Yet Another Math for DS Course
Посиделка №1: бесконечности бывают разные

Ульянкин Ппилиф [*]

> Икеевских карандашей никогда не наберешь достаточно!
>
> *народная мудрость*

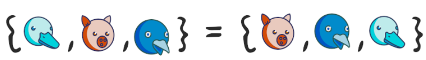
В этой посиделке мы поговорим про множества и их мощности. В теории вероятностей очень удобно описывать события через множества, а вероятности оценивать с помощью подсчёта их мощностей. Поэтому нам важно понимать, какие действия можно делать над множествами и какими могут быть их мощности.

1 Что такое множество

Что такое множество никто не знает. Это неопределимое понятие. Можно попробовать объяснить что это такое через следующее определение.

Определение. Под **множеством** понимается совокупность объектов любой природы, объединенных по какому либо признаку. Объекты, составляющие множество, называются его **элементами.**

Другими словами, множество — это неупорядоченная коллекция уникальных элементов. Что значит неупорядоченная? Это значит, что два множества эквивалентны, если содержат одинаковые элементы[1].



Элементы мноества должны быть уникальными, множество не может содержать одинаковых элементов. Добавление элементов, которые уже есть в множестве, не изменяет это множество.

Вроде что-то объяснили, но с помощью синонимов. На самом деле для нас не очень важно, что такое множество. Важно уметь отвечатать на вопрос, правда ли элемент лежит в множестве или нет

[*] https://github.com/FUlyankin/yet_another_math_for_DS

[1] Картинки спёр с Хабра: https://habr.com/ru/post/516858/

$$a \in A.$$

Если мы научимся отвечать на такой вопрос, проблем не будет. Давайте введём ещё несколько определений.

Определение. Множества, состоящие из конечного числа элементов, называются **конечными**, а остальные **бесконечными**. **Мощностью** конечного множества называют число его элементов. Мы будем обозначать мощность множества A как $|A|$. Понятие мощности для бесконечных множеств мы введём позже.

Определение. Если каждый элемент множества A является элементом множества B, говорят, что множество A является **подмножеством** множества B. Это записывают как $A \subset B$. Если множества потенциально могут совпадать, пишут $A \subseteq B$ (аналоги $<$ и $\leqslant$).

Если все элементы множества A являются подмножествами множества B и наоборот, это указывает на то, что A и B имеют одинаковые элементы и $A = B$.

$$A = B \Leftrightarrow A \subseteq B \text{ и } B \subseteq A.$$

Множество, не содержащее ни одного элемента — **пустое множество**. Его обычно обрзначают символом $\emptyset$. Считается, что пустое множество — подмножество любого множества.

Множество можно задать разными способами.

а. Перечислением его элементов, т.е. списком элементов $A = \{a, b, c, d\}$.

б. Указанием характеристических свойств его элементов $A = \{x \mid P(x)\}$.

в. Порождающей процедурой, т.е. множество задаётся совокупностью правил, по которым получаются элементы этого множества из уже известных элементов или других объектов.

Пример 1. Пусть A — множество целых чисел, являющихся степенями двойки. Можно задать его набором правил:

- $1 \in A$;
- если $a \in A$, то $2 \cdot a \in A$

Также его можно было бы записать как

$$A = \{2^{k-1} \mid k \in \mathbb{N},\}$$

где $\mathbb{N}$ — множество всех натуральных чисел, то есть $\mathbb{N} = \{1, 2, 3, 4, 5, ...\}$

Определение. Множество всех подмножеств множества A называется **булеаном** этого множества и обозначается как 2^A.

Пример 2. Выпишем булеан для множества $A = \{1, 2, 3\}$.

$$2^A = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}, \{\varnothing\}\}.$$

Обратите внимание, что пустое множество тоже входит в множество всех подмножеств A. Мощность множества A равна числу его элементов, так как это конечное множество. То есть $|A| = 3$. Мощность булеана оказывается $|2^A| = 2^{|A|}$.

2 Парадокс Рассела

Вроде бы красота. Но тут начинаются проблемы. Давайте рассмотри множество всех множеств, которые не содержат себя в качестве множества.

$$M = \{A \mid A \notin A\}.$$

Должно ли множество M включать себя? Представим себе множество парней студентов. Мы взяли одного из них и сказали, что он должен брить всех, кто не бреется сам. Должен ли он брить себя? С одной стороны да, с другой нет. Бедный парень сходит с ума. Эта штука называется **парадокс Рассела.**

Чтобы разобраться с парадоксами в начале XX века математики решили упорядочить способы задания множеств. В этой системе аксиом M, рассмотренное выше, это не множество. Нельзя заводить множество объектов, которые мне нравятся. Оно должно удовлетворять аксиомам, чтобы не возникало парадоксов.

Шире всего сегодня распространена аксиоматика Цермело-Френкеля. Все понятия, которые мы даём в этом конспекте как опредеелния, можно аккуратно вывести из этой аксиоматики. При этом никаких парадоксов не будет.

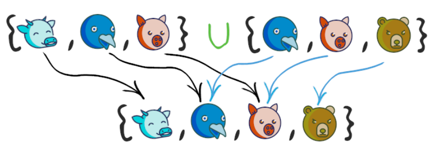
3 Операции над множествами

Определение. **Объединением** двух множеств A и B называют такое множество, которое состоит из тех и только тех элементов, которые принадлежат хотя бы одному множеству

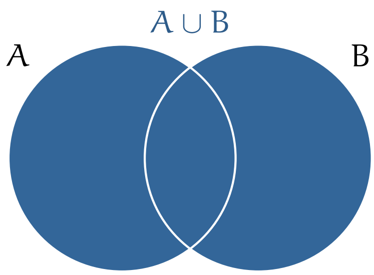
$$A \cup B = \{x : x \in A \text{ или } x \in B\}$$

.

Аналогично определяются объединения трёх и более множеств.

