12 Комбинаторика. Простейшие формулы. Примеры. Некоторые общие методы.

12.1 Простейшие базовые формулы

- **Формула включений и исключений**. Пусть $|A|$ обозначает число элементов в конечном множестве A. Тогда для (конечного) семейства конечных множеств A_i, $i = 1, 2, \ldots n$ выполнено:

$$|\bigcup_{i=1}^{n} A_i| = \sum_{i}^{n} |A_i| - \sum_{i_1 < i_2} |A_{i_1} \cap A_{i_2}| + \sum_{i_1 < i_2 < i_3} |A_{i_1} \cap A_{i_2} \cap A_{i_3}| - \ldots + (-1)^{n-1} |A_{i_1} \cap A_1 \cap A_2 \cap \ldots \cap A_n|$$

  Доказательство математической индукцией по подсчету вхождений каждого элемента из $\bigcup_{i=1}^{n} A_i$ в множества из правой части равенства.

- **Таблицы свойств**: пусть имеется k списков, в каждом таком списке перечислено N_k характеристик, надо сосчитать сколько возможных вариантов набора характеристик (по одной из каждого списка). Например, если списков два, то варианты удобно представлять в виде позиций в прямоугольной таблице – это дает ответ $N_1 \cdot N_2$. Многомерный аналог этого подхода дает ответ $N_1 \cdot N_2 \cdot \ldots N_k$.

- **Размещения**: сколько способов построить n солдат в шеренгу длины k. Для правого фланга есть n вариантов, следующее по порядку место заполняется $n-1$ способами итд. Всего вариантов получается $n \cdot (n-1) \cdot (n-2) \cdot \ldots (n-k+1)$, что традиционно обозначается как A_n^k.

- **Перестановки** Размещения с $n = k$ традиционно называются перестановками, обозначаются $\prod_n = n!$, где по определению $0! = 1$.

- **Сочетания**: сколько способов собрать из n солдат взвод в k человек. Ясно, что такой взвод можно построить в шеренгу $\prod_k$ способами, откуда вытекает число способов формирования взвода $\frac{A_n^k}{\prod_k} = \frac{n!}{k!(n-k)!}$, оно называется числом сочетаний или биномиальным коэффициентом, обозначалось ранее как C_n^k, но в современной литературе [8] как $\binom{n}{k}$.

12.2 Некоторые примеры

Много примеров с решениями задач элементарной комбинаторики можно отыскать в несложной книге Н.Виленкин "Комбинаторика". Мы рассмотрим здесь лишь самые основные.

12.2.1 Перестановки разнотипных объектов или перестановки с повторениями

Пусть имеются предметы k различных типов. Сколько различных перестановок можно сделать в последовательности из n_1 предметов первого типа, n_2 предметов второго типа, …, n_k предметов k -го типа? Число элементов в последовательности равно $n = \sum_{i=1}^{k} n_k$. Поскольку различимы элементы только по их типам, то общее число перестановок будет меньше $\prod_n$: некоторые перестановки надо отождествить. В самом деле, однотипные предметы можно переставлять между собой и это даст неотличимую от исходной последовательность. Для разных типов это можно делать независимо и таких «внутренних» перестановок будет соответственно $n_1!$ в первом типе, $n_2!$ во втором итд. Таким образом общее число действительно различных перестановок получится таким:

$$\frac{(n_1 + n_2 + \ldots + n_k)!}{n_1! \cdot n_2! \cdot \ldots \cdot n_k!}$$

12.3 Разложения шаров по ящикам. Различимые и неразличимые объекты

Пусть имеются n шаров и $k \leqslant n$ ящиков, мы собираемся сосчитать количество конфигураций, когда все шары разложены по ящикам. Совершенно очевидно, что прежде всего необходимо договориться, какие кофигурации следует считать разными, это приводит к следующщим задачам с разными условиями:

1. **Все яшики различимы между собой и все шары различимы между собой**. Условие означает, что можно пронумеровать предметы и тогда для каждого шара возникает ровно k возможностей, по формуле таблиц общее число раскладок будет k^n

[8] кроме того для любого действительного x полагают $\binom{x}{k} = \frac{x(x-1)(x-2)\ldots(x-k+1)}{k!}$

