

MAT220 第4讲：概率的数学基础

Contents

1 集合论	2
1.1 什么是集合?	2
1.2 集合的规则	2
1.3 基数	4
1.4 子集	4
1.5 幂集	5
2 集合的运算	6
2.1 构造新的集合	6
2.2 补集	6
2.3 交集	7
2.4 并集	8
2.5 组合运算	8
2.6 基数与集合运算	10
3 配置的计数	10
3.1 乘法原则	10
4 排列的计数	12
4.1 什么是排列?	12
4.2 排列的计数	13
4.3 用 8 名学生和 8 把椅子解释公式	14
5 组合的计数	16
5.1 帕斯卡三角形	16
5.2 用组合数计算子集	18

在第1讲到第3讲中，我们学习了统计学的基本语言、数据的可视化描述，以及均值、中位数、方差和标准差等数值概括。现在，我们将从描述统计转向概率论。不过，在真正进入概率论之前，我们需要先介绍一些在概率论中经常使用的数学概念。本讲将正是为此而展开，介绍本课程后续部分会用到的集合论和计数方法。

1 集合论

1.1 什么是集合？

为了用数学方式讨论一组事物，我们需要一种把它们一起写下来的方法。这就是集合：集合就是一个没有顺序的项目列表。

定义：集合

集合是一些事物的无结构集合。

我们可以把集合想象成一个数学盒子，用来把事物放在一起。这个盒子本身没有内在结构，除了它存在并且包含它所包含的东西之外。事实上，就是这样：集合只是把事物收集到同一个地方的数学方式。由于这个概念非常开放，我们基本上可以由任何对象形成集合。例如：

- 所有大于 2 的数的集合，
- 教室里所有棕色头发的人的集合，
- 今天乘电车到校的所有学生的集合，
- 所有名叫 Carl 的企鹅的集合。

集合中的对象称为集合的元素，它们就像集合的成员。在数学记号中，我们用花括号 { 和 } 来表示集合。例如，记号

$$A = \{1, 2, 3, 4, 5, 6, 7\}$$

表示集合 A ，它的元素是 1, 2, 3, 4, 5, 6, 7。需要强调的是，最好把集合理解为没有内在顺序。当然，我们在列出集合元素时必须按照某种顺序写出来，但这个顺序完全不重要：集合只由它的元素决定。例如，集合 A 也可以同样写成：

$$A = \{3, 4, 1, 7, 2, 5, 6\}, \quad \text{或} \quad A = \{3, 1, 4, 5, 7, 6, 2\}.$$

如果对象 x 是集合 A 的一个元素，那么我们写作

$$x \in A,$$

读作“ x 属于 A ”。如果 x 不是 A 的元素，那么我们会像写 $\neq$ 一样，在这个符号上加一条斜线。记号

$$x \notin A$$

读作“ x 不属于 A ”。例如，如果 $A = \{1, 3, 5, 7, 9\}$ ，那么 $3 \in A$ ，但 $4 \notin A$ 。

1.2 集合的规则

集合是几乎所有其他数学结构的基本组成部分，因为在讨论一组事物可能具有或不具有怎样的结构之前，我们首先必须把它们收集到同一个地方。从这个意义上说，集合论可以看作现代数学的基础之一。不过，虽然我们可以把集合看作一种“前结构”，它仍然有一些基本规则。

集合的基本规则

规则 1: 不能把同一个元素写两次。

规则 2: 如果两个集合拥有完全相同的元素，那么它们相等。

规则 3: 集合可以把其他集合作为元素，但集合不能包含它自己。

规则 4: 有一个完全不含任何东西的特殊集合，称为空集，记作 $\emptyset$ 。

规则 5: 不存在包含绝对一切事物的全集。

规则1说明，一个集合的元素只写一次。例如，这意味着下面两个集合相等：

$$\{1, 1, 2\} = \{1, 2\}$$

也就是说，在写出一个集合时，我们不写重复项。换句话说，集合不记录其成员出现的频率，只记录它们是否存在。第二条规则意味着

$$\{1, 2, 3\} = \{3, 2, 1\} = \{1, 3, 2\},$$

这也就是说，集合中元素的顺序不重要。规则3允许我们写出类似

$$A = \{1, 2, \{3, 4\}\},$$

这样的东西。这是一个包含三个元素的集合：数字 1、数字 2，以及集合 $\{3, 4\}$ 。但是，它不允许我们写 $A = \{A\}$ 这样的东西，因为这会导致自指悖论。用符号表示，规则3强制所有集合 A 都满足 $A \notin A$ 。第四条规则引入了空集，记作 $\emptyset$ ，它是不含任何元素的唯一集合。这个特殊集合可以看作集合论中的一种零元素，也就是一种形式化地表示“什么都没有”的方式。需要注意的是，根据规则3，空集并不等于包含空集的集合：

$$\emptyset \neq \{\emptyset\},$$

因为前者有零个元素，而后者有一个元素： $\emptyset$ 本身。最后，规则5说明，不存在包含绝对所有数学对象的集合。这个规则是出于技术原因而设定的：如果有一个集合 U 包含绝对一切，那么这个集合本身也是一个东西，所以按定义它应该包含它自己。但这会与规则3矛盾。

练习：集合的相等

判断下列集合是否相等。

- (a) $A = \{1, 3, 5, 7, 9\}$ and $B = \{9, 7, 5, 3, 1\}$.
- (b) $A = \{13, 6, 7, 9, 2\}$ and $B = \{6, 2, 13, 9\}$.
- (c) $A = \{2, 4, 6, 8\}$ and $B = \{2, 2, 2, 2, 4, 6, 8, 8\}$.
- (d) $A = \{14\}$ and $B = \{7 + 7\}$.

