Элементарная комбинаторика

1 Принцип Дирихле

1.1. Начнем мы данную главу с изложения очень простого принципа, который, однако, является весьма эффективным способом решения многих комбинаторных задач — принципа Дирихле (или в английском варианте — Pigeon-Hole Principle).

1.1.1. В своей простейшей формулировке он гласит следующее:

Утверждение 1.1. *Если в n ящиков положить $k > n$ предметов (в n клеток посадить $k > n$ голубей), то хотя бы в одном ящике будут лежать по крайней мере два предмета (будут сидеть по крайней мере два голубя).*

Доказательство. Несмотря на всю очевидность данного утверждения, все же дадим его формальное доказательство. Предположим, что утверждение неверно, то есть в каждом ящике находится не более одного предмета. Обозначим через m количество ящиков, в котором ничего не лежит. Очевидно, что $m \geqslant 0$. Тогда ровно по одному предмету лежит в $(n-m)$ ящиках. Это означает, собственно, что общее количество предметов равно $n - m \leqslant n < k$, что противоречит условию нашего утверждения. □

1.1.2. На первый взгляд совершенно непонятно, как такой настолько простой принцип может использоваться при решении каких-либо серьезных задач, ведь он кажется совершенно очевидным. Однако оказывается, что в конкретных задачах, использующих в своих решениях данный принцип, зачастую очень нелегко понять, где предметы, где ящики, и почему предметов больше, чем ящиков. Далее, далеко не всегда по формулировке самой задачи можно догадаться, что для ее решения следует воспользоваться данным принципом. И наконец, данный принцип дает нам неконструктивное доказательство какого-то факта — используя его, мы не можем сказать, в каком конкретно ящике находятся два предмета, мы знаем лишь, что такой ящик существует. Попытка же дать конструктивное доказательство, то есть попытка нахождения данного ящика, часто связана с очень большими трудностями. Тем не менее, данный принцип очень полезен, и с его помощью можно доказывать совершенно неочевидные на первый взгляд результаты.

1.1.3. Приведем вначале простейший пример на применение данного принципа.

Пример 1.2. Докажите, что в наборе из любого $(n+1)$-го положительного целого числа найдутся по крайней мере два числа, имеющих один и тот же остаток от деления на n.

Доказательство. Действительно, рассмотрим произвольный набор $\{a_1, a_2, \ldots, a_n, a_{n+1}\}$ положительных целых чисел. Поделив их с остатком на n, мы получим набор $\{r_1, r_2, \ldots, r_n, r_{n+1}\}$ из $(n+1)$-го остатка от этого деления — набор неотрицательных целых чисел, каждое из которых меньше n. Но мы знаем, что имеется только лишь n неотрицательных целых чисел, меньших n. Следовательно, по принципу Дирихле, хотя бы два остатка из полученного набора должны совпадать. □

1.1.4. Обычно в качестве примера применения принципа Дирихле дается чуть более сложная формулировка того же самого утверждения.

Пример 1.3. Докажите, что в последовательности чисел

$$7,\ 77,\ 777\ \ldots$$

один из первых 2014 членов данной последовательности делится на 2013.

Решение. Вначале ход рассуждений полностью повторяет рассуждения предыдущего примера. Именно, поделив с остатком первые 2014 членов выбранной числовой последовательности на 2013, мы получим 2014 остатков, значение каждого из которых меньше 2013. Следовательно, согласно принципу Дирихле существуют хотя бы два различных числа с одинаковыми остатками. Обозначим эти числа через a_i и a_j, и пусть для определенности $i < j$. Тогда

$$a_i = 2013 \cdot q_i + r, \qquad a_j = 2013 \cdot q_j + r,$$

а число $(a_j - a_i)$ делится без остатка на число 2013. Покажем теперь, что число a_{j-i} делится без остатка на 2013.

Действительно,

$$\begin{array}{rc} 77777777777777777777777777 & j \text{ цифр} \\ -77777777777777777 & i \text{ цифр} \\ = 77777777700000000000000000 & \begin{array}{c} j-i \text{ цифр, равных } 7, \\ i \text{ цифр, равных } 0, \end{array} \end{array} \qquad \Longrightarrow \qquad a_j - a_i = a_{j-i} \cdot 10^i.$$

Но 10^i и 2013 взаимно просты, поэтому a_{j-i} делится на 2013.

1.1.5. Рассмотрим еще один характерный пример.

Пример 1.4. Докажите, что среди нескольких находящихся в одной комнате человек хотя бы двое имеют одинаковое количество знакомых среди присутствующих в комнате.

Доказательство. Пусть в комнате находится n человек. Тогда для любого из них количество k с ним знакомых может принимать значения от нуля (ни с кем не знаком) до $n-1$ (со всеми знаком). Но отношение знакомства является взаимным: если a знаком с b, то и b знаком с a. Поэтому, к примеру, невозможен случай, когда в одной комнате имеются одновременно человек, у которого $(n-1)$ знакомых, и человек, у которого 0 знакомых: в первом случае у второго человека не может быть 0 знакомых, а во втором — у первого не может быть $(n-1)$ знакомый. Тем самым мы исключили вариант, когда у каждого находящегося в комнате свое, отличное от других, число знакомых. Иными словами, хотя бы для одного значения $k \in \{0, 1, \ldots, n-1\}$ в комнате нет человека с k знакомыми, и количество различных значений для k меньше числа n. Следовательно, согласно принципу Дирихле, хотя бы для двух из n находящихся в комнате человек значение k одно и то же. $\square$

Данная задача имеет многочисленные переформулировки. Например, если в комнату заходят n человек, и кто-то с кем-то здоровается за руку, то существует по крайней мере два человека, которые пожали одинаковое количество рук. Или если в шахматном турнире участвуют n человек, и любые два участника играют только по одной партии друг против друга, то в любой момент времени существует по меньшей мере два игрока, которые завершили одинаковое количество партий. Легче же всего эта задача формулируется на языке теории графов. Именно, с точки зрения теории графов все эти задачи равносильны утверждению о том, что в графе на n вершинах хотя бы две вершины имеют одинаковую степень.

1.1.6. Большое количество примеров на применение принципа Дирихле можно найти в разного рода геометрических задачах [**?**]. Приведем достаточно характерный пример такого рода задач.

Пример 1.5. В прямоугольнике со сторонами 6×8 сантиметров помещены пять точек. Докажите, что существуют хотя бы две точки, расстояние между которыми меньше или равно пяти сантиметрам.

Доказательство. Действительно, поделим исходный прямоугольник на четыре равные части размерами 3×4 сантиметра. Так как пять точек должны находиться либо внутри, либо на границах этих четырех прямоугольников, то, согласно принципу Дирихле, хотя бы две из них должны лежать либо внутри, либо на границах одного и того же прямоугольника. Осталось заметить, что расстояние между любыми такими точками меньше или равно пяти сантиметрам. $\square$

1.1.7. В заключение данного пункта приведем еще один интересный пример на принцип Дирихле, который предложили венгерские математики Эрдеш и Секереш.

Пример 1.6. Выделим в множестве $[2n]$ первых $2n$ целых положительных чисел подмножество $S \subset [2n]$ мощности $(n+1)$. Докажите, что в этом подмножестве существуют хотя бы два числа, одно из которых делит другое.

Решение. Для применения принципа Дирихле нам нужно каким-то образом связать голубей с элементами $(n+1)$-элементного подмножества S, а клетки — с элементами какого-то n-элементного множества N. Оказывается, что в рассматриваемом примере нам в качестве множества N нужно выбрать n-элементное подмножество всех нечетных чисел.

Действительно, ключевое для решения данной задачи наблюдение состоит в том, что любое число $s \in S$ может быть представлено в виде

$$s = 2^r \cdot q,$$

где q есть нечетное число, $1 \leqslant q \leqslant 2n-1$. Так как таких чисел s ровно $n+1$ штук, то, согласно принципу Дирихле, у нас обязательно найдутся среди них два числа с одинаковым нечетным делителем q. Как следствие, одно из этих чисел обязательно делит второе.

Данная задача допускает еще несколько решений. Одно из них базируется на построении дерева с корнем в вершине 1 по следующему правилу. Соединим все простые числа ребрами с единицей. Для любого другого числа соединим его ребром с наибольшим делителем этого числа. Заметим, что листьев в таком дереве не более n — это следует из того, что любое число i в диапазоне от 1 до n является делителем по крайней мере одного числа из диапазона от 1 до $2n$, а именно, числа $2i$. Так как нам задано $n+1$ число, то хотя бы два из них будут принадлежать ветке, имеющей общий лист, а значит, будут делиться одно на другое.

Еще одно решение является некоторой комбинацией двух предыдущих. Именно, разобъем множество чисел на цепочки

$$1, 2, 4, 8, \ldots, \qquad 3, 6, 12, 24, \ldots, \qquad 5, 10, 20, 40, \ldots, \qquad 7, 14, 28, \ldots$$

в каждой из которых любое число делится на предыдущее. Заметим, что любая такая цепочка может начинаться только лишь с нечетного числа — любое четное число $2a$ мы можем поделить на два и получить число a, делящее $2a$. Следовательно, количество таких цепочек совпадает с количеством нечетных чисел и равно n. Тогда какие-то два из $(n+1)$-го числа обязательно попадут в одну цепочку, а значит, одно из них будет делиться на второе.

1.2. Часто в задачах нужно определить минимальное количество k предметов, гарантированно обеспечивающих наличие хотя бы r предметов в одном из n ящиков.

1.2.1. Очевидно, что в случае $r = 2$ мы должны взять $n+1$ предмет — действительно, согласно принципу Дирихле, в этом случае хотя бы в одном ящике гарантированно окажется по крайней мере два предмета.

В более общем случае нам достаточно взять

$$k = n \cdot (r-1) + 1 \tag{1}$$

предметов. Действительно, $n \cdot (r-1)$ предмет мы еще можем распределить по n ящикам так, чтобы в каждом ящике находилось не более $(r-1)$-го предмета, а вот в случае $k = n \cdot (r-1) + 1$ у нас хотя бы в одном ящике гарантированно окажется хотя бы r предметов.

Последнее утверждение часто называют обобщенным принципом Дирихле и формулируют его так. Пусть у нас имеются k предметов, которые мы должны распределить по $n < k$ ящикам. Тогда существует по крайней мере один ящик, в котором содержится не менее чем

$$\lceil k/n \rceil$$

предметов, где $\lceil k/n \rceil$ — ближайшее целое число, большее или равное k/n.

1.2.2. В качестве примера на использование обобщенного принципа Дирихле рассмотрим следующую задачу.

Пример 1.7 (Эрдеш, Секереш)**.** Докажите, что любая последовательность из $n^2 + 1$ целых чисел содержит либо убывающую, либо возрастающую подпоследовательность, состоящую из не менее чем $(n+1)$-го числа.

Для доказательства данного утверждения введем количество s_i чисел в самой длинной возрастающей подпоследовательности, начинающейся с a_i. Например, для последовательности чисел

$$\{a_1, a_2, a_3, a_4, a_5\} = \{1, 5, 3, 2, 4\} \qquad s_1 = 3,\ s_2 = 1,\ s_3 = 2,\ s_4 = 2,\ s_5 = 1.$$

Сразу заметим, что если существует такое i, что $s_i \geqslant (n+1)$, то процесс завершен. В нашем случае $s_1 = 3 = (2+1)$, и поэтому в нашем случае все в порядке — эта последовательность содержит возрастающую подпоследовательность $\{1, 3, 4\}$ длины три. Для подпоследовательности же

$$\{a_1, a_2, a_3, a_4, a_5\} = \{1, 5, 4, 3, 2\} \qquad s_1 = 2,\ s_2 = 1,\ s_3 = 1,\ s_4 = 1,\ s_5 = 1,$$

и такого i не существует. Иными словами, в этом случае все (n^2+1) штук s_i принимают значения из диапазона $[1, n]$. Тогда, согласно обобщенному принципу Дирихле, существует по крайней мере

$$\left\lceil \frac{n^2+1}{n} \right\rceil = n + 1$$

чисел s_i, значения которых одинаковы и равны s. Все, что нам осталось понять — это то, что соответствующая этим значениям $s_i = s$ подпоследовательность $\{a_i\}$, состоящая из $(n+1)$-го числа, является убывающей.

Действительно, предположим, что в этой подпоследовательности существует такая пара индексов (i, j), что $j > i$ и $a_j > a_i$. Напомним, что по определению числа s_j, возрастающая подпоследовательность чисел, начинающаяся с a_j имеет длину $s_j = s$. Добавляя a_i в начало этой подпоследовательности, мы получим возрастающую последовательность длины $s+1$, начинающуюся с a_i. Это, в свою очередь, противоречит тому, что наибольшая числовая возрастающая последовательность, начинающаяся с a_i, имеет длину, равную $s_i = s$.

2 Основные понятия теории множеств

2.1. Перейдем теперь к описанию основных понятий теории множеств.

2.1.1. Само понятие множества — одно из самых фундаментальных понятий математики, и строго определить его достаточно сложно. Нам будет достаточно использовать следующее определение.

Определение 2.1. Множеством $X = \{x_1, x_2, \ldots, x_n\}$ называется совокупность различимых объектов x_i, $i = 1, \ldots, n$, объединенных по некоторому признаку. Иными словами, это совокупность каких-то объектов, которые мы полагаем принципиально различными и про которые мы можем четко сказать, входят они в данное множество или нет.

В качестве характерного примера можно рассмотреть, например, множество X всех студентов, находящихся в данной аудитории. Действительно, все студенты различимы, отличны друг от друга и объединены по признаку “собрались в данной аудитории”. Преподаватель, находящийся в той же аудитории, в множество X уже не входит — он студентом не является.

Объекты x_i, из которых состоит множество X, называются его *элементами.* Важно заметить, что множество, по сути, полностью определяется входящими в него элементами.

Если x — элемент множества X, то этот факт записывают как $x \in X$ и говорят, что x *принадлежит* множеству X. Если y множеству X не принадлежит, то это записывается как $y \notin X$. Для любого заданного объекта и любого заданного множества одно из этих утверждений (принадлежит или не принадлежит) обязательно имеет место.

Определение 2.2. *Пустым множеством* $\varnothing$ называется множество, не содержащее никаких элементов.

Определение 2.3. Говорят, что множество Y является *подмножеством* множества X, если каждый элемент $y \in Y$ является элементом множества X. Тот факт, что Y является подмножеством X, записывается как $Y \subseteq X$. Подмножество Y, отличное от X, называется *собственным* подмножеством множества X, что записывается как $Y \subset X$. Если Y также отлично и от $\varnothing$, то оно называется *нетривиальным* подмножеством.

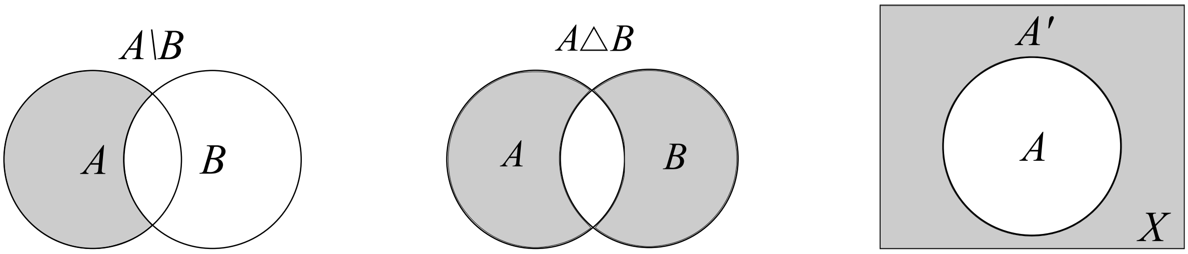
Определение 2.4. Говорят, что множество X является *одноэлементным* множеством ($X = \{a\}$), если $X \neq \varnothing$ и не имеет нетривиальных подмножеств. Множество X называется *двуэлементным* множеством, если оно содержит нетривиальное одноэлементное подмножество и любое нетривиальное подмножество является одноэлементным подмножеством. Множество X называется $(n+1)$-*элементным,* $n = 1, 2, \ldots$, если оно содержит нетривиальное n-элементное подмножество и любое его нетривиальное подмножество является k-элементным подмножеством, $k = 1, 2, \ldots, n$.

Определение 2.5. Мощностью $|X|$ множества X называется количество элементов в нем. Как правило, мы будем рассматривать конечные множества, в которых $|X| = n$, $n \in \mathbb{N}$, и называть их n-множествами.

2.1.2. Основные операции над множествами — это объединение $A \cup B$ (рис.1,a), пересечение $A \cap B$ (рис.1,b), разность $A \setminus B$ (рис.2,a) и симметрическая разность (рис.2,b) двух множеств A и B. В случае, если множество A является подмножеством некоторого более широкого множества X, удобно также рассматривать операцию дополнения $A' := X \setminus A$ множества A (рис.2,c).

(a) Объединение (b) Пересечение

Рис. 1



(a) Разность (b) Симметрическая разность (c) Дополнение

Рис. 2

Свойства операций над множествами удобно изучать графически, с использованием так называемых *диаграмм Эйлера-Венна* (смотри рисунки 1 и 2). Например, с их помощью достаточно просто проиллюстрировать справедливость законов де Моргана

$$A' \cap B' = (A \cup B)', \qquad A' \cup B' = (A \cap B)' \tag{2}$$

(смотри рис.3).

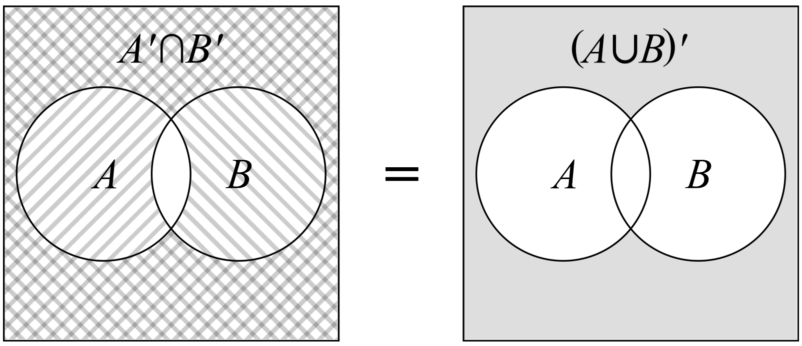


Рис. 3: Графическое доказательство закона де Моргана

2.1.3. В дальнейшем мы достаточно часто будем использовать понятие покрытия множества X семейством $\{X_1, X_2, \ldots, X_k\}$ множеств, а также всевозможные частные случаи этого понятия.

Определение 2.6. *Системой множеств* называется множество, элементами которого являются другие множества. Характерным примером системы множеств является множество всех подмножеств данного множества.

Определение 2.7. *Семейством множеств* $\{X_1, X_2, \ldots, X_k\}$ называется индексированный набор множеств, то есть набор множеств, каждому из которых приписано некоторое натуральное число i, $i = 1, \ldots, k$. Важно заметить, что разным числам может соответствовать одно и то же множество, то есть возможен случай $X_i = X_j$, $i \neq j$.

Определение 2.8. Семейство множеств $\{X_1, X_2, \ldots, X_k\}$ называется *покрытием* множества X, если их объединение дает нам все множество X:

$$X_1 \cup X_2 \cup \ldots \cup X_k = X.$$

Важным частным случаем покрытия является понятие разбиения множества.

Определение 2.9. Говорят, что семейство множеств $\{X_1, X_2, \ldots, X_k\}$ образует *разбиение* множества X, если

1. множества $X_i \neq \varnothing$, $i = 1, \ldots, n$;
2. $X_i \cap X_j = \varnothing$ $\forall i \neq j$;
3. $X_1 \cup X_2 \cup \ldots \cup X_k = X$.

Элементы X_i этого семейства называются *блоками* разбиения.

В качестве характерного примера можно рассмотреть разбиение студентов данного курса на группы. Студенческие группы являются при этом блоками данного разбиения.

Если по каким-то причинам оказывается важным порядок следования блоков, то говорят об *упорядоченном разбиении* $(X_1, X_2, \ldots, X_k)$ множества X. Например, если мы выводим группы на сцену для вручения им дипломов, то важен порядок, в котором они туда выходят. Следовательно, в данном случае мы получаем упорядоченное разбиение студентов данного курса.

Наконец, еще одним частным случаем покрытия X семейством множеств $\{X_1, X_2, \ldots, X_k\}$ является понятие *разделения* множества X. Разделение есть аналог упорядоченного разбиения, в котором допускаются пустые блоки. Точное определение таково:

Определение 2.10. Разделением множества X называется упорядоченная последовательность $(X_1, X_2, \ldots, X_k)$ возможно пустых, попарно непересекающихся множеств, объединение которых дает все множество X.

2.1.4. Еще одной часто использующейся в комбинаторике операцией над множествами является операция декартова произведения множеств.

Определение 2.11. Парой (или, более точно, упорядоченной парой) называется упорядоченный набор (x, y), состоящий из двух элементов x и y. Упорядоченность в этом определении отражает тот факт, что порядок, в котором появляются объекты x и y, важна — в случае, если x и y обозначают разные объекты, то (x, y) и (y, x) — это разные пары. Более того, если $x = y$, то (x, x) — это по-прежнему пара элементов, не равная $\{x\}$. Аксиоматически пару элементов мы можем определить как объект, для которого выполняется следующее свойство:

$$(x_1, y_1) = (x_2, y_2) \qquad \Longleftrightarrow \qquad x_1 = x_2 \quad \text{и} \quad y_1 = y_2. \tag{3}$$

Замечание 2.12. Существуют чуть более формальные и чуть менее интуитивные определения пары элементов. Например, Куратовский предложил следующее определение пары элементов в терминах теории множеств:

$$(x, y) = \{x, \{x, y\}\}.$$

Несложно убедиться в том, что при таком определении свойство (3) выполняется.

Замечание 2.13. Понятие пары можно обобщить, введя понятие тройки (x, y, z) элементов x, y, z, а также произвольной n-ки элементов $(x_1, \ldots, x_n)$ как упорядоченной последовательности из n элементов.

Определение 2.14. Декартовым произведением множеств A и B называется множество

$$A \times B := \{(a, b) \,|\, a \in A,\ b \in B\}$$

всех упорядоченных пар (a, b), таких, что $a \in A$, $b \in B$.

В качестве простейшего примера декартова произведения множеств обычно приводят шахматную доску. Любая клетка шахматной доски имеет координаты "буква-цифра", например, $e5$ или $h4$. Иными словами, координаты клеток шахматной доски являются элементами декартова произведения множеств $A = \{a, b, \ldots, h\}$ и $B = \{1, 2, \ldots, 8\}$.

В частном случае множества A и B могут совпадать. В этом случае декартово произведение $A \times A$ обозначается через A^2.

Определение 2.15. Декартовым произведением k множеств $X_1, X_2, \ldots, X_k$ называется множество

$$X_1 \times X_2 \times \ldots \times X_k := \{(x_1, x_2, \ldots, x_k) \,|\, x_i \in X_i,\ \forall\ i = 1, \ldots, k\}$$

всевозможных упорядоченных k-элементных последовательностей вида $(x_1, x_2, \ldots, x_k)$.

В частном случае $X_1 = X_2 = \ldots = X_k = X$ имеем декартову степень $X \times X \times \ldots \times X =: X^k$.

2.1.5. С понятием декартова произведения связано и еще одно очень важное понятие — понятие отношения. С этим понятием мы достаточно часто встречаемся в практической жизни. Например, я и ручка находятся в отношении собственности, а я и самолет — нет. Этот пример наталкивает нас на следующее определение: отношение — это есть некоторое свойство, определенное на множестве упорядоченных пар (x, y) элементов двух множеств X и Y, позволяющее нам сказать, состоит ли первый элемент этой пары со вторым элементом в рассматриваемом отношении или нет. Более формально это понятие определяется так.

Определение 2.16. Пусть X, Y — пара множеств. *Бинарным отношением* между X и Y называется некоторое подмножество ω декартова произведения $X \times Y$ этих множеств X и Y. При этом говорят, что x и y находятся в отношении ω, если пара $(x, y) \in \omega$. Часто для элемента x, находящегося в отношении ω к y, используется обозначение $x\omega y$. В случае $X = Y$ говорят про бинарное отношение на X.

Нам также достаточно часто будет встречаться важный частный случай бинарного отношения — отношение эквивалентности.

Определение 2.17. Бинарное отношение $\sim$ на X называется отношением эквивалентности, если оно

1) *рефлексивно,* то есть для любого $x \in X$ элемент x соответствует самому себе: $x \sim x$;

2) *симметрично,* то есть для любых $x, y \in X$ из условия $x \sim y$ следует $y \sim x$;

3) *транзитивно,* то есть для любых $x, y, z \in X$ из условий $x \sim y$ и $y \sim z$ следует $x \sim z$.

Пример 2.18. Отношение родства, отношение “жить в одном доме”, а также отношение “родиться в один и тот же день” являются отношениями эквивалентности. Еще одним тривиальным примером примером является отношение равенства на множестве целых чисел. Отношение же неравенства $\neq$ на множестве $\mathbb{Z}$ отношением эквивалентности не является — для него не выполняется свойство транзитивности ($1 \neq 2$, $2 \neq 1$, и при этом $1 = 1$). Не выполняется аксиома транзитивности и для отношения знакомства.

Определение 2.19. Подмножество

$$\widetilde{x} := \{y \in X \mid y \sim x\}$$

всех элементов y, эквивалентных данному элементу x, называется классом эквивалентности, порожденным элементом x. Любой элемент $y \in \widetilde{x}$ называется представителем этого класса $\widetilde{x}$.

Пример 2.20. Для отношения эквивалентности “жить в одном доме” класс эквивалентности — это множество жильцов одного и того же дома.

Из свойства транзитивности отношения эквивалентности следует, что в случае $y \in \widetilde{x}$ классы эквивалентности $\widetilde{x}$ и $\widetilde{y}$ совпадают. Давайте докажем чуть более сильный результат.

Утверждение 2.21. *Множество всех классов эквивалентности образует* разбиение множества X на блоки $\widetilde{x}_i$, *то есть представляет собой объединение непустых, попарно непересекающихся между собой подмножеств* $\widetilde{x}_i$, *объединение которых дает нам все множество* X.

Доказательство. Так как любой элемент x принадлежит классу эквивалентности $\bar{x}$, который он порождает, то

$$X = \bigcup_{x \in X} \widetilde{x}.$$

Далее, любой класс эквивалентности $\widetilde{x}$ не пуст — по крайней мере, он содержит элемент x. Наконец, предположим, что классы эквивалентности $\widetilde{x}$ и $\widetilde{y}$ имеют непустое пересечение:

$$\widetilde{x} \cap \widetilde{y} \neq \varnothing \qquad \Longleftrightarrow \qquad \exists z \in X \colon \quad z \in \widetilde{x} \cap \widetilde{y}.$$

Из условия $z \in \widetilde{x}$ следует, что $\widetilde{z} = \widetilde{x}$, а из условия $z \in \widetilde{y}$ — что $\widetilde{z} = \widetilde{y}$. Следовательно, если два смежных класса имеют непустое пересечение, то они совпадают: $\widetilde{x} = \widetilde{y}$. □

Замечание 2.22. Обратно, любое разбиение $\pi(X)$ множества X на блоки $\widetilde{x}_i$ порождает некоторое отношение эквивалентности $\sim$, по отношению к которому $\widetilde{x}_i$ будут классами эквивалентности.

Доказательство. Будем считать, что $x \sim y$ в случае, если элементы x и y относятся к одному и тому же блоку $\widetilde{x}_i$ разбиения $\pi(X)$. Ясно, что это отношение рефлексивно, симметрично и транзитивно, то есть является отношением эквивалентности, причем классы эквивалентности совпадают с $\widetilde{x}_i$. □

2.1.6. Любой элемент X^k есть упорядоченный набор из k элементов множества X, в котором некоторые элементы могут повторяться. Если же в таком k-множестве порядок элементов не важен, говорят о *мультимножестве над множеством* X. Формальное определение мультимножества таково.

Определение 2.23. Мультимножеством над n-элементным множеством X называется пара (X, φ), где $\varphi \colon X \to \mathbb{Z}_+$ есть функция, сопоставляющая любому элементу $x \in X$ количество $\varphi(x)$ его вхождений в мультимножество.

Любую функцию φ такого рода можно определить с помощью множества Ξ упорядоченных пар

$$\Xi := \{(x, \varphi(x)) \mid x \in X\}.$$

Поэтому, например, мультимножество над множеством $X = \{x, y\}$, состоящее из двух элементов x и одного элемента y, можно формально записывать в виде $(X, \{(x, 2), (y, 1)\})$. Однако чаще вместо такой формальной записи используют более наглядную форму записи вида $\{x, x, y\}$.

Самый простой и понятный пример мультимножества — это монеты в кошельке. В этом примере в качестве множества X выступает множество из девяти монет разного достоинства:

$$X = \{1\text{ копейка}, 5\text{ копеек}, 10\text{ копеек}, 50\text{ копеек}, 1\text{ рубль}, 2\text{ рубля}, 5\text{ рублей}, 10\text{ рублей}\}.$$

Любой набор из этих монет в количестве k штук образует k-мультимножество над множеством X.

3 Основные правила перечислительной комбинаторики

3.1. Теория множеств как раздел математики создавалась значительно позже комбинаторики. Поэтому некоторые наиболее важные понятия теории множеств исторически получили в комбинаторике свои, специфические названия. Именно,

1. k-сочетанием без повторений называется любое k-элементное подмножество n-элементного множества;
2. k-сочетанием с повторениями называется любое k-мультимножество над n-множеством;
3. k-перестановкой без повторений называется упорядоченное k-подмножество n-элементного множества;
4. k-перестановкой с повторениями называется любой элемент декартовой степени X^k.

3.2. Теперь перейдем к двум самым простым, но в то же время достаточно важным правилам перечислительной комбинаторики — правилу суммы и правилу произведения.

3.2.1. Начнем с простейшего примера: пусть на одном блюде лежат три яблока, а на втором — две груши; сколькими способами можно выбрать один фрукт? Ответ очевиден: пятью способами.

Обобщающее этот пример простейшее *правило суммы* можно сформулировать так: если некоторый объект из множества A можно выбрать k способами, и, вне зависимости от выбора этого объекта, можно n способами выбрать некоторый элемент множества B, то выбор объекта из множества A *или* из множества B можно осуществить $k + n$ способами.