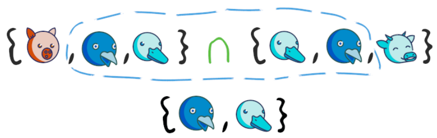


Операции над множествами удобно изображать на **диаграммах Эйлера-Венна:**



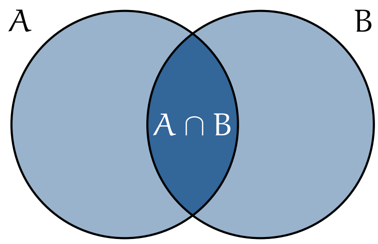
Определение. **Пересечением** множеств A и B называется такое множество, которое состоит из тех и только тех элементов, которые принадлежат и A, и B.

$$A \cap B = \{x : x \in A \text{ и } x \in B\}$$



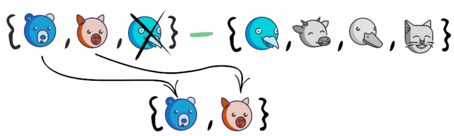
Иногда бывает так, что пересечение множеств оказывается пустым.

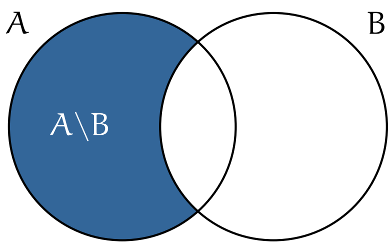
Пересечение множеств можно изобразить на диаграме Эйлера-Вена:



Определение. **Разностью** множеств A и B называется такое множество, которое состоит из тех и только тех элементов множества A, которые не содержатся в B

$$A \backslash B = \{x : x \in A \text{ и } x \notin B\}.$$



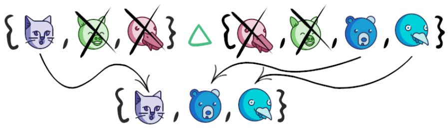


Определение. **Симметрической разностью (кольцевой суммой)** называются множество, которое включает в себя все те элементы, которые принадлежат только одному из множеств

$$A \Delta B = (A \backslash B) \cup (B \backslash A),$$

либо

$$A \Delta B = (A \cup B) \backslash (B \cap A).$$



Определение. Дополнением множества A называют множества всех элементов универсального множества U, не входящих во множество A

$$\bar{A} = \{x : x \notin A \text{ и } x \in U\}.$$

4 Свойства операций над множествами

Как и у любых операций, у операций над множествами есть куча свойств. Давайте приведём тут большой список хороших свойств и немного на него посмотрим.

а. От перестановки мест слагаемых сумма не меняется (коммутативность)

$$A \cup B = B \cup A$$
$$A \cap B = B \cap A$$

б. Можно расставлять и раскрывать скобки (ассоциативность)

$$(A \cup B) \cup C = A \cup (B \cup C)$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

в. Дистрибутивность

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$$
$$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

г. Законы двойственности(де Моргана):

$$\overline{\overline{A} \cup \overline{B}} = \overline{A} \cap \overline{B}$$
$$\overline{\overline{A} \cap \overline{B}} = \overline{A} \cup \overline{B}$$

д. Законы идемпотентности:

$$A \cup A = A$$
$$A \cup \emptyset = A$$
$$A \cup U = U$$
$$A \cap A = A$$
$$A \cap \emptyset = \emptyset$$
$$A \cap U = A$$

е. Законы поглощения:

$$A \cup (A \cap B) = A$$
$$A \cap (A \cup B) = A$$

ж. Закон включения:

$$A \subseteq B \Leftrightarrow \bar{B} \subseteq \bar{A}$$

з. Закон двойного дополнения:

$$\bar{\bar{A}} = A$$

и. Закон равенства:

$$A = B \Leftrightarrow (A \subseteq B) \cap (B \subseteq A)$$

Можно вывести ещё много разных свойств. Не надо запоминать их наизусть. Когда они нам будут полезны на практике, мы будем их выводить. Список нужен здесь просто для знакомства.

5 Декартово произведение множеств

Определение. Пусть даны множества A и B, пара (a, b), где $a \in A, b \in B$ называется **упорядоченной**.

Множество всех упорядоченных пар вида (a, b) называется **декартовым произведением** множеств A и B

$$A \times B = \{(a, b) \mid a \in A, b \in B\}.$$

Отметим, что

$$A \times B \neq B \times A.$$

Определение. Пусть даны множества $A_1, A_2, A_3, \dots, A_n$ **Декартовым произведением** этих множеств называется множества упорядоченных наборов $(a_1, a_2, a_3, \dots, a_n$, где $a_i \in A_i$. В частности, если $A_1 = A_2 = \dots = A_n$, то $A_1 \times A_2 \times \dots \times A_n = A^n$.

У декартового произведения тоже есть куча свойств. Вот некоторые из них:

