1.2 概率公理化

前面所提的"统计概率"是根据频率与概率的关系而定义的, 即通过大量的重复试验所得的频率近似值来逼近概率, 相关定义缺乏数学的严谨性. 为此, 前苏联数学家柯尔莫哥洛夫 (Andrey Kolmogorov) 于 1930s 给出了概率公理化体系, 即通过基本性质来定义概率, 建立媲美于欧氏几何公理化的理论体系.

定义1.1 (概率公理化定义)**.** *若随机试验 E 所对应的样本空间 Ω 中每一个随机事件 A, 均赋予一实数 $P(A)$, 且满足以下条件:*

1° *非负性: 对任意事件 A 有 $P(A) \geq 0$;*

2° *规范性: 对样本空间 Ω 有 $P(\Omega) = 1$;*

3° *可列可加性: 若 $A_1, A_2, \ldots, A_n, \ldots$ 是两两不相容事件, 即 $A_iA_j = \emptyset$ $(i \neq j)$, 有*

$$P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = P(A_1) + P(A_2) + \cdots + P(A_n) + \cdots;$$

*则称 $P(A)$ 为随机事件 A 的***概率***.*

根据概率公理化定义, 可推导概率的一系列重要性质.

性质1.1. $P(\emptyset) = 0$.

Proof. 令 $A_i = \emptyset$ $(i = 1, 2, \ldots)$, 则有 $\emptyset = \cup_{i=1}^{\infty} A_i$ 且 $A_i \cap A_j = \emptyset$ $(i \neq j)$. 根据公理 3° 有

$$P(\emptyset) = P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{\infty} P(\emptyset).$$

再根据公理 1° 可知 $P(\emptyset) = 0$. □

性质1.2 (有限可加性)**.** *若 $A_1, A_2, \cdots, A_n$ 是两两不相容事件, 则*

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i).$$

Proof. 令 $A_i = \emptyset$ $(i > n)$, 则 $A_1, A_2, \cdots, A_n, A_{n+1}, \ldots$ 是两两互不相容事件, 根据公理 3° 可知

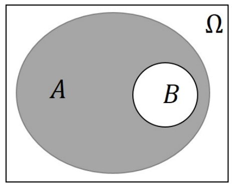
$$P\left(\bigcup_{i=1}^{n} A_i\right) = P\left(\bigcup_{i=\infty}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) + \sum_{i=n+1}^{\infty} P(\emptyset) = \sum_{i=1}^{n} P(A_i)$$

性质得证. □

性质1.3. *对任意事件 A, 有 $P(\overline{A}) = 1 - P(A)$.*

Proof. 由于 $\Omega = \overline{A} \cup A$, 以及事件 A 与 $\overline{A}$ 互不相容, 根据有限可加性有 $1 = P(\Omega) = P(A) + P(\overline{A})$. □

性质1.4. *若事件 $B \subset A$, 则有 $P(B) \leq P(A)$ 以及 $P(A - B) = P(A) - P(B)$.*



Proof. 若 $B \subset A$, 如右图所示有 $A = B \cup (A - B)$, 根据定义可知 B 与 $A - B$ 互不相容。由有限可加性有

$$P(A) = P(B) + P(A - B).$$

再根据公理 1° 有 $P(A - B) = P(A) - P(B) \geq 0$, 从而得到 $P(A) \geq P(B)$. □

性质1.5. 对任意事件 A 和 B, 有

$$P(A - B) = P(A) - P(AB) = P(A \cup B) - P(B).$$

Proof. 根据 $A = (A - B) \cup (AB)$, 以及 $A - B$ 与 AB 互斥, 可得 $P(A) = P(A - B) + P(AB)$. 再根据 $A \cup B = (A - B) \cup B$, 以及 $A - B$ 与 B 互斥, 有 $P(A - B) = P(A \cup B) - P(B)$. □