解答

- (a) 是。两个集合有完全相同的元素，所以 $A = B$ 。
- (b) 否。集合 A 包含 7，但集合 B 不包含。因此 $A \neq B$ 。
- (c) 是。我们忽略重复出现的元素，所以 $A = B$ 。
- (d) 是。因为 $7 + 7 = 14$ ，两个集合包含完全相同的元素。因此 $A = B$ 。

1.3 基数

由于集合没有内在结构，可以讨论的内容并不多。不过，我们可以数出集合中包含的元素个数，从而得到一种大小的概念。这个“集合的大小”有一个特殊名称：它称为集合的基数。

定义：基数

对于有限集合 A ， A 的**基数**就是 A 中元素的个数。它记作

$$|A|.$$

例如，

$$|\{a, b, c, d, e\}| = 5, \quad |\emptyset| = 0.$$

1.4 子集

顾名思义，集合 A 的子集是从 A 中取出的一种较小的事物集合。形式上，我们可以如下定义子集。

定义：子集

对于任意两个集合 A 和 B ，如果 A 的每一个元素也都是 B 的元素，那么我们说 A 是 B 的**子集**，记作

$$A \subseteq B,$$

。

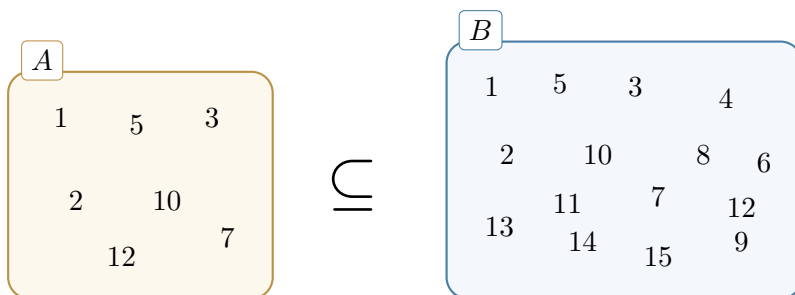
例如，集合

$$A = \{1, 2, 3, 5, 7, 10, 12\}$$

是更大集合

$$B = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\},$$

的子集，因为 A 中包含的每一个元素也都在 B 中。从无结构盒子的角度看，我们可以把子集想象成下面这样。



只要逐一检查 A 的每一个元素是否也出现在 B 中，就可以验证 A 是 B 的子集。由于 A 的大小是 7，我们可以一个一个地检查。

有两个关于子集的事实特别重要。对于任意集合 A ，总有

$$A \subseteq A \quad \text{且} \quad \emptyset \subseteq A.$$

第一个式子说明，每个集合都是它自己的子集。这在某种意义上总是显然成立的： A 中的任意元素 x 默认就在 A 中。第二个式子说明，空集是每个集合的子集，因为 $\emptyset$ 中没有任何元素可能不属于 A 。这种陈述常被称为空真。

练习：子集

设

$$A = \{1, 2, 3, 4\}, \quad B = \{2, 3, 5, 7, 9\},$$

并令 S 为所有素数的集合。判断下列命题是真还是假。

- (a) $A \subseteq B$.
- (b) $A \subseteq S$.
- (c) $B \subseteq S$.
- (d) $\{2, 3\} \subseteq B$.
- (e) $A \subseteq A$.
- (f) $\emptyset \subseteq A$.

解答

- (a) 假。元素 1 和 4 在 A 中，但不在 B 中。
- (b) 假。数字 1 和 4 不是素数。
- (c) 假。数字 9 在 B 中，但 9 不是素数。
- (d) 真。2 和 3 都是 B 的元素，所以 $\{2, 3\} \subseteq B$ 。
- (e) 真。每个集合都是它自己的子集。
- (f) 真。空集是每个集合的子集。

1.5 幂集

假设我们有一个集合 A ，它有 n 个元素：

$$A = \{x_1, x_2, \dots, x_n\}$$

从上一节我们知道，可以通过选择 A 中的一些元素并把它们放入一个集合，同时需要在需要时舍弃其他元素，来形成 A 的一个子集。从最小的情况开始，对于 A 中任意元素 x ，我们都可以形成只包含这个元素的子集： $\{x\} \subseteq A$ 。如果对 A 的每一个元素都这样做，那么我们会得到许多大小为 1 的子集：

$$\text{大小为 1 的子集：}\{x_1\}, \{x_2\}, \{x_3\}, \dots, \{x_n\}.$$

我们也可以尝试列出所有大小为 2 的子集。这些子集由所有元素对组成：

$$\text{大小为 2 的子集：}\{x_1, x_2\}, \{x_2, x_3\}, \{x_1, x_3\}, \dots$$

这样的子集似乎会更多，但原则上它们都可以写出来。事实上，我们可以重复这个过程，形成大小为 3, 4, 5, 等等的子集。最终这个过程会停止，因为 A 只有 n 个元素，这意味着大小为 n 的子集只有一个： A 本身。