а. $(A \cup B) \times C = (A \times C) \cup (B \times C)$

б. $(A \cap B) \times C = (A \times C) \cap (B \times C)$

в. $(A \backslash B) \times C = (A \times C) \backslash (B \times C)$

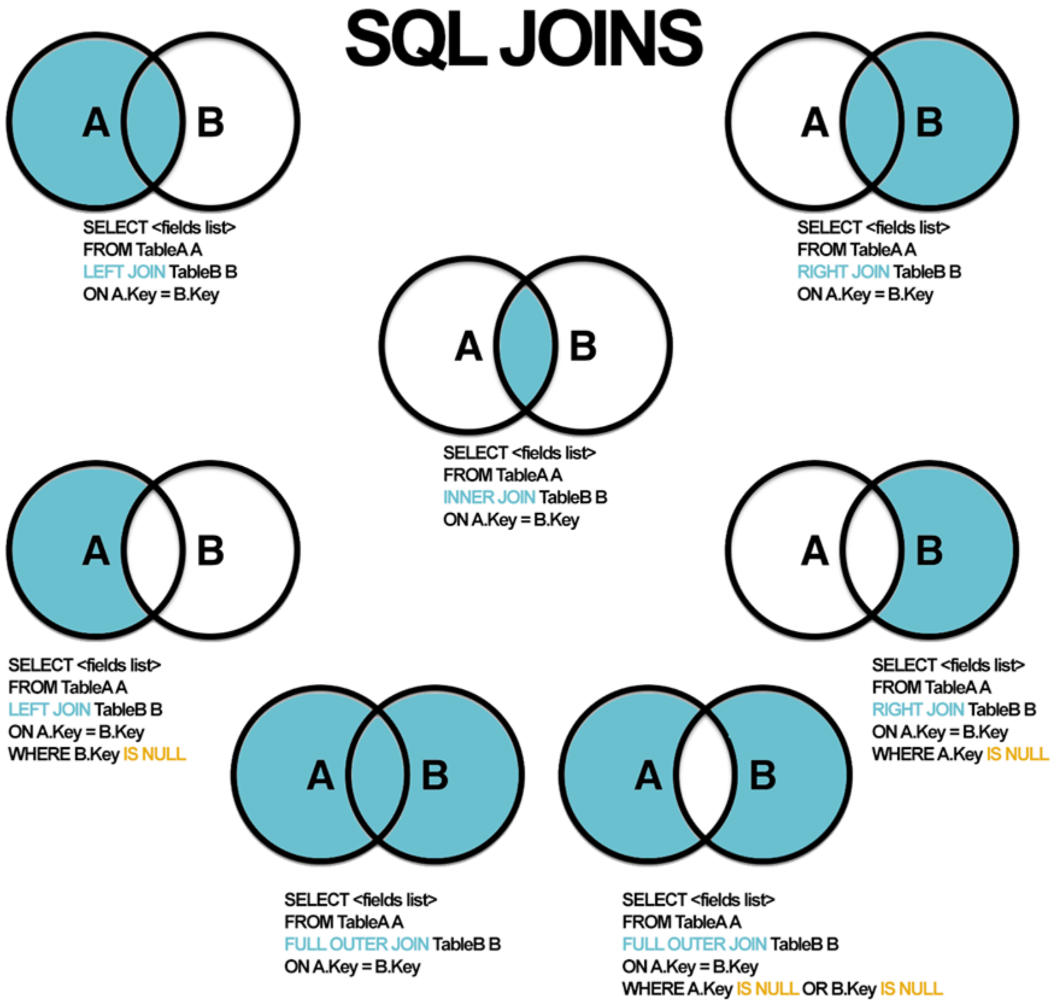
г. $C \times (A \backslash B) = (C \times A) \backslash (C \times B)$

Пусть A и B — конечные множества. Напомним, что **мощностью** конечного множества называется число его элементов.

Теорема 1. Мощность декартова произведения конечного числа конечных множеств равно произведению мощностей этих множеств, то есть

$$|A_1 \times A_2 \times \ldots \times A_n| = |A_1| \times |A_2| \times \ldots \times |A_n|.$$

6 SQL-картиночки



Часто при написании SQL-запросов надо объединять таблицы по какому-то полю, индексу. В документации можно встретить диаграммы Эйлера-Вена, илюстрирующие разные типы объединения (джойнов) таблиц.

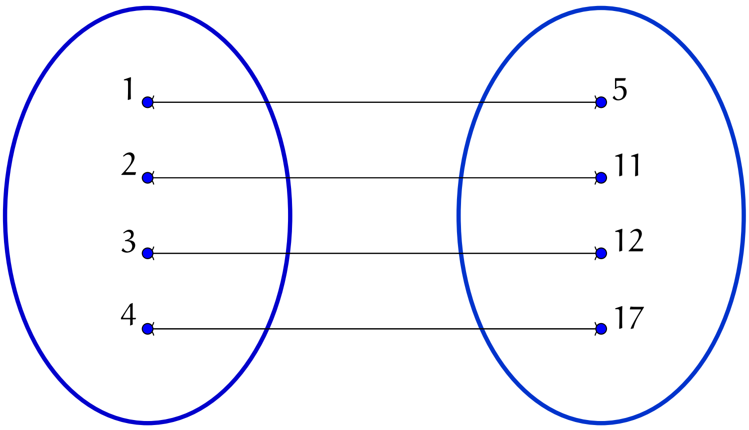
В центре картинки находится **INNER JOIN**. При нём в итоговой таблице остаются строки с ключами *key*, которые были в обеих таблицах. Например, если в колонке *key* записано имя, а Ратибор есть в обеих таблицах, он попадёт и в итоговую. Если Добронрав есть только в левой или в правой таблице, он не попадёт в итоговую. Все остальные операции можно проинтерпретировать аналогичным образом.

Все основные виды **JOIN** строят декартово произведение строк по совпадающим значениям ключей. Иногда аналитики случайно пытаются поджойнить таблицы по неуникальным ключам. Например, имя Маша повторяется в таблице A десять раз и в таблице B двадцать раз. SQL попытается сделать декартово произведение каждой строки с каждой и для имени Маша в итоговую таблицу запишет все возможные комбинации. То есть, 200 строк. Если ваш **JOIN** работает очень медленно, задумайтесь о своих декартовых произведениях.

7 Бесконечности бывают разными

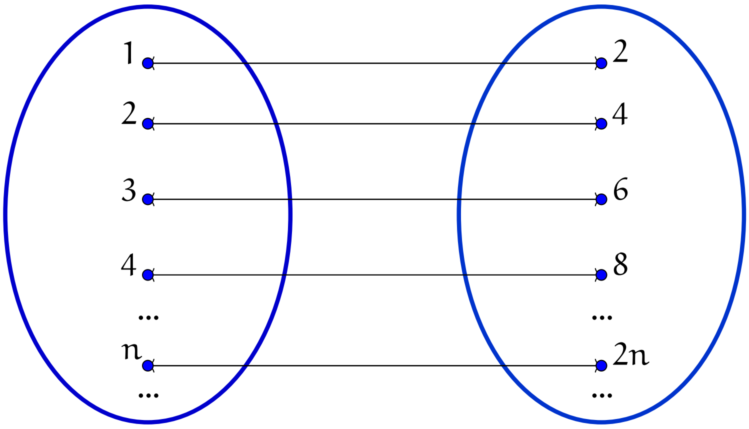
Определение. Множества[2] A и B называются **равномощными**, если между элементами этих множеств существует взаимно-однозначное соответствие. Мощность множества A обозначается $|A|$. Соответственно равномощность двух множеств можно записать как $|A| = |B|$.