Очевидна переформулировка этого правила на языке теории множеств: пусть пересечение двух множеств A и B пусто; тогда

$$|A \cup B| = |A| + |B|.$$

В частности, если $A \subset X$ и A' — дополнение множества A, то

$$|A| + |A'| = |X|. \tag{4}$$

В более общем случае, рассматривая произвольное разбиение множества X на блоки, имеем равенство вида
$$|X| = |X_1| + |X_2| + \ldots + |X_k|,$$
которое также называется правилом суммы в комбинаторике.

3.2.2. Под правилом произведения в комбинаторике понимается равенство
$$|X_1 \times X_2 \times \ldots \times X_k| = |X_1| \cdot |X_2| \cdot \ldots \cdot |X_k|.$$

Приведем простейший пример на применение этого правила в комбинаторике. Пусть в аудитории находятся 32 студента одной группы, 24 студента другой группы и 17 студентов третьей группы. В этом случае тройку, состоящую из представителей каждой группы, можно выбрать $32 \cdot 24 \cdot 17$ способами.

3.3. Наряду с правилом суммы, в элементарной комбинаторике также достаточно часто используется и несложное его обобщение — так называемый *принцип включения-исключения*. Если правило суммы связано с разбиением множества X, то принцип включения-исключения связан с некоторым произвольным покрытием этого n-множества семейством $\{X_1, X_2, \ldots, X_k\}$. Сформулируем его для самого простого случая двух множеств.

3.3.1. Рассмотрим два конечных множества A и B, пересечение которых может быть и непусто. Тогда количество элементов в объединении этих множеств, очевидно, равно
$$|A \cup B| = |A| + |B| - |A \cap B|. \tag{5}$$
Действительно, когда мы считаем количество $|A|$ элементов в множестве A и складываем его с количеством $|B|$ элементов в множестве B, мы любой элемент, принадлежащий как множеству A, так и множеству B, считаем дважды. Чтобы этот избыток убрать, нам нужно один раз вычесть количество элементов, содержащихся в пересечении этих двух множеств.

Равенство (5) можно называть обобщенным правилом суммы — оно обобщает правило суммы на случай, когда пересечение двух множеств не пусто.

3.3.2. Предположим теперь, что A и B являются подмножествами некоторого более широкого множества X. В этом случае у множества $A \subset X$ и множества $B \subset X$ имеются дополнения к ним — множества A' и B', причем $A \cup A' = B \cup B' = X$.

Рассмотрим теперь пересечение $A' \cap B'$ дополнений множеств A и B. Согласно одной из теорем де Моргана (2), $A' \cap B' = (A \cup B)'$. Следовательно, количество элементов в этом пересечении с учетом равенства (4) и обобщенного правила суммы (5) можно сосчитать так:
$$|A' \cap B'| = |(A \cup B)'| = |X| - |A \cup B| = |X| - |A| - |B| + |A \cap B|.$$
Равенство
$$|A' \cap B'| = |X| - |A| - |B| + |A \cap B|, \tag{6}$$
и называется в комбинаторике принципом включения-исключения.

3.3.3. Приведем простейший пример его использования. Пусть в аудитории находятся 30 человек, 20 человек из которых знают английский, 12 — французский, а 6 человек знают оба языка. Сколько человек не знает ни один из этих иностранных языков? Ответ, согласно принципу включения-исключения (6), следующий:
$$N = 30 - 20 - 12 + 6 = 4.$$

3.3.4. Несложно обобщить равенства (5) и (6) на случай большего количества множеств. Например, для трех множеств A, B, C соответствующие формулы выглядят так:

а) обобщенное правило суммы:

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|; \qquad (7)$$

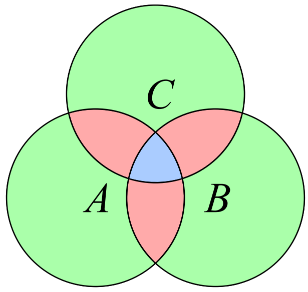


Рис. 4: Диаграмма Эйлера-Венна для трёх множеств

б) принцип включения-исключения:

$$|A' \cap B' \cap C'| = |X| - |A| - |B| - |C| + |A \cap B| + |A \cap C| + |B \cap C| - |A \cap B \cap C|. \qquad (8)$$

Действительно, рассмотрим, к примеру, левую часть равенства (7). Она подсчитывает количество элементов, принадлежащих объединению трех множеств (смотри рис.4). Если элемент x_1, принадлежащий этому объединению, содержится в множестве A, но не содержится в множествах B и C, то он один раз подсчитывается в правой части равенства (7) (слагаемое $|A|$, отвечающее зеленой подобласти на рис.4). Если элемент x_2 принадлежит множествам A и B, но не принадлежит C (красная подобласть на рис.4), то в правой части (7) этот элемент входит в слагаемые $|A|$, $|B|$ и $-|A \cap B|$, то есть также подсчитывается ровно один раз. Наконец, если x_3 принадлежит пересечению всех трех множеств (зеленая подобласть на рис.4), то за этот элемент отвечают все слагаемые в правой части (7). Так как четыре из них входят со знаком плюс, а три — со знаком минус, то этот элемент также считается в правой части (7) лишь однажды.

4 Подсчет k-сочетаний из n элементов. Биномиальные коэффициенты.

4.1. Основная задача данного параграфа состоит в подсчете количества всех k-сочетаний из n элементов. Начнем мы с подсчета количества k-сочетаний из n элементов без повторений.

4.1.1. Число k-сочетаний без повторений известно в литературе под названием биномиальных коэффициентов. Ранее в советской литературе такие числа обозначались через C_n^k. В настоящее время для этих коэффициентов используется обозначение $\binom{n}{k}$ (читается “из n по k”).

4.1.2. Обычно на вопрос, чему равны биномиальные коэффициенты, вспоминают формулу

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}.$$

Эта формула не очень удачна с двух точек зрения — с вычислительной и с идейной. С вычислительной точки зрения ее затруднительно использовать для достаточно больших значений n и k. Например, при $n = 38$ и $k = 19$ числитель и знаменатель могут просто не уместиться в диапазон изменения целых чисел для того или иного языка программирования. С идейной же точки зрения эта формула неудобна потому, что она не позволяет обобщить понятие биномиальных коэффициентов на случай целых, вещественных или комплексных значений n. В следующем параграфе мы получим более удобное выражение для этих коэффициентов, допускающее такое обобщение. Здесь же мы с помощью правила суммы выведем рекуррентное соотношение для биномиальных коэффициентов, позволяющее эти биномиальные коэффициенты эффективно вычислять для достаточно больших значений n и k.

4.1.3. Для получения рекуррентного соотношения введем множество Σ_k всех k-элементных подмножеств n-элементного множества X. Например, для $X = \{x_1, x_2, x_3\}$ множество $\Sigma_2 = \{\{x_1, x_2\}, \{x_1, x_3\}, \{x_2, x_3\}\}$. Заметим, что биномиальный коэффициент $\binom{n}{k}$ как раз и описывает мощность такого множества. Разобьем теперь множество Σ_k на два блока — блок $\Sigma_k^{(1)}$, k-элементные подмножества которого содержат элемент x_1, и блок $\Sigma_k^{(2)}$, подмножества которого этот элемент не содержат. Понятно, что это — непустые, непересекающиеся подмножества, объединение которых дает нам все множество Σ_k. Поэтому по правилу суммы мы получаем равенство вида

$$\binom{n}{k} = |\Sigma_k| = |\Sigma_k^{(1)}| + |\Sigma_k^{(2)}|.$$

Осталось выразить через биномиальные коэффициенты количество элементов в каждом из блоков $\Sigma_k^{(1)}$, $\Sigma_k^{(2)}$. А это делается довольно легко.

4.1.4. Действительно, во всех подмножествах первого блока элемент x_1 уже выбран, и нам из оставшегося $(n-1)$-элементного множества $X \setminus x_1$ остается выбрать $(k-1)$-элементные подмножества . Сделать это можно $\binom{n-1}{k-1}$ способами. Во втором блоке содержатся k-элементные подмножества множества $X \setminus x_1$, состоящего из $(n-1)$-го элемента. Их количество, очевидно, равно $\binom{n-1}{k}$. Таким образом окончательно имеем следующее рекуррентное соотношение:

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \qquad k \geqslant 1, \quad n \geqslant 1. \tag{9}$$

4.1.5. Соотношение (9) следует дополнить начальными и граничными условиями. Так как k-элементных подмножеств n-элементного множества в случае $k > n$ не существует, то

$$\binom{n}{k} = 0 \qquad \text{при} \quad k > n.$$

Далее, пустое подмножество можно выбрать всегда и только одним способом; поэтому

$$\binom{n}{0} = 1 \qquad \forall\ n \geqslant 0.$$

Используя эти условия, можно шаг за шагом вычислить коэффициенты $\binom{n}{k}$. Часто их записывают в виде так называемого треугольника Паскаля:

$$\begin{array}{ccccccccc} & & & & 1 & & & & \\ & & & 1 & & 1 & & & \\ & & 1 & & 2 & & 1 & & \\ & 1 & & 3 & & 3 & & 1 & \\ 1 & & 4 & & 6 & & 4 & & 1 \\ & & & & \dots & & & & \end{array}$$

4.1.6. Как видно, треугольник Паскаля симметричен, т.е. $\binom{n}{k} = \binom{n}{n-k}$. Комбинаторное доказательство этого факта очевидно. Действительно, выбирая любое k-элементное множество, мы тем самым однозначно выбираем и дополнение к нему, т.е. $(n-k)$-элементное множество. Следовательно, количество k-элементных и $(n-k)$-элементных подмножеств совпадает.

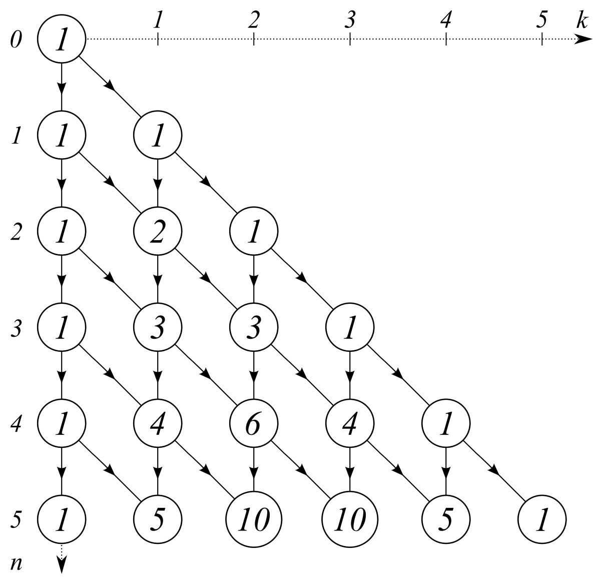


Рис. 5: Графическое представление чисел $\binom{n}{k}$ на координатной плоскости (n, k)

4.2. Наряду с треугольником Паскаля биномиальные коэффициенты допускают и еще одно очень удобное графическое представление — представление на координатной плоскости (n, k)

4.2.1. Данное представление связано со следующей довольно интересной комбинаторной задачей. Рассмотрим плоскость (n, k) и нарисуем на этой плоскости пути, исходящие из начала координат, приходящие в точку с координатами (n, k), $n \geqslant 0$, $k = 0, \dots, n$, и состоящие из диагональных и вертикальных отрезков (рис.5). В самих точках (n, k) отметим количество таких путей, приходящих в эту точку из начала координат. Заметим теперь, что попасть в точку с координатами (n, k) мы можем, пройдя только лишь через точку с координатами $(n-1, k)$ или через точку с координатами $(n-1, k-1)$. С комбинаторной точки зрения это означает, что количество путей $\binom{n}{k}$, приходящих в точку с координатами (n, k), равняется количеству $\binom{n-1}{k}$ путей, приходящих в точку с координатами $(n-1, k)$, плюс количество путей $\binom{n-1}{k-1}$, приходящих в точку с координатами $(n-1, k-1)$. Но равенство

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$$

есть основное тождество для биномиальных коэффициентов (9). А это означает, что мы получили новую комбинаторную интерпретацию таких коэффициентов. Именно, числа $\binom{n}{k}$ описывают количество путей, состоящих из диагональных $(1,1)$ и вертикальных $(1,0)$ отрезков, выходящих из начала координат — точки $(0,0)$, и оканчивающихся в точке с координатами (n,k).

4.2.2. Графическое представление чисел $\binom{n}{k}$ на плоскости (n,k) очень удобно для получения разного рода тождеств с биномиальными коэффициентами. В качестве первого примера зафиксируем какое-то конкретное значение параметра k и просуммируем биномиальные коэффициенты $\binom{m}{k}$ по m от k до некоторого фиксированного значения n. Например, выберем $k=1$, $n=4$. Складывая числа $1,2,3,4$, мы получим число 10, то есть биномиальный коэффициент, стоящий правее и ниже рассмотренной цепочки биномиальных коэффициентов. Есть подозрение, что данный факт, а именно, равенство вида

$$\binom{k}{k}+\binom{k+1}{k}+\ldots+\binom{n}{k}=\binom{n+1}{k+1} \tag{10}$$

выполняется для любых значений параметров n и k.

4.2.3. Для формального доказательства тождества

$$\sum_{m=0}^{n}\binom{m}{k}=\underbrace{\binom{0}{k}+\binom{1}{k}+\ldots+\binom{k-1}{k}}_{=0}+\binom{k}{k}+\binom{k+1}{k}+\ldots+\binom{n}{k}=\sum_{m=k}^{n}\binom{m}{k}=\binom{n+1}{k+1},$$

называемого формулой суммирования биномиальных коэффициентов по верхнему индексу, применим рекуррентное соотношение (9) к коэффициенту $\binom{m+1}{k+1}$:

$$\binom{m+1}{k+1}=\binom{m}{k}+\binom{m}{k+1} \qquad\Longrightarrow\qquad \binom{m}{k}=\binom{m+1}{k+1}-\binom{m}{k+1}.$$

Просуммируем теперь полученное равенство по m от k до n:

$$\begin{aligned}\sum_{m=k}^{n}\binom{m}{k}&=\binom{n+1}{k+1}+\sum_{m=k}^{n-1}\binom{m+1}{k+1}-\sum_{m=k}^{n}\binom{m}{k+1}=\\&=\binom{n+1}{k+1}+\sum_{m=k}^{n-1}\binom{m+1}{k+1}-\sum_{m=k+1}^{n}\binom{m}{k+1}=\\&=\binom{n+1}{k+1}+\sum_{m'=k+1}^{n}\binom{m'}{k+1}-\sum_{m=k+1}^{n}\binom{m}{k+1}=\binom{n+1}{k+1}.\end{aligned}$$

4.2.4. Комбинаторное доказательство тождества (10) основано на следующем общем подходе: мы разбиваем множество Σ_{k+1} всех $(k+1)$-элементных подмножеств $(n+1)$-элементного множества X на блоки, подсчитываем количество элементов в каждом блоке, а затем пользуемся правилом суммы для подсчета числа $|\Sigma_{k+1}|=\binom{n+1}{k+1}$.

Для реализации этого подхода возьмем $(n+1)$-элементное множество

$$X=\{x_1,x_2,\ldots,x_{k-1},x_k,x_{k+1},\ldots,x_{n-1},x_n,x_{n+1}\}$$

и будем разбивать множество всех $(k+1)$-элементных подмножеств множества X следующим образом. В первый блок поместим все $(k+1)$-элементные подмножества, содержащие элемент

x_{n+1}. Такие подмножества элемент x_{n+1} гарантированно содержат. Нам из оставшихся элементов $x_1, \ldots, x_n$ множества $X \setminus x_{n+1}$ нужно выбрать недостающие k элементов. По определению биномиального коэффициента, сделать мы это можем $\binom{n}{k}$ способами.

Теперь мы рассмотрим все $(k+1)$-элементные подмножества, которые не содержат элемент x_{n+1}, но обязательно содержат элемент x_n. Для того, чтобы любое такое подмножество сформировать, нам нужно из множества элементов $\{x_1, \ldots, x_{n-1}\}$ выбрать недостающие k элементов. Это можно сделать $\binom{n-1}{k}$ количеством способов.

Затем рассмотрим все $(k+1)$-элементные подмножества, которые содержат в обязательном порядке элемент x_{n-1} и не содержат элементов x_{n+1} и x_n. Эти подмножества получаются выбором недостающих k элементов из множества элементов $\{x_1, \ldots, x_{n-2}\}$, и выбрать такие подмножества можно $\binom{n-2}{k}$ способами.

Продолжая далее, мы дойдем когда-то до ситуации, в которой нам нужно выбрать все $(k+{}$ $+1)$-элементные подмножества, которые содержат элемент x_{k+1} и не содержат элементы со старшими индексами. Так как элемент x_{k+1} у нас уже выбран, то нам остается из множества $\{x_1, \ldots, x_k\}$ выбрать k-элементное подмножество. Сделать мы это можем, очевидно, $\binom{k}{k} = 1$ способом.

Складывая теперь количество элементов в каждом блоке, мы и получаем тождество (10).

Пример 4.1. Пусть $n = 4$, $X = \{x_1, x_2, x_3, x_4, x_5\}$, $k = 2$, $k+1 = 3$. Приведем список всех трехэлементных подмножеств этого множества:

$$\{x_1,x_2,x_3\}, \qquad \{x_1,x_2,x_4\}, \qquad \{x_1,x_2,x_5\}, \qquad \{x_1,x_3,x_4\}, \qquad \{x_1,x_3,x_5\},$$

$$\{x_1,x_4,x_5\}, \qquad \{x_2,x_3,x_4\}, \qquad \{x_2,x_3,x_5\}, \qquad \{x_2,x_4,x_5\}, \qquad \{x_3,x_4,x_5\}.$$

В первый блок разбиения этого множества подмножеств включим подмножества, содержащие элемент x_5; таковых имеется $\binom{4}{2} = 6$ штук. Из *оставшегося* списка выберем все подмножества, содержащие x_4; их $\binom{3}{2} = 3$ штуки. Наконец, у нас остается единственное подмножество элементов, не содержащих ни x_4, ни x_5, т.е. подмножество $\{x_1, x_2, x_3, \}$.

4.2.5. В заключение данного пункта докажем комбинаторно еще одно важное тождество для биномиальных коэффициентов — так называемое тождество Вандермонта

$$\binom{n+m}{k} = \sum_{i=0}^{k} \binom{n}{i} \cdot \binom{m}{k-i}. \tag{11}$$

Рассмотрим для этого группу, состоящую из n мужчин и m женщин. По определению, количество способов выбрать из нее команду, состоящую из k человек, равно биномиальному коэффициенту $\binom{n+m}{k}$. С другой стороны, мы можем выбирать эту команду так, чтобы в ней было ровно i мужчин и $(k-i)$ женщин. При фиксированном i количество способов подбора такой команды, согласно правилу произведения, равно $\binom{n}{i} \cdot \binom{m}{k-i}$. Меняя теперь i от нуля до k, мы вновь получим общее количество способов образовать требуемую команду.

Замечание 4.2. Приведенные выше рассуждения очень часто используются в комбинаторике. Они даже имеют специальное название — *принцип double counting* или правило подсчета двумя способами. Основная идея такого рода рассуждений состоит в следующем: если две формулы подсчитывают количество одних и тех же элементов, то эти формулы равны.

4.3. Название “биномиальные коэффициенты” связано с тем, что они, помимо всего прочего, встречаются в формуле бинома Ньютона

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k\, y^{n-k}. \tag{12}$$

4.3.1. Комбинаторное доказательство этой формулы довольно элементарно: нужно просто расписать $(x+y)^n$ в виде произведения n сомножителей

$$(x+y)^n = \underbrace{(x+y)}_{1} \cdot \underbrace{(x+y)}_{2} \cdot \ldots \cdot \underbrace{(x+y)}_{n}$$

и пометить каждый из таких сомножителей числом в диапазоне от единицы до n. В результате мы имеем множество X, состоящее из n *различных* экземпляров сомножителей вида $(x+y)$.

После перемножения этих n скобок получается определенный набор слагаемых вида $x^k y^{n-k}$, $k = 0, 1, \ldots, n$. Для подсчета количества этих слагаемых при фиксированном значении параметра k заметим, что любое слагаемое $x^k y^{n-k}$ можно получить так: выбрать в n-элементном множестве X k-элементное подмножество, взять в этом подмножестве в качестве сомножителей переменные x, а в оставшемся $(n-k)$-элементном подмножестве выбрать в качестве сомножителей переменные y. Как следствие, количество слагаемых $x^k y^{n-k}$ совпадает с количеством способов выбрать k-элементное подмножество n-элементного множества X и равно $\binom{n}{k}$.

4.3.2. Формула (12) оказывается чрезвычайно полезной для вывода разного рода соотношений, связанных с биномиальными коэффициентами. Например, полагая в ней $x = y = 1$, получаем тождество

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n. \tag{13}$$

Заметим теперь, что, суммируя биномиальные коэффициенты $\binom{n}{k}$ по k, мы подсчитываем *все* подмножества n-элементного множества. Иными словами, мы формально доказали тот факт, что количество всех подмножеств данного n-множества равно 2^n.

4.3.3. Для комбинаторного доказательства равенства (13) воспользуемся чрезвычайно полезным и часто использующимся в комбинаторике *принципом биекции*. Формально этот принцип можно сформулировать следующим образом: пусть X, Y — пара конечных множеств, и пусть существует биекция $f\colon X \to Y$, т.е. такое отображение, что

$$\forall\, y \in Y \qquad \exists!\, x \in X : \qquad y = f(x).$$

Тогда количество элементов в множествах X и Y совпадают: $|X| = |Y| = n$.

Неформально использование принципа биекции можно проиллюстрировать на следующем примере. Предположим, что вы устраиваете вечеринку и приглашаете на нее довольно много друзей. Как гостеприимный хозяин, вы встречаете всех своих друзей на входе в дом, но запоминаете только пришедших к вам девушек. В какой-то момент вы решаете подсчитать, сколько парней пришло к вам на вечеринку. Вы знаете количество пришедших к вам девушек, и вам кажется, что количество девушек и парней одинаково. Как вам быстро проверить это предположение? Ответ достаточно очевиден: попросить каждую девушку взять ровно одного парня за руку. Если в результате этой процедуры все множество гостей разбилось на пары, то ваше предположение окажется верным. Тем самым вы сильно упростили себе жизнь — вам не пришлось проделывать довольно утомительную работу по пересчету пришедших к вам парней, вы просто воспользовались для их подсчета результатом уже проделанной работы по пересчету пришедших к вам девушек.

Вернемся теперь к равенству (13). Заметим, что любое подмножество A множества X мы можем закодировать бинарной строкой $f(A)$ длины n, то есть строкой над алфавитом $\{0, 1\}$. Единице на i-м месте строки будет при этом отвечать ситуация, при которой $x_i \in A$, а нулю — вариант, при котором $x_i \notin A$. Так, если множество X имеет вид $\{x_1, x_2, x_3, x_4\}$, а подмножество A записывается в виде $A = \{x_2, x_4\}$, то соответствующая этому подмножеству строка длины 4 равна

$$f(A) = (0, 1, 0, 1).$$

Очевидно, что построенное отображение f взаимно-однозначно. Следовательно, количество подмножеств данного n-множества X совпадает с количеством бинарных строк длины n. А таких строк у нас имеется ровно 2^n штук — действительно, на любое из n мест мы двумя способами можем поставить либо ноль, либо единицу.

4.3.4. Полагая в (12) $x = -1$, $y = 1$, мы получаем еще одно важное тождество

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0, \qquad n > 0. \tag{14}$$

Для того, чтобы понять комбинаторный смысл равенства (14), перепишем его в следующем виде:

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \ldots = 0 \qquad \Longleftrightarrow \qquad \binom{n}{0} + \binom{n}{2} + \ldots = \binom{n}{1} + \binom{n}{3} + \ldots$$

В левой части последнего равенства стоит количество всех подмножеств, в которых содержится четное количество элементов, а в правой — число подмножеств, содержащих нечетное количество элементов. Иными словами, следствием равенства (14) является тот факт, что количество четных подмножеств любого множества, отличного от $\varnothing$, равняется количеству его нечетных подмножеств.

Наконец, из (14) мы также можем заключить, что количество, например, всех нечетных подмножеств ровно в два раза меньше общего количества всех подмножеств. С учетом (13) это означает, что количество всех нечетных подмножеств равно 2^{n-1}.

4.3.5. Наконец, продифференцируем (12) по x:

$$n\,(x + y)^{n-1} = \sum_{k=1}^{n} k \binom{n}{k} x^{k-1}\, y^{n-k}. \tag{15}$$

Подставляя в это равенство $x = y = 1$, получим еще одно полезное равенство:

$$\sum_{k=1}^{n} k \binom{n}{k} = n\, 2^{n-1}. \tag{16}$$

4.4. Перейдем теперь к задачам, связанным с подсчетом k-сочетаний с повторениями.

4.4.1. Начнем с примера. Пусть множество X состоит из двух чисел 1 и 2. Выпишем все 3-сочетания с повторениями из 2-множества X:

$$\{1,1,1\}, \qquad \{1,1,2\}, \qquad \{1,2,2\}, \qquad \{2,2,2\}.$$

Как видно, таковых оказалось 4 штуки. Для подсчета количества $\left(\!\binom{n}{k}\!\right)$ таких мультимножеств в общем случае нам будет удобнее вначале конкретизировать n-множество X, а именно, взять в качестве X множество $[n] := \{1, 2, \ldots, n\}$ первых n натуральных чисел. Любое k-мультимножество такого множества можно записать, очевидно, в следующем виде:

$$1 \leqslant a_1 \leqslant a_2 \leqslant \ldots \leqslant a_k \leqslant n.$$

Например, 3-мультимножество $\{1,1,2\}$ над 2-множеством $X = \{1,2\}$ можно записать так:

$$1 \leqslant (a_1 = 1) \leqslant (a_2 = 1) \leqslant (a_3 = 2) \leqslant (n = 2).$$

Теперь превратим в этой цепочке все нестрогие неравенства в строгие. Для этого мы к a_2 прибавим единицу, к a_3 — двойку, к a_4 — тройку, и так далее. К последнему числу a_n мы, таким образом, добавим число $(k-1)$. В результате получим цепочку строгих неравенств вида

$$1 \leqslant a_1 < a_2 + 1 < a_3 + 2 < a_4 + 3 < \ldots < a_k + (k-1) \leqslant n + (k-1).$$

В нашем примере

$$1 \leqslant (a_1 = 1) < (a_2 + 1 = 2) < (a_3 + 2 = 4) \leqslant (n + (3-1) = 4).$$

Заметим, что в результате этой операции мы получили некоторое k-элементное подмножество *различных* чисел вида $a_i + (i-1)$ множества $\widetilde{X} = [n+k-1]$ всех чисел от единицы до $n+k-1$. Иными словами, мы сопоставили любому k-мультимножеству над множеством $X = [n]$ вполне определенное k-подмножество множества $\widetilde{X} = [n+k-1]$. Очевидно, что это сопоставление взаимно-однозначно. Но количество всех k-подмножеств данного множества мы знаем — оно равно $\binom{n+k-1}{k}$. Следовательно, этому числу равно, по принципу биекции, и количество $\left(\!\binom{n}{k}\!\right)$ всех k-мультимножеств над множеством $X = [n]$:

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k}.$$

Справедливость этого равенства для произвольного n-элементного множества X вновь следует из принципа биекции.

5 k-перестановки из n элементов. Урновые схемы и схемы раскладки предметов по ящикам.

5.1. Перейдем теперь к подсчету количества k-перестановок из n элементов.

5.1.1. Напомним, что k-перестановкой из n элементов называется *упорядоченный* набор

$$(a_1, a_2, \ldots, a_k)$$

элементов, в котором все a_i принадлежат одному и тому же n-элементному множеству X.

Элементы a_i в наборе $(a_1, a_2, \dots, a_k)$ могут как повторяться, так и не повторяться. В первом случае говорят о k-перестановках с повторениями, во втором — о k-перестановках без повторений.

Номер паспорта — это типичный пример k-перестановки с повторениями над множеством из десяти цифр $X = \{0, 1, \dots, 9\}$. Классическим примером 3-перестановки без повторений является упорядоченный список спортсменов, занявших призовые места в любых спортивных соревнованиях.

5.1.2. В литературе встречается довольно много альтернативных названий для данного объекта. Именно, упорядоченный набор $(a_1, a_2, \dots, a_k)$, $a_i \in X$, также иногда называется

– k-размещением из n элементов;
– кортежем из k элементов множества X;
– упорядоченной k-выборкой из n элементов;
– k-мерным вектором над множеством X;
– k-элементным словом над n-элементным алфавитом.

5.1.3. Сосчитаем количество k-перестановок с повторениями.

Утверждение 5.1. *Количество k-перестановок с повторениями из n элементов равно n^k.*

Для доказательства можно либо просто сослаться на правило произведения, либо рассмотреть $(a_1, a_2, \dots, a_k)$, $a_i \in X$ как слово из k элементов над алфавитом из $n = |X|$ букв. На первое место в слове мы можем поставить любую из n букв, на второе — также любую из n букв и так далее. Всего же получаем n^k вариантов записать данное слово.

5.1.4. Перейдем теперь к подсчету количества перестановок без повторений.

Утверждение 5.2. *Количество $P(n, k)$ k-перестановок из n элементов без повторений равно*

$$P(n, k) = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) =: (n)_k.$$

Доказательство очевидно — на первое место в строке длины k я могу поставить любой из n элементов, на второе — любой из оставшихся $(n-1)$ элементов и так далее.