2. **Все ящики различимы между собой, а все шары неразличимы между собой**. Можно закодировать каждую такую раскладку схематической картинкой: палочки, обозначающие стенки плотно стоящих ящиков слева направо и шары между ними. Достаточно укзать только $k-1$ стенку – самая левая и самая правая стенки ничего не прибавляют к знанию о раскладке. А теперь, если шары тоже нарисовать палочками, но поменьше размером, то получится такая, например, картинка, кодирующая раскладку 3,7,2,0,4 шестнадцати одинаковых шаров по пяти различимым (по их порядку) ящикам:

   

   В общем случае всего палочек $n+k-1$ и из них $k-1$ длинных. Всего таких картинок (и раскладок!) получается $\binom{n+k-1}{k-1}$.

3. **Все ящики неразличимы между собой и все шары неразличимы между собой**. Здесь все кодируется числом способов разбиения числа n на не более, чем $k \leqslant n$ ненулевых слагаемых, порядок которых неважен. Поэтому всякому такому разбиению $n = n_1 + n_2 + \ldots + n_j$ можно сопоставить картинку: $j \leqslant k$ выровненных по левой границе горизонтальных полосок (каждая из некоторого числа клеточек, всего клеточек n), нарисованных друг под другом в порядке (нестрогого) убывания длин [9]. Такая картинка см. Рис.10 из клетчатых полосок называется диаграммой Юнга. Нас интересует подсчет всех таких табличек с общим числом клеток n и числом строк не более k.

   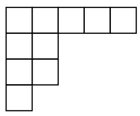

   Рис. 10: Пример диаграммы Юнга

   Первым делом заметим, что каждую такую табличку можно *транспонировать* – строки по порядку выписать столбцами – тогда опять получится диаграмма Юнга из n клеток уже с любым (естественно, не превосходящим n) количеством строк, но у которой длина любой строки теперь не превосходит k. Мы свели исходную задачу к подсчету $p(n,k)$ – количества разбиений числа n на слагаемые, *каждое из которых* не превосходит k. Продолжение вычислений с диаграммами Юнга см. в разделе методов 12.4.3.

4. **Все ящики неразличимы между собой, а все шары различимы между собой**. То есть речь идет о подсчете количества неупорядоченных разбиений множества $\{1,2,\ldots n\}$ на не более чем k подмножеств. Количество разбиений $\{1,2,\ldots n\}$ *в точности на i непустых множеств* может быть явно указано, оно называется *числом Стирлинга второго рода* и обычно обозначается как $\left\{ n \atop i \right\}$. Поэтому в нашей задаче про разложение различимых шаров в неразличимые ящики окончательный ответ дается суммой чисел Стирлинга $\left\{ n \atop i \right\}$ по i от единицы до k. Явное вычисление чисел Стирлинга см. далее разделе методов 12.4.4.

12.3.1 Связь со статистической физикой

Вопрос о том, какие объекты в природных процессах ведут себя как различимые, а какие нет, изучается в физике (и более широко – науками, которые используют физические модели для своих нужд).

В классической статистической физике, созданной Максвеллом и Больцманом, объекты считаются различимыми друг от друга. Так себя ведут, например, молекулы газа. Выше объяснено число способов для разложения n различных шаров по k различимым ящиками. Это приводит к формулам Максвелла-Больцмана для распределения частиц с заданными энергиями.

В атомной и субатомной физике микромира все оказалось вовсе не так просто: например, фотоны и атомные ядра подчинены иной статистике, разработанной Эйнштейном и Бозе, в ее основе лежит выражение для числа разложения n неразличимых частиц-шаров по k различимым областям-ящиками, говоря короче, в статистике Бозе-Эйнштейна частицы считаются неразличимыми друг от друга. Другие частицы микромира — например, электроны— оказались и неразличимы и вдобавок подчиняются ограничительному правилу, что в одном ящике

[9]французские математики предпочитают рисовать их по возрастанию

(тут, конечно же, надо еще объяснять что именно в теории элементарных частиц считается «ящиком», но эта тема в нашем вовсе не физическом курсе не будет рассматриваться) не может находиться более одной частицы – эти правила определяют статистику Ферми-Дирака.