Множества $\{1, 2, 3, 4\}$ и $\{5, 11, 12, 17\}$ равномощны, так как есть взаимно-однозначное соответствие:

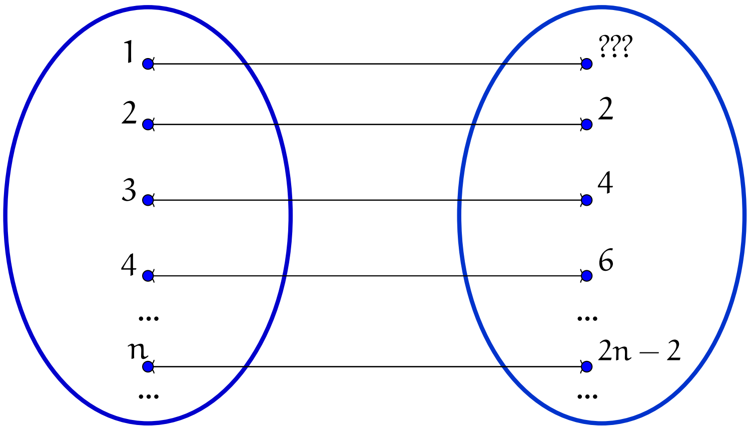


Множество $\mathbb{N}$ натуральных чисел и множество чётных натуральных чисел равномощны в силу соответствия:

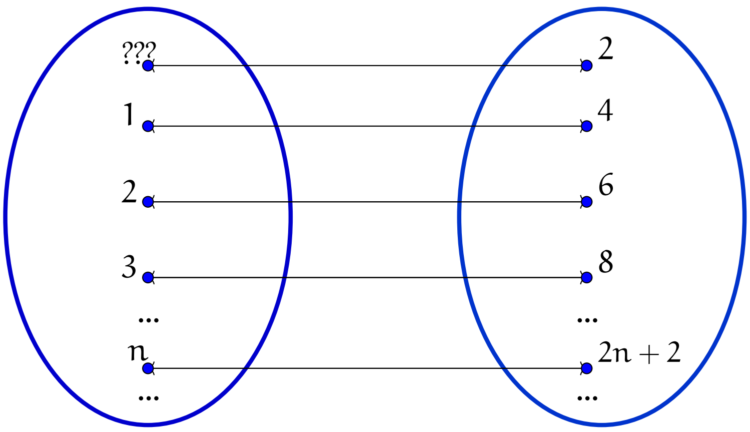
[2]Все следующие разделы я взял из нашей с Борисом Демешевым попытки написать книгу https://github.com/bdemeshev/sc_book



Заметим, что в определении требуется только существование взаимно-однозначного соответствия. Это не противоречит тому, что могут существовать другие, не взаимно-однозначные соответствия. Например, можно добиться и того, что среди натуральных чисел останутся «лишние»:



Можно добиться того, что среди чётных останутся «лишние», а именно:

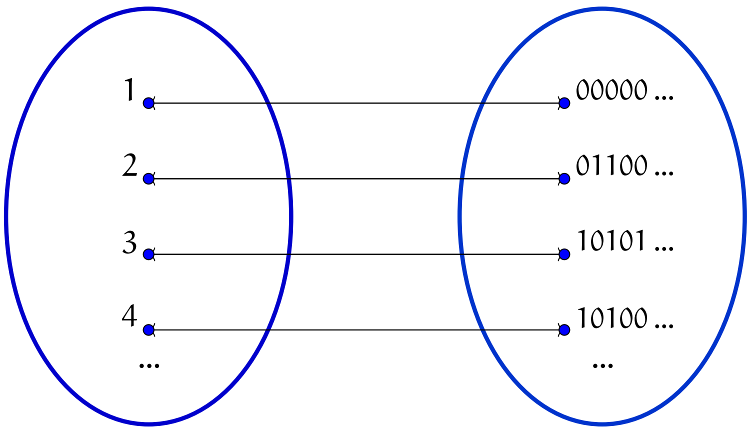


Множества $\{1, 2, 3, 4\}$ и множество $\mathbb{N}$ неравномощны. При любой попытке построить взаимно-однозначное соответствие среди натуральных чисел останутся «лишние». Возникает естественный вопрос, все ли бесконечные множества равномощны?

Пусть S множество всех бесконечных вправо последовательностей из 0 и 1. Например, одним из элементов S является последовательность $1010101010\ldots$

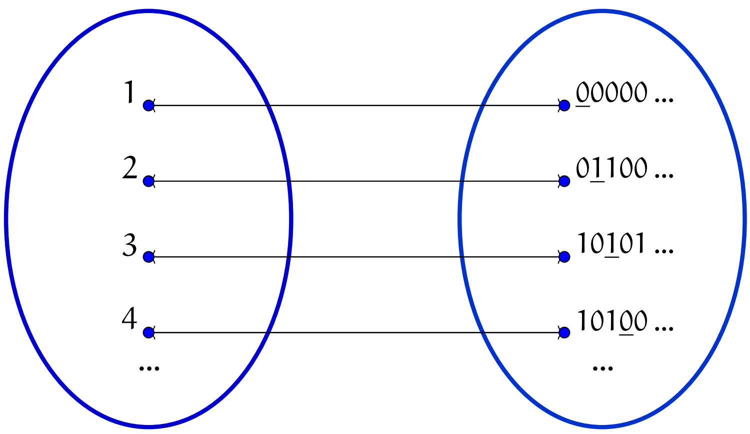
Теорема 2. Множество S бесконечно, но не равномощно множеству $\mathbb{N}$.