假设我们完成了这个过程，那么我们就已经写出了 A 的所有子集。毫不意外，这个集合也可以构成另一个集合。我们把这个“ A 的所有子集构成的集合”称为 A 的幂集。

定义：幂集

集合 A 的**幂集**，记作 $\mathcal{P}(A)$ ，是 A 的所有子集构成的集合。

例如，如果

$$A = \{1, 2, 3\},$$

那么

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

可以看到，幂集似乎比原来的集合大得多。在上面的例子中，集合 $A = \{1, 2, 3\}$ 有 3 个元素，而它的幂集有 8 个元素。这并不是巧合：实际上 $2^3 = 8$ 。一般来说，我们有如下规则。

幂集的大小

如果有限集合 A 有 n 个元素，那么

$$|\mathcal{P}(A)| = 2^n.$$

我们会在本讲后面说明这个规则。

2 集合的运算

2.1 构造新的集合

由旧集合构造新集合有几种方法。在概率论的语境中，我们会重点关注三种主要运算：

1. 集合的补集，
2. 集合的交集，
3. 集合的并集。

正如我们将在下一讲看到的，这三种运算使我们能够恰当地表达并计算包含“非”“且”“或”等词语的事件概率。下面我们简要回顾这些运算。

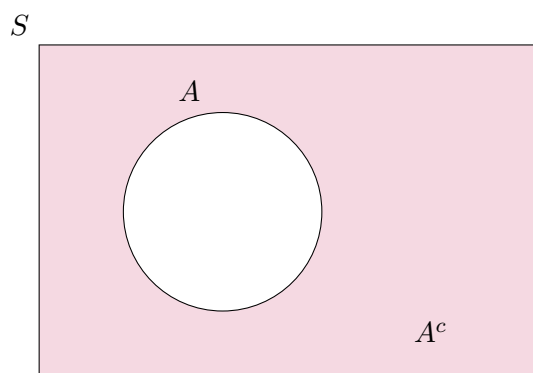
2.2 补集

A 的补集有时写作 A^c ，就是所有不在 A 中的事物的集合。根据规则5，我们不能简单地考察所有不在 A 中的事物。相反，我们需要相对于某个更大的集合 S 来取补集。

定义：补集

设 $A \subseteq S$ 。 A 在 S 中的**补集**是 S 中所有不属于 A 的元素 x 构成的集合。

从图形上看，补集可以想象成下面阴影区域：



例如，考虑集合 $S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ 。如果令 A 为 S 中所有偶数的集合，即 $A = \{2, 4, 6, 8, 10\}$ ，那么 A 在 S 中的补集就是 S 中那些不在 A 中的数。我们可以通过写出 S 并删去 A 中的所有元素来找到这些数。剩下的就是补集。写出 S 并删去 A 的元素，我们得到

$$\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\} \rightarrow \{1, 3, 5, 7, 9\}.$$

因此，补集 A^c 等于集合 $\{1, 3, 5, 7, 9\}$ ，即 S 中的奇数集合。

2.3 交集

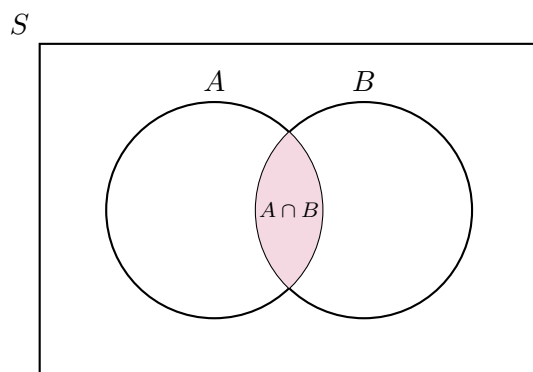
A 和 B 的交集有时写作 $A \cap B$ ，是所有同时在 A 和 B 中的事物的集合。

定义：交集

A 和 B 的交集是

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

从图形上看，两个集合的交集可以像下面的阴影区域：



不过，需要注意的是，两个集合完全可能有空交集。如果上图中的两个圆没有重叠，就会发生这种情况。

例如，令 $A = \{1, 2, 3, 4\}$ 且 $B = \{2, 3, 5, 7, 9\}$ 。对于交集，我们要寻找同时在 A 和 B 中的元素。观察可知，

$$A \cap B = \{2, 3\},$$

因为 2 和 3 是唯一同时属于两个集合的元素。

2.4 并集

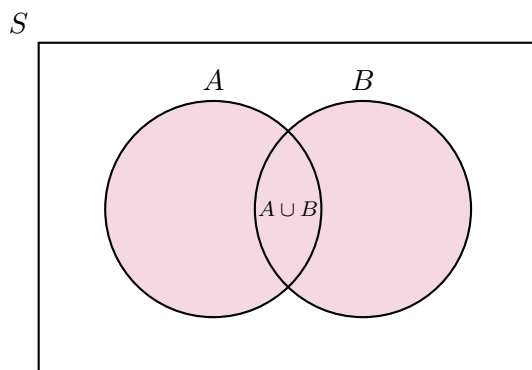
A 和 B 的并集有时写作 $A \cup B$ ，是所有在 A 中、在 B 中，或同时在二者中的事物的集合。我们可以把它想象成把 A 和 B 合并成一个集合后得到的集合。