12.4 Методы вычислений

12.4.1 Рекуррентные соотношения. Треугольник Паскаля

Биномиальные коэффициенты имеют много тождеств и рекуррентных соотношений, например:

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

Это соотношение порождает возможность вычислять биномиальные коэффициенты последовательно выписывая значения, выдаваемые рекуррентным соотношением, в треугольную таблицу (треугольник Паскаля), строки которой принято нумеровать от нуля.

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

Из треугольника Паскаля возможно вычислить и множество других очень полезных комбинаторных формул. Индукцией по номеру строки несложно проверить, что в треугольнике Паскаля k стоят коэффициенты многочлена $(a+b)^n$ последовательно возрастающие как раз в соответствии с рекуррентным соотношением. В частности, при $a = b = 1$ возникает тождество:

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{n} = 2^n$$

В докомпьютерную эпоху последовательные вычисления при больших n, r биномиальных коэффициентов $\binom{n}{r}$ даже и с помощью треугольника Паскаля представляли трудность, поэтому возникли асимптотические формулы для факториалов, дающие приблизительные вычисления по формуле $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. Часто встречается асимптотическая формула Стирлинга, объяснение которой приведено в разделе про биномиальное распределение.

12.4.2 Производящие функции

В математике часто встречается следующий прием для вычислений. Пусть $a_0, a_1, a_2, \ldots$ — произвольная числовая последовательность $\{a_n\}$. Формальный степенной ряд $\sum_{n=0}^{\infty} a_n q^n$ от переменной q называется производящей функцией для этой последовательности. Например, для последовательности из одних единиц ее производящая функция геометрическая прогрессия — формальный ряд для $\frac{1}{1-q}$, а для (конечной) последовательности биномиальных коэффициентов $\binom{n}{0}, \binom{n}{1} \ldots \binom{n}{n}$ производящая функция есть $(1+q)^n$ итп. Явное знание вида производящей функции позволяет дифференциированием находить ее ряд Маклорена, коэффициенты которого дадут исходную последовательность. Мы применим производящие функции к поставленным выше задачам о разложениях шаров в ящики, больше примеров найдется, например, в книге С.Ландо "Лекции по комбинаторике".

12.4.3 Диаграммы Юнга. Подсчет числа $p(n, k)$ для разбиений

Мы сосредоточимся на вычислении производящей функции $P_k(q)$ для последовательности $\{p(n, k)\}$ с фиксированным k. Ясно что $P_1(q) = \frac{1}{1-q}$ потому что каждое число единственным образом разбивается в сумму единиц. Теперь заметим, что количество способов разбить число n в сумму слагаемых, *каждое из которых* равно двум — это либо 1, если n четно, либо 0, если n нечетно. Чередование единиц и нулей отвечает производящей функции $\frac{1}{1-q^2}$, а $P_2(q) = \frac{P_1(q)}{1-q^2}$. Действительно, раскроем скобки, но не будем пока приводить подобные члены в произведении рядов

$$(1 + q + q^2 + q^3 + \ldots)(1 + q^2 + q^4 + q^6 + \ldots)$$

Каждое слагаемое после раскрытия скобок имеет вид $q^r q^{2s}$ и каждому такому слагаемому можно сопоставить разбиение числа $r + 2s$ в сумму r единиц и s двоек. А после приведения подобных членов коэффициент при

q^n окажется как раз $p(n,2)$. Рассуждая аналогичным образом получаем, что производящая функция $P_k(q) = \sum_n p(n,k)q^n$ равна

$$\frac{1}{(1-q)(1-q^2)\dots(1-q^k)} = \prod_{m=1}^{k}(1-q^m)^{-1}$$

Последовательным дифференциированием правой части мы найдем ее ряд Маклорена и тем самым необходимый коэффициент $p(n,k)$ — достаточно сложный путь!