Доказательство. Допустим противоположное, что S и $\mathbb{N}$ равномощны. Тогда существует взаимно-однозначное соответствие между натуральными числами и последовательностями. К примеру оно могло бы выглядеть так:



Оказывается какое бы соответствие ни было создано, всегда существует последовательность, которой не сопоставлено ни одно число!

Создадим последовательность a по следующему принципу: возьмем первую цифру из первой последовательности, затем вторую из второй, затем третью из третьей и т.д.



Получаем последовательность $a = 0110\ldots$ Затем построим последовательность b заменив единицы на нули, а нули на единицы в последовательности a. В нашем примере $b = 1001\ldots$

Вне зависимости от того, какое соответствие мы взяли, последовательность b не может идти в нём ни под каким номером! Она не может идти под номером 1, так как отличается от первой последовательности первой цифрой. Она не может идти под номером 2, так как отличается от второй последовательности второй цифрой и т.д.

Мы пришли к противоречию, в S есть «лишняя» незанумерованная последовательность b. Значит S и $\mathbb{N}$ неравномощны.

Аналогичным способом можно построить «лишнюю» последовательность для любого другого предложенного соответствия. □

Определение. Мы говорим, что множество A имеет мощность **континуум**, если оно равномощно множеству S бесконечных вправо последовательностей из 0 и 1.

Определение. Множество A называется **счётным**, если оно конечно или равномощно множеству $\mathbb{N}$ натуральных чисел.

Будьте бдительны при чтении других источников: некоторые авторы определяют счётные как равномощные натуральным числам, но таких авторов меньшинство. Из данного определения следует, что **несчётные** множества — это бесконечные множества не равномощные множеству $\mathbb{N}$ натуральных чисел.

Оказывается, что любые бесконечные множества можно сравнивать: либо они равномощные, либо одно из них «больше».

Теорема 3. Если A и B — два произвольных множества, то возможна одна и только одна из трех ситуаций:

1) Множества A и B равномощны.

2) Множества A и B неравномощны, но A равномощно какому-нибудь подмножеству множества B.

3) Множества A и B неравномощны, но B равномощно какому-нибудь подмножеству множества A.

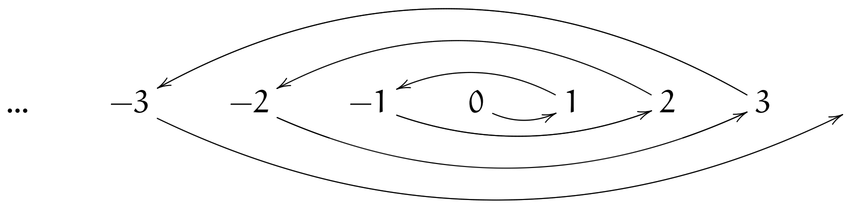
Замечание. Доказательство можно найти в [1]. Может возникнуть вопрос, а что тут собственно доказывать, всё же «очевидно»? Доказывать нужно два утверждения. Во-первых, что ситуации 2) и 3) не могут произойти одновременно. Во-вторых, что невозможна гипотетическая ситуация «несравнимости», когда A и B неравномощны, в A нет части равномощной B, и в B нет части равномощной A.

Из этой теоремы следует важное следствие:

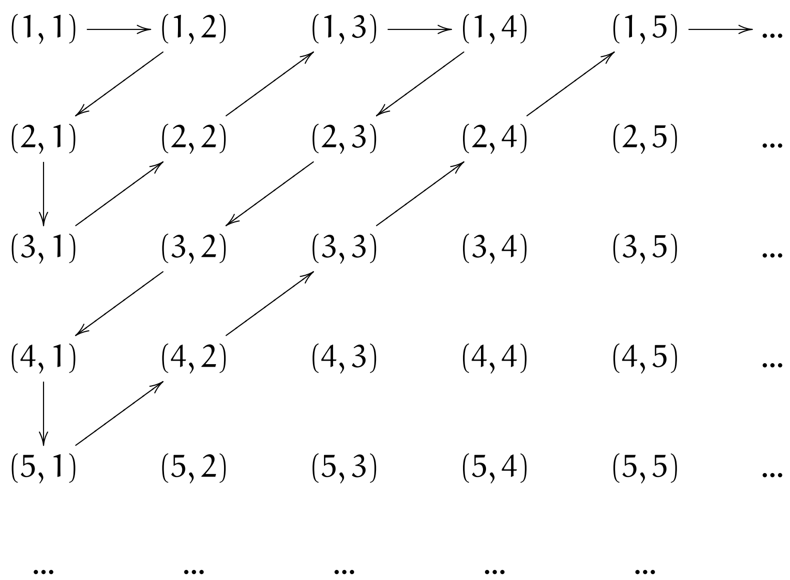
Теорема 4. Если A равномощно подмножеству B, то $|A| \leqslant |B|$, т.е. либо мощность A меньше мощности B, либо A и B равномощны.

Давайте посмотрим на несколько важных примеров. **Начнём с бесконечных счётных множеств.**

- Множество целых чисел $\mathbb{Z}$ — счётное! Между множествами $\mathbb{Z}$ и $\mathbb{N}$ существует взаимно-однозначное соответствие. Его можно установить, например, следующим образом. $1 \leftrightarrow 0$, $2 \leftrightarrow 1$, $3 \leftrightarrow -1$, $4 \leftrightarrow 2$ и т.д. Этот способ нумерации элементов множества $\mathbb{Z}$ показан на рисунке.



- Множество пар натуральных чисел $\mathbb{N}^2$ — счётное! Все элементы множества $\mathbb{N}^2$ можно пересчитать. Один из способов сделать это показан на рисунке.



Фактически было доказано, что объединение счётного количества счётных множеств счётно[3]. А именно, множество $\mathbb{N}^2$ можно разбить на счётное множество компонент: $\mathbb{N}^2 = \{0\} \times \mathbb{N} \cup \{1\} \times \mathbb{N} \cup \{2\} \times \mathbb{N} \cup \ldots$ Элементы каждой компоненты можно выписать в отдельную строку, а после занумеровать.