5.1.5. В частном случае $k = n$ k-перестановки из n элементов без повторений называются просто перестановками n-элементного множества X. Их количество равно

$$P_n \equiv P(n) = n!, \qquad P(0) = 0! = 1.$$

5.1.6. Любую k-перестановку из n элементов без повторений можно рассматривать и как упорядоченное k-подмножество n-множества. Количество таких подмножеств мы можем сосчитать следующим образом: мы можем $\binom{n}{k}$ способами выбрать k-подмножество n-элементного множества, а затем $k!$ способами его упорядочить. Действительно, на первое место мы можем поставить любой из k элементов подмножества, на второе — любой из оставшихся $k-1$ элементов и так далее. Таким образом, мы получаем некоторое новое выражение для количества всех упорядоченных k-элементных подмножеств n-множества, то есть количества k-перестановок из n элементов без повторений, равно

$$(n)_k = k! \cdot \binom{n}{k} \qquad \Longrightarrow \qquad \binom{n}{k} = \frac{(n)_k}{k!}.$$

Последняя формула часто используется как некомбинаторное определение биномиальных коэффициентов $\binom{n}{k}$ в случае, когда $k \in \mathbb{Z}$, а n принадлежит $\mathbb{Z}$, $\mathbb{R}$ или даже $\mathbb{C}$. Именно, по определению,

$$\binom{q}{k} := \begin{cases} \dfrac{q\,(q-1)\,\ldots\,(q-k+1)}{k!} =: \dfrac{(q)_k}{k!}, & \text{если } k \in \mathbb{N}, \\ 1, & \text{если } k = 0, \\ 0, & \text{если } k < 0, \end{cases} \qquad \forall\, q \in \mathbb{C}.$$

Например,

$$\binom{-1}{3} = \frac{(-1)\cdot(-2)\cdot(-3)}{3\cdot 2\cdot 1} = -1.$$

5.1.7. Функцию $(q)_k$ часто также обозначают через $q^{\underline{k}}$ и называют *убывающей факториальной степенью* [?]. Наряду с убывающей можно ввести и так называемую *возрастающую факториальную степень*

$$q^{(k)} \equiv q^{\overline{k}} := q\cdot(q+1)\cdot\ldots\cdot(q+k-1).$$

В частности, с ее помощью получается удобное для вычислений выражение для количества $\left(\!\binom{n}{k}\!\right)$ сочетаний с повторениями:

$$\left(\!\!\binom{n}{k}\!\!\right) = \binom{n+k-1}{k} = \frac{n^{(k)}}{k!}. \tag{17}$$

5.2. Итак, мы получили простые соотношения для подсчета количества четырех основных объектов элементарной комбинаторики — k-сочетаний и k-перестановок из n элементов с повторениями и без повторений. Эти объекты встречаются в огромном количестве внешне не очень похожих друг на друга задач элементарной комбинаторики. Оказывается, однако, что большинство этих задач можно свести к одной из двух простейших схем — либо к так называемой урновой схеме, либо к схеме раскладки предметов по ящикам.

5.2.1. В урновой схеме имеется урна, в которой находятся *n различимых* предметов. Из урны последовательно вытаскивается k предметов. Задача состоит в подсчете количества различных способов выбора этих предметов, или, как еще говорят, в подсчете различных k-элементных выборок из n предметов, находящихся в урне.

На практике наиболее часто встречаются четыре модификации этой задачи, различающиеся способами формирования k-элементной выборки. Прежде всего, мы можем возвращать или не возвращать вытаскиваемые предметы обратно в урну. В первом случае говорят о *выборке с повторениями*, во втором — о *выборке без повторений.* Далее, в некоторых задачах нам важен порядок вытаскиваемых предметов. В этом случае имеем так называемые *упорядоченные* выборки. В противном случае выборки называются *неупорядоченными.*

Нетрудно понять, что задачи о подсчете k-элементных выборок представляют собой, по сути, те же самые задачи о подсчете k-перестановок или k-сочетаний из n элементов. Действительно, любая неупорядоченная k-элементная выборка представляет собой либо k-элементное подмножество n-множества, либо k-мультимножество над n-элементным множеством в зависимости от того, возвращаем мы вытаскиваемые предметы обратно в урну или не возвращаем. Следовательно, количество таких неупорядоченных выборок совпадает с коэффициентами $\binom{n}{k}$ или $\left(\!\binom{n}{k}\!\right)$. Очевидно также, что любая упорядоченная k-элементная выборка есть просто некоторая k-перестановка n-элементного множества. Поэтому количество таких выборок равно n^k или $(n)_k$ в зависимости от того, говорим ли мы о выборке с повторениями или без повторений.

В результате получаем следующую таблицу решений задач, связанных с урновыми схемами:

Предметы на выходе	с возвращением	без возвращения
упорядоченные	n^k	$(n)_k$
неупорядоченные	$\left(\!\binom{n}{k}\!\right)$	$\binom{n}{k}$

5.2.2. Приведем несколько характерных примеров, достаточно естественно сводящихся к одной из описанных выше урновых схем.

Пример 5.3. Предположим, что у нас имеется некоторое общество, состоящее из двадцати членов. Сколькими способами можно выбрать президента, вице-президента, секретаря и казначея этого общества?

Решение. Очевидно, что любой способ выбора представляет собой упорядоченное 4-элементное подмножество 20-элементного множества. Следовательно, существует ровно $(20)_4$ различных способов выбора членов общества на эти должности.

Пример 5.4. Для того, чтобы открыть сейф, нужно набрать код из пяти символов с помощью вращающихся дисков. На каждом из этих дисков нанесено 12 символов, одинаковых для каждого из дисков. Сколько вариантов различных кодов существует?

Решение. Любой код представляет собой упорядоченную 5-элементную выборку с повторениями или, иначе, строку из пяти символов над алфавитом из 12 букв. Следовательно, имеется 12^5 вариантов различных кодов.

Пример 5.5. На почте продаются открытки десяти различных видов. Сколькими способами можно купить восемь открыток? А восемь открыток разных видов?

Решение. Понятно, что в первом случае любой набор из восьми открыток представляет собой неупорядоченную выборку с повторениями, а во втором — выборку без повторений из 10 элементов. Следовательно, в первом случае имеем $\left(\!\binom{10}{8}\!\right)$, а во втором — $\binom{10}{8}$ способов покупки восьми открыток.

5.2.3. Второй, не менее популярной в элементарной комбинаторике схемой, связанной с подсчетом количества k-перестановок и k-сочетаний, является схема раскладки предметов по ящикам. В этой схеме имеется n *различимых* ящиков, по которым нужно разложить k *различимых* или *неразличимых* предметов. При этом мы можем накладывать определенные ограничения на количество предметов в каждом ящике.

Рассмотрим, к примеру, задачу о подсчете количества способов раскладки k различимых предметов по n различимым ящикам при условии, что в любой ящик можно класть любое количество предметов. Количество способов совершить эти действия равно, очевидно, n^k. Действительно, любой предмет мы можем положить в любой из n ящиков вне зависимости от того, куда мы положили оставшиеся предметы. Поэтому, согласно правилу произведения, это количество равно n^k. Иными словами, данная задача представляет собой переформулировку задачи о подсчете количества k-перестановок из n элементов с повторениями.

Теперь предположим, что в той же схеме мы не имеем права класть более одного предмета в один ящик. Тогда первый предмет мы можем поместить в любой из n ящиков, второй — в любой из оставшихся свободными $(n-1)$ ящиков и так далее. Всего же получаем $(n)_k$ способов раскладки. Следовательно, данная схема соответствует подсчету k-перестановок без повторений.

5.2.4. Пусть теперь у нас имеются n различимых ящиков и k неразличимых предметов. Тогда подсчет количества различных способов раскладки этих предметов по ящикам сводится к задаче о подсчете количества k-сочетаний из n элементов.

Действительно, в данной схеме в качестве n-элементного множества выступает множество, состоящее из n различимых ящиков. В случае, когда в каждый ящик можно класть не более одного предмета, мы, раскладывая предметы по ящикам, выделяем в этом множестве некоторое k-элементное подмножество. Следовательно, количество таких раскладок совпадает с количеством различных k-элементных подмножеств n-множества и равно $\binom{n}{k}$.

В случае же, когда никаких ограничений на количество предметов в ящике не накладывается, мы, раскладывая по ящикам k неразличимых предметов, задаем тем самым некоторое k-мультимножество n-множества X. Поэтому количество различных способов такой раскладки равно количеству $\left(\!\binom{n}{k}\!\right)$ k-сочетаний из n элементов с повторениями.

Подводя итоги, построим таблицу рассмотренных схем раскладок k предметов по n ящикам:

Предметы	Ящики	Произвольное количество предметов в ящике	Не более одного предмета в ящике
различимые	различимые	n^k	$(n)_k$
неразличимые	различимые	$\left(\!\binom{n}{k}\!\right)$	$\binom{n}{k}$

5.2.5. Проиллюстрируем некоторые характерные примеры задач, которые довольно естественно сводятся к схеме схем раскладки предметов по ящикам.

Пример 5.6. Сколькими способами можно разложить по двум *различимым* карманам (например, левому и правому) девять монет *различного* достоинства?

Решение. Рассматриваемый пример является типичной задачей, которая естественным образом сводится к схеме раскладки предметов по ящикам. Действительно, в роли ящиков здесь выступают левый и правый карман, а в роли предметов — девять различных монет. Поэтому ответ в этой задаче — 2^9 способов.

Замечание 5.7. Рассмотренная задача, однако, не всегда решается верно: в качестве ответа иногда выдают 9^2 способов. Путаница, как правило, происходит потому, что эту задачу пытаются свести к урновой схеме, считая, что имеются урна, в которой расположены 9 различных предметов, а также 2 различимые позиции на выходе.

Для того, чтобы этой путаницы избежать, полезно сформулировать следующие основные отличия схемы раскладки по ящикам от соответствующей ей урновой схемы. Во-первых, в схеме раскладки предметов по ящикам предметы обратно не возвращаются, они остаются в ящике. Во-вторых, в этой схеме в любой ящик можно класть любое количество предметов. В аналогичной урновой схеме на любую позицию помещается ровно один предмет.

Приведем теперь два характерных примера, связанных с раскладкой неразличимых предметов по различимым ящикам.

Пример 5.8. У отца имеется 5 (неразличимых) апельсинов, которые он может раздать восьми своим сыновьям. Если его задача состоит в том, чтобы раздать их максимальному количеству сыновей, то он должен поставить дополнительное условие — любой из его сыновей не должен получить более одного апельсина. В этом случае количество способов, которыми он может раздать своим сыновьям эти пять апельсинов, равно, очевидно, $\binom{8}{5}$. Если же он раздает их по каким-то заслугам, и может, таким образом, любому сыну отдать любое количество апельсинов, то количество способов это сделать равно $\left(\!\binom{8}{5}\!\right) = \binom{12}{5}$.

Пример 5.9. В физике встречаются задачи, в которых имеются n различных уровней энергии и k неразличимых элементарных частиц. Если эти частицы — фермионы, то для них действует так называемый принцип запрета Паули, согласно которому на любом энергетическом уровне может находиться не более одной элементарной частицы. Как следствие, количество различных распределений k фермионов по n энергетическим уровням равно $\binom{n}{k}$. Наряду с фермионами существуют и частицы иного сорта — бозоны, для которых не существует ограничений на количество частиц, занимающих один и тот же уровень энергии. Для бозонов количество таких распределений равно, очевидно, $\left(\!\binom{n}{k}\!\right)$.

5.2.6. К задачам раскладки неразличимых предметов по различимым ящикам, связанным с подсчетом количества k-сочетаний, сводятся также задачи о так называемом *разбиении* натурального числа k на n слагаемых. Данная задача формулируется следующим образом: сколькими способами можно представить натуральное число k в виде суммы n слагаемых вида

$$a_1 + a_2 + \ldots + a_n = k$$

при условии, что порядок слагаемых важен, то есть при условии, что разбиения вида

$$1 + 3 + 3 + 3 = 10 \qquad \text{и} \qquad 3 + 1 + 3 + 3 = 10$$

считаются различными?

Если на числа a_i накладывается единственное условие вида $a_i \geqslant 0$, то количество разбиений равно количеству $\left(\!\binom{n}{k}\!\right)$ k-мультимножеств над n-элементным множеством. Действительно, в упорядоченной сумме $a_1 + a_2 + \ldots + a_n$ любой индекс i слагаемого a_i можно рассматривать как i-й ящик, в который мы складываем a_i единиц. Следовательно, эту задачу можно трактовать как задачу о раскладке k "неразличимых" единиц по n различимым ящикам.

Пример 5.10. Подсчитать количество разбиений числа $k = 4$ на два слагаемых:

$$a_1 + a_2 = 4, \qquad a_1, a_2 \geqslant 0.$$

Ответ: $\left(\!\binom{2}{4}\!\right) = \binom{5}{4} = 5$ разбиений: $0 + 4 = 1 + 3 = 2 + 2 = 3 + 1 = 4 + 0 = 4$.

К подсчету числа k-сочетаний из n элементов без повторений задача о разбиении числа k сводится в случае, когда на числа a_i накладываются следующие условия:

$$a_i = 0 \qquad \text{или} \qquad a_i = 1.$$

В этом случае индекс i также можно трактовать как i-й ящик; его можно выбрать (положив $a_i = 1$) или не выбрать (положив $a_i = 0$). Всего же нужно выбрать k таких ящиков. Это можно сделать $\binom{n}{k}$ способами.

Пример 5.11. Подсчитать количество разбиений числа 2 на три слагаемых при условии, что любое слагаемое может принимать значения 0 или 1.

Ответ: $\binom{3}{2} = 3$ разбиения: $1 + 1 + 0 = 1 + 0 + 1 = 0 + 1 + 1 = 2$.

5.2.7. Достаточно часто на практике встречаются ситуации, когда одну и ту же задачу можно свести и к урновой схеме, и к схеме раскладки предметов по ящикам.

Пример 5.12. В кондитерском магазине продаются пирожные трех разных видов. Сколькими различными способами можно купить семь пирожных?

Решение. Ответ в этой задаче, очевидно, равен $\left(\!\binom{3}{7}\!\right) = \binom{9}{7}$. Этот ответ можно, например, получить, представляя себе коробку с тремя отделениями, в каждое из которых кладется пирожное только одного вида; в этом случае мы сводим задачу к подсчету количества раскладок семи неразличимых предметов по трем различимым ящикам. Другой способ получить тот же ответ — это представлять себе урну, в которой находятся три разных пирожных, и считать количество способов выбора из урны семи пирожных с возвращениями любого выбранного пирожного обратно в урну. Наконец, можно вообще забыть о любых схемах, если понимать, что любые семь купленных пирожных трех различных видов представляют собой 7-мультимножество над 3-элементным множеством различных видов пирожных.

6 Подсчет количества отображений конечных множеств. Числа Стирлинга второго рода

6.1. Оказывается, задачи о раскладке различимых предметов по различимым же ящикам имеют и еще одну, чрезвычайно важную комбинаторную интерпретацию — они эквивалентны задачам о подсчете количества отображений конечных множеств.

6.1.1. Напомним определение произвольного отображения $f\colon X \to Y$.

Определение 6.1. Пусть X, Y — пара конечных множеств. Отображением f из X в Y называется правило, согласно которому любому элементу $x \in X$ ставится в соответствие единственный элемент $y \in Y$:

$$\forall\, x \in X \qquad \exists!\, y \in Y\colon \qquad y = f(x).$$

С комбинаторной точки зрения любое отображение f из n-элементного множества X в k-элементное множество Y можно рассматривать как некоторый вариант раскладки n различимых предметов по k различимым ящиками при отсутствии каких-либо ограничений на количество предметов в каждом ящике.

Пример 6.2. Рассмотрим отображение f из трехэлементного множества X в четырехэлементное множество Y вида

$$f(x_1) = y_2, \qquad f(x_2) = y_2, \qquad f(x_3) = y_4$$

(смотри рис.6). Этому отображению отвечает раскладка трех различимых предметов по четырем различимым ящикам, при которой первые два предмета размещаются во втором ящике, а третий предмет — в четвертом ящике.

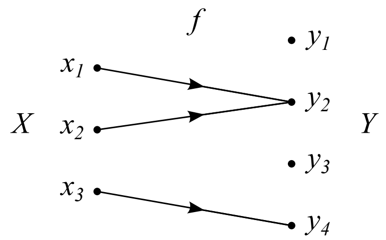


Рис. 6

Как следствие, общее количество всех отображений n-элементного множества X в k-элементное множество Y равно k^n.

6.1.2. Напомним теперь определение *инъективного* отображения.

Определение 6.3. Отображение $f\colon X \to Y$ называется инъективным, если из условия

$$f(x_1) = f(x_2) \qquad \Longrightarrow \qquad x_1 = x_2.$$

Иными словами, отображение называется инъективным, если у любого элемента $y \in Y$ имеется не более одного прообраза, т.е. элемента $x \in X$, такого, что $y = f(x)$.

Понятно, что любому инъективному отображению $f\colon X \to Y$ отвечает такая раскладка n элементов множества X, при которой в каждом из k ящиков находится не более одного элемента. Как следствие, количество всевозможных инъективных отображений равно $(k)_n$.

6.1.3. Наконец, рассмотрим случай биективного и сюръективного отображений.

Определение 6.4. Отображение $f\colon X \to Y$ называется *биективным,* если

$$\forall\, y \in Y \qquad \exists!\, x \in X\colon \qquad y = f(x).$$

Количество таких отображений равно, очевидно, $n!$, где $n = |X| = |Y|$.

Определение 6.5. Отображение $f\colon X \to Y$ называется сюръективным, если

$$\forall\, y \in Y \qquad \exists\, x \in X\colon \qquad y = f(x).$$

Другими словами, отображение f сюръективно, если у любого элемента $y \in Y$ найдется хотя бы один прообраз $x \in X$, то есть такой x, что $f(x) = y$.

Комбинаторная интерпретация сюръективного отображения такова: это есть некоторая раскладка n различимых предметов по k различимым ящикам при условии, что в каждом ящике находится хотя бы один предмет. Количество таких раскладок при $k > n$ равно, очевидно, нулю. Задача следующего пункта данного параграфа — сосчитать количество этих раскладок для случая $0 \leqslant k \leqslant n$.

6.2. Обозначим через $\widehat{S}(n,k)$ количество всех сюръективных отображений n-элементного множества X в k-элементное множество Y. Сосчитаем $\widehat{S}(n,k)$ для случая $n \geqslant 0$, $0 \leqslant k \leqslant n$.

6.2.1. Рассмотрим множество *всех* отображений из n-множества X в k-множество Y. Как мы знаем, количество таких отображений равно k^n. Наша задача состоит в том, чтобы подсчитать это количество по-другому, выразив k^n через числа $\widehat{S}(n,k)$.

6.2.2. Заметим, что *любое* отображение $f\colon X \to Y$ можно рассматривать как *сюръективное* отображение множества X на множество

$$\mathrm{Im}(f) = \{y \in Y \mid \exists x\colon\ y = f(x)\},$$

являющееся образом множества X при отображении f.

Так, для отображения f из примера 6.2 образ $\mathrm{Im}(f) = \{y_2, y_4\}$, а отображение $f\colon X \to Y$ является сюръективным отображением множества X на подмножество $\mathrm{Im}(f) \subset Y$.

6.2.3. Разобъем теперь все множество отображений $f\colon X \to Y$ на блоки, включив в i-й блок все отображения, образ $\mathrm{Im}(f)$ которых содержит ровно i элементов: $|\,\mathrm{Im}(f)| = i$, $i = 1, \ldots, k$. Все, что нам остается — это сосчитать количество элементов в каждом блоке, а затем воспользоваться правилом суммы для того, чтобы получить общее количество k^n всех отображений.

6.2.4. Заметим, что существует $\binom{k}{i}$ способов выбрать i-элементное подмножество k-множества Y. Для каждого из этих подмножеств имеется $\widehat{S}(n,i)$ различных сюръективных отображений из n-элементного множества X в выбранное i-элементное подмножество множества Y. Таким образом, по правилу произведения, общее количество элементов в i-м блоке равно

$$\binom{k}{i} \cdot \widehat{S}(n,i).$$

Тогда по правилу суммы можем записать, что

$$k^n = \sum_{i=0}^{k} \binom{k}{i} \cdot \widehat{S}(n,i). \tag{18}$$

При этом мы суммируем не от 1 до k, а от 0 до k, учитывая, что $\widehat{S}(n,0) = 0$ для всех $n > 0$.

Замечание 6.6. Формулу (18) полезно иногда записывать в виде

$$k^n = \sum_{i=0}^{n} \binom{k}{i} \cdot \widehat{S}(n,i). \tag{19}$$

Несложно убедиться в том, что эта формула непосредственно следует из (18), а также в том, что она оказывается справедливой как для случая $n \geqslant k$, так и для случая $n < k$.

6.2.5. Мы выразили количество всех отображений n-элементного множества X в k-элементное множество Y через количество $\widehat{S}(n,i)$ сюръективных отображений. Нам же нужна обратная формула, выражающая количество $\widehat{S}(n,k)$ сюръективных отображений через число i^n всех отображений. Для ее получения воспользуемся так называемыми *формулами обращения.*

Утверждение 6.7. *Пусть* $(f_0, f_1, f_2, \ldots)$ *и* $(g_0, g_1, g_2, \ldots)$ *— две числовые последовательности, и пусть одна из них выражается через вторую по формулам*

$$f_k = \sum_{i=0}^{k} \binom{k}{i} g_i, \qquad k \geqslant 0. \tag{20}$$

Тогда справедлива следующая формула обращения:

$$g_k = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} f_i, \qquad k \geqslant 0. \tag{21}$$

С учетом этих формул обращения можно, считая n параметром, из соотношения (18) получить следующую явную формулу для вычисления чисел $\widehat{S}(n,k)$:

$$\widehat{S}(n,k)=\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}\cdot i^n.$$

6.3. Задачи подсчета количества отображений n-элементного множества X в k-элементное множество Y имеют еще одну важную комбинаторную интерпретацию.

6.3.1. Начнем с простого примера.

Пример 6.8. Для трехэлементного множества $X=\{x_1,x_2,x_3\}$ и двухэлементного множества $Y=\{y_1,y_2\}$ имеется, как мы знаем, $2^3=8$ различных отображений множества X в множество Y. Запишем все эти отображения как упорядоченные пары подмножеств множества X:

$$(\{x_1,x_2,x_3\},\varnothing),\qquad(\{x_1,x_2\},\{x_3\}),\qquad(\{x_1,x_3\},\{x_2\}),\qquad(\{x_2,x_3\},\{x_1\}),$$

$$(\{x_1\},\{x_2,x_3\}),\qquad(\{x_2\},\{x_1,x_3\}),\qquad(\{x_3\},\{x_1,x_2\}),\qquad(\varnothing,\{x_1,x_2,x_3\}).$$

Видно, что записанное в таком виде решение представляет собой не что иное, как список всех возможных *разделений* множества X, то есть упорядоченных разбиений X на два блока, один из которых может быть и пустым.

6.3.2. Очевидно, что данный результат справедлив и в общем случае. Именно, любое отображение $f\colon X\to Y$ задает нам некоторое разделение множества X, то есть разбиение этого множества на k упорядоченных блоков, часть из которых могут быть пустыми. Как следствие, количество таких разделений совпадает с количеством всех отображений f и равно k^n.

Аналогичные рассуждения показывают, что любое сюръективное отображение $f\colon X\to Y$ задает нам некоторое *упорядоченное разбиение* множества X на блоки. Поэтому количество всех упорядоченных разбиений n-элементного множества X на k блоков равно числу $\widehat{S}(n,k)$.

6.3.3. Рассмотрим теперь некоторый специальный вид k-разделений множества X, а именно, такие k-разделения, в которых в первом блоке содержится a_1 элемент, во втором блоке — a_2 элемента, в k-м блоке — a_k элементов. Очевидно, что при этом общая сумма всех элементов должна быть равна мощности $|X|=n$:

$$a_1+a_2+\ldots+a_k=n,\qquad a_i\geqslant 0.$$

Утверждение 6.9. *Количество всех таких k-разделений n-множества X равно*

$$P(n;a_1,a_2,\ldots,a_k):=\binom{n}{a_1}\cdot\binom{n-a_1}{a_2}\cdot\ldots\cdot\binom{n-a_1-a_2-\ldots-a_{k-1}}{a_k}=\frac{n!}{a_1!\cdot a_2!\cdot\ldots\cdot a_k!}.\quad(22)$$

Доказательство. Действительно, из любого n-элементного множества X мы $\binom{n}{a_1}$ способами можем выбрать a_1 элементов и положить их в первый ящик (отнести к первому блоку разбиения). Затем для каждого такого выбора мы $\binom{n-a_1}{a_2}$ способами можем из оставшегося $(n-$ $-a_1)$-элементного множества выбрать a_2 элементов и положить их во второй ящик (отнести ко второму блоку разбиения), и так далее. Формула (22), описывающая общее количество способов совершить все эти действия, следует теперь из правила произведения. $\square$

Следствие 6.10. *Общее количество k^n всех k-разделений n-множества X выражается через числа $P(n; a_1, a_2, \ldots, a_k)$ по формуле*

$$k^n = \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ a_i \geqslant 0}} P(n; a_1, a_2, \ldots, a_k) = \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ a_i \geqslant 0}} \frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}. \qquad (23)$$

Замечание 6.11. Если в условии рассматриваемой задачи заменить нестрогие неравенства $a_i \geqslant 0$ на строгие, то есть на неравенства $a_i > 0$, то вместо разделения мы получим упорядоченное разбиение специального вида. Количество таких упорядоченных разбиений также описывается формулой (22), а вместо формулы (23) получается не менее полезное соотношение вида

$$\widehat{S}(n, k) = \sum_{\substack{a_1+a_2+\ldots+a_k=n \\ a_i > 0}} \frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}. \qquad (24)$$

6.4. Числа $P(n; a_1, a_2, \ldots, a_k)$ имеют и еще один важный комбинаторный смысл. Именно, они перечисляют так называемые *перестановки n-множества X с повторениями.*

6.4.1. Рассмотрим в качестве элементарного примера следующую задачу: на полке имеются 15 различных книг по математике, 16 по информатике и 12 по физике; каково количество способов перестановки этих книг на полке? Ответ очевиден: $(15 + 16 + 12)! = 43!$ способов.

Предположим теперь, что мы перестали различать книги, посвященные одному и тому же предмету. В этом случае количество различных способов перестановки таких книг уменьшится. Обозначим это количество через λ_n. Так как существует $15!$ способов упорядочить книги по математике, $16!$ – по информатике и $12!$ – по физике, то по правилу произведения мы можем записать, что

$$43! = \lambda_n \cdot 16! \cdot 15! \cdot 12! \qquad \Longrightarrow \qquad \lambda_n = \frac{43!}{16! \cdot 15! \cdot 12!} = P(43; 15, 16, 12).$$

6.4.2. Аналогичные рассуждения справедливы и в общем случае. Именно, пусть среди n переставляемых предметов имеется a_1 неразличимых предметов первого сорта, a_2 неразличимых предметов второго сорта и так далее, причем $a_1 + a_2 + \ldots + a_k = n$. Тогда для количества перестановок таких предметов с повторениями получаем уже знакомую нам формулу

$$\frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!} = P(n; a_1, a_2, \ldots, a_k).$$

6.4.3. В частном случае перестановки n предметов двух различных сортов получаем

$$P(n; k, n-k) = \frac{n!}{k! \cdot (n-k)!} = \binom{n}{k}.$$

Иными словами, количество таких перестановок совпадает с количеством различных k-элементных подмножеств n-элементного множества. Для комбинаторного доказательства данного факта можно воспользоваться рассуждениями, которые мы проводили при подсчете количества всех подмножеств данного множества. Напомним, что там мы любое подмножество кодировали упорядоченной битовой строкой длины n, состоящей из k единиц и $(n - k)$ нулей. Осталось заметить, что любая такая строка представляет собой некоторую перестановку k элементов первого сорта (единиц) и $(n - k)$ элементов второго сорта (нулей).

6.4.4. В заключение отметим еще одну полезную биекцию, позволяющую несколько по-другому сосчитать количество k-мультимножеств n-элементного множества. Мы знаем, что любому k-мультимножеству над n-множеством отвечает некоторая раскладка k неразличимых предметов по n различимым ящикам. В свою очередь, любую такую раскладку можно рассматривать как упорядоченный набор, состоящий из k неразличимых предметов одного сорта (например, k неразличимых шаров), и $(n-1)$-го предмета второго сорта ($(n-1)$-й неразличимой перегородки между этими шарами). Как следствие, количество всех k-мультимножеств

$$\left(\!\binom{n}{k}\!\right) = P(k+n-1; k, n-1) = \frac{(n+k-1)!}{(n-1)!\cdot k!} = \binom{n+k-1}{k}.$$

6.5. Вернемся к числам $\widehat{S}(n,k)$, описывающим, в частности, количество всех упорядоченных разбиений n-множества на k блоков. Обозначим теперь через $S(n,k)$ количество обычных, неупорядоченных разбиений n-множества на k блоков. Так как для любого неупорядоченного разбиения существует $k!$ способов упорядочить k его блоков, то

$$\widehat{S}(n,k) = k!\, S(n,k).$$

Отсюда с учетом (18) и (24) получаем следующие две явные формулы, позволяющие вычислять числа $S(n,k)$:

$$S(n,k) = \frac{1}{k!}\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n = \frac{1}{k!}\sum_{\substack{a_1+a_2+\ldots+a_k=n\\ a_i>0}}\frac{n!}{a_1!\cdot a_2!\cdot\ldots\cdot a_k!}. \qquad (25)$$

Числа $S(n,k)$ называются *числами Стирлинга второго рода* и встречаются в большом количестве комбинаторных приложений. Исследуем эти числа поподробнее.

6.5.1. Для практического расчета чисел $S(n,k)$ удобно использовать рекуррентные соотношения.

Утверждение 6.12. *Числа Стирлинга второго рода удовлетворяют следующим рекуррентным соотношениям:*

$$S(n,k) = S(n-1,k-1) + k\cdot S(n-1,k), \qquad k=1,\ldots,n; \qquad (26)$$

$$S(0,0)=1; \qquad S(n,0)=0 \quad \forall\ n>0; \qquad S(n,k)=0 \quad \forall\ k>n.$$

Доказательство. Граничные условия $S(n,0)=0$ для всех $n>0$ и $S(n,k)=0$ для всех $k>n$ очевидны — n-элементное множество в случае $n>0$ нельзя разбить на 0 блоков, а также на k блоков в случае, когда $k>n$. Равенство $S(0,0)=1$ введено просто для удобства.

Докажем теперь соотношение (26). Разобьем для этого множество всех k-разбиений на два блока. К первому блоку Σ_1 отнесем все разбиения, содержащие одноэлементное подмножество $\{x_1\}$. Ко второму блоку Σ_2 отнесем все оставшиеся k-разбиения, т.е. разбиения, в которых элемент x_1 входит в подмножества, содержащие как минимум два элемента.