定义：并集

A 和 B 的并集是

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

从图形上看，两个集合的并集可以像下面这样：



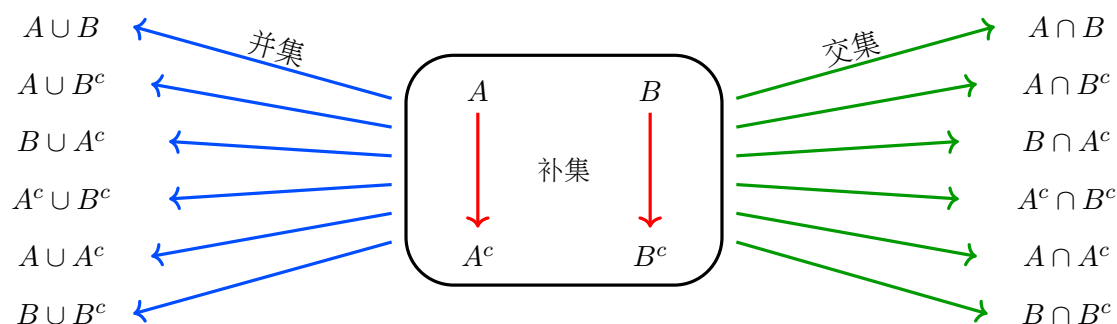
例如，假设 $A = \{1, 2, 3\}$ ， $B = \{2, 3, 4, 5\}$ 。那么并集就是把所有这些东西收集在一起：

$$A \cup B = \{1, 2, 3, 4, 5\}.$$

注意，根据第1.2节的规则1，我们不把元素 2 和 3 写两次。

2.5 组合运算

利用补集、交集和并集这三种运算，我们可以从少量初始信息构造出许多其他集合。只从两个集合 A 和 B 出发，并把它们都看作某个更大集合 S 的子集，我们就可以应用这三种运算来创造许多新集合：



练习：集合运算

令

$$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\},$$

并考虑三个子集

$$A = \{2, 4, 6, 8, 10\}, \quad B = \{2, 3, 5, 7, 9\}, \quad C = \{1, 3, 5, 7, 9\}.$$

写出下列集合。

- (a) A^c
- (b) $A \cap B$
- (c) $(A \cap B) \cup C$
- (d) $A \cap C$
- (e) $A \cap B \cap C$

解答

(a) 因为 A 是 S 中偶数的集合，所以它的补集是

$$A^c = \{1, 3, 5, 7, 9\}.$$

(b) A 和 B 中唯一共同的元素是 2，所以

$$A \cap B = \{2\}.$$

(c) 因为 $A \cap B = \{2\}$ ，所以

$$(A \cap B) \cup C = \{2\} \cup \{1, 3, 5, 7, 9\} = \{1, 2, 3, 5, 7, 9\}.$$

(d) A 和 C 中没有共同元素，所以

$$A \cap C = \emptyset.$$

(e) 因为 $A \cap C = \emptyset$ ，所以不可能有元素同时在三个集合中。因此

$$A \cap B \cap C = \emptyset.$$

2.6 基数与集合运算

集合运算以有用的方式与基数相互作用。一些基本事实是：

$$|A \cap B| \leq |A| \quad \text{且} \quad |A \cap B| \leq |B|.$$

这是因为交集只包含原本同时在两个集合中的元素。

类似地，

$$|A| \leq |A \cup B| \quad \text{且} \quad |B| \leq |A \cup B|.$$

这是因为并集包含 A 中的所有元素，也包含 B 中的所有元素。

不过，在计算 $A \cup B$ 的大小时，我们必须避免重复计算重叠部分。因此得到公式

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

例。令

$$A = \{1, 2, 3, 4\}, \quad B = \{3, 4, 5, 6, 7\}.$$

则

$$|A| = 4, \quad |B| = 5, \quad A \cap B = \{3, 4\}.$$

所以

$$|A \cap B| = 2.$$

因此

$$|A \cup B| = |A| + |B| - |A \cap B| = 4 + 5 - 2 = 7.$$

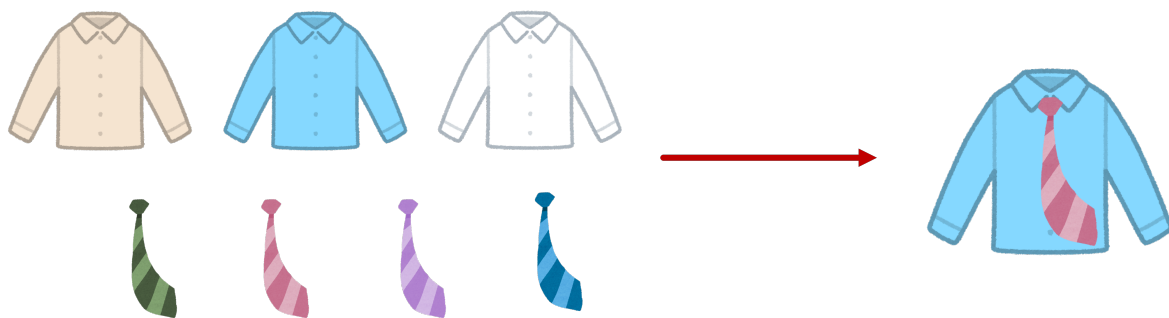
事实上，

$$A \cup B = \{1, 2, 3, 4, 5, 6, 7\}.$$

3 配置的计数

3.1 乘法原则

假设你有 3 件衬衫和 4 条领带，并且今天早上上班时想各穿戴一件。例如，一种可能的选择是下面这样的蓝色衬衫和红色领带：



我们可以搭配出多少种可能的服装？选择一套服装的问题可以分解成两个独立的选择：选择一件衬衫，然后选择一条领带。先选一件衬衫后，就有四种可能的服装：这件衬衫加上四条领带中的任意一条。对所有衬衫重复这个选择，我们得到下面的结构化表格，它展示了所有可能的衬衫和领带配置。