性质1.6 (容斥原理 Inclusion-Exclusion Principle)**.** 对任意事件A和B, 有

$$P(A \cup B) = P(A) + P(B) - P(AB).$$

Proof. 因 $A \cup B = (A - B) \cup (AB) \cup (B - A)$, 以及 $A - B$, $B - A$, AB两两互不相容, 由有限可加性可知

$$P(A \cup B) = P(A - B) + P(B - A) + P(AB).$$

再将 $P(A - B) = P(A) - P(AB)$ 和 $P(B - A) = P(B) - P(AB)$ 代入上式即可完成证明. □

推论1.1. 对任意事件A, B, C, 有

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$

Proof. 将事件 $A \cup B$ 看作一个随机事件, 根据容斥原理有

$$P(A \cup B \cup C) = P(A \cup B) + P(C) - P((A \cup B) \cap C) = P(A \cup B) + P(C) - P(AC \cup BC).$$

再次利用容斥原理有

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P((AB), \\ P(AC \cup BC) &= P(AC) + P(BC) - P(AC \cap BC) = P(AC) + P(BC) - P(ABC). \end{aligned}$$

代入上式即可完成证明. □

推论1.2. 对事件 $A_1, A_2, \ldots, A_n$, 有

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 \cdots A_n).$$

课题练习(Matching问题)**.** 有 n 对夫妻参加一次聚会, 现将所有参会人员随机分成 n 组, 每组 1 男 1 女, 问至少有一对夫妻被分到同一组的概率是多少?

性质1.7 (Union Bound)**.** 对事件 $A_1, A_2, \dots A_n$, 有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \le P(A_1) + P(A_2) + \cdots + P(A_n).$$

Proof. 利用数学归纳法证明, 当$n = 2$时, 由容斥原理有

$$P(A \cup B) = P(A) + P(B) - P(AB) \le P(A) + P(B). \tag{1}$$

假设当$n = k$时性质成立, 对$n = k + 1$有

$$\begin{aligned} P(A_1 \cup \cdots \cup A_{k+1}) &= P((A_1 \cup \cdots \cup A_k) \cup A_{k+1}) \\ &\le P(A_1 \cup \cdots \cup A_k) + P(A_{k+1}) \\ &\le P(A_1) + \cdots + P(A_k) + P(A_{k+1}), \end{aligned}$$

上式中第一个不等式根据式 (1), 而第二个不等式根据归纳假设. □

推广(Bonferroni不等式)**.** 对事件 $A_1, A_2, \dots A_n$, 有

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) &\le \sum_{i=1}^{n} P(A_i); \\ P\left(\bigcup_{i=1}^{n} A_i\right) &\ge \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i A_j); \\ P\left(\bigcup_{i=1}^{n} A_i\right) &\le \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k); \\ \cdots \quad &\cdots \quad \cdots \end{aligned}$$

例1.4. *设 $P(A) = p$, $P(B) = q$, $P(AB) = r$, 用 p, q, r 分别表示事件的概率: 1) $P(\overline{A} \cup \overline{B})$, 2) $P(\overline{A}B)$; 3) $P(\overline{A} \cup B)$; 4) $P(\overline{A} \cap \overline{B})$.*

解. 对问题 1), 根据事件的对偶律有

$$P(\overline{A} \cup \overline{B}) = P(\overline{AB}) = 1 - r.$$

对问题 2), 根据差事件的定义

$$P(\overline{A}B) = P(B - A) = P(B) - P(AB) = q - r.$$

对问题 3), 根据容斥原理有

$$P(\overline{A} \cup B) = P(\overline{A}) + P(B) - P(\overline{A}B) = 1 - p + q - (q - r) = 1 - p + r.$$

对问题 4), 根据对偶律与容斥原理有

$$P(\overline{A} \cap \overline{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 + r - p - q.$$

□

课题练习. *1) 已知$P(AB) = 0$, 求证: $P(ABC) = 0$.*
2) 设$P(A) = P(B) = P(C) = \frac{1}{4}$, $P(AB) = 0$, $P(AC) = P(BC) = \frac{1}{16}$, 问事件$A, B, C$至少有发生一个的概率.

1.3 古典概型与几何概型

本节介绍两种经典的概率模型: 古典概型与几何概型

1.3.1 古典概型

古典概型相对简单, 是概率论早期最重要的研究对象, 在概率论中具有重要的意义.