Рассмотрим, к примеру, все 2-разбиения множества $X=\{x_1,x_2,x_3,x_4\}$:

$$\{\{x_1,x_2,x_3\},\{x_4\}\}, \quad \{\{x_1,x_2,x_4\},\{x_3\}\}, \quad \{\{x_1,x_3,x_4\},\{x_2\}\}, \quad \{\{x_2,x_3,x_4\},\{x_1\}\},$$

$$\{\{x_1,x_2\},\{x_3,x_4\}\}, \quad \{\{x_1,x_3\},\{x_2,x_4\}\}, \quad \{\{x_1,x_4\},\{x_2,x_3\}\}.$$

Для этого примера к первому блоку относится единственное разбиение вида $\{\{x_2, x_3, x_4\}, \{x_1\}\}$. Остальные шесть разбиений относятся в данном случае ко второму блоку.

Довольно очевидно, что количество элементов в первом блоке равно количеству $S(n-1, k-1)$ всех $(k-1)$-разбиений оставшегося $(n-1)$-элементного множества. В примере это число равно $S(3,1) = 1$ — любое множество можно только одним способом разбить на один блок.

Для подсчета количества элементов во втором блоке заметим, что, по сути дела, это число равно количеству всевозможных разбиений $(n-1)$-элементного множества на k непустых подмножеств с поочередным добавлением элемента x_1 в каждое из этих подмножеств. Действительно, для разобранного выше примера имеется три разбиения трехэлементного множества $\{x_2, x_3, x_4\}$ на два непустых подмножества, а именно, разбиения вида

$$\{\{x_2\}, \{x_3, x_4\}\}, \quad \{\{x_3\}, \{x_2, x_4\}\}, \quad \{\{x_4\}, \{x_2, x_3\}\}.$$

Добавляя к каждому из этих подмножеств элемент x_1, получим $2 \cdot 3 = 6$ выписанных выше разбиений четырехэлементного множества X на блоки требуемого вида. Количество таких разбиений в общем случае равно, очевидно, $k \cdot S(n-1, k)$. □

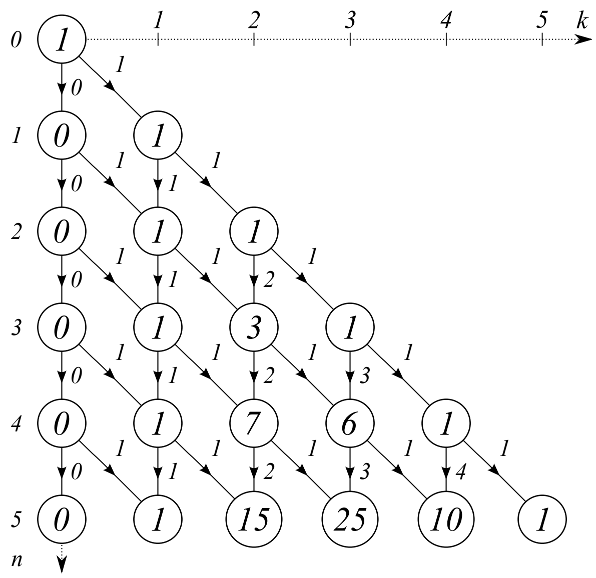


Рис. 7: Графическое представление чисел $S(n, k)$

6.5.2. Как и биномиальные коэффициенты, числа Стирлинга второго рода удобно представлять в виде треугольного массива на плоскости (n, k). На рис.7 показаны первые несколько строк такого массива, вычисленных с помощью рекуррентных соотношений (26). Такой рисунок дает нам еще одну комбинаторную интерпретацию чисел $S(n, k)$ — это есть количество *взвешенных* путей, идущих из начала координат в точку с координатами (n, k). Любой участок такого пути, имеющий вес i, можно рассматривать как набор из i путей, соединяющих соответствующие точки на плоскости.

6.5.3. Числа Стирлинга второго рода встречается в очень большом количестве самых разнообразных задач. В качестве примера вернемся к формуле (19) и перепишем ее через числа Стирлинга второго рода:

$$k^n = \sum_{i=0}^{n} \binom{k}{i} \cdot i! \cdot S(n, i) = \sum_{i=0}^{n} (k)_i \cdot S(n, i). \tag{27}$$

Оказывается, эта формула справедлива для любых вещественных и даже комплексных значений k. Именно, справедливо равенство

$$x^n = \sum_{i=0}^{n} (x)_i \cdot S(n,i).$$

Эта формула активно используется в теории конечных операторов — она позволяет перейти от базиса x^n к базису $(x)_n$.

6.5.4. С точки зрения схемы раскладки предметов по ящикам числа Стирлинга второго рода описывают количество способов разложить n различимых предметов по k неразличимым ящикам так, чтобы в любом ящике содержался хотя бы один предмет. Сосчитаем, зная $S(n,k)$, количество $B(n,k)$ различных раскладок n различимых предметов по k неразличимым ящикам в случае, когда ограничения на количество предметов в любом ящике отсутствуют.

Для этого вновь разобьем множество всех таких раскладок на блоки, поместив в i-й блок все раскладки, в которых занято ровно i ящиков. В случае различимых ящиков в процессе аналогичного разбиения множества на блоки мы в i-м блоке имели $\binom{k}{i} \cdot \widehat{S}(n,i)$ элементов. В случае неразличимых ящиков мы i из k ящиков выбираем ровно одним способом, а затем $S(n,i)$ способами заполняем эти ящики n предметами. Используя правило суммы, получаем отсюда следующее выражение для чисел $B(n,k)$:

$$B(n,k) = \sum_{i=1}^{k} S(n,i).$$

6.5.5. Числа $B(n,k)$ в случае $n = k$ называются числами Белла $B(n)$. Эти числа перечисляют количество *всех* возможных разбиений n-элементного множества X.

Заметим, что любое разбиение множества X можно получить, введя на этом множестве некоторое отношение эквивалентности. Как следствие, количество всех возможных отношений эквивалентности на n-элементном множестве X описывается числами Белла B_n.

6.5.6. Покажем, что для чисел Белла $B(n)$ справедливо следующее рекуррентное соотношение:

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k. \tag{28}$$

Действительно, рассмотрим блок разбиения $(n+1)$-элементного множества, содержащий число $n+1$. Все возможные разбиения мы можем разбить на $n+1$ группу (блок) в зависимости от того, сколько чисел содержится вместе с $n+1$. Предоположим, что вместе с $n+1$ находятся i чисел, $i = 0, \ldots, n$. Эти числа мы можем выбрать $\binom{n}{i} = \binom{n}{n-i}$ способами. Оставшиеся $k = n - i$ чисел, $k = 0, \ldots, n$, мы можем $B(k)$ способами разбить на блоки. Суммируя по всем k, мы и получим формулу (28).

6.6. Подведем итоги, построив таблицу различных схем раскладок n предметов по k ящикам:

Элементы множества X (предметы)	Элементы множества Y (ящики)	Произвольное количество предметов в ящике	Не более 1 предмета в ящике	Как минимум 1 предмет в ящике
различимые	различимые	k^n	$(k)_n$	$\widehat{S}(n,k)$
неразличимые	различимые	$\left(\!\binom{k}{n}\!\right)$	$\binom{k}{n}$	$\binom{n-1}{k-1}$
различимые	неразличимые	$B(n,k)$	$0, \quad n > k$ $1, \quad n \leqslant k$	$S(n,k)$

Как видно, остался еще один неразобранный вариант — схема размещения n неразличимых предметов по k неразличимым ящикам. Для количества таких размещений явных аналитических формул не существует. Для того, чтобы их перечислить, необходимо использовать аппарат производящих функций.

7 Понятие дискретной вероятности

7.1. Многие задачи элементарной комбинаторики часто формулируются на языке подсчета вероятностей наступления тех или иных случайных событий. Цель настоящего параграфа — изложить основные понятия элементарной теории вероятностей, а также показать связь данной науки с элементарной комбинаторикой.

7.1.1. Начнем, как всегда, с самых простых, базовых понятий теории вероятности.

Под случайным экспериментом будем понимать математическую модель некоторого реального эксперимента, результат которого невозможно точно предсказать. Простейшим и наиболее хрестоматийным примером случайного эксперимента является подбрасывание идеальной монетки. В результате любого такого эксперимента у нас обязательно выпадает или орел, или решка (на ребро идеальная монетка упасть не может). Заранее предсказать, что именно выпадет, мы не можем. Другим простейшим примером случайного эксперимента является подбрасывание игральной кости, результатом которого можно считать выпадение одного из шести чисел, нанесенных на грани куба.

7.1.2. Любой результат ω случайного эксперимента называется *элементарным событием* или *исходом*. Множество всех возможных исходов обозначим через Ω. Далее мы всегда будем считать, что множество Ω конечно или счетно.

Случайным событием или просто *событием* A называется любое подмножество множества Ω. Так, в примере с подбрасыванием игральной кости событиями являются, например, выпадение четного числа или выпадение числа, меньшего тройки.

Говорят, что в результате случайного эксперимента *произошло событие* A, если элементарный исход эксперимента является элементом множества A. Так как один и тот же исход может быть элементом нескольких разных событий, то в результате одного и того же случайного эксперимента возможно появление нескольких случайных событий. *Исход же любого случайного эксперимента может быть только один.*

Так, в примере с игральной костью элементарный исход — это выпадение конкретного числа, например, двойки. Этот элементарный исход, однако, отвечает появлению сразу двух описанных выше событий — события A_1, описывающего выпадение четного числа, и события A_2, соответствующего выпадению числа, меньшего тройки.

7.1.3. Предположим теперь, что у нас для заданного случайного эксперимента имеется некоторый набор $\mathcal{A}_0$ событий. С помощью теоретико-множественных операций объединения, пересечения и дополнения можно из этого набора $\mathcal{A}_0$ построить некоторую новую, более полную систему множеств, также являющихся событиями. Присоединяя к этой системе так называемое невозможное ($A = \varnothing$) и достоверное ($A = \Omega$) события, мы получаем систему множеств $\mathcal{A}$, называемую *алгеброй событий,* то есть такую систему подмножеств множества исходов Ω, что, во-первых, само $\Omega \in \mathcal{A}$, и во-вторых, для любой пары $A, B \in \mathcal{A}$ их объединение, пересечение и дополнение также принадлежат $\mathcal{A}$.

Как правило, в качестве $\mathcal{A}_0$ выбирают некоторый набор подмножеств, образующий разбиение множества Ω. В этом случае $\mathcal{A}$ называется алгеброй, порожденной данным разбиением $\mathcal{A}_0$.

В случае конечного множества Ω в качестве $\mathcal{A}$ чаще всего выбирается множество 2^Ω всех подмножеств данного множества Ω.

7.1.4. Припишем теперь любому элементарному событию $\omega \in \Omega$ некоторое вещественное число $\mathrm{Pr}(\omega)$ из диапазона $[0, 1]$. Это отображение $\mathrm{Pr} : \Omega \to \mathbb{R}_+$ называется *вероятностью,* если для него выполняется следующее условие нормировки:

$$\sum_{\omega \in \Omega} \mathrm{Pr}(\omega) = 1.$$

Зная вероятность $\mathrm{Pr}(\omega)$ любого элементарного исхода, можно по формуле

$$\mathrm{Pr}(A) = \sum_{\omega \in A} \mathrm{Pr}(\omega) \tag{29}$$

определить вероятность случайного события A. Иными словами, мы с помощью этой формулы можем продолжить отображение Pr на все элементы A заданной алгебры событий $\mathcal{A}$ и считать, что функция Pr задана не на Ω, а на некоторой алгебре событий $\mathcal{A}$.

Зафиксируем теперь некоторую алгебру событий $\mathcal{A}$. Из определения (29) сразу же вытекают следующие простейшие свойства вероятности:

$$\mathrm{Pr}(\varnothing) = 0, \qquad \mathrm{Pr}(\Omega) = 1;$$
$$\mathrm{Pr}(A \cup B) = \mathrm{Pr}(A) + \mathrm{Pr}(B) - \mathrm{Pr}(A \cap B) \qquad \forall\ A, B \in \mathcal{A}.$$

События называются *несовместными,* если $A \cap B = \varnothing$. Для таких событий

$$\mathrm{Pr}(A \cup B) = \mathrm{Pr}(A) + \mathrm{Pr}(B). \tag{30}$$

Следствием формулы (30) является следующее полезное равенство:

$$\mathrm{Pr}(\bar{A}) = 1 - \mathrm{Pr}(A), \qquad \text{где} \qquad \bar{A} = \Omega \setminus A. \tag{31}$$

Формулы (30) и (31) легко обобщаются на случай набора $\{A_1, A_2, \ldots, A_n\}$ несовместных событий, а также на случай набора $\mathcal{A}_0 = \{A_1, A_2, \ldots, A_n\}$ подмножеств, образующих разбиение множества Ω. В первом случае имеем равенство

$$\mathrm{Pr}(A_1 \cup A_2 \cup \ldots \cup A_n) = \mathrm{Pr}(A_1) + \mathrm{Pr}(A_2) + \ldots + \mathrm{Pr}(A_n), \tag{32}$$

называемое *формулой сложения вероятностей несовместных событий.* Во втором случае получаем равенство

$$\Pr(A_1) + \Pr(A_2) + \ldots + \Pr(A_n) = 1.$$

Набор $\mathcal{A}_0 = \{A_1, A_2, \ldots, A_n\}$ называется при этом *полной группой несовместных событий.*

7.1.5. Приведенное в предыдущем пункте определение вероятности годится только лишь для случая конечного или счетного множества Ω. В случае несчетного Ω вероятность любого элементарного исхода ω, как правило, равняется нулю. Как следствие, формула (29) перестает работать. Поэтому в случае несчетного множества Ω вероятность $\Pr$ определяется не как функция на множестве событий Ω, а сразу как функция на некоторой алгебре событий $\mathcal{A}$. При таком подходе равенства $\Pr(\Omega) = 1$ и (32) становятся аксиомами теории вероятности.

Иными словами, в дискретном случае мы можем ограничиться множеством Ω элементарных событий и заданной на нем функцией $\Pr : \Omega \to [0, 1]$, удовлетворяющей условию нормировки. В непрерывном же случае нам необходимо рассматривать тройку $(\Omega, \mathcal{A}, \Pr)$, где Ω — множество исходов, $\mathcal{A}$ — некоторая алгебра подмножеств Ω, а $\Pr$ — заданная на $\mathcal{A}$ вероятность. При этом говорят, что тройка $(\Omega, \mathcal{A}, \Pr)$ определяет *вероятностную модель* или *вероятностное пространство* некоторого случайного эксперимента с множеством исходов Ω и алгеброй событий $\mathcal{A}$.

Как мы уже сказали, мы в данном курсе будем рассматривать только дискретную вероятность. Однако для единообразия мы все же будем и в дискретном случае использовать тройку $(\Omega, \mathcal{A}, \Pr)$ для описания случайных экспериментов.

Пример 7.1. Рассмотрим случайный эксперимент с однократным подбрасыванием монетки. Его можно описать, задав пространство элементарных исходов

$$\Omega = \{\text{решка}, \text{орел}\}$$

и введя на нем вероятность $\Pr : \Omega \to [0, 1]$ по формулам

$$\Pr(\{\text{решка}\}) = p, \qquad \Pr(\{\text{орел}\}) = q,$$

где p, q — вещественные положительные числа, такие, что $p + q = 1$. Этот же случайный эксперимент можно описать чуть более сложно, рассмотрев вероятностное пространство вида $(\Omega_1, \mathcal{A}_1, \Pr_1)$, в котором

$$\Omega_1 = \{\text{решка}, \text{орел}\}, \qquad \mathcal{A}_1 = 2^{\Omega_1} = \{\varnothing, \{\text{решка}\}, \{\text{орел}\}, \{\text{орел или решка}\}\}, \tag{33}$$

а вероятность $\Pr_1$ на $\mathcal{A}_1$ описывается соотношениями

$$\Pr_1(\varnothing) = 0, \qquad \Pr_1(\{\text{решка}\}) = p, \qquad \Pr_1(\{\text{орел}\}) = q, \qquad \Pr_1(\{\text{орел или решка}\}) = 1. \tag{34}$$

7.1.6. Вообще говоря, вероятности элементарных событий можно определять достаточно произвольно. Так, если мы считаем монетку сделанной идеально, то вероятность q выпадения орла совпадает с вероятностью p выпадения решки и равна одной второй. В случае же монетки со смещенным центром тяжести вероятность этих элементарных событий может различаться. Так, ничто не мешает нам, например, определить $p = 1/4$, а $q = 3/4$.

Однако на практике все же часто удобно считать, что все различные элементарные события являются равновероятными. В этом случае для любого $\omega \in \Omega$

$$\Pr(\omega) = \frac{1}{|\Omega|}, \qquad \text{и, как следствие,} \qquad \Pr(A) = \frac{|A|}{|\Omega|}.$$

Такой способ задания вероятности называется классическим. В этом случае подсчет вероятности $\Pr(A)$ события A сводится к подсчету количества исходов, к этому событию приводящих. Иными словами, в этом случае задача становится чисто комбинаторной.

Пример 7.2. Вернемся к одной из классических урновых схем, а именно, к задаче о подсчете упорядоченных выборок с повторениями. Переформулируем эту задачу на языке дискретной вероятности.

Пусть в урне имеется n различимых предметов. Будем поочередно вытаскивать k предметов, записывать, какой из предметов и в каком порядке мы вытащили, а затем возвращать каждый предмет обратно в урну. Данная последовательность действий представляет собой случайный эксперимент, элементарным исходом которого является упорядоченный набор $(a_1, a_2, \ldots, a_k)$, $a_i \in X$ для любого $i = 1, \ldots, n$. Множество Ω всех возможных исходов выглядит в этом случае так:

$$\Omega = \{\omega : \ \omega = (a_1, a_2, \ldots, a_k), \quad a_i \in X\}.$$

Считая, что любые элементарные исходы данного случайного эксперимента являются равновероятными, мы получаем следующую вероятность любого элементарного исхода в этом эксперименте:

$$\Pr(\omega) = \frac{1}{|\Omega|} = \frac{1}{n^k}.$$

Рассмотрим теперь событие A, заключающееся в отсутствии повторений элементов a_i в наборе $(a_1, a_2, \ldots, a_k)$:

$$A = \{\omega : \ a_1 \neq a_2 \neq \ldots \neq a_k\}.$$

Из элементарной комбинаторики известно, что $|A| = (n)_k$ — количеству k-перестановок из n элементов без повторений. Как следствие,

$$\Pr(A) = \frac{|A|}{|\Omega|} = \frac{(n)_k}{n^k} = \left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \ldots \left(1 - \frac{k-1}{n}\right).$$

Рассмотренный пример является достаточно характерным — многие задачи элементарной комбинаторики могут быть переформулированы на языке теории вероятности. Например, задача об определении количества трехзначных чисел, содержащих цифры 3 и 6, может рассматриваться и как задача о нахождении вероятности того, что произвольно выбранное трехзначное число содержит цифры 3 и 6.

7.2. Перейдем теперь к важному понятию условной вероятности.

7.2.1. Начнем с достаточно характерного примера.

Пример 7.3. Рассмотрим случайный эксперимент, связанный с подбрасыванием игральной кости. Обозначим через A событие, заключающееся в том, что у нас выпало число, большее трех, а через B — событие, состоящее в том, что у нас выпало четное число. Ясно, что

$$\Pr(A) = \frac{3}{6} = \frac{1}{2}, \qquad \Pr(B) = \frac{3}{6} = \frac{1}{2}.$$

Теперь предположим, что мы получили некоторую дополнительную информацию, а именно, нам стало известно, что в результате случайного эксперимента у нас произошло событие A. Мы не знаем, произошло ли у нас событие B, однако появившаяся у нас дополнительная информация позволяет нам утверждать, что вероятность наступления события B увеличилась.

Действительно, тот факт, что у нас произошло событие A, сужает множество возможных исходов случайного эксперимента до подмножества $A = \{4, 5, 6\}$. Два исхода из этих трех — выпадение чисел 4 и 6 — являются благоприятными для наступления события B. Иными словами, вероятность $\Pr(B|A)$ наступления события B при условии, что событие A произошло, возрастает и становится равной $2/3$.

7.2.2. Данный пример удобно иллюстрировать на диаграммах Эйлера-Венна (рис.8). Как мы знаем, в случае, когда все элементарные исходы равновероятны, вероятности наступления событий A и B равны

$$\Pr(A) = \frac{|A|}{|\Omega|}, \qquad \Pr(B) = \frac{|B|}{|\Omega|}.$$

Информация о том, что у нас произошло событие A, сужает для B пространство возможных исходов с Ω до A. При этом все исходы, принадлежащие $A \cap B$, являются благоприятными для наступления события B, так что вероятность $\Pr(B|A)$ наступления события B при условии, что событие A произошло, становится равной

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

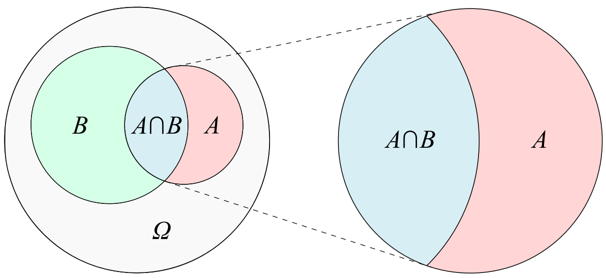


Рис. 8

Теперь мы можем сформулировать формальное определение понятия условной вероятности.

Определение 7.4. Условной вероятностью $\Pr(B|A)$, то есть вероятностью наступления события B при условии того, что событие A произошло, называется величина

$$\Pr(B|A) = \frac{\Pr(A \cap B)}{\Pr(A)}. \tag{35}$$

Заметим, что так как событие A произошло, то гарантированно $\Pr(A) > 0$, а значит, данное выше определение корректно.

7.2.3. Отметим некоторые простейшие свойства условной вероятности. Прежде всего, из определения $\Pr(B|A)$ с очевидностью следует, что

$$\Pr(A|A) = 1, \qquad \Pr(\varnothing|A) = 0, \qquad \Pr(B|A) = 1 \quad \text{в случае, если } A \subset B.$$

Далее, в случае, если $B_1 \cap B_2 = \varnothing$ выполняется равенство

$$\Pr(A \cap (B_1 \cup B_2)) = \Pr((A \cap B_1) \cup (A \cap B_2)) = \Pr(A \cap B_1) + \Pr(A \cap B_2),$$

а значит, в этом случае

$$\Pr((B_1 \cup B_2)|A) = \Pr(B_1|A) + \Pr(B_2|A).$$

Следствием этого свойства является важное равенство

$$\Pr(B|A) + \Pr(\bar{B}|A) = 1.$$

7.2.4. Вернемся к примеру с подбрасыванием игральной кости. Предположим теперь, что нам кто-то сообщил о том, что в результате случайного эксперимента событие A не произошло (или, что то же самое, произошло событие $\bar{A} = \Omega \backslash A$). В этом случае шансы на наступление события B уменьшились — из множества $\bar{A} = \{1, 2, 3\}$ возможных исходов только один исход — выпадение числа 2 — является для B благоприятным. Иными словами, вероятность $\Pr(B|\bar{A})$ при условии, что произошло событие $\bar{A}$, равна $1/3$. Заметим, что у нас при этом выполняется любопытное равенство вида

$$\Pr(A) \cdot \Pr(B|A) + \Pr(\bar{A}) \cdot \Pr(B|\bar{A}) = \frac{1}{2} \cdot \frac{2}{3} + \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{2} = \Pr(B). \tag{36}$$

Полученная формула называется *формулой полной вероятности.* В ее справедливости легко убедиться, проанализировав рис.8.

7.2.5. Формулу полной вероятности обычно записывают в несколько более общем виде. Именно, рассмотрим полную группу несовместных событий $A_1, A_2, \ldots, A_n$. Ясно, что

$$B = B \cap \Omega = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n),$$

откуда на основании свойства (30) следует, что

$$\Pr(B) = \Pr(B \cap A_1) + \Pr(B \cap A_2) + \ldots + \Pr(B \cap A_n).$$

С учетом формулы (35) вероятности $\Pr(B \cap A_i)$ можно выразить через условные вероятности $\Pr(B|A_i)$:

$$\Pr(B \cap A_i) = \Pr(B|A_i) \cdot \Pr(A_i). \tag{37}$$

Отсюда окончательно получается следующая обобщенная формула полной вероятности:

$$\Pr(B) = \sum_{i=1}^{n} P(B|A_i) \cdot \Pr(A_i). \tag{38}$$

7.2.6. Приведем достаточно характерный пример использования формулы (38).

Пример 7.5. Пусть в магазине имеется 100 лампочек, 60 из которых сделаны производителем номер 1, 25 — производителем номер 2, и 15 — производителем номер 3. Вероятность выхода из строя лампочки в первую неделю ее работы у первого производителя равна 0.02, у второго — 0.01, и у третьего — 0.03. Какова вероятность того, что случайно купленная в магазине лампочка выйдет из строя в течение первой недели?

Решение. Обозначим через A_i событие, заключающееся в том, что купленная лампочка принадлежит i-му производителю. Очевидно, что событие Ω, отвечающее покупке лампочки, является достоверным событием (то есть $\Pr(\Omega) = 1$), а $A_1 \cup A_2 \cup A_3 = \Omega$, то есть события A_i образуют полную группу несовместных событий. При этом $\Pr(A_1) = 0.6$, $\Pr(A_2) = 0.25$, $\Pr(A_3) = 0.15$.

Далее, вероятность выхода из строя лампочки в первую неделю ее работы у i-го производителя является, очевидно, условной вероятностью $\Pr(B|A_i)$, где B — событие, состоящее в выходе из строя лампочки в первую неделю ее работы. Следовательно, согласно формуле (38) полной вероятности,

$$\Pr(B) = 0.6 \cdot 0.02 + 0.25 \cdot 0.01 + 0.03 \cdot 0.15 = 0.019.$$

7.2.7. В примере 7.5 постановка задачи была в определенном смысле прямой: у нас было известно, что хотя бы одно из трех событий A_i произошло (лампочка была куплена), и мы искали вероятность наступления события B, состоящего в том, что купленная нами в магазине лампочка в первую неделю перегорела. На практике, однако, нас может интересовать и такая постановка задачи: пусть лампочка у нас в первую неделю все же перегорела; какова вероятность того, что в этом случае (то есть при наступлении события B) эта лампочка принадлежит, к примеру, 2-му производителю?

С формальной точки зрения речь идет о вычислении условной вероятности $\Pr(A_i|B)$: событие B произошло, плюс произошло и какое-то из трех возможных событий A_i; нас же интересует вероятность того, что в этом случае до наступления события B случилось именно событие A_2, а не два других события.

Оказывается, на такой вопрос также достаточно легко ответить. Именно, предположим, что вероятность $\Pr(B)$ события B строго больше нуля. Заметим, что тогда наряду с формулами

$$\Pr(B|A_i) = \frac{\Pr(B \cap A_i)}{\Pr(A_i)} \qquad \Longleftrightarrow \qquad \Pr(B \cap A_i) = \Pr(B|A_i) \cdot \Pr(A_i)$$

для любого $i = 1, 2, \dots, n$ мы можем написать и аналогичные равенства

$$\Pr(B \cap A_i) = \Pr(A_i|B) \cdot \Pr(B) \qquad \Longleftrightarrow \qquad \Pr(A_i|B) = \frac{\Pr(B \cap A_i)}{\Pr(B)}.$$

Вспоминая теперь, что $\Pr(B \cap A_i)$ вычисляется по формуле (37), а также то, что для $\Pr(B)$ справедлива формула полной вероятности (7.5), мы для $\Pr(A_i|B)$ получаем соотношение

$$\Pr(A_i|B) = \frac{\Pr(A_i) \cdot \Pr(B|A_i)}{\Pr(A_1) \cdot \Pr(B|A_1) + \Pr(A_2) \cdot \Pr(B|A_2) + \ldots + \Pr(A_n) \cdot \Pr(B|A_n)}. \tag{39}$$

В частности, в нашей задаче вероятность того, что лампочка была изготовлена вторым производителем, равна

$$\Pr(A_2|B) = \frac{0.25 \cdot 0.01}{0.6 \cdot 0.02 + 0.25 \cdot 0.01 + 0.03 \cdot 0.15} = \frac{0.01 \cdot 0.25}{0.019} = \frac{5}{38} = 0.13.$$

7.2.8. Формула (39) носит название *теоремы Байеса* и играет достаточно важную роль в разного рода практических задачах. В этих задачах события A_i часто называют *гипотезами,* вероятность $\Pr(A_i)$ — *априорной* вероятностью гипотезы A_i, а вероятность $\Pr(A_i|B)$ трактуется как *апостериорная* вероятность наступления события A_i, то есть вероятность этого события *после* наступления события B.

Одна из наиболее популярных задач в этой области — это нахождение так называемой *наиболее вероятной гипотезы,* то есть события A_i, для которого $\Pr(A_i|B)$ будет наибольшим среди всех A_i. В нашей задаче таковой будет, очевидно, гипотеза, состоящая в том, что перегоревшая лампочка была изготовлена первым предприятием:

$$\Pr(A_1|B) = \frac{0.6 \cdot 0.02}{0.019} = \frac{12}{19} = 0.63.$$

Так как знаменатель в формуле Байеса (39) равен $\Pr(B)$ и не зависит от A_i, то общем случае для определения наиболее вероятной гипотезы следует найти такую гипотезу A_i, для которой величина $\Pr(B|A_i) \cdot \Pr(A_i)$ будет максимальной.

Довольно часто в такого рода задачах все априорные вероятности считаются одинаковыми и равными $1/n$. В этом случае наиболее вероятной гипотезой будет, очевидно, событие A_i с наибольшей условной вероятностью $\Pr(B|A_i)$.

7.3. Следующее понятие — понятие независимых событий — является одним из самых важных и ключевых в теории вероятности.

7.3.1. Пусть у нас имеются два события A и B. Логично считать, что событие B не зависит от события A, если вероятность $\Pr(B)$ не зависит от того, наступило ли событие A или не наступило. Иными словами, B не зависит от A, если $\Pr(B|A) = \Pr(B)$.

Вспоминая теперь определение условной вероятности, мы получаем, что в случае $\Pr(A) > 0$ независимость B от A влечет равенство

$$\Pr(B) = \Pr(B|A) = \frac{\Pr(B \cap A)}{\Pr(A)} \qquad \Longrightarrow \qquad \Pr(B \cap A) = \Pr(B) \cdot \Pr(A).$$

Последнее равенство как раз и кладется в определение независимости случайных событий.