可以看到，一共有 12 种可能的服装，这正好是衬衫的数量 3 乘以领带的数量 4。

除了画漂亮的衣服图，我们也可以把这个问题转化成数学语言。如果把三件衬衫记作 s_1, s_2, s_3 ，并把四条领带记作 t_1, t_2, t_3, t_4 ，那么一套服装就是一个有序对 (s_i, t_j) 。上面的图可以改写为：

	t_1	t_2	t_3	t_4
s_1	(s_1, t_1)	(s_1, t_2)	(s_1, t_3)	(s_1, t_4)
s_2	(s_2, t_1)	(s_2, t_2)	(s_2, t_3)	(s_2, t_4)
s_3	(s_3, t_1)	(s_3, t_2)	(s_3, t_3)	(s_3, t_4)

3 种衬衫选择 $\times$ 4 种领带选择 = 12 种总配置。

这种由有序对构成的网格状结构称为笛卡尔积。如果我们有二个集合：由衬衫组成的 $S = \{s_1, s_2, s_3\}$ ，以及由领带组成的 $T = \{t_1, t_2, t_3, t_4\}$ ，那么我们可以把这两个集合“相乘”，得到所有形如 (s_i, t_j) 的有序对组成的集合。这个新集合，即 S 和 T 的笛卡尔积，记作 $S \times T$ 。我们刚刚从衬衫和领带的例子中发现的是， $|S \times T| = |S| \times |T|$ 。在计数配置的语境中，我们有如下通用规则。

配置的乘法原则

假设一个选择分阶段完成。如果第一阶段有 m 个选项，第二阶段有 n 个选项，那么配置的总数是

$$m \times n.$$

如果有两个以上的阶段，就把每个阶段的选项数量相乘。

练习：配置的计数

你去最喜欢的咖喱餐厅。那里有一个套餐，需要选择：

- 5 种前菜中的一种，
- 15 种主菜中的一种，
- 5 种甜点中的一种，
- 10 种饮料中的一种。

一共可以创造多少种不同的餐食？

解答

前菜有 5 种选择，主菜有 15 种选择，甜点有 5 种选择，饮料有 10 种选择。根据乘法原则，不同餐食的总数是

$$5 \times 15 \times 5 \times 10 = 3750.$$

因此，一共有 3750 种不同的餐食。这意味着你可以每天去一次，超过 10 年都不吃重复的餐食。

4 排列的计数

接下来我们将介绍排列，并开始计算排列的数量。不过，在此之前，我们需要先介绍一个新的记号。

定义：阶乘

对于正整数 n ， n 的阶乘是

$$n! = n(n-1)(n-2) \cdots 3 \cdot 2 \cdot 1.$$

我们还定义

$$0! = 1.$$

例如，

$$4! = 4 \times 3 \times 2 \times 1 = 24,$$

并且

$$10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3628800.$$

我们在 1 处停止，因为如果乘以 0，那么每个阶乘都会等于 0，这就没有用了。

4.1 什么是排列？

正如我们在第1节看到的，对象集合 $A = \{a_1, a_2, \dots, a_n\}$ 并不关心我们写出元素的顺序。不过，我们可以决定关心顺序，并把集合 A 的元素按照不同顺序写出来。为了区别于集合使用的花括号（集合不关心顺序），我们将使用圆括号（和）来表示有顺序的列表。因此，例如：

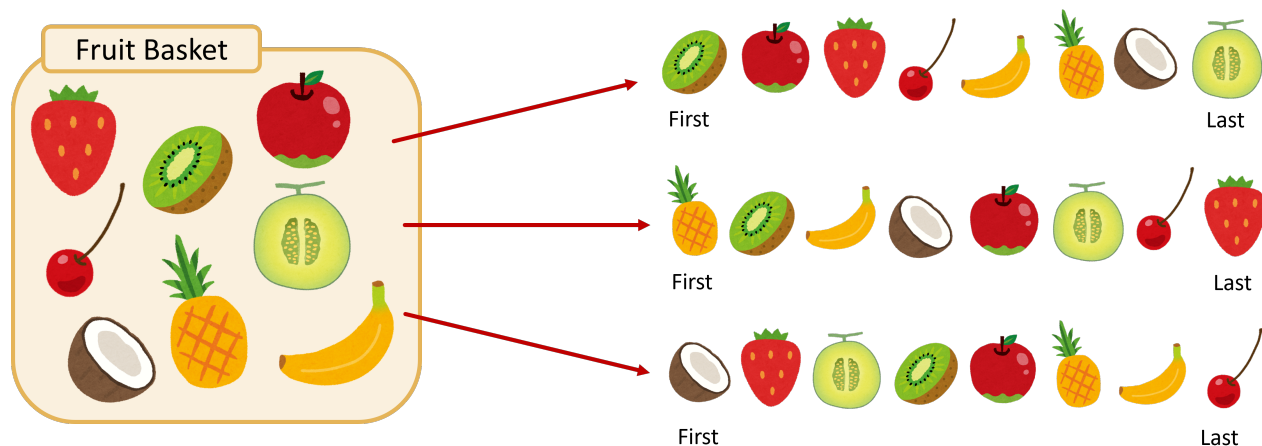
$$(a_1, a_2, \dots, a_n) \quad \text{和} \quad (a_2, a_1, \dots, a_n)$$