- Множество $\mathbb{Z}^n$ векторов из n целых чисел — счётное! Доказательство проведём по индукции. Мы уже доказали, что $\mathbb{Z}^1$ — счётное множество. Теперь предположим, что $\mathbb{Z}^{n-1}$ — счётное. Получаем цепочку: $|\mathbb{Z}^n| = |\mathbb{Z}^{n-1} \times \mathbb{Z}^1| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$.

- Множество рациональных чисел $\mathbb{Q}$ — счётное! Приведём два способа занумеровать все рациональные числа.

  Пусть $f(q)$ — это сумма модулей числителя и знаменателя дроби после сокращения, например, $f(-2/3) = 5$. Сначала нумеруем дроби с $f(q) = 1$, потом — дроби с $f(q) = 2$, потом — дроби с $f(q) = 3$ и т.д. В результате каждая дробь получает свой номер.

---

[3] Для знатоков может быть интересен тот факт, что в этом доказательстве неявно используется аксиома выбора. Без неё можно представить $\mathbb{R}$ как счётное объединение счётных множеств [2].

Рассмотрим другой, более наглядный способ нумерации, который изображен на рисунке ниже. В первой строке поместим все целые числа, во второй строке — все дроби со знаменателем 2, в третьей — со знаменателем 3 и т.д. После пронумеруем все элементы двигаясь по стрелочкам. Если элемент повторяется, например, мы уже пронумеровали 0, но снова наталкиваемся на него или уже пронумеровали 1 и наталкиваемся на $\frac{2}{2}$, то пропускаем его.

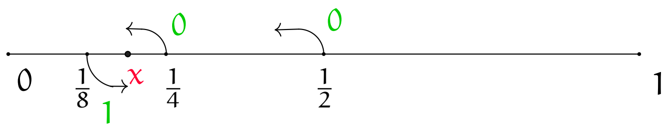
$$
\begin{tikzcd}
0 \arrow[r] & 1 \arrow[dl] & -1 \arrow[r] & 2 \arrow[dl] & -2 \arrow[r] & \dots \\
0 \arrow[d] & \frac{1}{2} \arrow[ur] & -\frac{1}{2} \arrow[dl] & \frac{2}{2} \arrow[ur] & -\frac{2}{2} & \dots \\
0 \arrow[ur] & \frac{1}{3} \arrow[dl] & -\frac{1}{3} \arrow[ur] & \frac{2}{3} & -\frac{2}{3} & \dots \\
0 \arrow[d] & \frac{1}{4} \arrow[ur] & -\frac{1}{4} & \frac{2}{4} & -\frac{2}{4} & \dots \\
0 \arrow[ur] & \frac{1}{5} & -\frac{1}{5} & \frac{2}{5} & -\frac{2}{5} & \dots \\
\dots & \dots & \dots & \dots & \dots &
\end{tikzcd}
$$

Теперь посмотрим на **множества мощности континуум**.

- Отрезок $[0;1]$ — множество мощности континуум!

  Покажем, что множество $[0;1]$ равномощно множеству S бесконечных вправо последовательностей из 0 и 1.

  Любое число $x \in [0;1]$ можно записать в виде бесконечной двоичной дроби. Первый знак этой дроби равен 1 или 0 в зависимости от того, попадает ли число x в левую или правую половину отрезка. Чтобы выбрать следующий знак, надо снова поделить выбранную половину пополам и посмотреть, куда попадет x, и т.д.



  Точка $1/2$ на рисунке ?? является серединой отрезка $[0;1]$. Точка x лежит левее $1/2$, то есть попадает в левую половину отрезка. Первый знак в элементе из S, который соответствует

x будет 0. Точка $1/4$ является серединой левой половины отрезка $[0;1]$. Точка x снова находится левее $1/4$, значит второй знак в элементе из S также будет 0. Точка $1/8$ является серединой отрезка $[0;1/4]$. Точка x находится правее $1/8$, значит третий знак в элементе из S будет 1. Далее посмотрим в какой части отрезка $[1/8;1/4]$ будет лежать точка x и получим четвертый знак элемента, затем пятый и так далее.

Это же соответствие можно описать в другую сторону: последовательности из нулей и единиц $x_0x_1x_2\dots$ соответствует число, являющееся суммой ряда

$$\frac{x_0}{2} + \frac{x_1}{2^2} + \frac{x_2}{2^3} + \dots$$

Например, последовательности $010100\dots$ соответствует число

$$\frac{0}{2} + \frac{1}{2^2} + \frac{0}{2^3} + \frac{1}{2^4} + \frac{0}{2^5} + \frac{0}{2^6} + \dots = \frac{1}{4} + \frac{1}{16} = \frac{5}{16}.$$

Описанное соответствие пока что не совсем взаимно-однозначное: дроби вида $m/2^n$ имеют два представления. Например, число $3/8$ можно записать в виде $0011000\dots$ и в виде $0010111\dots$ Соответствие станет однозначным, если отбросить последовательности с бесконечным хвостом из единиц, кроме последовательности $01111\dots$ Таких дробей счётное число и на мощность это никак не повлияет.

Блиц-вопрос 1. Какому числу соответствует последовательность $011111\dots$?

Последовательность $011111\dots$ соответствует единице.

$$\frac{0}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \frac{1}{2^4} + \frac{1}{2^5} + \dots = \frac{1/2}{1-1/2} = 1.$$

Именно поэтому последовательность $011111\dots$ нельзя отбросить.

Блиц-вопрос 2. Обычно в памяти компьютера числа хранятся в двоичной системе счисления.

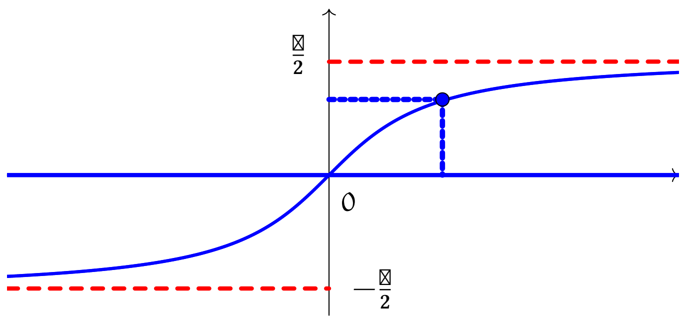
- Можно ли данным способом хранить в памяти число 0.15?
- А правда ли что с точки зрения компьютера $0.4 - 0.3$ равно 0.1?