Определение 7.6. События A и B являются независимым, если их совместная вероятность равна произведению этих вероятностей:

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B). \tag{40}$$

В противном случае события A и B называются зависимыми.

Последнее определение хорошо тем, что, во-первых, оно симметрично относительно A и B, а во-вторых, работает даже в случаях, когда $\Pr(A)$ или $\Pr(B)$ равны нулю.

Пример 7.7. Вернемся к примеру 7.3 с подбрасыванием кубика. Мы уже поняли, что события A и B зависимы — вероятность $\Pr(B|A)$ отлична от вероятности $\Pr(B)$. Этот же вывод дает нам равенство (40). Действительно, $\Pr(A \cap B) = 1/3$, тогда как $\Pr(A) \cdot \Pr(B) = 1/4$.

Рассмотрим теперь событие C, состоящее в том, что в процессе бросания кубика выпало число, кратное трем. Вероятность $\Pr(C) = 1/3$; при этом

$$\Pr(A \cap C) = \frac{1}{6} = \Pr(A) \cdot \Pr(C), \qquad \Pr(B \cap C) = \frac{1}{6} = \Pr(B) \cdot \Pr(C).$$

Следовательно, события A и C, равно как и события B и C попарно независимы.

7.3.2. Понятие независимости довольно естественно обобщается на случай нескольких случайных событий.

Определение 7.8. События $A_1, A_2, \ldots, A_n$ называются независимыми (иногда добавляют в совокупности), если

$$\Pr(A_{i_1} \cap A_{i_2} \cap \ldots A_{i_k}) = \Pr(A_{i_1}) \cdot \Pr(A_{i_1}) \cdot \ldots \Pr(A_{i_k}) \qquad \text{для любых } 1 \leqslant i_1 < i_2 < \ldots < i_k \leqslant n.$$

Важно заметить, что из попарной независимости событий независимость в совокупности не следует.

Пример 7.9. Предположим, что мы бросаем два игральных кубика. Обозначим через A событие, состоящее в том, что на первом кубике выпало нечетное количество очков, через B — событие, состоящее в том, что на втором кубике выпало нечетное количество очков, а через C — событие, состоящее в том, что сумма выпавших на обоих кубиков очков четна. Несложно убедиться, что вероятности всех этих событий равны $1/2$, а также то, что эти события попарно независимы. Например, $\Pr(A \cap B) = 1/4 = \Pr(A) \cdot \Pr(B)$. Однако

$$A \cap B \cap C = A \cap B \qquad \Longrightarrow \qquad \Pr(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = \Pr(A) \cdot \Pr(B) \cdot \Pr(C).$$

Как следствие, эти три события независимыми в совокупности не являются.

7.3.3. Обычно определить, являются ли какие-то два отдельных события A и B независимыми, довольно затруднительно. Например, не проводя соответствующих вычислений, не очень понятно, являются ли независимыми события “выпадение на кубике нечетного числа” и “выпадение на кубике числа, меньшего или равного трем”. Однако существует довольно распространенная конструкция, в которой пары независимых событий возникают довольно естественно, по построению. Для того, чтобы эту конструкцию проще было понять, мы опишем вначале схему ее построения на достаточно простом, но в то же время важном примере — на так называемой схеме Бернулли.

Рассмотрим случайный эксперимент, состоящий из последовательности n одинаковых независимых случайных испытаний, то есть таких испытаний, результаты каждого из которых никак не зависят от результатов прочих испытаний. Предположим, что в каждом таком испытании возможны ровно два исхода, называемые успехом и неудачей, причем вероятность успеха в каждом эксперименте одинакова и равна $p \in (0, 1)$, а вероятность q неудачи равна, соответственно, $1 - p$. Такого рода случайный эксперимент и называется *схемой Бернулли*.

Данная схема моделирует, например, случайный эксперимент, заключающийся в n-кратном подбрасывании монетки. Если монетка правильная, то $p = q = 1/2$. В противном случае имеем так называемую несимметричную монетку, для которой $p \neq q$. Кроме этого простейшего примера, схема Бернулли моделирует множество других случайных экспериментов [**?**].

7.3.4. С целью формального описания схемы Бернулли вернемся к простейшему случайному эксперименту, заключающемуся в подбрасывании монетки. Напомним, что вероятностное пространство для такого эксперимента описывается соотношениями (33)–(34).

Теперь несколько усложним его, а именно, предположим, что мы n раз подбрасываем данную монету. Результат любого такого эксперимента можно записать в виде битовой строки вида $(a_1, a_2, \ldots, a_n)$, в которой $a_i = 1$ в случае, если, например, у нас выпала решка, и $a_i = 0$ в случае выпадения орла. При таком подходе пространство всех исходов Ω_n и алгебра событий $\mathcal{A}_n$ имеют вид

$$\Omega_n = \{\omega : \ \omega = (a_1, a_2, \ldots, a_n)\}, \qquad \mathcal{A}_n = 2^{\Omega_n}. \tag{41}$$

Покажем, что вероятность $\Pr_n$ любого элементарного события ω можно в этом случае задать с помощью соотношения

$$\Pr_n(\omega) = p^k q^{n-k}, \qquad k = \sum_{i=1}^{n} a_i, \qquad p, q > 0, \qquad p + q = 1. \tag{42}$$

Действительно, очевидно, что $\Pr_n(\omega) \in (0,1)$. Кроме того, количество всех элементарных исходов, для которых $\sum_i a_i = k$, совпадает с количеством $\binom{n}{k}$ битовых строк длины n, содержащих k единиц. Поэтому

$$\sum_{\omega\in\Omega_n} \Pr_n(\omega) = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1.$$

Следовательно, описываемая приведенными выше соотношениями (41)–(42) тройка $(\Omega_n, \mathcal{A}_n, \Pr_n)$ определяет вероятностную модель, описывающую n-кратное подбрасывание монеты.

Обозначим через A_k событие, состоящее в том, что в результате серии n испытаний произошло ровно k успехов. Как мы только что показали, вероятность такого события равна

$$\Pr_n(A_k) = \binom{n}{k} p^k\, q^{n-k}, \qquad\qquad k = 0, 1, \ldots, n. \tag{43}$$

Очевидно, что набор $\{A_0, A_1, \ldots, A_n\}$ образует полную группу несовместных событий. Соответствующий этим событиям набор вероятностей $\{\Pr_n(A_0), \Pr_n(A_1), \ldots, \Pr_n(A_n)\}$ называется *биномиальным распределением* количества успехов в n испытаниях.

7.3.5. Перейдем теперь к описанию независимых событий в схеме Бернулли.

Определение 7.10. Говорят, что событие B_k зависит только от k-го испытания, если появление этого события определяется только лишь значением a_k.

Самым простым и самым важным событием такого рода является событие B_k вида

$$B_k = \{\omega : \;\; a_k = 1\}, \qquad\qquad \bar{B}_k = \{\omega : \;\; a_k = 0\}.$$

Иными словами, событие B_k заключается в появлении успеха, а $\bar{B}_k$ — неудачи на k-м испытании. Ясно, что

$$\Pr(B_k) = p, \qquad\qquad \Pr(\bar{B}_k) = q.$$

Несложно с помощью формальных выкладок также показать, что для всех $k \neq l$

$$\Pr(B_k \cap B_l) = p^2, \qquad \Pr(B_k \cap \bar{B}_l) = p \cdot q, \qquad \Pr(\bar{B}_k \cap \bar{B}_l) = q^2.$$

Следовательно, все такие события являются попарно независимыми. Более того, оказывается, что эти события являются также независимыми в совокупности, то есть такими событиями, для которых равенство

$$\Pr(B_{k_1} \cap B_{k_2} \cap \ldots \cap B_{k_i}) = \Pr(B_{k_1}) \cdot \Pr(B_{k_2}) \cdot \ldots \cdot \Pr(B_{k_i})$$

выполняется для любых $i = 1, \ldots, n$, $1 \leqslant k_1 < k_2 < \ldots < k_i \leqslant n$. Независимость событий B_k и дает формальное основание называть построенную вероятностную модель $(\Omega_n, \mathcal{A}_n, \Pr_n)$ моделью, отвечающую n независимым испытаниям с двумя исходами, или схемой Бернулли.

7.3.6. Заметим теперь, что формально та же самая схема Бернулли может быть построена и несколько по-другому. Именно, вернемся к элементарной вероятностной модели (33)–(34), описывающей однократное подбрасывание монетки, и рассмотрим декартовое произведение n множеств Ω_1:

$$\Omega_1 \times \Omega_1 \times \ldots \times \Omega_1 := \Omega.$$

Элементами ω этого множества Ω будут, как и в предыдущем случае, упорядоченные наборы вида $(a_1, a_2, \ldots, a_n)$, в которых $a_i \in \Omega_1$. Алгебра $\mathcal{A}$ ее подмножеств строится из множеств вида $A_1 \times \ldots \times A_1$, $A_1 \in \mathcal{A}_1$. Наконец, для любого $\omega = (a_1, a_2, \ldots, a_n)$ положим

$$\Pr(\omega) = \Pr_1(a_1) \cdot \Pr_1(a_2) \cdot \ldots \cdot \Pr_1(a_n).$$

Несложно убедиться, что получаемая в итоге тройка $(\Omega, \mathcal{A}, \Pr)$ образует вероятностную модель, полностью эквивалентную введенной выше модели $(\Omega_n, \mathcal{A}_n, \Pr_n)$, то есть также описывающую схему Бернулли.

7.3.7. Вернемся теперь к началу данного пункта и опишем общую схему построения такой вероятностной модели, в которой достаточно естественно возникают пары независимых событий. Сразу заметим, что эта схема легко обобщается на случай n независимых событий.

Пусть $(\Omega_1, \mathcal{A}_1, \Pr_1)$ и $(\Omega_2, \mathcal{A}_2, \Pr_2)$ — пара вероятностных моделей. Образуем декартово произведение пары множеств элементарных событий $\Omega_1 \times \Omega_2$, которое обозначим через Ω. Затем из набора подмножеств вида $A_1 \times A_2$, $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$ стандартным образом построим алгебру событий $\mathcal{A}$. Наконец, вероятность на этой алгебре событий определим так: для любых $A_1 \in \mathcal{A}_1$, $A_2 \in \mathcal{A}_2$

$$\Pr(A_1 \times A_2) = \Pr_1(A_1) \cdot \Pr_2(A_2).$$

Полученная тройка $(\Omega, \mathcal{A}, \Pr)$ в этом случае будет представлять собой некоторую вероятностную модель, в которой любые события вида (A_1, Ω_2) и (Ω_1, A_2) гарантированно являются независимыми. На практике большинство примеров независимых событий имеют именно такую природу.

8 Случайные величины

8.1. Во многих практических задачах нам не очень интересна конкретная природа пространства Ω элементарных событий — как правило, нас больше интересуют какие-то числовые характеристики того или иного случайного эксперимента. Например, в случайном эксперименте с n-кратным подбрасыванием монетки мы чаще следим за количеством успехов в n испытаниях, чем за результатами того или иного элементарного исхода. С формальной точки зрения описание таких числовых характеристик проводится на языке случайных величин.

8.1.1. Начнем с определения случайной величины.

Определение 8.1. Случайной величиной называется произвольная вещественная функция $\xi = \xi(\omega)$, определенная на (конечном) пространстве Ω элементарных событий.

Так как обычно в каждой конкретной задаче имеется лишь одно пространство Ω, то вместо $\xi(\omega)$ мы, как правило, будем просто писать ξ.

Пример 8.2. Рассмотрим случайный эксперимент, состоящий в подбрасывании двух игральных костей. Случайную величину $\xi = \xi(\omega)$ для такого эксперимента можно определить, например, как арифметическую сумму очков на выпавших гранях.

8.1.2. Предположим, что множество Ω элементарных исходов конечно ($|\Omega| = m$). Тогда конечным оказывается и множество $X = \{x_1, x_2, \ldots, x_n\}$ возможных значений, которые может принимать заданная на Ω случайная величина. Обозначим через

$$A_k = \{\omega \in \Omega : \ \xi(\omega) = x_k\} \tag{44}$$

событие, заключающееся в том, что ξ принимает заданное значение x_k. Вычислим для каждого $k = 1, \dots, n$ вероятность такого события:

$$\Pr(A_k) = \sum_{\omega:\ \xi(\omega)=x_k} \Pr(\omega).$$

Определение 8.3. Набор вероятностей $\{\Pr(A_1), \Pr(A_2), \dots, \Pr(A_n)\}$ случайных событий (44) называется *распределением вероятностей* случайной величины ξ.

Очевидно, что при любом исходе случайная величина какое-то значение обязательно примет. Как следствие, сумма всех вероятностей из данного набора обязана быть равной единице:

$$\sum_{k=1}^{n} \Pr(A_k) = 1.$$

Пример 8.4. Случайная величина ξ из примера 8.2 принимает 11 возможных значений $x_k \in X$, $X = \{2, 3, \dots, 12\}$. Соответствующее ξ распределение вероятностей здесь таково:

Значение x_k случайной величины ξ:	2	3	4	5	6	7	8	9	10	11	12
Вероятность $\Pr(\xi = x_k)$:	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Заметим, что отдельно взятая случайная величина ξ полностью характеризуется множеством X своих значений, а также распределением вероятностей этих значений. Детальная информация о структуре множества Ω всех элементарных событий нам в этом случае уже не важна. По сути, мы можем определить новое множество $\widetilde{\Omega} = X$ элементарных исходов, множество $\mathcal{B}$ всех его подмножеств, и ввести на $(X, \mathcal{B})$ вероятность $\Pr_\xi(B)$ по формуле

$$\Pr_\xi(B) = \Pr(\xi \in B) = \sum_{x_k \in B} \Pr(\xi = x_k).$$

Тройка $(X, \mathcal{B}, \Pr_\xi)$ в этом случае полностью определяет нам соответствующую вероятностную модель, в которой событие A_k вида "случайная величина ξ принимает заданное значение $x_k \in X$" является элементарным событием.

8.1.3. Довольно часто на практике распределения вероятностей случайной величины ξ сводятся к трем важным частным случаям — биномиальному распределению, геометрическому распределению и гипергеометрическому распределению.

Биномиальное распределение связано с разобранной в конце предыдущего параграфа схемой Бернулли. Именно, пусть n — количество испытаний в схеме Бернулли. Введем случайную величину ξ, принимающую значения из множества $X = \{0, 1, \dots, n\}$ и описывающую количество успехов и неудач в схеме Бернулли. Распределение вероятностей этой случайной величины совпадает с биномиальным распределением (43) и равно

$$\Pr(\xi = k) = \binom{n}{k} p^k q^{n-k}. \tag{45}$$

Определение 8.5. Распределение вероятностей (45) случайной величины ξ, значениями которой являются все возможные значения количества k успехов в схеме Бернулли, носит название *биномиального распределения* случайной величины ξ.

8.1.4. Со схемой Бернулли связано и так называемое геометрическое распределение дискретной случайной величины ξ. Именно, пусть в схеме Бернулли случайный эксперимент прекращается сразу, как только в этом эксперименте впервые случился успех. Ясно, что вероятность наступления такого события A_k равна

$$\Pr(A_k) = p \cdot q^{k-1}.$$

Введем случайную величину ξ, равную количеству проведенных в этом эксперименте испытаний.

Определение 8.6. Распределение вероятностей

$$\Pr(\xi = k) = p \cdot q^{k-1}, \qquad k = 1, 2, \ldots, \tag{46}$$

носит название *геометрического распределения вероятностей* случайной величины ξ.

Важно заметить, что в данной модели значение параметра k может быть любым натуральным числом. Следовательно, в этой модели дискретная случайная величина ξ принимает счетное число значений. При этом, как и следовало ожидать, сумма вероятностей

$$\sum_{k=1}^{\infty} \Pr(\xi = k) = \sum_{k=1}^{\infty} p \cdot q^{k-1} = \frac{p}{1-q} = 1.$$

8.1.5. Наконец, рассмотрим еще одно распределение случайной величины ξ, носящее название гипергеометрического распределения. Этому распределению отвечает следующая схема случайного эксперимента.

Пусть у нас имеется множество X элементов мощности $|X| = n + m$, состоящее из n элементов первого типа и m элементов второго типа. Предположим, что мы выбираем k элементов этого множества. Как мы знаем, это можно сделать $\binom{n+m}{k}$ количеством способов. Пусть теперь i — количество элементов первого вида, оказавшихся в выборке объема k. Введем случайную величину ξ, принимающую значения i, $i = 0, 1, \ldots, k$. Ясно, что вероятность

$$\Pr(\xi = i) = \frac{\binom{n}{i} \cdot \binom{m}{k-i}}{\binom{n}{k}}. \tag{47}$$

В силу тождества Вандермонта (11),

$$\left[\binom{n}{k}\right]^{-1} \sum_{i=0}^{k} \binom{n}{i} \cdot \binom{m}{k-i} = 1.$$

Определение 8.7. Распределение вероятностей (47) случайной величины ξ носит название *гипергеометрического распределения.*

8.2. Предположим теперь, что на одном и том же множестве Ω задана не одна, а две случайные величины ξ_1 и ξ_2.

8.2.1. Вернемся к примеру 8.2 с подбрасыванием двух кубиков.

Пример 8.8. Введем наряду с ξ случайную величину η, равную произведению чисел, выпавших на двух кубиках. Распределение вероятностей данной случайной величины таково:

Значение y_j:	1	2	3	4	5	6	8	9	10	12	15	16	18	20	24	25	30	36
$\Pr(\eta = y_j)$:	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Для того, чтобы охарактеризовать их поведение в совокупности, не зная ничего о структуре исходного вероятностного пространства, нам нужно, вообще говоря, задать, как говорят, их совместное распределение вероятностей, то есть сосчитать вероятности вида

$$\Pr(\xi_1 = x_k \text{ и } \xi_2 = y_j) \qquad \text{для всех} \quad x_k \in X, \quad y_j \in Y.$$

Иными словами, нам в данном случае в качестве множества значений нужно взять декартово произведение $X \times Y$ множества значений случайных величин ξ_1 и ξ_2, а затем определить вероятность для каждого элемента этого декартова произведения.

Пример 8.9. Для случайных величин ξ и η из примеров 8.2 и 8.8 мы в случае $\xi = x_1 = 2$, $\eta = y_1 = 1$ имеем $\Pr(\xi = 2, \eta = 1) = 1/36$. Действительно, имеется лишь один из 36 элементарных исходов, при котором сумма значений равна 2, а произведение равно 1 — элементарный исход, при котором на гранях двух кубиков выпали единицы. В случае $\xi = 2$, $\eta > 1$ вероятность $\Pr(\xi, \eta) = 0$.

8.2.2. В общем случае сосчитать совместное распределение вероятностей непросто. Однако существует важный частный случай, когда вычисление этого распределения сводится к задаче вычисления распределения вероятности каждой отдельно взятой случайной величины.

Определение 8.10. Говорят, что ξ_1 и ξ_2 являются *независимыми* случайными величинами, если

$$\Pr(\xi_1 = x_k \text{ и } \xi_2 = y_j) = \Pr(\xi_1 = x_k) \cdot \Pr(\xi_2 = y_j) \qquad \text{для всех} \quad x_k \in X, \quad y_j \in Y. \tag{48}$$

Пример 8.11. В примере с двумя кубиками мы можем ввести случайную величину ξ_1 как число очков на первом кубике, и случайную величину ξ_2 как количество очков на втором кубике. Ясно, что эти случайные величины являются независимыми.

Пример 8.12. Заметим, что случайные величины ξ и η из примеров 8.2 и 8.8 независимыми не являются. Это видно и чисто интуитивно: если нам сказали, что $\xi = 2$, то мы тут же можем сказать, что $\eta = 1$. Для формального доказательство данного факта достаточно, например, заметить, что произведение вероятностей $\Pr(\xi = x_k)$ и $\Pr(\eta = y_j)$ для любых допустимых значений x_k и y_j строго больше нуля, тогда как вероятность того, что $\xi = x_k$ и одновременно $\eta = y_j$, достаточно часто равна нулю — например, если $x_k = 2$, а $y_j > 1$.

8.2.3. В более общем случае рассмотрим некоторый набор $\{\xi_1, \xi_2, \ldots, \xi_m\}$ случайных величин, заданных на одном и том же множестве Ω и принимающих значения из некоторого конечного или счетного множества X.

Определение 8.13. Случайные величины $\xi_1, \xi_2, \ldots, \xi_m$ со значениями в одном и том же множестве X называются *независимыми (в совокупности)*, если для любых $x_i \in X$

$$\Pr(\xi_1 = x_{i_1}, \xi_2 = x_{i_2}, \ldots, \xi_m = x_{i_m}) = \Pr(\xi_1 = x_{i_1}) \cdot \Pr(\xi_2 = x_{i_2}) \cdot \ldots \cdot \Pr(\xi_m = x_{i_m}).$$

Пример 8.14. Простейший пример независимых в совокупности случайных величин можно получить, рассматривая схему Бернулли. Именно, пусть

$$\Omega = \{\omega:\ \omega = (a_1, a_2, \ldots, a_n), \quad a_i \in \{0,1\}\}, \qquad \Pr(\omega) = p^k q^{n-k}, \quad k = \sum_{i=0}^{n} a_i.$$

Мы заметили, что события вида

$$B_1 = \{\omega:\ a_1 = 1\}, \qquad B_2 = \{\omega:\ a_2 = 1\}, \qquad \ldots \qquad B_n = \{\omega:\ a_n = 1\}$$

являются независимыми в совокупности. Введем множество $X = \{0,1\}$ и рассмотрим n случайных величин вида $\xi_k(\omega) = a_k$, принимающих значения в этом множестве. Каждая такая случайная величина ξ_k характеризует результат случайного испытания в схеме Бернулли на k-м шаге. Свойство независимости событий B_k влечет тогда тот факт, что случайные величины $\xi_1, \ldots, \xi_n$ также являются независимыми в совокупности. Кроме того, каждая из этих n случайных величин имеют одинаковое распределение вероятностей

$$\Pr(\xi_k = 1) = p, \qquad \Pr(\xi_k = 0) = q.$$

Введенная таким образом совокупность $\xi_1, \xi_2, \ldots, \xi_n$ независимых и одинаково распределенных случайных величин носит название последовательности бернулиевских случайных величин.

8.2.4. Вернемся к примеру с двумя кубиками. Понятно, что любое значение случайной величины ξ из примера 8.2 представляет собой сумму значений случайных величин ξ_1 и ξ_2. Возникает вопрос — возможно ли по известным *распределениям вероятностей* случайных величин ξ_1 и ξ_2 вычислить *распределение вероятностей* случайной величины $\xi = \xi_1 + \xi_2$.

Итак, предположим, что мы имеем пару случайных величин ξ_1 и ξ_2, таких, что ξ_1 принимает значения из множества $X = \{x_1, x_2, \ldots, x_n\}$, а ξ_2 — из множества $Y = \{y_1, y_2, \ldots, y_m\}$. Введем новую случайную величину ξ как арифметическую сумму случайных величин ξ_1 и ξ_2. Ясно, что случайная величина $\xi = \xi_1 + \xi_2$ принимает значения из множества

$$Z = \{z \mid \text{z — всевозможные различные суммы } x_k + y_j, \quad k = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, m\}.$$

Распределение же вероятностей случайной величины ξ выражается через совместное распределение вероятностей случайных величин ξ_1 и ξ_2:

$$\Pr(\xi = z) = \Pr(\xi_1 + \xi_2 = z) = \sum_{(k,j):\, x_k + y_j = z} \Pr(\xi_1 = x_k \text{ и } \xi_2 = y_j).$$

Теперь понятно, что в случае, когда случайные величины ξ_1 и ξ_2 являются независимыми, последнее соотношение можно упростить и выразить распределение вероятностей ξ через известные распределения вероятностей случайных величин ξ_1 и ξ_2:

$$\Pr(\xi = z) = \Pr(\xi_1 + \xi_2 = z) = \sum_{(k,j):\, x_k + y_j = z} \Pr(\xi_1 = x_k) \cdot \Pr(\xi_2 = y_j) = \sum_{k=1}^{n} \Pr(\xi_1 = x_k) \cdot \Pr(\xi_2 = z - x_k).$$

В последней сумме вероятность $\Pr(\xi_2 = z - x_k)$ считается равной нулю, если $z - x_k \notin Y$.

8.2.5. В качестве характерного примера рассмотрим случайную величину $\xi = \xi_1 + \xi_2$, где ξ_1, ξ_2 — пара независимых бернуллиевских случайных величин. В этом случае $Z = \{0, 1, 2\}$, а распределение вероятностей случайной величины ξ описывается следующими соотношениями:

$$\Pr(\xi = 0) = \Pr(\xi_1 = 0) \cdot \Pr(\xi_2 = 0) = q \cdot q = q^2,$$

$$\Pr(\xi = 1) = \Pr(\xi_1 = 0) \cdot \Pr(\xi_2 = 1) + \Pr(\xi_1 = 1) \cdot \Pr(\xi_2 = 0) = 2\,p\,q,$$
$$\Pr(\xi = 2) = \Pr(\xi_1 = 1) \cdot \Pr(\xi_2 = 1) = p \cdot p = p^2.$$

Данный пример естественным образом обобщается на случай n независимых бернуллиевских случайных величин $\xi_1, \xi_2, \ldots, \xi_n$. Именно, сумма ξ_i представляет собой случайную величину $\xi = \xi_1 + \xi_2 + \ldots + \xi_n$, принимающую значения из множества $Z = \{0, 1, \ldots, n\}$ и описывающую количество успехов и неудач в схеме Бернулли. Распределение вероятностей этой случайной величины описывается биномиальным распределением (45).

8.3. Перейдем теперь к таким важным характеристикам случайной величины, как математическое ожидание и дисперсия.

8.3.1. Рассмотрим произвольный случайный эксперимент. Предположим, что на множестве Ω элементарных исходов этого случайного эксперимента задана случайная величина ξ, принимающая значения из множества X, $|X| = n$. Будем повторять этот случайный эксперимент N раз, и следить за значениями, которые принимает случайная величина ξ в этих экспериментах. Интуитивно ясно, что каждое значение $x_k \in X$ случайной величины ξ появляется в этих испытаниях $N \cdot p_k$ раз, где $p_k = \Pr(\xi = x_k)$. Тогда среднее значение, которое принимает случайная величина ξ в результате N случайных экспериментов, должно быть примерно равно

$$\frac{(N\,p_1)\,x_1 + (N\,p_2)\,x_2 + \ldots + (N\,p_n)\,x_n}{N} = \sum_{k=1}^{n} p_k \cdot x_k.$$

Данные рассуждения приводят нас к следующему определению.

Определение 8.15. Математическим ожиданием или средним значением случайной величины ξ называется сумма вида

$$\mathrm{E}(\xi) = \sum_{k=1}^{n} p_k \cdot x_k. \tag{49}$$

Пример 8.16. Колесо рулетки имеет равномерно расположенные по колесу ячейки, которые обозначены числами от 0 до 36. Предположим, что игрок ставит на некоторое число сумму, равную одному доллару. Если эта цифра выигрывает, то он получает 36 долларов, если проигрывает — то не получает ничего. Пусть ξ есть случайная величина, равная количеству денег, которые игрок выигрывает. Подсчитаем, сколько денег в среднем за n испытаний игрок может получить обратно.

Для этого заметим, что множество X возможных значений в данном случайном эксперименте есть $X = \{0, 36\}$. Первое значение случайная величина ξ принимает с вероятностью, равной $36/37$, а второе — с вероятностью $1/37$. Следовательно, по формуле (49) имеем

$$\mathrm{E}(\xi) = 36 \cdot \frac{1}{37} + 0 \cdot \frac{36}{37} = \frac{36}{37} = 0.973.$$

Как видим, в среднем за каждую попытку игрок теряет 0.027 долларов.

Пример 8.17. В случае однократного подбрасывания монетки

$$\mathrm{E}(\xi) = 1 \cdot p + 0 \cdot q = p.$$

Пример 8.18. В случае случайной величины ξ_1 из примера 8.11 имеем

$$\mathrm{E}(\xi_1) = (1 + 2 + 3 + 4 + 5 + 6)\,\frac{1}{6} = \frac{7}{2}.$$

8.3.2. Часто на практике помимо математического ожидания приходится использовать и другие числовые величины, в той или иной степени характеризующие среднее значение, принимаемое случайной величиной ξ. Одна из таких величин — это так называемая медиана распределения.

Определение 8.19. Медианой случайной величины ξ называется вещественное число m, такое, что

$$\Pr(\xi \leqslant m) \geqslant \frac{1}{2} \qquad \text{и} \qquad \Pr(\xi \geqslant m) \geqslant \frac{1}{2}.$$

Очень часто медиана используется в экономике для оценки среднего значения случайной величины при сильно неравномерных распределениях, то есть в случаях, когда матожидание дает не сильно содержательный с точки зрения здравого смысла результат. Приведем классический пример такой ситуации.

Пример 8.20. Предположим, что зарплата преподавателя в университете составляет 20 тысяч рублей, а зарплата члена ректората — 2 миллиона рублей. Пусть в вузе работают 96 преподавателей и 6 членов ректората. Возьмем случайную величину ξ, равную зарплате случайно выбранного работника вуза. Предполагая, что вероятность выбрать любого работника одинакова и равна $1/100$, мы получаем, что матожидание случайной величины ξ равно

$$\mathrm{E}(\xi) = \frac{1}{100}\left(6 \cdot 2 \cdot 10^6 + 96 \cdot 2 \cdot 10^4\right) = 139200.$$

Однако такую сумму, конечно же, не слишком разумно считать среднестатистической зарплатой работников этого учреждения. При этом медиана рассматриваемого случайного распределения в этом случае равна 20000 рублей. Действительно,

$$\Pr(\xi \leqslant 2 \cdot 10^4) = \Pr(\xi = 2 \cdot 10^4) = \frac{96}{100} \geqslant \frac{1}{2}, \qquad \Pr(\xi \geqslant 2 \cdot 10^4) = 1 \geqslant \frac{1}{2}.$$

При этом, конечно же, сумма в 20 тысяч рублей значительно точнее характеризует среднее значение заработной платы в данном университете.

