрически это равенство можно истолковать так: $F_\xi(x)$ есть вероятность того, что случайная величина ξ примет значение, которое изображается на числовой оси точкой на полуинтервале $(-\infty, x]$.

Свойство 1. *Значение функции распределения принадлежит отрезку* $[0, 1]$*:*

$$\forall \xi \ 0 \leqslant F_\xi(x) \leqslant 1.$$

Доказательство. Свойство возникает из определения функции распределения как вероятности: вероятность всегда есть неотрицательное число, не превышающее единицы. □

Свойство 2. $F_\xi(x)$ *— неубывающая функция, то есть*

$$\forall \xi \ F_\xi(x_2) \geqslant F_\xi(x_1), \ \textit{если} \ x_2 > x_1.$$

Доказательство. Пусть $x_2 > x_1$. Событие, состоящее в том, что ξ примет значение, меньшее или равное x_2, можно подразделить на следующие два несовместимых события: 1) ξ примет значение, меньшее или равное x_1, с вероятностью $P(\xi \leqslant x_1)$; 2) ξ примет значение, удовлетворяющее неравенству $x_1 < \xi \leqslant x_2$, с вероятностью $P(x_1 < \xi \leqslant x_2)$.
Тогда имеем:

$$P(\xi \leqslant x_2) = P(\xi \leqslant x_1) + P(x_1 < \xi \leqslant x_2).$$

Отсюда

$$P(\xi \leqslant x_2) - P(\xi \leqslant x_1) = P(x_1 < \xi \leqslant x_2)$$

или

$$F_\xi(x_2) - F_\xi(x_1) = P(x_1 < \xi \leqslant x_2).$$

Так как любая вероятность есть число неотрицательное, то $F_\xi(x_2) - F_\xi(x_1) \geqslant 0$, или $F_\xi(x_2) \geqslant F_\xi(x_1)$, что и требовалось доказать. □

Свойство 3. *Если возможные значения случайной величины* ξ *принадлежат интервалу* (a, b)*, то:*

(1) $F_\xi(x) = 0$ *при* $x \leqslant a$,
(2) $F_\xi(x) = 1$ *при* $x \geqslant b$.

Доказательство. (1) Пусть $x_1 \leqslant a$. Тогда событие $\xi \leqslant x_1$ невозможно (так как значений, меньших или равных x_1, величина ξ по условию не принимает) и, следовательно, вероятность такого события равна нулю.
(2) Пусть $x_2 \geqslant b$. Тогда событие $\xi \leqslant x_2$ достоверно (так как все возможные значения ξ меньше x_2) и, следовательно, вероятность такого события равна единице. □

Следствие 1. *Если возможные значения случайной величины* ξ *расположены на всей оси* x*, то справедливы следующие предельные соотношения:*

$$\lim_{x \to -\infty} F_\xi(x) = 0, \ \lim_{x \to +\infty} F_\xi(x) = 1.$$

Замечание 5. Доказанные свойства позволяют представить, как выглядит график распределения случайной величины.
График расположен в полосе, ограниченной прямыми $y = 0,\ y = 1$ (первое свойство).
При возрастании x в интервале (a, b), в котором заключены все возможные значения случайной величины, график "поднимается вверх" (второе свойство).
При $x \leqslant a$ ординаты графика равны нулю; при $x \geqslant b$ ординаты графика равны единице (третье свойство).

Замечание 6. График функции распределения дискретной случайной величины имеет ступенчатый вид.

Теорема 1. *(Теорема без доказательства)*
F_ξ однозначно определяет μ_ξ.

Определение 4. *Функция плотности* называется такая $\rho_\xi(x)$, что

$$F_\xi(x) = \int\limits_{-\infty}^{x} \rho_\xi(x)$$

Пример 2.
(Первый пример)
ξ имеет *равномерное распределение* на $[a, b]$, если

$$\rho_\xi(x) = \begin{cases} \frac{1}{b-a}, x \in [a, b] \\ 0, x \notin [a, b] \end{cases}$$

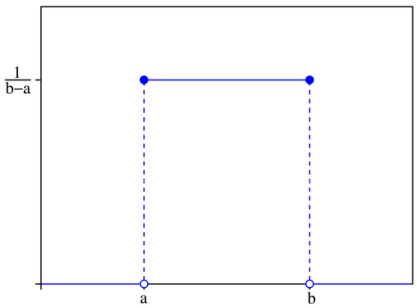


Рис. 4. График функции плотности равномерного распределения

$$F_\xi(x) = \begin{cases} 0, x < a \\ \frac{x-a}{b-a}, a \leqslant x < b \\ 1, x \geqslant b \end{cases}$$

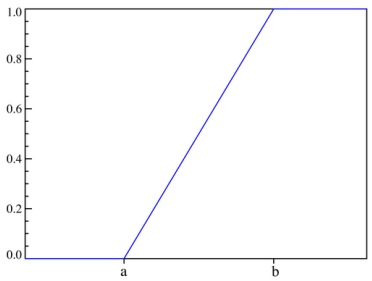


Рис. 5. График функции равномерного распределения

(Второй пример)
ξ имеет *показательное распределение*, если

$$\rho_\xi(x) = \begin{cases} 0, x < 0 \\ \lambda e^{-\lambda x}, x \geqslant 0 \end{cases}$$

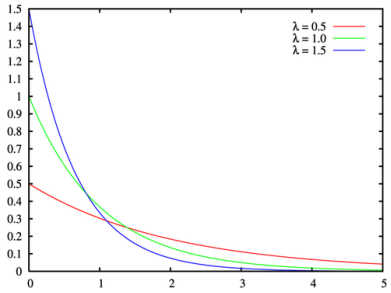


Рис. 6. График функции плотности показательного распределения

$$\rho_\xi(x) = \begin{cases} 0, x < 0 \\ 1 - \lambda e^{-\lambda x}, x \geqslant 0 \end{cases}$$

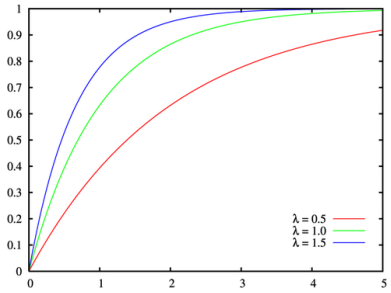


Рис. 7. График функции показательного распределения

(Третий пример)
ξ имеет *нормальное распределение* с параметрами μ и σ, если

$$\rho_\xi(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma}}$$

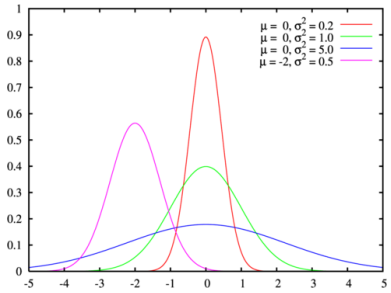


Рис. 8. График функции плотности нормального распределения

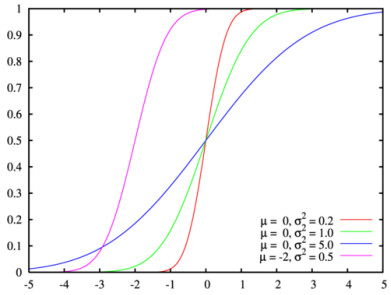


Рис. 9. График функции нормального распределения