Если попытаться записать 0.15 в виде последовательности из 0 и 1, мы получим периодическую бесконечную последовательность $0,1010101\dots$, которую компьютер не сможет запомнить, так как объем памяти в нём ограничен.

Если выполнить сравнение $0.4 - 0.3 == 0.1$ в большинстве языков программирования (R, Python, Julia, C++, ...) то результатом будет FALSE.

- Прямая $\mathbb{R}$ — множество мощности континуум! Показать, что прямая $\mathbb{R}$ — множество мощности континуум можно с помощью следующей двухходовочки.

  Ход первый: найдем отображение, которое ставит каждому элементу числовой прямой во взаимно-однозначное соответствие элементы интервала $(0;1)$. Функция $y = \arctan x$ ставит каждому элементу $x \in (-\frac{\pi}{2}; \frac{\pi}{2})$ в соответствие элемент числовой прямой $y \in \mathbb{R}$.

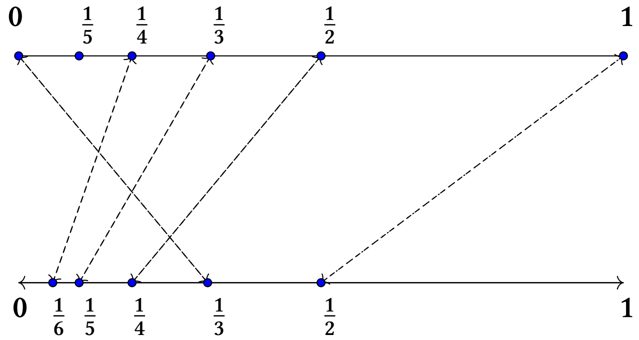
  

  Если к каждому элементу $x \in (-\frac{\pi}{2}; \frac{\pi}{2})$ прибавить $-\frac{\pi}{2}$, то можно получить интервал $(0;\pi)$. Если после этого в качестве бонуса разделить каждый элемент нового интервала на π, то можно получить интервал $(0;1)$. То есть отображение $x = -\frac{\pi}{2} + z\pi$ для каждого $z \in (0;1)$ находит $x \in (-\frac{\pi}{2}; \frac{\pi}{2})$.

  Таким образом взаимно-однозначное соответствие между $(0;1)$ и $\mathbb{R}$ устанавливается по формуле $y = \operatorname{arctg}(-\frac{\pi}{2} + z\pi)$.

  Аналогично можно построить взаимно-однозначное соответствие между $(0;1)$ и произвольным интервалом $(a;b)$ с помощью отображения $y = a + (b-a)x$. Действительно, если мы хотим пронормировать все элементы множества $(a;b)$ к интервалу $[0;1]$, мы должны из каждого элемента вычесть начало интервала a и поделить каждый элемент на длину интервала. То есть $\forall y \in [a;b] \quad x \in [0;1]$, если $x = \dfrac{y-a}{b-a}$.

  Ход второй: покажем что интервал $(0;1)$ и отрезок $[0;1]$ равномощны. Всем точкам отрезка $[0;1]$, кроме точек, имеющих вид $\frac{1}{n}$ поставим во взаимно-однозначное соответствие себя же. Точке 1 поставим в соответствие точку $1/2$, точке 0 точку $1/3$. Далее каждой точке вида $\frac{1}{n}$ будем ставить в соответствие точку $\frac{1}{n+2}$.

- Пространство $\mathbb{R}^n$ — множество мощности континуум!

  Разберём доказательство на примере плоскости $\mathbb{R}^2$. Для общего случая доказательство будет проводиться аналогичным образом.

  Доказывать взаимно-однозначное соответствие между $\mathbb{R}^2$ и S будем с помощью новой многоходовочки.

  Ход первый. Покажем, что множество $[0;1] \times [0;1]$ равномощно множеству $[0;1]$. Мы уже знаем, что вместо чисел на отрезке можно говорить о последовательностях нулей и единиц. Заметим, что каждой паре (x, y) из квадрата $[0;1] \times [0;1]$ соответствует пара последовательностей $(x_0x_1x_2x_3 \dots, y_0y_1y_2y_3 \dots)$. Каждой из таких последовательностей можно поставить единственным образом в соответствие последовательность $x_0y_0x_1y_1x_2y_2x_3y_3 \dots$. То есть множество точек отрезка $[0;1]$ равномощно множеству $[0;1] \times [0;1]$.

  Ход второй. Покажем, что можно построить отображение, которое перевело бы множество $[0;1] \times [0;1]$ в множество $\mathbb{R}^2$. Поиску такого отображения посвящено упражнение **??**.

  Для любого множества $\mathbb{R}^n$ рассуждение можно провести аналогичным образом.

  Полученный результат, как и многие другие результаты этой главы, противоречит интуиции! Квадрат же намного больше, чем отрезок! Квадрат — двумерен! В квадрате должно быть на порядок больше точек, чем на отрезке! Однако это не так. Впервые этот факт был доказан Георгом Кантором в 1877 году. Он очень удивил Кантора.

- Функции, непрерывные на отрезке $[0;1]$ — множество мощности континуум!

  Пусть $C[0;1]$ — множество функций, непрерывных на отрезке $[0;1]$.