8.3.3. Поговорим теперь об основных свойствах математического ожидания. Прежде всего, заметим, что иногда нам все же полезно вспомнить об исходном вероятностном пространстве $(\Omega, \mathcal{A}, \Pr)$ и переписать формулу (49) в следующем эквивалентном виде:

$$\mathrm{E}(\xi) = \sum_{\omega \in \Omega} \xi(\omega)\ \Pr(\omega). \tag{50}$$

Действительно, для всех ω, для которых $\xi(\omega) = x_k$, имеем

$$\sum_{\omega:\ \xi(\omega) = x_k} \Pr(\omega) = \Pr(\xi = x_k),$$

и мы от (50) приходим к (49).

Пусть у нас теперь имеются две произвольные случайные величины ξ_1 и ξ_2, определенные на одном и том же вероятностном пространстве $(\Omega, \mathcal{A}, \Pr)$. В этом случае, согласно формуле (50), математическое ожидание суммы $\xi = \xi_1 + \xi_2$ этих двух случайных величин равно сумме математических ожиданий $\mathrm{E}(\xi_1)$ и $\mathrm{E}(\xi_2)$:

$$\mathrm{E}(\xi_2 + \xi_2) = \sum_{\omega \in \Omega} (\xi_1(\omega) + \xi_2(\omega))\ \Pr(\omega) = \mathrm{E}(\xi_1) + \mathrm{E}(\xi_2).$$

Кроме того, из тех же соображений очевидно, что $\mathrm{E}(c\xi) = c\,\mathrm{E}(\xi)$ для любой константы $c \in \mathbb{R}$. Как следствие,

$$\mathrm{E}(c_1\xi_2 + c_2\xi_2) = c_1\,\mathrm{E}(\xi_1) + c_2\,\mathrm{E}(\xi_2) \qquad\qquad \forall\ c_1, c_2 \in \mathbb{R}.$$

Последнее равенство выражает собой свойство линейности математического ожидания.

Пример 8.21. Математическое ожидание $\mathrm{E}(\xi)$ случайной величины $\xi = \xi_1 + \xi_2$ из примера 8.2 равно 7.

В случае произведения случайных величин простая формула, связывающая математическое ожидание $\mathrm{E}(\xi_1 \cdot \xi_2)$ с математическими ожиданиями $\mathrm{E}(\xi_1)$ и $\mathrm{E}(\xi_2)$, имеет место лишь в том случае, когда ξ_1 и ξ_2 являются независимыми случайными величинами. В этом случае

$$\mathrm{E}(\xi_1 \cdot \xi_2) = \mathrm{E}(\xi_1) \cdot \mathrm{E}(\xi_2).$$

Действительно,

$$\mathrm{E}(\xi_1 \cdot \xi_2) = \sum_{\omega\in\Omega} \xi_1(\omega)\,\xi_2(\omega)\ \mathrm{Pr}(\omega) = \sum_{k,j} x_k y_j\ \mathrm{Pr}(\xi_1 = x_k \text{ и } \xi_2 = y_j) =$$

$$= \sum_{k,j} x_k y_j\ \mathrm{Pr}(\xi_1 = x_k)\ \mathrm{Pr}(\xi_2 = y_j) = \sum_{k} x_k\,\mathrm{Pr}(\xi_1 = x_k) \cdot \sum_{j} y_j\ \mathrm{Pr}(\xi_2 = y_j) = \mathrm{E}(\xi_1) \cdot \mathrm{E}(\xi_2).$$

Пример 8.22. Математическое ожидание $\mathrm{E}(\eta)$ случайной величины $\eta = \xi_1 \cdot \xi_2$ из примера 8.8 равно $\mathrm{E}(\xi_1) \cdot \mathrm{E}(\xi_2) = 7/2 \cdot 7/2 = 49/4$.

Пример 8.23. Для любой из n бернуллиевских случайных величин $\xi_1, \xi_2, \ldots, \xi_n$ математическое ожидание $\mathrm{E}(x_k)$ равно, очевидно, p. Следовательно, математическое ожидание случайной величины

$$\xi = \xi_1 + \xi_2 + \ldots + \xi_n$$

оказывается равным $\mathrm{E}(\xi) = n \cdot p$.

8.3.4. Перейдем теперь к понятию дисперсии $\mathrm{Var}(\xi)$ случайной величины ξ. Она характеризует степень разброса значений ξ относительно ее математического ожидания и определяется как средний квадрат отклонения от среднего:

$$\mathrm{Var}(\xi) := \mathrm{E}((\xi - \mathrm{E}(\xi))^2).$$

Так как

$$\mathrm{E}((\xi - \mathrm{E}(\xi))^2) = \mathrm{E}(\xi^2 - 2\xi\,\mathrm{E}(\xi) + (\mathrm{E}(\xi))^2) = \mathrm{E}(\xi^2) - 2\,\mathrm{E}(\xi) \cdot \mathrm{E}(\xi) + (\mathrm{E}(\xi))^2,$$

то

$$\mathrm{Var}(\xi) = \mathrm{E}(\xi^2) - (\mathrm{E}(\xi))^2. \tag{51}$$

Из последнего равенства, в частности, следует, что для произвольной константы $c \in \mathbb{R}$ дисперсия $\mathrm{Var}(c\xi) = c^2\,\mathrm{Var}(\xi)$.

Посмотрим теперь, чему равна дисперсия суммы случайных величин ξ_1 и ξ_2:

$$\mathrm{Var}(\xi_1 + \xi_2) = \mathrm{E}((\xi_1 + \xi_2)^2) - (\mathrm{E}(\xi_1 + \xi_2))^2 = \mathrm{E}(\xi_1^2 + 2\xi_1\xi_2 + \xi_2^2) - (\mathrm{E}(\xi_1) + \mathrm{E}(\xi_2))^2 =$$

$$= \mathrm{E}(\xi_1^2) + 2\,\mathrm{E}(\xi_1\xi_2) + \mathrm{E}(\xi_2^2) - \mathrm{E}(\xi_1)^2 - 2\,\mathrm{E}(\xi_1)\,\mathrm{E}(\xi_2) - \mathrm{E}(\xi_2)^2 = \mathrm{Var}(\xi_1) + \mathrm{Var}(\xi_2) + 2\mathrm{cov}(\xi_1, \xi_2),$$

где

$$\operatorname{cov}(\xi_1, \xi_2) := \mathrm{E}(\xi_1 \xi_2) - \mathrm{E}(\xi_1)\,\mathrm{E}(\xi_2) \tag{52}$$

есть так называемая *корреляция* двух случайных величин ξ_1 и ξ_2. Заметим, что в случае независимых случайных величин функция $\operatorname{cov}(\xi_1, \xi_2) = 0$, так что в этом случае дисперсия суммы равна сумме дисперсий:

$$\operatorname{Var}(\xi_1 + \xi_2) = \operatorname{Var}(\xi_1) + \operatorname{Var}(\xi_2).$$

Пример 8.24. Для случайной величины ξ_1 из примера 8.11 согласно (51) имеем

$$\operatorname{Var}(\xi_1) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) - \left(\frac{7}{2}\right)^2 = \frac{35}{12},$$

поэтому дисперсия случайной величины ξ из примера 8.2 равна

$$\operatorname{Var}(\xi) = \frac{35}{12} + \frac{35}{12} = \frac{35}{6}.$$

Пример 8.25. В случае бернуллиевской случайной величины ξ_1, принимающей два значения 0 и 1 с вероятностями q и p соответственно,

$$\operatorname{Var}(\xi_1) = \mathrm{E}(\xi_1^2) - (\mathrm{E}(\xi_1))^2 = (p \cdot 1^2 + q \cdot 0^2) - p^2 = p - p^2 = p(1-p) = pq.$$

Как следствие, если $\xi_1, \xi_2, \ldots, \xi_n$ есть последовательность бернуллиевских случайных величин и ξ есть их сумма, то $\operatorname{Var}(\xi) = npq$.

8.3.5. Теперь мы можем более строго обосновать тот факт, что дисперсия есть мера разброса случайной величины ξ.

Лемма 8.26. *Для любого вещественного $\alpha > 0$ справедливо неравенство*

$$\Pr\left((\xi - \mathrm{E}(\xi))^2 \geqslant \alpha\right) \leqslant \frac{\operatorname{Var}(\xi)}{\alpha}, \tag{53}$$

называемое неравенством Чебышева.

Доказательство. Обозначим через $\mu := \mathrm{E}(\xi)$. По определению,

$$\operatorname{Var}(\xi) = \mathrm{E}((\xi-\mu)^2) = \sum_{\omega \in \Omega} \Pr(\omega)\,(\xi(\omega) - \mu)^2 \geqslant \sum_{\omega \in \Omega\colon (\xi(\omega)-\mu)^2 \geqslant \alpha} \Pr(\omega)\,(\xi(\omega) - \mu)^2 \geqslant$$

$$\geqslant \sum_{\omega \in \Omega\colon (\xi(\omega)-\mu)^2 \geqslant \alpha} \Pr(\omega) \cdot \alpha = \alpha \cdot \Pr\left((\xi - \mu)^2 \geqslant \alpha\right).$$

□

Для того, чтобы лучше понять смысл неравенства (53), перепишем его в несколько ином виде. Введем наряду с $\operatorname{Var}(\xi)$ так называемое стандартное отклонение $\sigma = \sqrt{\operatorname{Var}(\xi)}$ случайной величины ξ и перейдем от параметра α к новому параметру c по формуле $\alpha = c^2 \operatorname{Var}(\xi) = c^2\sigma^2$. Условие $(\xi - \mu)^2 \geqslant \alpha$ в этом случае принимает вид $(\xi - \mu)^2 \geqslant c^2\sigma^2$, что равносильно тому, что $|\xi - \mu| \geqslant c\sigma$. В результате неравенство Чебышева (53) записывается так:

$$\Pr\left(|\xi - \mu| \geqslant c\sigma\right) \leqslant \frac{1}{c^2}.$$

Данное неравенство полезно в случае $c > 1$. В этом случае его можно трактовать так: ξ отклоняется от своего среднего значения μ больше, чем на c стандартных отклонений σ с вероятностью меньшей или равной $1/c^2$. Например, в случае $c = 2$ имеем $1/c^2 = 0.25$, и поэтому отклонение случайной величины ξ от μ не превосходит 2σ по крайней мере для 75% испытаний. В случае $c = 10$ этот процент возрастает до 99%.

9 Рекуррентные соотношения

9.1. Ранее мы уже несколько раз получали решения комбинаторных задач, записанные в виде тех или иных рекуррентных соотношений. Настало время поговорить о них немного поподробнее.

9.1.1. Начнем с простого примера [**?**].

Пример 9.1. Популяция лягушек в озере увеличивается в четыре раза каждый год. В последний день каждого года 100 лягушек отлавливают и переправляют на другие озера. Предполагая, что в начале первого года в озере было 50 лягушек, найти количество лягушек в начале любого последующего года.

Решение. Обозначим через a_n количество лягушек в начале $(n+1)$-го года. По условию задачи, $a_0 = 50$. Тогда, очевидно,

$$a_1 = 4 \cdot 50 - 100 = 100, \qquad a_2 = 4 \cdot 100 - 100 = 300,$$

а в общем случае

$$a_{n+1} = 4 \cdot a_n - 100, \qquad n = 0, 1, 2, \ldots$$

Полученное равенство является простейшим примером *рекуррентного соотношения.*

9.1.2. Перейдем теперь к формальным определениям.

Определение 9.2. Пусть $a_0, a_1, a_2, \ldots$ — произвольная числовая последовательность. Если для любого $n \geqslant 0$ число a_{n+m} является некоторой функцией от m предыдущих членов последовательности, т.е.

$$a_{n+m} = f_n(a_n, a_{n+1}, \ldots, a_{n+m-1}), \tag{54}$$

то такая последовательность называется рекуррентной последовательностью, а соотношение (54) — рекуррентным соотношением m-го порядка.

В частном случае линейной функции f имеем так называемое *линейное рекуррентное соотношение*

$$a_{n+m} = b_1(n)\, a_{n+m-1} + b_2(n)\, a_{n+m-2} + \ldots + b_{m-1}(n)\, a_{n+1} + b_m(n)\, a_n + u(n). \tag{55}$$

В случае $u(n) = 0$ оно называется *однородным,* в противном случае — *неоднородным.*

Самый простой случай рекуррентного соотношения (55) — это *линейное однородное рекуррентное соотношение с постоянными коэффициентами*

$$a_{n+m} = b_1\, a_{n+m-1} + b_2\, a_{n+m-2} + \ldots + b_{m-1}\, a_{n+1} + b_m\, a_n. \tag{56}$$

Очевидно, что для однозначного определения всех a_n необходимо наряду с самим рекуррентным соотношением (54) задать и первые m членов $a_0, a_1, \ldots, a_{m-1}$ данной последовательности, то есть, как говорят, *задать начальные условия* для рекуррентного соотношения.

9.1.3. Итак, наличие рекуррентного соотношения и начальных условий позволяет нам последовательно, шаг за шагом, определить любое наперед заданное количество n членов рекуррентной числовой последовательности. Зачастую это все, что нам требуется от задачи. Иными словами, получение рекуррентного соотношения для искомой числовой последовательности a_n является вполне приемлемым, а зачастую — и наиболее удобным с вычислительной точки зрения ответом для поставленной комбинаторной задачи.

Однако иногда нам полезно иметь явную формулу, которая для любого n позволяет вычислять значение a_n. Такую формулу называют *решением* рассматриваемого рекуррентного соотношения. Мы в данном параграфе покажем, как строить такие решения для случая линейного однородного рекуррентного соотношения с постоянными коэффициентами (56).

9.2. Прежде чем рассматривать общий случай уравнения (56), рассмотрим наиболее простой его вариант — линейное однородное рекуррентное соотношение первого порядка

$$a_{n+1} = b_1 \cdot a_n, \qquad n = 0, 1, 2, \ldots; \qquad a_0 \text{ — заданное число.} \tag{57}$$

9.2.1. Решение этого уравнения построить легко. Действительно,

$$a_1 = b_1 \cdot a_0; \qquad a_2 = b_1 \cdot a_1 = b_1^2 \cdot a_0; \qquad \ldots \qquad a_n = b_1^n \cdot a_0.$$

9.2.2. Предположим теперь, что нам кто-то сразу подсказал вид решения, а именно, что решение нашего уравнения (57) степенным образом зависит от n:

$$a_n = r^n \qquad \text{для некоторого } r.$$

Воспользовавшись этой подсказкой, подставим это выражение в исходное рекуррентное соотношение (57). В результате получается равенство вида

$$r^{n+1} = b_1 \cdot r^n,$$

из которого сразу же следует, что $r = b_1$; при этом a_n оказывается равным b_1^n. Это означает, что при $n = 0$ число $a_0 = b_1^0 = 1$. Иными словами, $a_n = b_1^n$ есть решение исходной задачи (57) в частном случае $a_0 = 1$. Поэтому такое решение называется *частным решением* уравнения (57).

9.2.3. Теперь попытаемся, зная частное решение, построить общее решение задачи (57). Для этого заметим, что в силу однородности уравнения (57) любое его частное решение, умноженное на произвольную постоянную c, по-прежнему этому уравнению удовлетворяет:

$$c \cdot b_1^{n+1} \equiv c \cdot b_1 \cdot b_1^n.$$

Решение же вида $c \cdot b_1^n$ позволяет удовлетворить любому начальному условию. Действительно, подставляя его в начальное условие для уравнения (57), получим:

$$c \cdot b_1^0 = a_0 \qquad \Longrightarrow \qquad c = a_0 \qquad \Longrightarrow \qquad a_n = a_0 \cdot b_1^n.$$

По этой причине решение вида $c \cdot b_1^n$ называется *общим решением* уравнения (57).

9.3. Подведем предварительные итоги. Так как исходное уравнение (57) было очень простым, то нам удалось сразу построить его решение и выяснить, что оно степенным образом зависит от параметра n. Затем мы заметили, что если бы нам кто-то заранее подсказал степенной характер решения уравнения (57), то нам хватило бы этой информации для построения как общего, так и частного решения нашего рекуррентного уравнения.

Возникает вопрос: зачем же нам нужен для столь простого уравнения столь сложный алгоритм построения его решения? Оказывается, что этот алгоритм практически без изменений работает и для построения решения линейного однородного уравнения произвольного порядка.

9.3.1. Рассмотрим в качестве примера линейное однородное рекуррентное соотношение второго порядка

$$a_{n+2} = b_1 \cdot a_{n+1} + b_2 \cdot a_n, \qquad n = 0, 1, 2, \ldots, \qquad a_0,\ a_1 \text{ — заданные числа,} \tag{58}$$

и попытаемся применить к этому уравнению алгоритм, описанный в конце предыдущего пункта.

Для этого предположим, что частное решение уравнения (58) по прежнему степенным образом зависит от параметра n, то есть предположим, что существуют такие значения параметров a_0 и a_1, при которых $a_n = r^n$. Подставляя это выражение в рекуррентное соотношение, получим

$$r^{n+2} = b_1 \cdot r^{n+1} + b_2 \cdot r^n \qquad \Longrightarrow \qquad r^2 - b_1\, r - b_2 = 0, \tag{59}$$

т.е. квадратное уравнение на r. Любое его решение r_0 дает нам некоторое частное решение уравнения (58). По этой причине данное уравнение называется *характеристическим уравнением* для рекуррентного соотношения (58).

9.3.2. Предположим, что характеристическое уравнение (59) имеет два различных вещественных корня r_1 и r_2. Покажем, что в таком случае выражение

$$a_n = c_1\, r_1^n + c_2\, r_2^n, \tag{60}$$

где c_1 и c_2 — произвольные постоянные, является *общим решением* соотношения (58) в том смысле, что любое решение (58) с заданными начальными условиями в нем содержится.

Действительно, покажем вначале, что выражение (60) действительно удовлетворяет рекуррентному соотношению (58):

$$\begin{gathered}
c_1\, r_1^{n+2} + c_2\, r_2^{n+2} = b_1\, c_1\, r_1^{n+1} + b_1\, c_2\, r_2^{n+1} + b_2\, c_1\, r_1^n + b_2\, c_2\, r_2^n \qquad \Longleftrightarrow \\
\Longleftrightarrow \qquad c_1\, r_1^n\,(r_1^2 - b_1\, r_1 - b_2) + c_2\, r_2^n\,(r_2^2 - b_1\, r_2 - b_2) \equiv 0.
\end{gathered}$$

Покажем теперь, что это — действительно общее решение, т.е. что мы всегда можем подобрать константы c_1 и c_2 так, чтобы решение вида (60) удовлетворяло любым заданным начальным условиям. Для этого рассмотрим это выражение при $n = 0$ и $n = 1$:

$$\begin{aligned}
a_0 &= c_1\, r_1^0 + c_2\, r_2^0 = c_1 + c_2, \\
a_1 &= c_1\, r_1^1 + c_2\, r_2^1 = c_1\, r_1 + c_2\, r_2.
\end{aligned}$$

Эти выражения следует рассматривать как систему линейных уравнений для определения неизвестных постоянных c_1 и c_2. Определитель этой системы

$$\begin{vmatrix} 1 & 1 \\ r_1 & r_2 \end{vmatrix} = r_2 - r_1 \neq 0,$$

поэтому система всегда имеет единственное решение.

9.3.3. Пожалуй, самым известным и важным примером рекуррентного соотношения (58) является соотношение вида

$$F_{n+2} = F_{n+1} + F_n, \qquad n = 0, 1, 2, \ldots, \qquad F_0 = 0,\ \ F_1 = 1,$$

определяющее так называемые числа Фибоначчи F_n. Характеристическое уравнение для этого рекуррентного уравнения имеет вид

$$r^2 = r + 1 \qquad \Longrightarrow \qquad r_{1,2} = \frac{1 \pm \sqrt{5}}{2} \qquad \Longrightarrow \qquad F_n = c_1\, \left(\frac{1 + \sqrt{5}}{2}\right)^n + c_2\, \left(\frac{1 - \sqrt{5}}{2}\right)^n.$$

Константы c_1 и c_2 определим из начальных условий:

$$F_0 = 0 = c_1 + c_2$$
$$F_1 = 1 = c_1 \frac{1+\sqrt{5}}{2} + c_2 \frac{1-\sqrt{5}}{2} = c_1 \left[\frac{1+\sqrt{5}-1+\sqrt{5}}{2}\right] = c_1 \sqrt{5} \qquad \Longrightarrow$$

$$\Longrightarrow \qquad c_1 = +\frac{1}{\sqrt{5}}, \qquad c_2 = -\frac{1}{\sqrt{5}}.$$

Таким образом, окончательно получаем следующее явное выражение для чисел Фибоначчи:

$$F_n = \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n. \tag{61}$$

9.3.4. Предположим теперь, что характеристическое уравнение (59) имеет ровно один кратный корень $r_1 = r_2 =: \rho$. Покажем, что в этом случае общее решение уравнения (58) имеет вид

$$a_n = c_1 \rho^n + c_2 n \rho^n.$$

Мы уже показали, что частное решение вида ρ^n удовлетворяет уравнению (58). С учетом линейности и однородности этого уравнения нам осталось показать, что этому уравнению удовлетворяет и частное решение вида $n \rho^n$. Подставляя в уравнение (58) выражение $n \rho^n$, получим

$$(n+2)\, \rho^{n+2} = b_1 (n+1)\, \rho^{n+1} + b_2\, n\, \rho^n \qquad \Longrightarrow \qquad n\,(\rho^2 - b_1 \rho - b_2) + \rho\,(2\rho - b_1) = 0.$$

Первое слагаемое в левой части этого выражения равно нулю в силу характеристического уравнения. Для того, чтобы понять, что равно нулю и последнее слагаемое, заметим, что в случае совпадающих корней $r_1 = r_2 = \rho$ имеем

$$\rho = \frac{b_1 \pm 0}{2} \qquad \Longrightarrow \qquad 2\rho - b_1 = 0.$$

Осталось показать, что такой вид решения может удовлетворить любым начальным условиям. Подставив начальные условия в выражение для a_n, получим следующую систему двух линейных алгебраических уравнений относительно параметров c_1 и c_2:

$$c_1 = a_0,$$

$$\rho\,(c_1 + c_2) = a_1.$$

Понятно, что такая система имеет решение при любых a_0, a_1 и $\rho \neq 0$. Случай же $\rho = 0$ тривиален — в этом случае $b_1 = b_2 = 0$, и потому $a_n = 0$ для любых n.

9.3.5. Наконец, давайте рассмотрим случай, когда уравнение (59) имеет два комплексно сопряжённых корня

$$r_1 = x + iy = \rho\, e^{i\vartheta}, \qquad r_2 = x - iy = \rho\, e^{-i\vartheta}, \qquad \rho, \vartheta \neq 0.$$

Покажем, что в таком случае общее решение линейного однородного рекуррентного соотношения с постоянными коэффициентами второго порядка (58) можно записать так:

$$a_n = \widetilde{c}_1\, \rho^n\, \cos(n\,\vartheta) + \widetilde{c}_2\, \rho^n\, \sin(n\,\vartheta). \tag{62}$$

Действительно, рассуждения, аналогичные проведенным для случая двух различных вещественных корней, показывают, что линейная комбинация частных решений вида

$$a_n = c_1\, r_1^n + c_2\, r_2^n$$

удовлетворяет рекуррентному соотношению (58) при любых c_1 и c_2. Рассмотрим теперь систему линейных уравнений

$$\begin{aligned} a_0 &= c_1 + c_2, \\ a_1 &= c_1\, \rho\, e^{i\vartheta} + c_2\, \rho\, e^{-i\vartheta} = \rho\,(c_1 + c_2) \cos\vartheta + \rho\, i\,(c_1 - c_2) \sin\vartheta. \end{aligned} \tag{63}$$

Определитель этой системы по-прежнему отличен от нуля, так что мы с помощью общего решения по-прежнему можем удовлетворить любым начальным условиям. Из (63) также следует, что коэффициенты $\widetilde{c}_1 = c_1 + c_2$ и $\widetilde{c}_2 = i\,(c_1 - c_2)$ являются вещественными числами. Действительно, первое уравнение (63) гарантирует нам, что сумма $c_1 + c_2$ является вещественным числом. Но тогда вещественным числом обязано быть и выражение $i\,(c_1 - c_2)$.

Заметим теперь, что линейную комбинацию частных решений мы в рассматриваемом случае можем переписать так:

$$a_n = c_1\, r_1^n + c_2\, r_2^n = c_1\, \rho^n\, e^{i\,\vartheta\, n} + c_2\, \rho^n\, e^{-i\,\vartheta\, n} =$$

$$= (c_1 + c_2)\, \rho^n\, \cos(n\,\vartheta) + i\,(c_1 - c_2)\, \rho^n\, \sin(n\,\vartheta).$$

Обозначив через $\widetilde{c}_1 = c_1 + c_2$ и $\widetilde{c}_2 = i\,(c_1 - c_2)$, мы получаем для a_n формулу (62).

9.3.6. Описанная выше техника решения линейных однородных рекуррентных соотношений с постоянными коэффициентами обобщается и на случай соотношений более высокого порядка. Именно, рассмотрим линейное однородное рекуррентное соотношение m-го порядка (56)

$$a_{n+m} = b_1 \cdot a_{n+m-1} + b_2 \cdot a_{n+m-2} + \ldots + b_m \cdot a_n, \qquad a_0, \ldots, a_{m-1} \text{ — заданы.}$$

Предположим, что частное решение этого соотношения имеет вид $a_n = r^n$. Подставляя это выражение в исходное рекуррентное соотношение, получаем для параметра r следующее характеристическое уравнение:

$$r^m = b_1 \cdot r^{m-1} + b_2 \cdot r^{m-2} + \ldots + b_m \cdot r.$$

Пусть $r_1, \ldots, r_k$ есть корни этого уравнения. Каждый такой корень дает свой вклад в общее решение рассматриваемого рекуррентного соотношения. В частности, если r_j есть корень кратности q_j, то вклад этого корня в общее решение задается выражением вида

$$c_{1,j} \cdot r_j^n + c_{2,j} \cdot n \cdot r_j^n + \ldots + c_{q_j,j} \cdot n^{q_j - 1} \cdot r_j^n.$$

Так, если характеристическое уравнение для некоторого рекуррентного соотношения шестого порядка имеет вид

$$(r-1)^3 \cdot (r-2)^2 \cdot (r-3) = 0,$$

то общее решение такого соотношения записывается так:

$$a_n = c_1 + c_2 \cdot n + c_3 \cdot n^2 + c_4 \cdot 2^n + c_5 \cdot n \cdot 2^n + c_6 \cdot 3^n.$$

9.4. Вернемся к примеру о лягушках, рассмотренному вначале параграфа. В отличие от разобранных выше примеров, рекуррентное соотношение, полученное для этой задачи, было неоднородным. Постараемся понять, как нам решать такие соотношения.

9.4.1. Рекуррентное соотношение

$$a_{n+1} = 4 \cdot a_n - 100, \qquad\qquad a_0 = 50$$

представляет собой частный случай линейного неоднородного рекуррентного соотношения с постоянными коэффициентами первого порядка. В общем случае такое соотношение имеет следующий вид:

$$a_{n+1} = b_1 \cdot a_n + u, \qquad\qquad a_0 \text{ — заданное число.} \tag{64}$$

Как и в случае однородного рекуррентного соотношения (57), решение рекуррентного соотношения (64) довольно легко построить, последовательно выражая a_n через a_0 и u. Именно,

$$\begin{gathered} a_1 = b_1 \cdot a_0 + u, \\ a_2 = b_1 \cdot a_1 + u = b_1 \cdot (b_1 \cdot a_0 + u) + u = b_1^2 \cdot a_0 + b_1 \cdot u + u, \\ a_3 = b_1 \cdot a_2 + u = b_1 \cdot (a_0 \cdot b_1^2 + b_1 \cdot u + u) + u = b_1^3 \cdot a_0 + b_1^2 \cdot u + b_1 \cdot u + u, \\ \ldots \\ a_n = b_1^n \cdot a_0 + (b_1^{n-1} + b_1^{n-2} + \ldots + b_1 + 1) \cdot u. \end{gathered}$$

В зависимости от значений параметра b_1 мы имеем два варианта развития событий. В случае $b_1 \neq 1$ мы выражение для a_n можем переписать так:

$$a_n = b_1^n \cdot a_0 + \frac{b_1^n - 1}{b_1 - 1} \cdot u = \left(a_0 - \frac{u}{1 - b_1}\right) \cdot b_1^n + \frac{u}{1 - b_1}.$$

В случае $b_1 = 1$ мы получаем

$$a_n = a_0 + n \cdot u.$$

9.4.2. Из приведенных выше рассуждений мы можем сделать два важных наблюдения. Первое из них заключается в том, что в обоих случаях решение a_n неоднородного рекуррентного соотношения может быть записано в виде

$$a_n = c_1 \cdot r^n + q_n,$$

где $c_1 \cdot r^n$ есть общее решение соответствующего (64) однородного рекуррентного соотношения

$$a_{n+1}^{(0)} = b_1 \cdot a_n^{(0)},$$

а q_n есть некоторое частное решение неоднородного рекуррентного соотношения (64), то есть решение, при подстановке которого в рекуррентное соотношение последнее превращается в тождество. Оказывается, что этот факт справедлив и в общем случае. Именно, справедлива следующая

Теорема 9.3. *Пусть q_n есть частное решение рекуррентного соотношения*

$$a_{n+m} = b_1\, a_{n+m-1} + b_2\, a_{n+m-2} + \ldots + b_{m-1}\, a_{n+1} + b_m\, a_n + u(n).$$

Тогда любое решение этого рекуррентного соотношения имеет вид $a_n = p_n + q_n$, где p_n есть общее решение соответствующего однородного рекуррентного соотношения

$$a_{n+m}^{(0)} = b_1\, a_{n+m-1}^{(0)} + b_2\, a_{n+m-2}^{(0)} + \ldots + b_{m-1}\, a_{n+1}^{(0)} + b_m\, a_n^{(0)}.$$

Второе важное наблюдение состоит в том, что мы легко можем построить решение рекуррентного соотношения в случае, если нам кто-то подскажет конкретный вид как общего решения однородного соотношения, так и частного решения неоднородного соотношения. Предположим, например, что нам кто-то подсказал, что в случае $b_1 \neq 1$ частное решение (64) следует искать в виде $q(n) = c_0 \cdot u$, а общее решение соответствующего (64) однородного соотношения — в виде $a_n = c_1 \cdot r^n$. Подставляя в (64) вместо q_n выражение $c_0 \cdot u$, получаем

$$c_0 \cdot u = b_1 \cdot c_0 \cdot u + u \qquad \Longrightarrow \qquad c_0 = \frac{1}{1 - b_1}.$$

Далее, подставляя вместо $a_n^{(0)}$ в соответствующее (64) однородное рекуррентное соотношение

$$a_{n+1}^{(0)} = b_1 \cdot a_n^{(0)}$$

выражение $c_1 \cdot r^n$, мы находим, что $r = b_1$. Осталось определить константу c_1. Для этого воспользуемся начальным условием:

$$c_1 \cdot b_1^0 + \frac{u}{1 - b_1} = a_0 \qquad \Longrightarrow \qquad c_1 = a_0 - \frac{u}{1 - b_1}.$$

Аналогичный алгоритм проходит и в случае $b_1 = 1$.