被认为是不同的，尽管它们由同样的 n 个元素组成。非常简单地，排列可以定义如下。

定义：排列

排列是对象的一种安排方式，其中顺序是重要的。

例如，假设我们有一个水果篮，里面有 8 个不同的水果。如果我们特别饿，可能想把篮子里的每一个水果都吃掉。由于必须一个一个地吃，所以有许多不同的可能顺序。下面列出了三个可能的顺序。



8 个水果的三种可能排列。

4.2 排列的计数

为了理解一个给定集合有多少种排列，我们可以从小的数字开始，看看是否能发现规律。因此，从 $n = 1$ 开始并逐渐增加：

$$n = 1: \quad 1 \text{ 种,}$$

$$n = 2: \quad 2 \times 1 = 2 \text{ 种,}$$

$$n = 3: \quad 3 \times 2 \times 1 = 6 \text{ 种,}$$

$$n = 4: \quad 4 \times 3 \times 2 \times 1 = 24 \text{ 种.}$$

由此可见，一个明显的规律出现了：我们应该用 $n!$ 来计算 n 个对象的排列数。事实上，在本课程中，我们会使用一个稍微更一般的公式，如下所示。

排列的一般公式

从 n 个不同对象中取出 r 个并按顺序排列的方法数是

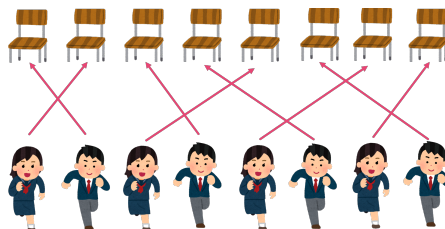
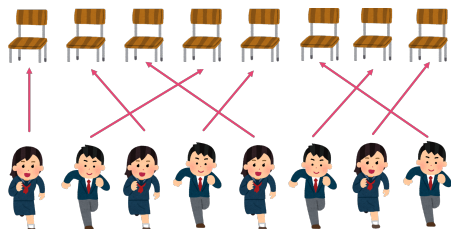
$$P_{n,r} = \frac{n!}{(n-r)!},$$

其中 n 和 r 是整数，并且 $n \geq r$ 。

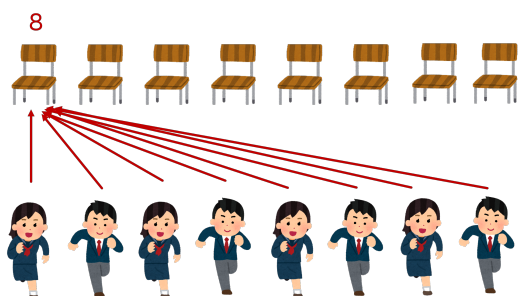
乍一看，这个公式可能显得相当突然。因此，我们现在通过一个详细的例子来仔细说明它。

4.3 用 8 名学生和 8 把椅子解释公式

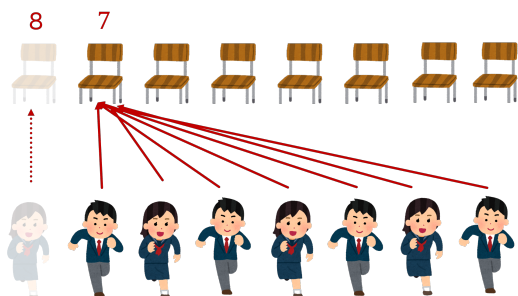
假设一个班级有 8 名学生，并且有 8 把椅子。我们想通过把这 8 名学生分配到房间里的 8 把椅子上来创造一种“座位安排”。例如，下图包含两种可能的座位安排。



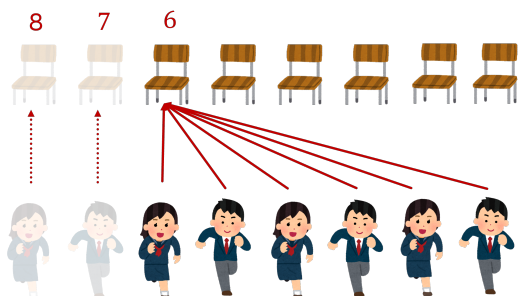
实际上，每一种分配都等同于一种排列：我们取学生集合，并说“你先坐，你第二个坐”，依此类推。因此，座位安排的总数应该是 $8! = 40320$ 。下面我们逐步解释。