  Интуиция сразу же подсказывает нам, что любая константа — непрерывная функция, а непрерывных функций на $[0;1]$ гораздо больше чем констант, значит отрезок $[0;1]$ равномощен подмножеству множества $C[0;1]$.

  $$0 \leftrightarrow f(x) = 0,\ 0.1 \leftrightarrow f(x) = 0.1,\ 0.05 \leftrightarrow f(x) = 0.05, \dots$$

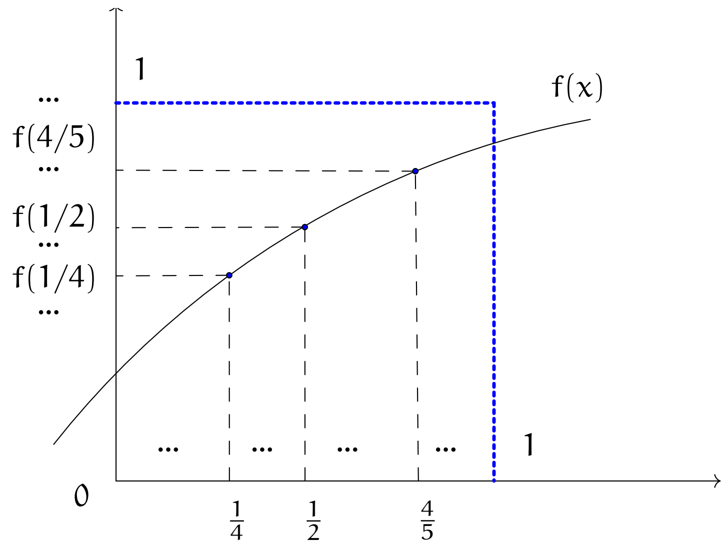
  Это означает, что либо мощность отрезка $[0;1]$ меньше мощности множества $C[0;1]$, либо они равномощны. Чтобы показать, что эти множества равномощны нужно найти подмножество множества $[0;1]$, которому будет равномощно множество $C[0;1]$.

  Пусть $x_1, x_2, \dots$ — последовательность всех рациональных точек отрезка $[0;1]$. Поставим в соответствие каждой непрерывной функции $f(x)$ последовательность вещественных чисел — значений функции $f(x)$ в точках $x_1, x_2, \dots$

  $$f(x) \leftrightarrow (f(x_i))_{i=1}^{\infty}$$

  .

При этом двум различным функциям $f(x)$ и $g(x)$ будут отвечать различные последовательности $(f(x_i))_{i=1}^{\infty}$ и $(g(x_i))_{i=1}^{\infty}$, поскольку две непрерывные функции, совпадающие во всех рациональных точках, совпадают всюду.



Таким образом, множество функций, непрерывных на отрезке $[0;1]$ можно считать эквивалентным некоторой части множества всех числовых последовательностей.

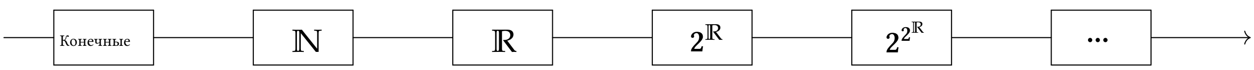
Множество всех числовых последовательностей будет множеством мощности континуум. Доказательству этого факта посвящено упражнение из листочка с задачками к лекции.

Для целей стохастического анализа нам хватит этих сведений про мощности множеств, однако для общего развития полезно знать еще пару фактов.

Факт 1. Бесконечные множества не исчерпываются равномощными множеству $\mathbb{N}$ и множеству последовательностей S. Бесконечности бывают разные, и их бесконечно много. Если есть одно бесконечное множество A, то найдется бесконечное множество B, где элементов ещё больше, чем в A!

Теорема 5. Если A — произвольное множество, то множество всех подмножеств множества A, обозначаемое 2^A, не равномощно множеству A.

Иными словами, подмножеств действительных чисел больше, чем действительных чисел. Списков подмножеств действительных чисел больше, чем подмножеств действительных чисел. И так далее ...



Факт 2. Невинный вопрос: а есть ли мощности промежуточные между континуумом и мощностью множества натуральных чисел? оказывается неожиданно сложным! Еще более неожиданно то, что любой ответ на него, и «да», и «нет» оказывается верным. Для ответа на этот вопрос интуитивных представлениях о множествах не хватает и нужно аксиоматизировать теорию множеств.

Так вот, разные системы аксиом приводят к разным ответам. В данном курсе мы это обсуждать не будем. Но заинтересовавшиеся могут посмотреть книгу Верещагин Шень, Начала теории множеств, [1].

Мощность множества натуральных чисел $\mathbb{N}$ обозначается символом $\aleph_0$ — «алеф-ноль». Счётные множества — это самые маленькие из бесконечных множеств. Множество вещественных чисел $\mathbb{R}$ имеет мощность континуума. Мощность таких множеств обозначается символом $\mathfrak{c}$. Предположение о том, что промежуточных мощностей нет, называется **континуум-гипотезой** и записывается как $\mathfrak{c} = \aleph_1$. В общем виде $2^{\aleph_\alpha} = \aleph_{\alpha+1}$.

$\aleph_0$	$2^{\aleph_0} = \aleph_1$	$2^{\aleph_1} = \aleph_2$	...	$2^{\aleph_\alpha} = \aleph_{\alpha+1}$	...

Авторам больше нравится система аксиом, где промежуточных мощностей нет, но это уже вопрос вкуса, а о вкусах не спорят.

В данной книжке мы придерживаемся системы аксиом ZFC. Хотя бы просто потому, что авторы (по крайней мере пока) не умеет думать без аксиомы выбора. Подробности можно найти в [3]. Также особо любопытные могут прочитать всего лишь сорок страниц заметок о теории множеств в [4].

Список литературы

[1] *Николай Константинович Верещагин и Александр Шень.* (1999). Начала теории множеств. // http://www.mccme.ru/free-books/

[2] *David Williams и David Williams.* (Springer, 2001). Weighing the odds: a course in probability and statistics.

[3] *Horst Herrlich.* (Springer, 2006). Axiom of choice. // Книга, которая чуть более, чем полностью посвящена аксиоме выбора.

[4] *Ingo Lutkebohle* (Springer, 2001). Set theory. // Всего 40 страниц заметок «Set theory». // http://alpha.math.uga.edu/~pete/expositions2012.html