9.4.3. Мы уже знаем, в каком виде следует искать общее решение однородного рекуррентного соотношения. Осталось понять, как определять вид частного решения неоднородного рекуррентного соотношения. Ранее мы отмечали, что общее решение соответствующего однородного рекуррентного соотношения представляет собой сумму функций $f_j(n)$ вида

$$f_j(n) = (c_{1,j} + c_{2,j} \cdot n + c_{3,j} \cdot n^2 + \ldots + c_{q_j,j} \cdot n^{q_j - 1}) \cdot r_j^n.$$

Предположим, что неоднородность $u_n \equiv u(n)$ имеет вид

$$u(n) = (d_0 + d_1 \cdot n + \ldots + d_l \cdot n^l) \cdot r^n =: P_l(n) \cdot r^n.$$

Тогда частное решение исходного неоднородного уравнения следует искать в виде

$$(c_0 + c_1 \cdot n + \ldots + c_l \cdot n^l) \cdot n^q \cdot r^n,$$

где $q = 0$ в случае, если r отличен от корней r_j характеристического уравнения, и равно кратности корня r_j в случае, если r совпадает с одним из корней r_j. В случае, когда неоднородность представляет собой сумму описанных выше функций $u_t(n)$, частное решение рекуррентного соотношения можно искать в виде суммы частных решений неоднородных рекуррентных соотношений вида

$$a_{n+m} = b_1 \cdot a_{n+m-1} + b_2 \cdot a_{n+m-2} + \ldots + b_m \cdot a_n + u_t(n).$$

9.4.4. Разберем несколько примеров решения неоднородных рекуррентных соотношений.

Пример 9.4. Построить общее решение неоднородного рекуррентного соотношения второго порядка

$$a_{n+2} = 4a_{n+1} - 4a_n + 3 \cdot 2^n.$$

Решение. Характеристическое уравнение для соответствующего однородного рекуррентного соотношения имеет следующий вид:

$$r^2 - 4r + 4 = 0 \qquad \Longleftrightarrow \qquad (r - 2)^2 = 0.$$

Как следствие, общее решение соотоветствующего однородного рекуррентного соотношения имеет вид

$$c_1 \cdot 2^n + c_2 \cdot n \cdot 2^n.$$

Так как $r = 2$ есть корень кратности 2 характеристического уравнения, то частное решение исходного неоднородного рекуррентного соотношения следует искать в виде

$$c_0 \cdot n^2 \cdot 2^n.$$

Подставим это выражение в рекуррентное соотношение:

$$c_0 \cdot (n+2)^2 \cdot 2^{n+2} = 4c_0 \cdot (n+1)^2 \cdot 2^{n+1} - 4c_0 \cdot n^2 \cdot 2^n + 3 \cdot 2^n \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \qquad 4c_0(n^2+4n+4) = 8c(n^2+2n+1) - 4cn^2 + 3 \qquad \Longleftrightarrow \qquad 8c_0 = 3.$$

Следовательно, $c_0 = 3/8$, так что частное решение заданного рекуррентного соотношения имеет вид

$$3 \cdot n^2 \cdot 2^{n-3}.$$

Пример 9.5. Построить общее решение следующего линейного неоднородного рекуррентного соотношения второго порядка:

$$a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot n + 3^n \cdot \cos(\pi\, n/2).$$

Решение. Для построения общего решения соответствующего однородного рекуррентного соотношения запишем его характеристическое уравнение:

$$r^2 = 6r - 9 \qquad \Longleftrightarrow \qquad (r-3)^2 = 0.$$

Как видно, это уравнение имеет корень $r = 3$ кратности два. Следовательно, общее решение однородного рекуррентного соотношения записывается в виде

$$(d_1 + d_2 \cdot n) \cdot 3^n.$$

Неоднородная часть рекуррентного соотношения состоит из двух слагаемых:

$$u_1(n) = 3^n \cdot n \qquad \text{и} \qquad u_2(n) = 3^n \cdot \cos(\pi\, n/2).$$

Так как

$$\cos(\pi\, n/2) = \frac{e^{i\,\pi\, n/2} + e^{-i\,\pi\, n/2}}{2},$$

то выражение для $u_2(n)$ можно переписать так:

$$u_2(n) = \frac{1}{2}\left(r_1^n + r_2^n\right), \qquad \text{где} \qquad r_1 = 3 \cdot e^{i\,\pi/2}, \qquad r_2 = 3 \cdot e^{-i\,\pi/2}.$$

Иными словами, неоднородная часть рекуррентного соотношения состоит из трех слагаемых вида $r^n \cdot P_i(n)$, и только в первом из них показатель совпадает с корнем характеристического уравнения. Как следствие, мы можем искать отдельно частное решение, связанное с первым слагаемым, и частное решение, связанное с оставшимися слагаемыми.

Рассмотрим вначале рекуррентное соотношение

$$a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot n.$$

Частное его решение следует искать в виде $(c_1 + c_2 \cdot n) \cdot n^2 \cdot 3^n$. Подставляя это выражение в рекуррентное соотношение, получим равенство

$$54n \cdot c_2 + 54 \cdot c_2 + 18 \cdot c_1 = n.$$

Приравнивая коэффициенты при одинаковых степенях n, получаем значения констант $c_2 = = 1/54$, $c_1 = -1/18$.

Теперь запишем рекуррентное соотношение

$$a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot \cos(\pi\, n/2) \quad \Longleftrightarrow \quad a_{n+2} = 6a_{n+1} - 9a_n + 3^n \cdot e^{i\,\pi\, n/2}/2 + 3^n \cdot e^{-i\,\pi\, n/2}/2.$$

Частное его решение можно искать в виде умноженной на 3^n линейной комбинации экспонент с показателями $\pm\, i\,\pi\, n/2$, либо, что более удобно, в виде линейной комбинации синуса и косинуса, умноженной на 3^n:

$$c_3 \cdot 3^n \cdot \cos(\pi\, n/2) + c_4 \cdot 3^n \cdot \sin(\pi\, n/2).$$

Подставляя это выражение в записанное выше рекуррентное соотношение, получаем равенство

$$18 \cdot c_3 \cdot \sin(\pi\, n/2) - (18 \cdot c_4 + 1) \cdot \cos(\pi\, n/2) = 0,$$

из которого следует, что $c_3 = 0$, $c_4 = -1/18$.

Окончательно получаем следующее общее решение исходного рекуррентного соотношения:

$$(d_1 + d_2 \cdot n) \cdot 3^n + (n - 3) \cdot n^2 \cdot 3^{n-3}/2 - 3^{n-2} \cdot \sin(\pi\, n/2)/2.$$

10 Производящие функции и линейные рекуррентные соотношения

10.1. Читатели, хорошо знакомые с математическим анализом, уже, видимо, заметили, насколько похожи методы решения линейных рекуррентных соотношений и методы решения линейных обыкновенных дифференциальных уравнений с постоянными коэффициентами.

10.1.1. Действительно, рассмотрим линейное однородное дифференциальное уравнение первого порядка вида

$$\frac{dy}{dx} = b_1 \cdot y(x), \qquad y(x_0) = y_0.$$

Такое уравнение легко решается с помощью разделения переменных:

$$\frac{dy}{y} = b_1 \cdot dx \qquad \Longleftrightarrow \qquad d\ln(y) = b_1 \cdot dx \qquad \Longleftrightarrow \qquad \int\limits_{y_0}^{y} d\ln(y) = b_1 \cdot \int\limits_{x_0}^{x} dx \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \qquad \ln(y) - \ln(y_0) = \ln\frac{y}{y_0} = b_1 \cdot (x - x_0) \qquad \Longleftrightarrow \qquad y = y_0\, e^{b_1 \cdot (x - x_0)}.$$

Как и в случае рекуррентных соотношений, мы могли бы решить это уравнение и без использования метода разделения переменных — для его решения достаточно знать характер решения такого дифференциального уравнения. Именно, предположим, что решение $y(x)$ экспоненциально зависит от x:

$$y(x) = e^{\lambda \cdot x}.$$

Подставим это выражение в исходное дифференциальное уравнение:

$$y'(x) = \lambda \cdot e^{\lambda \cdot x} = b_1 \cdot e^{\lambda \cdot x} \qquad \Longleftrightarrow \qquad \lambda = b_1.$$

Заметим теперь, что решение $y(x) = e^{b_1 \cdot x}$ удовлетворяет начальным условиям $y(x_0) = e^{b_1 \cdot x_0}$. Для того, чтобы удовлетворить произвольным начальным условиям, достаточно воспользоваться однородностью данного уравнения и рассмотреть вместо частного общее решение вида $y(x) = c \cdot e^{b_1 \cdot x}$. Константа c легко находится из начальных условий:

$$y(x_0) = y_0 = c \cdot e^{b_1 \cdot x_0} \qquad \Longrightarrow \qquad c = y_0 \cdot e^{-b_1 \cdot x_0} \qquad \Longrightarrow \qquad y = y_0 \cdot e^{b_1 (x - x_0)}.$$

10.1.2. Теперь рассмотрим линейное однородное дифференциальное уравнение второго порядка с постоянными коэффициентами

$$y''(x) = b_1 \cdot y'(x) + b_2 \cdot y(x), \qquad\qquad y(0) = x_0, \quad y'(0) = x_1.$$

Предполагая, что характер решения этого уравнения остался тем же, что и в предыдущем случае, а именно, полагая, что $y(x) = e^{\lambda \cdot x}$, мы в результате подстановки этого выражения в исходное дифференциальное уравнение получаем абсолютно то же самое характеристическое уравнение на параметр λ, что и в случае линейного однородного рекуррентного соотношения второго порядка с постоянными коэффициентами:

$$y''(x) = \lambda^2 \cdot e^{\lambda \cdot x} = b_1 \cdot \lambda \cdot e^{\lambda \cdot x} + b_2 \cdot e^{\lambda \cdot x} \qquad \Longrightarrow \qquad \lambda^2 - b_1 \cdot \lambda - b_2 = 0.$$

Если теперь это уравнение имеет пару различных вещественных корней λ_1 и λ_2, то общее решение обыкновенного дифференциального уравнения в силу его линейности и однородности записывается в виде

$$y(x) = c_1 \cdot e^{\lambda_1 \cdot x} + c_2 \cdot e^{\lambda_2 \cdot x}.$$

В случае единственного вещественного корня λ характеристического уравнения решение попрежнему ищется в виде

$$y(x) = c_1 \cdot e^{\lambda \cdot x} + c_2 \cdot x \cdot e^{\lambda \cdot x}.$$

Наконец, в случае комплексных корней $\lambda_{1,2} = \alpha \pm i \cdot \beta$ решение обыкновенного дифференциального уравнения можно искать в виде

$$y(x) = c_1 \cdot e^{\alpha \cdot x} \cos(\beta \cdot x) + c_2 \cdot e^{\alpha \cdot x} \sin(\beta \cdot x).$$

10.1.3. Отмеченная аналогия наводит нас на мысль о существовании взаимно-однозначного отображения, позволяющего сопоставить линейному однородному рекуррентному соотношению с постоянными коэффициентами аналогичное линейное однородное дифференциальное уравнение с постоянными коэффициентами. Оказывается, что такое отображение построить достаточно просто. Рассмотрим для определенности линейное однородное рекуррентное соотношение с постоянными коэффициентами второго порядка:

$$a_{n+2} = b_1 \cdot a_{n+1} + b_2 \cdot a_n, \qquad\qquad a_0,\ a_1 \text{ — заданные числа.}$$

Домножим это уравнение на $x^n/n!$ и просуммируем по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} a_{n+2} \frac{x^n}{n!} = b_1 \sum_{n=0}^{+\infty} a_{n+1} \frac{x^n}{n!} + b_2 \sum_{n=0}^{+\infty} a_n \frac{x^n}{n!}.$$

Предположим теперь, что ряд

$$a_0 + a_1\frac{x^1}{1!} + a_2\frac{x^2}{2!} + \ldots + a_n\frac{x^n}{n!} + \ldots = \sum_{n=0}^{+\infty} a_n\frac{x^n}{n!}$$

представляет собой ряд Тейлора некоторой функции $y(x)$, построенный в окрестности точки $x = 0$:

$$y(x) = \sum_{n=0}^{+\infty} a_n\frac{x^n}{n!}, \qquad a_n = y^{(n)}(0).$$

Тогда

$$y'(x) = a_1 + a_2\frac{x^1}{1!} + a_3\frac{x^2}{2!} + \ldots + a_n\frac{x^{n-1}}{(n-1)!} + a_{n+1}\frac{x^n}{n!} + \ldots = \sum_{n=0}^{+\infty} a_{n+1}\frac{x^n}{n!},$$

$$y''(x) = a_2 + a_3\frac{x^1}{1!} + a_4\frac{x^2}{2!} + \ldots + a_n\frac{x^{n-2}}{(n-2)!} + a_{n+1}\frac{x^{n-1}}{(n-1)!} + +a_{n+2}\frac{x^n}{n!} + \ldots = \sum_{n=0}^{+\infty} a_{n+2}\frac{x^n}{n!},$$

так что исходное рекуррентное соотношение превращается в обыкновенное дифференциальное уравнение вида

$$y''(x) = b_1 \cdot y'(x) + b_2 \cdot y(x)$$

с начальными условиями $y(0) = a_0$, $y'(0) = a_1$. Обратно, раскладывая функцию $y(x)$ в ряд Тейлора в окрестности точки $x = 0$, подставляя это разложение в обыкновенное дифференциальное уравнение и приравнивая коэффициенты при одинаковых степенях $x^n/n!$, мы из исходного обыкновенного дифференциального уравнения получаем рекуррентное соотношение на коэффициенты $a_n = y^{(n)}(0)$.

10.1.4. Полученное нами взаимно-однозначное соответствие полезно как для решения обыкновенных дифференциальных уравнений, так и для решения рекуррентных соотношений. Так, с вычислительной точки зрения иногда имеет смысл переходить от обыкновенного дифференциального уравнения к соответствующему рекуррентному соотношению. Быстрое вычисление чисел a_n позволяет эффективно вычислять коэффициенты в разложении искомой функции в ряд Тейлора, а следовательно, и значение этой функции в произвольной точке x. Напротив, методы, разработанные для решения обыкновенных дифференциальных уравнений, часто оказываются полезными и для решения аналогичных рекуррентных соотношений.

10.2. Оказывается, что переход от числовой последовательности a_n к некоторой функции $y(x)$ полезен не только при анализе линейных рекуррентных соотношений. Как мы увидим позже, работа с такими функциями является одной из центральных и наиболее плодотворных идей современной перечислительной комбинаторики. При этом сами такие функции имеют специальное название — они называются производящими функциями для числовой последовательности a_n.

10.2.1. Начнем мы с определения экспоненциальной производящей функции.

Определение 10.1. Пусть $a_0, a_1, \ldots, a_n, \ldots$ — произвольная числовая последовательность. *Экспоненциальной производящей функцией* для этой последовательности называется выражение вида

$$F(z) = a_0 + a_1\frac{z^1}{1!} + a_2\frac{z^2}{2!} + \ldots + a_n\frac{z^n}{n!} + \ldots \tag{65}$$

Как мы отмечали выше, с точки зрения математического анализа экспоненциальную производящую функцию можно рассматривать как ряд Тейлора с коэффициентами $a_n = F^{(n)}(0)$. Однако такая точка зрения довольно сильно ограничивает возможность использования производящих функций. В частности, довольно часто входящие в (65) коэффициенты a_n не отвечают никакой аналитической в окрестности $z = 0$ функции $F(z)$. При этом с выражением (65) все равно можно вполне успешно работать и получать содержательные с комбинаторной точки зрения результаты. При этом, однако, возникает вполне естественный вопрос — как дать математически строгое определение выражению (65) и как ввести формальные правила работы с ним. Ответ на этот вопрос можно получить, введя понятие *формальных степенных рядов.* О том, что это за объекты и как с ними работать, мы поговорим в конце этой главы. Сейчас же мы просто ограничимся данным выше определением и будем работать с $F(z)$ как с обычным функциональным рядом.

10.2.2. Оказывается, при анализе линейных рекуррентных соотношений наряду с экспоненциальными производящими функциями полезно рассматривать и так называемые *обыкновенные производящие функции.* Как мы увидим несколько позже, использование этих функций позволяет довольно существенным образом упростить решение рекуррентных соотношений.

Определение 10.2. Пусть $a_0, a_1, \ldots, a_n, \ldots$ — произвольная числовая последовательность. *Обыкновенной производящей функцией* для этой последовательности называется выражение вида

$$F(z) = a_0 + a_1 z^1 + a_2 z^2 + \ldots + a_n z^n + \ldots \tag{66}$$

Как и в случае экспоненциальной производящей функции, обыкновенная производящая функция представляет собой некоторый формальный степенной ряд. Однако нам сейчас это не очень важно — мы пока что будем рассматривать ее как некоторый степенной ряд, сходящийся в некоторой окрестности точки $z = 0$. Наша ближайшая задача — показать, как можно использовать этот объект для решения линейных рекуррентных соотношений.

10.2.3. Начнем мы с конкретного примера, рассмотренного в начале данной главы и описывающего изменение популяции лягушек в озере. Напомним, что для этой задачи нами было получено следующее линейное неоднородное рекуррентное соотношение первого порядка:

$$a_{n+1} = 4\,a_n - 100, \qquad n = 0, 1, 2, \ldots, \qquad a_0 = 50. \tag{67}$$

Покажем, как можно построить решение этого рекуррентного соотношения с помощью обыкновенной производящей функции (66).

Домножим для этого левую и правую часть уравнения (67) на z^{n+1} и просуммируем полученный результат по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} a_{n+1}\, z^{n+1} = 4 \sum_{n=0}^{+\infty} a_n\, z^{n+1} - 100 \sum_{n=0}^{+\infty} z^{n+1}. \tag{68}$$

Заметим, что левая часть этого равенства представляет собой обыкновенную производящую функцию $f(z)$ рассматриваемой числовой последовательности за вычетом коэффициента a_0:

$$\sum_{n=0}^{+\infty} a_{n+1}\, z^{n+1} = a_1 z^1 + a_2 z^2 + \ldots + a_n z^n + \ldots = f(z) - a_0.$$

Первая сумма в правой части равенства (68), очевидно, равна

$$4\sum_{n=0}^{+\infty} a_n\, z^{n+1} = 4\, z \sum_{n=0}^{+\infty} a_n\, z^n = 4\, z\, f(z).$$

Наконец, последнюю сумму в (68) можно свернуть и записать в виде дроби

$$\sum_{n=0}^{+\infty} z^{n+1} = z \sum_{n=0}^{+\infty} z^n = \frac{z}{1-z}.$$

Поэтому окончательно равенство (68) переписывается в следующем виде:

$$f(z) - a_0 = 4\, z\, f(z) - 100\frac{z}{1-z}.$$

Как видно, в результате этих преобразований мы получили достаточно элементарное линейное алгебраическое уравнение на функцию $f(z)$, решение которого имеет следующий вид:

$$f(z) = \frac{a_0}{1-4z} - \frac{100\, z}{(1-z)(1-4z)}. \tag{69}$$

Последнее, что нам осталось — это получить отсюда явное аналитическое выражение для коэффициентов a_n этой производящей функции. А это сделать довольно легко — для этого нам остается разложить правую часть (69) в ряд по степеням z.

Проще всего разобраться с первым слагаемым в правой части (69). Из математического анализа известно, что функция

$$h(z) = \frac{1}{1-\alpha z}.$$

в области $|z| < 1/\alpha$ раскладывается в равномерно сходящийся ряд вида

$$1 + \alpha z + \alpha^2 z^2 + \ldots + \alpha^n z^n + \ldots$$

Воспользовавшись этим обстоятельством, мы можем для функции $g(z) = a_0/(1-4z)$ получить

$$g(z) = \frac{a_0}{1-4z} = a_0\left(1 + 4z + 4^2 z^2 + \ldots\right) = a_0 \sum_{n=0}^{\infty} 4^n z^n.$$

Для того, чтобы проделать аналогичные операции со вторым слагаемым, нам предварительно следует разложить эту дробь на простейшие:

$$\frac{z}{(1-z)(1-4z)} = \frac{A}{1-z} + \frac{B}{1-4z} = \frac{A - 4Az + B - Bz}{(1-z)(1-4z)} \implies \begin{cases} A + B = 0 \\ 4A + B = -1 \end{cases} \implies \begin{cases} A = -1/3 \\ B = 1/3 \end{cases}$$

Как следствие,

$$-\frac{100\, z}{(1-z)(1-4z)} = -\frac{100}{3}\left[\frac{1}{1-4z} - \frac{1}{1-z}\right] = -\frac{100}{3}\left[\sum_{n=0}^{+\infty} (4^n - 1)\, z^n\right].$$

Теперь мы можем записать явное аналитическое выражение для чисел a_n:

$$a_n = a_0\, 4^n - \frac{100}{3}\left[4^n - 1\right] = \frac{50}{3}\left[4^n + 2\right], \qquad n = 0, 1, 2, \ldots.$$

Итак, переход от числовой последовательности a_n, задаваемой линейным рекуррентным соотношением (67), к обыкновенной производящей функции $f(z)$ позволил нам получить для $f(z)$ довольно простое линейное алгебраическое уравнение, решение которого строится совершенно элементарно. Обратный переход от $f(z)$ к a_n позволил нам получить явные аналитические выражения для чисел a_n.

10.2.4. Оказывается, описанный выше подход проходит и для произвольного линейного рекуррентного соотношения m-го порядка с постоянными коэффициентами. Перепишем такое соотношение в следующем виде:

$$b_0 a_{n+m} + b_1 a_{n+m-1} + \ldots + b_m a_n = u_n, \qquad n = 0, 1, 2, \ldots, \tag{70}$$

$$b_0 \neq 0, \qquad\qquad a_0, a_1, \ldots, a_{m-1} \text{ – заданы.}$$

Введем теперь для числовой последовательности $a_0, a_1, a_2, \ldots$, отвечающей этому рекуррентному соотношению, обыкновенную производящую функцию

$$f(z) = a_0 + a_1\, z + a_2\, z^2 + \ldots = \sum_{n=0}^{+\infty} a_n z^n,$$

домножим (70) на z^{n+m} и просуммируем полученное уравнение по n от 0 до $+\infty$. Так как

$$b_0 \sum_{n=0}^{+\infty} a_{n+m} z^{n+m} = b_0\,[f(z) - a_0 - a_1 z - \ldots - a_{m-1} z^{m-1}],$$

$$b_1 \sum_{n=0}^{+\infty} a_{n+m-1} z^{n+m} = b_1 z\,[f(z) - a_0 - a_1 z - \ldots - a_{m-2} z^{m-2}],$$

$$\ldots$$

$$b_{m-2} \sum_{n=0}^{+\infty} a_{n+2} z^{n+m} = b_{m-2} z^{m-2}\,[f(z) - a_0 - a_1 z],$$

$$b_{m-1} \sum_{n=0}^{+\infty} a_{n+1} z^{n+m} = b_{m-1} z^{m-1}\,[f(z) - a_0],$$

$$b_m \sum_{n=0}^{+\infty} a_n z^{n+m} = b_m z^m f(z),$$

то получающееся в результате этих преобразований равенство мы можем переписать в виде

$$g(z) \cdot f(z) - h(z) = z^m \cdot u(z).$$

Здесь

$$g(z) := b_0 + b_1 z + b_2 z^2 + \ldots + b_{m-1} z^{m-1} + b_m z^m,$$

$$h(z) := b_0 a_0 + (b_0 a_1 + b_1 a_0) z + \ldots + (b_0 a_{m-1} + b_1 a_{m-2} + \ldots + b_{m-2} a_1 + b_{m-1} a_0) z^{m-1},$$

$$u(z) := \sum_{n=0}^{+\infty} u_n z^n.$$

Как видно, мы вновь получили на $f(z)$ линейное алгебраическое уравнение. Решение этого уравнения можно записать в виде

$$\Longrightarrow \qquad f(z) = [h(z) + z^m u(z)] \cdot g^{-1}(z).$$

Несложно убедиться (см. также параграф, посвященный формальным степенным рядам), что условие $b_0 \neq 0$ гарантирует нам существование функции $g^{-1}(z)$.

В случае *однородного* линейного рекуррентного соотношения с постоянными коэффициентами m-го порядка $u(z) = 0$, так что производящая функция $f(z)$ для рекуррентной последовательности $a_0, a_1, a_2, \ldots$ представляет собой правильную рациональную дробь вида

$$f(z) = \frac{h(z)}{g(z)} = \frac{\sum\limits_{n=0}^{m-1} c_n z^n}{\sum\limits_{n=0}^{m} b_n z^n}.$$

По сути, этим устанавливается взаимно-однозначное соответствие между всеми линейными однородными рекуррентными соотношениями с постоянными коэффициентами и рациональными производящими функциями.

10.2.5. Как правило, при вычислении явного вида коэффициентов a_n вместо формального деления степенных рядов целесообразно разложить дробь на простейшие — элементарные дроби вида

$$\frac{A}{1-\alpha\, z}, \qquad \frac{B}{(1-\alpha\, z)^2}, \qquad \ldots \qquad \frac{D}{(1-\alpha\, z)^k},$$

а затем воспользоваться формулами для разложения этих функций в ряд в малой окрестности точки $z = 0$. Для этого, в частности, можно воспользоваться известной из курса математического анализа формулой бинома Ньютона

$$(1+z)^q = \sum_{n=0}^{+\infty} \frac{(q)_n}{n!} z^n = 1 + \sum_{n=1}^{+\infty} \frac{q\,(q-1)\, \ldots\, (q-n+1)}{n\,(n-1)\, \ldots\, 1} z^n, \tag{71}$$

справедливой для любого $q \in \mathbb{C}$ и $|z| < 1$. На основании этой формулы,

$$\frac{1}{(1-\alpha\, z)^k} = 1 + \sum_{n=1}^{+\infty} \frac{(-k)(-k-1)(-k-2)\ldots(-k-n+1)}{n!} (-\alpha)^n z^n =$$

$$= 1 + \sum_{n=1}^{+\infty} \frac{k\,(k+1)(k+2)\ldots(k+n-1)}{n!} \alpha^n z^n = \sum_{n=0}^{+\infty} \left(\!\!\binom{k}{n}\!\!\right) \alpha^n z^n.$$

Формулой

$$\frac{1}{(1-\alpha\, z)^k} = \sum_{n=0}^{+\infty} \left(\!\!\binom{k}{n}\!\!\right) \alpha^n z^n \tag{72}$$

мы довольно часто будем пользоваться в дальнейшем.

10.2.6. Подводя итоги, мы можем сформулировать следующий общий алгоритм использования аппарата обыкновенных производящих функций для решения линейных рекуррентных соотношений m-го порядка с постоянными коэффициентами вида (70):

(1) ввести обыкновенную производящую функцию $f(z)$ для рекуррентной числовой последовательности $a_0, a_1, a_2, \ldots$;

(2) трансформировать (70) в уравнение для $f(z)$, домножив рекуррентное соотношение на z^{n+m}, просуммировав полученное выражение по n от 0 до $+\infty$ и выразив каждую из полученных таким образом сумм через $f(z)$;

(3) разрешить полученное уравнение относительно $f(z)$;

(4) определить числа a_n как коэффициенты при z^n в разложении $f(z)$ по степеням z.

10.3. Перейдем теперь к анализу линейных рекуррентных соотношений с переменными коэффициентами. Начнем, как всегда, с простого примера.

10.3.1. Рассмотрим числовую последовательность, записанную в следующем явном виде:

$$a_n = \binom{2n}{n}, \qquad n = 0, 1, 2, \ldots$$

Так как

$$a_{n+1} = \binom{2n+2}{n+1} = \frac{(2n+2)!}{(n+1)! \cdot (n+1)!} = \frac{2(2n+1)(2n)!}{(n+1) \cdot n! \cdot n!} = \frac{2(2n+1)}{n+1}\, a_n,$$

то числовая последовательность a_n удовлетворяет рекуррентному соотношению

$$(n+1)a_{n+1} = 4na_n + 2a_n, \qquad n = 0, 1, 2, \ldots, \qquad a_0 = 1.$$

10.3.2. Итак, мы решили в каком-то смысле обратную задачу — мы из явной формулы для коэффициентов a_n вывели рекуррентное соотношение, которому эти коэффициенты удовлетворяют. Постараемся теперь с использованием обыкновенных производящих функций решить прямую задачу, а именно, по заданному рекуррентному соотношению определить явный вид чисел a_n.