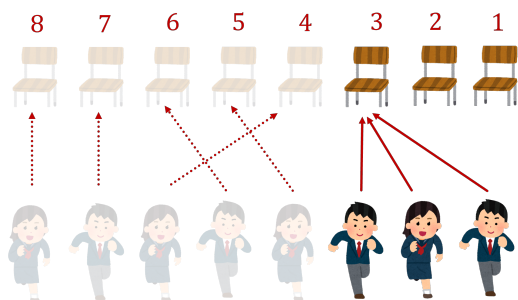
第1步：第一把椅子。 由于一开始有 8 名学生，所以第一把椅子可以由谁来坐有 8 种可能选择。



第2步：第二把椅子。 当我们将第二把椅子重复同样过程时，现在只剩下 7 名学生和 7 把空椅子。因此，不再有 8 种选择，而只有 7 种选择。



第3步：第三把椅子。 在从 7 名学生中选择一人坐到第二把椅子之后，我们可以继续考虑第三把椅子。此时，我们已经分配了两名学生，所以可能坐到第三把椅子上的学生只剩下 6 名。换句话说，第三把椅子现在有 6 种可能选择。



第4步：继续这个模式。 我们可以不断重复这个推理，得到 5 种、然后 4 种、然后 3 种、然后 2 种，最后剩余椅子有 1 种可能选择。

为了得到可能配置的总数，我们可以对这个 8 步过程使用乘法原则，得出座位安排的总数为

$$8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 8!.$$

这就恢复了前面提到的阶乘公式。那么，前面提到的公式

$$P_{n,r} = \frac{n!}{(n-r)!}$$

又是怎么来的呢？我们通过修改设置来解释。

现在假设我们从教室里移走 3 把椅子，因此只剩下 5 把椅子。所以，如果我们试图让 8 名学生坐下，那么只有其中 5 人能坐下，另外 3 人必须站着。接下来的问题是：我们有多少种方式可以让 8 名学生中的 5 人坐下？下面展示了两种这样的安排。



我们可以重复之前的推理，因此第一把椅子有 8 种选择，第二把椅子有 7 种选择，依此类推。不过，由于现在只有 5 把椅子需要坐满，我们必须在 5 把椅子之后提前停止这个过程。这意味着答案不是 $8!$ ，而是前面推导中的前 5 项：

$$P_{8,5} = 8 \times 7 \times 6 \times 5 \times 4.$$

这里 $r = 5$ ，我们观察到，同一个公式也可以写成：

$$P_{8,5} = \frac{8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1}{3 \times 2 \times 1}.$$

这与前面提到的公式相同。这里的思想是， $P_{n,r}$ 的分子告诉我们如何排列 n 个总对象，而分母 $(n-r)!$ 告诉我们如何排列没有被选中的 $(n-r)$ 个对象。我们除以这个数，是为了把它们从计数中移除。在上面的椅子例子中， $(n-r)!$ 项对应于仍然站着的 3 名学生的 6 种排列方式。

5 组合的计数

讨论排列时，我们关心学生被安排的顺序。也就是说，我们把所有椅子按顺序排好，并通过“你第一个，你第二个，...”来分配不同学生。我们可以想象还有另一种做法：更一般地说，也许我们可以只是选择 r 名学生，而不关心他们坐在椅子上的顺序。这相当于从 8 名学生中选择 5 名。下面展示了两组这样的选择。



事实上，我们也可以数出“从 n 个对象中选择 r 个对象的所有可能方式”。这些选择不是排列，因为在这种情况下我们不关心顺序。相反，这些选择只是从群体中挑选一些成员。

事实上，对于从 n 个对象的集合中选择 r 个对象的所有可能方法，有一个公式。在数学中，我们把这些不考虑顺序的选择称为“组合”，并用记号 $C_{n,r}$ 表示从大小为 n 的集合中选择 r 个对象的方法数。 $C_{n,r}$ 的公式与 $P_{n,r}$ 的公式非常相似，如下所示。

组合的一般公式

从 n 个对象中每次取 r 个对象的组合数是

$$C_{n,r} = \frac{n!}{r!(n-r)!},$$

其中 n 和 r 是整数，并且 $n \geq r$ 。

注意，这里我们有 $P_{n,r}$ 的公式，但还额外除以一个 $r!$ 项。为什么要这样做？在 $P_{n,r}$ 的情况下，我们数的是有序选择，而同样一组 r 个对象可以按 $r!$ 种不同方式排列。然而，对于无序组合，我们并不关心这 $r!$ 种排列，因此我们把整体公式除以 $r!$ ，从计数中去掉这些选择。

为了说明这个公式，我们可以计算从 8 名学生中选择 5 名学生的方法数：

$$C_{8,5} = \frac{8 \cdot 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{(5 \cdot 4 \cdot 3 \cdot 2 \cdot 1)(3 \cdot 2 \cdot 1)}.$$