12.4.4 Вычисление чисел Стирлинга второго рода

Рассмотрим Y – все *эпиморфные* отображения f множества $S_1 = \{1,2,\dots n\}$ на множество $S_2 = \{1,2,\dots k\}$. Каждое такое отображение делит S_1 в точности на k кусков P_i так, что $f(P_i) = i$. Ясно, что поскольку порядок кусков нам не важен, то искомых отображений будет $k!\left\{{n \atop k}\right\} = |Y|$. С другой стороны по формуле таблиц все (то есть уже не обязательно эпиморфные) такиие функции составляют множество X из в точности k^n элементов, обозначим множество тех из них, которые в образе не содержат j через X_j. Понятно, что $Y = \bigcap_{j=1}^{k}(X \setminus X_j)$, а потому

$$|Y| = |X - \bigcup_{j=1}^{k} X_j| = k^n - |\bigcup_{j=1}^{k} X_j|$$

При этом ясно, что $|X_j| = (k-1)^n$ и опять-таки по формуле таблиц

$$\sum_{1\leqslant i_1<i_2<\dots<i_j\leqslant k} |X_{i_1} \cap X_{i_2} \cap \dots X_{i_j}| = \binom{k}{j}(k-j)^n$$

Применяя формулу включений и исключений получаем

$$k!\left\{{n \atop k}\right\} = |Y| = k^n - \left[\sum_{j=1}^{k}(-1)^{j+1}\binom{k}{j}(k-j)^n\right] = \sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n$$

откуда уже и получается итоговая формула

$$\left\{{n \atop k}\right\} = |Y| = \frac{1}{k!}\sum_{j=0}^{k}(-1)^j\binom{k}{j}(k-j)^n$$

12.5 Пример использования методов: этапы решения одной задачи. Числа Каталана

12.5.1 Задача

Пусть у нас есть n пар скобок: n левых и столько же правых. Расстановку этих скобок назовем правильной, если она задает, например, порядок действий в длинной последовательности сложений. Например, такая $(())()()$ расстановка четырех пар скобок правильная, а такая $)()((())$ нет. Общее число правильных расстановок $\mathcal{W}[n]_k$ из n пар скобок называется *числом Каталана* C_n, к этой задаче сводится некоторое количество других комбинаторных (и важных для физики) задач. Как найти формулу для числа Каталана C_n?

Примем, что , что $C_0 = 1$, и попробуем отыскать рекуррентное соотношение между числами Каталана.

Выберем самую левую открывающую скобку и найдем парную ей закрывающую. В зависимости от ее положения мы получим представление $\mathcal{W}[n]_k = (\mathcal{W}[m]_i)\mathcal{W}[n-m-1]_j$ при $0 \leqslant m < n$. Занимаясь подсчетами конфигураций $\mathcal{W}[n]_k$ с учетом таких представлений получаем рекуррентную формулу $C_n = \sum_{i=0}^{n-1} C_i C_{n-i-1}$.

Теперь рассмотрим производящую функцию $C(q) = \sum_0^\infty C_k q^k$ чисел Каталана. Возводя этот ряд в квадрат получим явно

$$\begin{aligned}[C(q)]^2 &= C_0C_0 + (C_0C_1 + C_1C_0)q + (C_0C_2 + C_1C_1 + C_2C_0)q^2 + \dots + \left(\sum_{i=0}^{n-1} C_iC_{n-i-1}\right)q^{n-1} + \dots \\ &= C_1 + C_2q + C_3q^2 + \dots + C_nq^{n-1} + \dots = \frac{C(q) - C_0}{q}\end{aligned}$$

Таким образом, возникает квадратное уравнение относительно производящей функции $C(q)$:

$$q([C(q)]^2 - C(q) + 1 = 0 \qquad \text{откуда берем положительное решение} \qquad C(q) = \frac{1-\sqrt{1-4q}}{2q}$$

Ряд Маклорена для $\sqrt{1-x}$ можно выписать явно (упражнение по математическому анализу!), используя биномиальные коэффициенты:

$$\sqrt{1-x} = 1 - 2\sum_{k=1}^{\infty} \frac{1}{k}\binom{2k-2}{k-1} x^k$$

$$\frac{1-\sqrt{1-4q}}{2q} = \sum_{k=1}^{\infty} \frac{1}{k}\binom{2k-2}{k-1} q^{k-1} \qquad \Longrightarrow \qquad C(q) = \sum_{n=0}^{\infty} \frac{1}{n+1}\binom{2n}{n} q^n$$

и потому число Каталана $C_n = \frac{1}{n+1}\binom{2n}{n}$