Для этого домножим наше рекуррентное соотношение на z^n и просуммируем его по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty}(n+1)a_{n+1}z^n = 4\sum_{n=0}^{+\infty} n\, a_n z^n + 2\sum_{n=0}^{+\infty} a_n z^n = 4z\sum_{n=1}^{+\infty} n\, a_n z^{n-1} + 2\sum_{n=0}^{+\infty} a_n z^n.$$

По определению производной обыкновенной производящей функции,

$$\sum_{n=1}^{+\infty} n\, a_n z^{n-1} = \sum_{n=0}^{+\infty}(n+1)a_{n+1}z^n = f'(z).$$

Поэтому предыдущее равенство можно записать в следующем компактном виде:

$$(1-4z)\, f'(z) = 2f(z), \qquad f(0) := a_0 = 1. \tag{73}$$

10.3.3. Итак, с помощью обыкновенных производящих функций мы свели задачу решения линейного рекуррентного соотношения первого порядка с переменными коэффициентами к задаче Коши для обыкновенного дифференциального уравнения первого порядка. В рассматриваемом примере задача (73) легко решается методом разделения переменных:

$$\frac{df}{f} = \frac{2\,dz}{1-4z} = -\frac{1}{2}\frac{d(1-4z)}{(1-4z)} \iff d\ln(f) = -\frac{1}{2}d\ln(1-4z) \implies f(z) = \frac{1}{\sqrt{1-4z}}.$$

Для того, чтобы разложить полученную функцию в ряд, воспользуемся формулой бинома Ньютона (71). Полагая в (71) в качестве $q = -1/2$ и заменяя в ней z на $4z$, получаем следующую цепочку равенств:

$$f(z) = (1-4z)^{-1/2} = 1 + \sum_{n=1}^{+\infty} \frac{(-1/2)(-1/2-1)\dots(-1/2-n+1)}{n!}(-4)^n z^n =$$

$$= 1 + \sum_{n=1}^{+\infty} \frac{1\cdot 3\cdot 5\cdot \ldots \cdot (2n-1)}{n!} 2^n z^n = \Big|_{n!2^n=(1\cdot 2)(2\cdot 2)(3\cdot 2)\dots(n\cdot 2)=2\cdot 4\cdot 6\cdot\ldots\cdot 2n}\Big| =$$

$$= 1 + \sum_{n=1}^{+\infty} \frac{1\cdot 2\cdot 3\cdot \ldots \cdot (2n-1)(2n)}{(n!)^2} z^n = \sum_{n=0}^{+\infty} \binom{2n}{n} z^n.$$

10.3.4. Заметим, что условие аналитичности функции $f(z)$ в окрестности нуля является в данном алгоритме существенным. Оно обычно выполняется лишь в том случае, если соответствующие коэффициенты a_n растут не слишком быстро. В противном случае описанный выше алгоритм решения рекуррентных соотношений с переменными коэффициентами может перестать работать.

В качестве характерного примера рассмотрим следующее несложное линейное рекуррентное соотношение с переменными коэффициентами:

$$a_{n+1} = (n+1)\, a_n, \qquad n = 0, 1, 2, \dots, \qquad a_0 = 1. \tag{74}$$

Попытаемся решить его с помощью обыкновенных производящих функций. Домножая рекуррентное соотношение на z^{n+1} и суммируя по n от 0 до $+\infty$, имеем

$$\sum_{n=0}^{+\infty} a_{n+1} z^{n+1} = \sum_{n=0}^{+\infty} n\, a_n z^{n+1} + \sum_{n=0}^{+\infty} a_n z^{n+1} = z^2 \sum_{n=0}^{+\infty} n\, a_n z^{n-1} + z \sum_{n=0}^{+\infty} a_n z^n \qquad \Longleftrightarrow$$

$$\Longleftrightarrow \qquad f(z) - 1 = z^2 f'(z) + z f(z), \qquad f(0) = 1.$$

Полученная задача Коши не имеет решения, аналитического в окрестности начала координат. Этот результат в данном случае легко объясним. Действительно, исходное рекуррентное соотношение (74) настолько простое, что мы легко можем получить явное выражение для коэффициентов a_n:

$$a_{n+1} = (n+1)\, a_n = (n+1)\, n\, a_{n-1} = \ldots = (n+1)!\, a_0 = (n+1)!$$

Отвечающий этой числовой последовательности степенной ряд

$$f(z) = a_0 + a_1 z + a_2 z^2 + \ldots + n! z^n + \ldots = 1 + z + 2z^2 + 6z^3 + \ldots + n!\, z^n + \ldots,$$

как уже отмечалось ранее, расходится при любых $z > 0$, если рассматривать его с точки зрения обычного математического анализа.

10.3.5. В принципе, полученное нами дифференциальное уравнение можно преобразовать так, чтобы его решение выражалось через гипергеометрические ряды, и построить искомое решение $a_n = n!$ (см., например, [**?**]). Однако сама процедура получения такого решения оказывается крайне сложной по сравнению со сложностью исходного рекуррентного соотношения. Оказывается, однако, что ситуацию можно подправить, решая это рекуррентное соотношение не с помощью обыкновенных, а с помощью экспоненциальных производящих функций.

Действительно, заметим, что числовой последовательности $a_n = n!$ отвечает экспоненциальная производящая функция

$$F(z) = 1 + 1! \cdot \frac{z^1}{1!} + 2! \cdot \frac{z^2}{2!} + \ldots + n! \cdot \frac{z^n}{n!} + \ldots = 1 + z + z^2 + \ldots + z^n + \ldots$$

Последний ряд сходится к функции $1/(1-z)$ при всех значениях $|z| < 1$. Следовательно, есть надежда, что заменяя в алгоритме обыкновенную производящую функцию на экспоненциальную, мы сможем добиться успеха.

Именно, введем для числовой последовательности a_n, описываемой рекуррентным соотношением (74), экспоненциальную производящую функцию

$$F(z) = a_0 + a_1 \frac{z}{1!} + a_2 \frac{z^2}{2!} + \ldots + a_n \frac{z^n}{n!} + \ldots$$

Домножим (74) на $z^{n+1}/(n+1)!$ и просуммируем его по n:

$$\sum_{n=0}^{+\infty} a_{n+1} \frac{z^{n+1}}{(n+1)!} = \sum_{n=0}^{+\infty} a_n \frac{z^{n+1}}{n!} \qquad \Longleftrightarrow \qquad F(z) - 1 = z\,F(z) \qquad \Longrightarrow \qquad F(z) = \frac{1}{1-z}.$$

Итак, в случае, когда коэффициенты a_n, отвечающие линейному рекуррентному соотношению с переменными коэффициентами, растут слишком быстро, разумно для решения этого соотношения использовать экспоненциальные производящие функции.

10.3.6. В связи с последним утверждением возникает естественный вопрос: а что, если числовая последовательность a_n будет расти столь быстро, что и отвечающий ей ряд $F(z)$, понимаемый в смысле математического анализа, будет расходиться всюду в окрестности нуля? В принципе такие примеры придумать можно. Однако на практике такие задачи, как правило, все же не встречаются.

10.3.7. Разумеется, не все линейные рекуррентные соотношения с переменными коэффициентами сводятся к линейным алгебраическим уравнениям на экспоненциальные производящие функции. Часто соответствующие этим соотношениям уравнения содержат производные экспоненциальных производящих функций.

Рассмотрим, к примеру, рекуррентное соотношение

$$B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_k, \qquad B_0 = 1 \tag{75}$$

для чисел Белла B_n, описывающих количество всевозможных разбиений n-элементного множества. Домножим это равенство на $z^n/n!$ и просуммируем по n от нуля до $+\infty$:

$$\sum_{n=0}^{+\infty} B_{n+1} \frac{z^n}{n!} = \sum_{n=0}^{+\infty} \sum_{k=0}^{n} \binom{n}{k} B_k \frac{z^n}{n!}.$$

В левой части этого равенства стоит производная $B'(z)$ рассматриваемой экспоненциальной производящей функции. Правая его часть представляет собой произведение пары экспоненциальных функций — $B(z)$ и функции

$$1 + 1 \cdot \frac{z}{1!} + 1 \cdot \frac{z^2}{2!} + \ldots + 1 \cdot \frac{z^n}{n!} + \ldots := e^z.$$

Следовательно, в терминах экспоненциальных производящих функций равенство (75) записывается так:

$$B'(z) = e^z B(z), \qquad B(0) := B_0 = 1.$$

Рассмотрим теперь последние равенства как задачу Коши для функции $B(z)$ комплексного аргумента z. Эта задача легко решается:

$$d\ln B(z) = de^z, \qquad B(0) = 1 \qquad \Longrightarrow \qquad B(z) = e^{e^z - 1}.$$

Заметим, что получившаяся в результате решения функция $B(z)$ с точки зрения математического анализа представляет собой хоть и быстрорастущую, но аналитическую функцию комплексного аргумента при любом $z \in \mathbb{C}$.

11 Числа Каталана. Нелинейные рекуррентные соотношения

11.1. До сих пор мы рассматривали лишь линейные рекуррентные соотношения. В практических задачах, однако, довольно часто встречаются и более сложные, нелинейные рекуррентные соотношения. К сожалению, для таких соотношений общих алгоритмов построения решений не существует. Однако многие задачи, приводящие к нелинейным рекуррентным соотношениям, по-прежнему допускают решения с помощью производящих функций.

В качестве самого простого, и в то же время, пожалуй, самого известного в перечислительной комбинаторике примера рассмотрим задачу подсчета так называемых правильных скобочных последовательностей.

11.1.1. Как известно, порядок вычислений в любом арифметическом выражении можно однозначно задать расстановкой скобок. Давайте возьмем какое-то достаточно произвольное арифметическое выражение, например,

$$(3 - 1) \cdot (4 + (15 - 9) \cdot (2 + 6)),$$

и сотрем в нем все числа и знаки арифметических операций. В результате такого действия мы получим последовательность открывающихся и закрывающихся скобок

$$()(()()),$$

представляющую собой так называемую правильную скобочную последовательность.

Определение 11.1. Правильная скобочная последовательность — это строка, состоящая из n открывающихся и n закрывающихся скобок, обладающая следующим свойством: при проходе вдоль этой структуры слева направо количество открывающихся скобок всегда больше или равно количеству закрывающихся скобок.

Перечислим все правильные скобочные последовательности с числом *пар* скобок $n = 1, 2, 3$:

$$\begin{array}{lll}
n = 1: & () & \text{– 1 последовательность;} \\
n = 2: & ()(),\ (()) & \text{– 2 последовательности;} \\
n = 3: & ()()(),\ ()(()),\ (())(),\ (()()),\ ((())) & \text{– 5 последовательностей.}
\end{array}$$

Числа C_n, описывающие количество таких последовательностей, называются *числами Каталана.* Как мы увидим несколько позднее, удобно по определению положить $C_0 = 1$.

Последовательность чисел Каталана

$$1, 1, 2, 5, 14, 42, 132, 429, 1430, \ldots$$

(последовательность $A000108$ в `OEIS` (`oeis.org`)) встречается в огромном количестве различных комбинаторных задач. В книге [?] приведено порядка 100 задач, в которых эти числа появляются. Приведем лишь несколько наиболее важных из них.

11.1.2. Рассмотрим вначале очень простую интерпретацию правильной скобочной последовательности — задачу об очереди в кассу (см. [?]). Предположим, что у кассы кинотеатра стоит очередь, состоящая из $2n$ человек. У половины из них имеется по 100 рублей, у второй половины — по 50 рублей. Билет в кино стоит 50 рублей. В начале продажи билетов касса кинотеатра пуста. Спрашивается, сколькими способами можно расставить людей в очереди правильно, т.е. так, чтобы никому не пришлось ждать у кассы сдачу.

Очевидно, что между этой задачей и задачей о количестве правильных скобочных последовательностей имеется биекция: любому человеку, имеющему 50 рублей, отвечает открывающаяся скобка в правильной скобочной последовательности, а человеку со 100 рублями — закрывающаяся скобка.

11.1.3. Задачу об очереди в кассу часто формулируют более формально, а именно, как задачу о подсчете количества различных *слов Дика* длины $2n$. Словом Дика называют строку длины $2n$ над алфавитом, состоящим из двух символов (например, U и D), в которой количество символов U и D совпадает, и в которой никакой начальный сегмент строки не содержит символов D больше, чем символов U. В этой формулировке биекция с задачей о подсчете правильных скобочных последовательностей имеет, очевидно, вид

$$(\;\longleftrightarrow\; U, \qquad\qquad) \;\longleftrightarrow\; D.$$

11.1.4. К задаче о подсчете правильных скобочных последовательностей сводится, очевидно, и задача о перечислении путей на плоскости, выходящих из начала координат, состоящих из отрезков $(1, 1)$ и $(1, -1)$, заканчивающихся в точке $(2n, 0)$ на оси абсцисс и нигде не пересекающих эту ось. Такие пути на плоскости называются *путями Дика* (рис 9). Биекция с правильными скобочными последовательностями здесь такова:

$$(\;\longleftrightarrow\; \text{вектор } (1, 1), \qquad\qquad) \;\longleftrightarrow\; \text{вектор } (1, -1).$$

11.1.5. Одной из основных дискретных структур, использующихся в теории алгоритмов, является плоское корневое бинарное дерево, иногда называющаяся просто бинарным или двоичным деревом. Неформально корневое дерево — это дерево, в котором любая вершина имеет ровно двух (возможно, пустых) потомков — левого и правого (рис. 10). В упражнении **??** предлагается доказать, что количество таких деревьев на n вершинах описывается числами Каталана C_n.

11.1.6. Рассмотрим выражение вида

$$a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_n \cdot a_{n+1}, \tag{76}$$

где элементы a_i принадлежат множеству S с введенной на нем неассоциативной бинарной операцией $' \cdot '$. Очевидно, что это выражение не имеет смысла до тех пор, пока мы не расставим

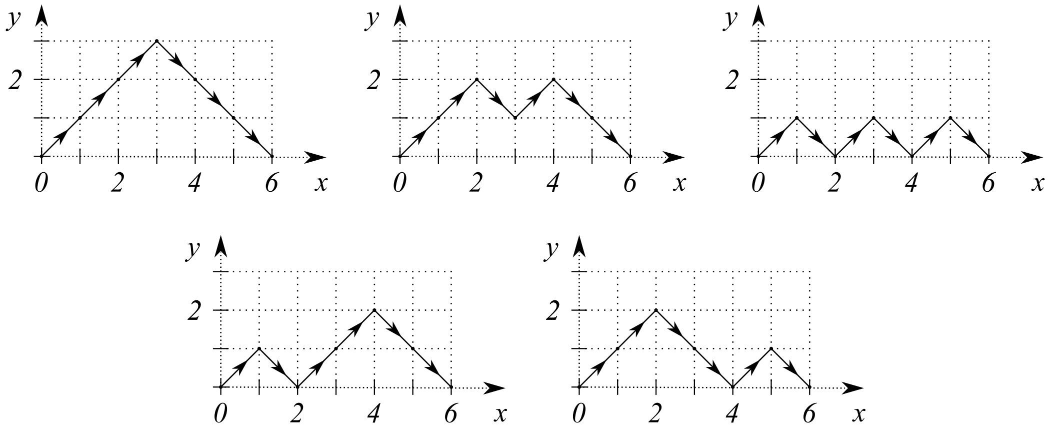


Рис. 9: Все пути Дика из трёх пар шагов.

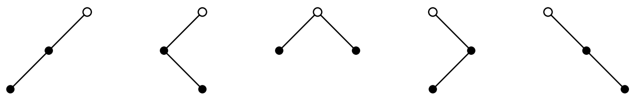


Рис. 10: Все бинарные деревья на трёх вершинах.

скобки так, чтобы указать на последовательность проводимых операций. Несложно убедиться, что количество расстановки скобок в описанном выше выражении есть число Каталана C_n (смотри упражнение **??**). Так, для случая $n = 3$ мы имеем следующие 5 различных вариантов расстановки скобок:

$$((a_1 \cdot a_2) \cdot a_3) \cdot a_4, \quad (a_1 \cdot (a_2 \cdot a_3)) \cdot a_4, \quad (a_1 \cdot a_2) \cdot (a_3 \cdot a_4), \quad a_1 \cdot ((a_2 \cdot a_3) \cdot a_4), \quad a_1 \cdot (a_2 \cdot (a_3 \cdot a_4)).$$

11.1.7. Еще одна важная комбинаторная интерпретация чисел Каталана появилась впервые в работах Леонарда Эйлера (и, кстати сказать, задолго до работ самого Эжена Каталана). Эйлер рассмотрел количество разбиений выпуклого $(n+2)$-угольника с занумерованными (т.е. различимыми) вершинами на треугольники непересекающимися между собой диагоналями этого $(n+2)$-угольника (рис. 11).

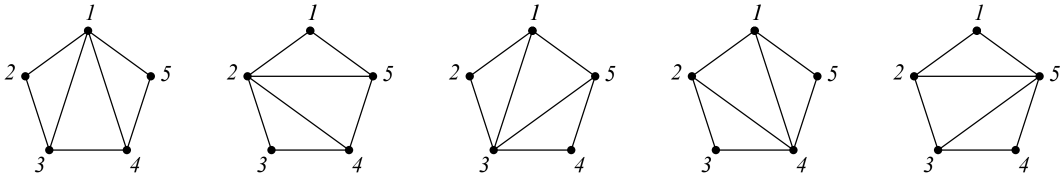


Рис. 11: Все триангуляции пятиугольника.

В упражнении **??** предлагается доказать, что это количество описывается числами Каталана C_n, установив биекцию между всеми триангуляциями выпуклого $(n+2)$-угольника и плоскими корневыми бинарными деревьями, построенными на n вершинах.

11.1.8. Рассмотрим теперь все плоские корневые деревья (рис 12), построенные на $(n+1)$-й вершине, $n = 0, 1, 2, \ldots$. Количество таких деревьев также равно числу Каталана C_n. Для того, чтобы это понять, осуществим в любом таком дереве поиск в глубину. При движении вдоль некоторого ребра вниз, т.е. от корня, сопоставим этому ребру левую открывающуюся скобку. При движении в обратном направлении сопоставим этому же ребру закрывающуюся скобку. Тем самым мы устанавливаем биекцию между всеми такими деревьями и всеми правильными скобочными последовательностями.

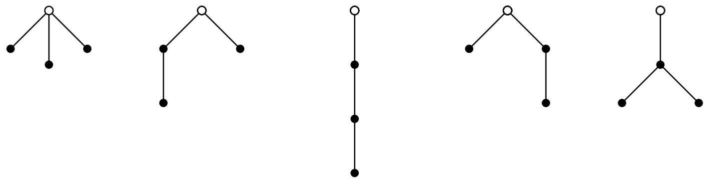


Рис. 12: Все плоские деревья на четырёх вершинах.

11.1.9. Наконец, перейдем к еще одному полезному понятию — стек-сортируемой последовательности. Согласно Кнуту [**?**], стек-сортируемая последовательность — это линейно упорядоченный набор из n чисел, который можно получить из линейно упорядоченной последовательности $(1, 2, \ldots, n)$ с помощью единственного стека.

В качестве примера рассмотрим последовательность чисел $(1, 2, 3, 4)$. Разместим ее справа от стека и будем сдвигать ее влево, помещая числа в стек и вынимая их из этого стека с помощью следующей последовательности действий:
– поместить число 1 в стек;
– поместить число 2 в стек;
– извлечь число 2 из стека;
– поместить число 3 в стек;
– поместить число 4 в стек;
– извлечь число 4 из стека;
– извлечь число 3 из стека;
– извлечь число 1 из стека.

В результате этих действий мы получим линейно упорядоченный набор чисел $(2, 4, 3, 1)$.

Возникает вопрос: сколько различных наборов чисел можно получить из последовательности $(1, 2, \ldots, n)$ с помощью подобных операций? Ответ, конечно же, вполне ожидаем — это количество равно числу Каталана C_n.

Для того, чтобы убедиться в этом, достаточно установить следующую биекцию между парой используемых в алгоритме операций и парой скобок:

$$(\longleftrightarrow \text{ поместить число в стек}, \qquad\qquad) \longleftrightarrow \text{ извлечь число из стека.}$$

В случае $n = 3$ имеется пять стек-сортируемых последовательностей:

$$(1, 2, 3), \qquad (3, 2, 1), \qquad (2, 1, 3), \qquad (1, 3, 2), \qquad (3, 1, 2).$$

Единственной перестановкой, которую нельзя перевести в последовательность $(1, 2, 3)$ с помощью стека, является перестановка вида $(2, 3, 1)$.

11.2. Теперь, после стольких примеров, пора научиться вычислять числа Каталана. Начнем с вывода рекуррентного соотношения для этих чисел. В качестве основного объекта мы выберем множество всех правильных скобочных последовательностей. Рекуррентное соотношение для чисел C_n получим, воспользовавшись хорошо нам уже знакомым принципом разбиения множества на блоки и подсчетом количества элементов в каждом из блоков в отдельности.

11.2.1. Рассмотрим произвольную правильную скобочную последовательность. Для любой открывающейся скобки в такой последовательности можно ввести понятие парной ей закрывающейся скобки. Для этого будем идти от открывающейся скобки вправо и для каждой закрывающейся скобки будем проверять условие “количество закрывающихся скобок равно количеству открывающихся скобок". Первая закрывающаяся скобка, для которой это правило выполнится, и будет парной для нашей открывающейся скобки.

11.2.2. Разобъем теперь множество всех правильных скобочных последовательностей на блоки. Для этого возьмем крайнюю левую открывающуюся скобку в правильной скобочной последовательности и поместим в k-й блок все правильные скобочные последовательности, для которых парная ей закрывающаяся скобка стоит на $2k$-м месте.

Так, в случае $n = 3$ имеем разбиение множества, состоящего из пяти различных правильных скобочных последовательностей, на три блока:

$$\underbrace{()()(), \qquad ()(())}_{k=1}; \qquad \underbrace{(())()}_{k=2}; \qquad \underbrace{(()()), \qquad ((()))}_{k=3}.$$

11.2.3. Рассмотрим крайнюю левую открывающуюся и парную ей закрывающуюся скобки — выделенную пару скобок в нашей последовательности. Основное наблюдение здесь состоит в следующем: как внутри, так и снаружи указанной пары скобок стоят правильные скобочные подпоследовательности.

Действительно, рассмотрим вначале подпоследовательность, ограниченную выделенной парой скобок. Количество открывающихся скобок в этой подпоследовательности, очевидно, совпадает с количеством закрывающихся скобок. Более того, такая подпоследовательность является правильной — если бы это было не так, то условие правильности было бы нарушено и для всей последовательности скобок в целом.

Проверка того, что и правая подпоследовательность является правильной, проводится с помощью аналогичных рассуждений.

11.2.4. Подсчитаем теперь количество элементов в k-м блоке. Для этого заметим, что количество способов построить правильную скобочную последовательность внутри выделенной пары скобок равно, очевидно, C_{k-1}. Вне зависимости от выбора этой последовательности мы C_{n-k} способами можем построить правильную скобочную подпоследовательность справа от выделенной пары скобок. Следовательно, по правилу произведения в каждом блоке существует $C_{k-1} \cdot C_{n-k}$ способов построить правильную скобочную последовательность длины $2n$. Общее же число способов получить такую последовательность согласно правилу суммы равно

$$C_n = \sum_{k=1}^{n} C_{k-1}\, C_{n-k}, \qquad n = 1, 2, \ldots; \qquad C_0 = 1. \tag{77}$$

Заметим, что рекуррентное соотношение для чисел Каталана является нелинейным. Кроме того, n-й член последовательности C_n зависит от всех n предыдущих членов этой последовательности.

11.3. Постараемся решить рекуррентное соотношение (77).

11.3.1. Для этого введем обыкновенную производящую функцию для последовательности чисел Каталана:

$$f(z) = C_0 + C_1 z + C_2 z^2 + \ldots + C_n z^n + \ldots$$

Нам будет удобно переписать рекуррентное соотношение (77) в следующем виде:

$$C_{n+1} = C_0 \cdot C_n + C_1 \cdot C_{n-1} + \ldots + C_n \cdot C_0 = \sum_{k=0}^{n} C_k C_{n-k}.$$

Домножим это рекуррентное соотношение на z^{n+1} и просуммируем полученное равенство по n от 0 до $+\infty$:

$$\sum_{n=0}^{+\infty} C_{n+1} z^{n+1} = \sum_{n=0}^{+\infty} z^{n+1} \Bigl(\sum_{k=0}^{n} C_k\, C_{n-k} \Bigr) \quad \Longleftrightarrow \quad f(z) - 1 = z \Biggl(\sum_{n=0}^{+\infty} z^n \Bigl(\sum_{k=0}^{n} C_k\, C_{n-k} \Bigr) \Biggr) = z \cdot f^2(z).$$

Следовательно, функция $f(z)$ определяется из следующего равенства:

$$f(z) = 1 + z\, f^2(z). \tag{78}$$

11.3.2. Как видно, нелинейное рекуррентное соотношение (77) для чисел C_n приводит к нелинейному же уравнению (78) на производящую функцию $f(z)$. Как правило, решение такого рода уравнений строится с помощью формулы обращения Лагранжа, о которой подробно будет рассказано в следующей главе. Мы же сейчас найдем эту функцию, используя связь обыкновенных производящих функций с функциональными рядами из математического анализа.

Именно, предположим, что имеется функция $f(z)$ комплексного аргумента z, аналитическая в окрестности точки $z = 0$ и удовлетворяющая уравнению (78). Тогда эта функция единственным образом раскладывается в окрестности точки $z = 0$ в степенной ряд

$$f(z) = C_0 + C_1 z + C_2 z^2 + \ldots$$

Итак, все, что нам остается сделать — это найти функцию $f(z)$, аналитическую в окрестности точки $z = 0$ и удовлетворяющую уравнению (78), а затем разложить ее по степеням z^n в окрестности этой точки. Для этого разрешим уравнение (78) относительно функции $f(z)$:

$$f(z) = \frac{1 \pm \sqrt{1 - 4z}}{2z}.$$

Решение

$$f(z) = \frac{1 + \sqrt{1 - 4z}}{2z}$$

расходится в окрестности точки $z = 0$, поэтому его следует исключить из рассмотрения. Второе решение

$$f(z) = \frac{1 - \sqrt{1 - 4z}}{2z} = \frac{1}{2z} - \frac{1}{2z}\,(1 - 4z)^{1/2} \tag{79}$$

разложим в ряд в окрестности точки $z = 0$, используя формулу бинома Ньютона:

$$f(z) = \frac{1}{2z} - \frac{1}{2z} \Bigl(1 + \sum_{n=1}^{+\infty} \frac{(1/2)(1/2 - 1)(1/2 - 2) \ldots (1/2 - n + 1)}{n!} (-4z)^n \Bigr) =$$

$$= -\frac{1}{2z} \sum_{n=1}^{+\infty} (-1)^{n-1} \frac{1 \cdot 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n - 3)}{2^n n!} (-4)^n z^n = \sum_{n=1}^{+\infty} \frac{2^{n-1} \cdot 1 \cdot 3 \cdot 5 \cdot \ldots \cdot (2n - 3)}{n!} z^{n-1} =$$

$$= \sum_{n=1}^{+\infty} \frac{(2n-2)!}{n\,((n-1)!)^2}\, z^{n-1} = \bigl|_{n'=n-1;\, n=n'+1;\, n'=0,1,2,\ldots}\bigr| = \sum_{n'=0}^{+\infty} \frac{(2n')!}{(n'+1)\,(n'!)^2}\, z^{n'}.$$

Следовательно, числа Каталана могут быть выражены через биномиальные коэффициенты по формуле

$$C_n = \frac{1}{n+1} \binom{2n}{n}, \qquad n = 0, 1, 2, \ldots \tag{80}$$

11.4. Явную аналитическую формулу (80) для чисел Каталана можно получить, используя и прямые комбинаторные рассуждения (A.André, 1878).

11.4.1. Перепишем для этого (80) в виде

$$C_n = \binom{2n}{n} - \binom{2n}{n+1}, \qquad n = 0, 1, 2, \ldots, \tag{81}$$

и заметим, что первое слагаемое в этой формуле описывает общее количество всех возможных расстановок n открывающихся и n закрывающихся скобок в строке из $2n$ символов, равное

$$\binom{2n}{n} = \frac{(2n)!}{n! \cdot n!} = P(2n; n, n),$$

где $P(2n; n, n)$ — количество перестановок с повторениями в строке длины $2n$, состоящей из n неразличимых символов первого сорта и n неразличимых символов второго сорта (см. главу 1). Осталось из этого количества вычесть "неправильные" скобочные последовательности, т.е. такие, в которых нарушается условие "при проходе строки слева направо количество открывающихся скобок больше или равно количеству закрывающихся скобок".

11.4.2. Для этого рассмотрим любую такую "неправильную" скобочную последовательность. Найдем в этой строке первую закрывающуюся скобку, в которой условие "правильности" нарушается. Эта скобка будет стоять на позиции $2k+1$ для некоторого $k = 0, 1, 2, \ldots$. В стоящей слева от нее подстроке длины $2k$ количество открывающихся скобок, равное k, в точности равно количеству закрывающихся скобок. Заменим теперь в получившейся подстроке длины $2k+1$ все открывающиеся скобки закрывающимися и наоборот. В результате получим некоторую строку длины $2n$, содержащую $n+1$ открывающуюся скобку и $n-1$ закрывающуюся скобку.

11.4.3. Теперь возьмем *любую* строку длины $2n$, содержащую $n+1$ открывающуюся скобку и $n-1$ закрывающуюся. Оказывается, ее всегда можно превратить в *неправильную* скобочную последовательность обратным преобразованием. Действительно, будем идти вдоль такой строки слева направо и проверять условие "количество закрывающихся скобок больше или равно количеству открывающихся скобок". Так как общее количество открывающихся скобок в такой строке строго больше общего количества закрывающихся скобок, то обязательно найдется открывающаяся скобка, для которой это условие нарушится. Она будет стоять на позиции $2k+$ $+\,1$, а слева от нее будет стоять подстрока длины $2k$, в которой количество закрывающихся скобок (равное k) совпадает с количеством открывающихся скобок. Меняя в такой подстроке открывающиеся скобки на закрывающиеся и наоборот, мы получаем строку, состоящую из n открывающихся и n закрывающихся скобок, в которой условие правильности скобочной последовательности нарушается на позиции $2k+1$.

11.4.4. Итак, мы установили взаимно однозначное соответствие между множеством всех неправильных скобочных последовательностей и множеством всех строк длины $2n$, содержащих $n+1$ открывающуюся скобку и $n-1$ закрывающуюся скобку. Мощность последнего множества равна, очевидно,

$$P(2n; n+1, n-1) = \frac{(2n)!}{(n+1)! \cdot (n-1)!} = \binom{2n}{n+1}.$$

Согласно принципу биекции, этому же числу равна и мощность множества всех неправильных скобочных последовательностей длины $2n$. Как следствие, количество всех правильных скобочных последовательностей равно

$$P(2n; n, n) - P(2n; n+1, n-1) = \binom{2n}{n} - \binom{2n}{n+1} = \frac{1}{n+1}\binom{2n}{n} = C_n,$$

что и требовалось доказать.