当然，这可以直接输入计算器。不过，我们也可以观察到，这个分数的分子和分母中都有一个 $5!$ 因子。因此，我们可以约去它们，得到：

$$C_{8,5} = \frac{8 \cdot 7 \cdot 6}{3 \cdot 2 \cdot 1} = \frac{336}{6} = 56.$$

所以，从 8 名学生中选择 5 名学生一共有 56 种方法。

5.1 帕斯卡三角形

事实上，公式 $C_{n,r}$ 满足一些有趣的数学性质。我们先从两个非常简单且值得记住的观察开始：

$$C_{n,0} = 1 \quad \text{且} \quad C_{n,n} = 1.$$

这些等式可以通过代数方式检查公式而轻易验证。我们有

$$C_{n,0} = \frac{n!}{0!(n-0)!} = \frac{n!}{0!n!} = 1,$$

因为 $0! = 1$ 。类似地,

$$C_{n,n} = \frac{n!}{n!(n-n)!} = \frac{n!}{n!0!} = 1.$$

从直观上说, 这有一个清楚的解释: 从 n 个对象中什么都不选恰好有一种方法, 而选择全部对象也恰好有一种方法。

更有趣的是, 对于 $1 \leq r \leq n$, 这些 $C_{n,r}$ 公式还满足恒等式

$$C_{n,r-1} + C_{n,r} = C_{n+1,r},$$

这通常称为帕斯卡恒等式。对于不太相信的人, 我们可以用代数方法验证它。

代数验证。 我们计算:

$$C_{n,r-1} + C_{n,r} = \frac{n!}{(r-1)!(n-r+1)!} + \frac{n!}{r!(n-r)!}.$$

为了相加这两个分数, 使用公分母

$$r!(n-r+1)!.$$

于是

$$\begin{aligned} C_{n,r-1} + C_{n,r} &= \frac{rn!}{r!(n-r+1)!} + \frac{(n-r+1)n!}{r!(n-r+1)!} \\ &= \frac{rn! + (n-r+1)n!}{r!(n-r+1)!} \\ &= \frac{(n+1)n!}{r!(n-r+1)!} \\ &= \frac{(n+1)!}{r!((n+1)-r)!} \\ &= C_{n+1,r}. \end{aligned}$$

还有一种简单的计数解释。要从 $n+1$ 个对象中选择 r 个对象, 关注其中一个特殊对象。这个特殊对象要么被选中, 要么没有被选中。

- 如果特殊对象被选中, 那么还需要从其他 n 个对象中选择 $r-1$ 个对象。这给出 $C_{n,r-1}$ 种可能。
- 如果特殊对象没有被选中, 那么需要从其他 n 个对象中选择全部 r 个对象。这给出 $C_{n,r}$ 种可能。

把这两种情况合在一起, 就得到

$$C_{n,r-1} + C_{n,r} = C_{n+1,r}.$$

对于较小的 n 和 r , 我们可以把 $C_{n,r}$ 的值排列成一个三角形。三角形顶部是 $C_{0,0} = 1$, 下一行是 $C_{1,0} = 1$ 和 $C_{1,1} = 1$ 。不过, 当我们继续增加行时, 事情开始变得有趣。下图中, $C_{n,r}$ 的值被写成三角形的形式。

第 0 行:			1														$C_{0,0}$
第 1 行:			1		1												$C_{1,0}$ $C_{1,1}$
第 2 行:			1		2		1										$C_{2,0}$ $C_{2,1}$ $C_{2,2}$
第 3 行:			1		3		3		1								$C_{3,0}$ $C_{3,1}$ $C_{3,2}$ $C_{3,3}$
第 4 行:		1		4		6		4		1							$C_{4,0}$ $C_{4,1}$ $C_{4,2}$ $C_{4,3}$ $C_{4,4}$
第 5 行:	1		5		10		10		5		1						$C_{5,0}$ $C_{5,1}$ $C_{5,2}$ $C_{5,3}$ $C_{5,4}$ $C_{5,5}$
⋮																	⋮

注意，三角形中的每一个内部数字都由它上方两个数字的和给出，这正是帕斯卡恒等式在起作用。例如，第4行中的“6”（用红色标出）由第3行中它上方的“3”和“3”相加得到。通过应用这个加法规则，我们可以形成任意多的新行。

注意，每一行由 $C_{n,r}$ 中的“ n ”编号，而该行的各项则由 r 从 0 到 n 的不同取值给出。这意味着我们不需要直接计算 $C_{n,r}$ 的值，而可以直接从三角形中读出它的值。

5.2 用组合数计算子集

组合与幂集之间有一个重要联系。设 A 是一个有 n 个元素的集合。如果我们想计算 A 中恰好有 r 个元素的子集数量，那么这就是从 n 个元素中选择 r 个元素，并且顺序不重要。因此，大小为 r 的子集数量是

$$C_{n,r}.$$

例。令

$$A = \{a, b, c, d\}.$$

这个集合的大小是 4。我们可以按照子集的大小来计数：

- $C_{4,0} = 1$ 计算大小为 0 的唯一子集，即 $\emptyset$ 。
- $C_{4,1} = 4$ 计算子集 $\{a\}, \{b\}, \{c\}, \{d\}$ 。
- $C_{4,2} = 6$ 计算子集 $\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}, \{c, d\}$ 。
- $C_{4,3} = 4$ 计算子集 $\{a, b, c\}, \{a, b, d\}, \{a, c, d\}, \{b, c, d\}$ 。
- $C_{4,4} = 1$ 计算大小为 4 的唯一子集，即 A 本身。

把这些加起来得到

$$1 + 4 + 6 + 4 + 1 = 16 = 2^4.$$

这与幂集规则一致：

$$|\mathcal{P}(A)| = 2^4 = 16.$$

这个联系也解释了为什么帕斯卡三角形每一行的和都是 2 的幂。

[illegible]