定义1.2 (古典概型). *如果试验E满足*

- *样本空间只包含有限个元素*
- *每个基本事件发生的可能性相同*

这类试验称为 ***古典概型****, 又称* ***等可能概型****.*

古典概型要求每个基本事件发生的可行性相同. 例如在相同条件下连续两次抛一枚均匀硬币, 此试验观察的结果: A) 两正面, B) 两反面, C)一正一反. 根据古典概型可知

$$P(A) = P(B) = P(C) = 1/3.$$

这种结论不正确, 因为这三个事件发生的可能性不同. 正确的理解是事件 $C = \{C_1, C_2\}$, 其中 C_1 表示先正后反的事件, C_2 表示先反后正的事件, 从而有

$$P(A) = P(B) = P(C_1) = P(C_2) = 1/4.$$

假设古典概型的样本空间 $\Omega = \{w_1, w_2, \ldots, w_n\}$, 其中 w_i 为基本事件. 若事件 A 包含 k 个基本事件, 则

$$P(A) = k/n = |A|/|\Omega|,$$

这里 $|A|$ 表示事件 A 包含的事件的个数, 由此可知古典概型的本质是计数 (Counting). 下面介绍一些基本的计数原理.

排列: 从 n 个不同的元素中取出 r 个元素进行排列, 有 $(n)_r = n(n-1)\cdots(n-r+1)$ 种不同的方法.

组合: 从 n 个不同的元素中取出 r 个元素, 取出的元素之间无顺序关系, 共有 $\binom{n}{r}$ 种方法, 其中

$$\binom{n}{r} = \frac{n!}{r!(n-r)!} = \frac{(n)_r}{r!}.$$

设过程 $\mathcal{A}_1$ 有 n_1 种方法, 过程 $\mathcal{A}_2$ 有 n_2 种方法, 可定义如下加法与乘法原理:

- 加法原理: 若一项工作可用 $\mathcal{A}_1$ 或 $\mathcal{A}_2$ 来完成, 则完成该任务有 $n_1 + n_2$ 种可能.
- 乘法原理: 若一项工作需分别通过 $\mathcal{A}_1$ 和 $\mathcal{A}_2$ 两过程, 则完成该任务需 $n_1 \times n_2$ 种可能.

下面看一些排列组合的案例.

例1.5. *将 n 个不同的球随机放入 N $(N \geq n)$ 个不同的盒子, A: 恰有 n 个盒子且每盒一球; B: 指定的n个盒子中各有一球; C: 指定一盒子恰有m个球. 求事件发生的概率.*

解. 将 n 只不同的球随机放入 N 个不同的盒子, 共有 N^n 种方法. 对事件 A, 有 $(N)_n = \binom{N}{n} n!$ 中不同的方法, 因此

$$P(A) = \frac{(N)_n}{N^n} = \frac{N!}{N^n r!}.$$

对事件 B, 由于有 $n!$ 种不同的方法方法, 因此

$$P(B) = \frac{n!}{N^n}.$$

对事件 C, 可分为两步, 第一步在指定的盒子里放 m 个球, 有 $\binom{n}{m}$ 种不同的方法; 第二步将剩下的 $n-m$ 个球放入 $N-1$ 个盒子, 共有 $(N-1)^{n-m}$ 种不同的方法, 因此

$$P(C) = \frac{\binom{n}{m}(N-1)^{n-m}}{N^n}.$$

□

下面的例子分析抽签的先后顺序不同是否会对抽签结果产生影响?

例1.6. *袋中有 a 个不同的白球, b 个不同的红球, 随机将球取出并依次排成一列, 问第 k 次取出红球的概率是多少?*

解. 令 A 表示第 k 次取到红球的事件, 将 a 个不同的白球和 b 个不同的红球排列成一排, 有 $(a+b)!$ 种不同的方法. 第 k 次取到红球有 b 种不同的方法, 排列其它的球有 $(a+b-1)!$ 种不同的方法, 因此

$$P(A) = \frac{b(a+b-1)!}{(a+b)!} = \frac{b}{a+b}.$$

由此可知第 k 次取出红球的概率与 k 无关, 任何一人拿到红球的概率均是 $\frac{b}{a+b}$, 抽签先后顺序对抽签的结果没有影响. □

课题练习. *袋中有 a 个相同的白球, b 个相同的红球, 随机将球取出并依次排成一列, 问第 k 次取出红球的概率是多少?*

例1.7. *有 k 个人 $(k < 365)$, 每个人的生日等可能地出现于 365 天中的任意一天, 求至少两人生日相同的概率.*

解. 令事件 A 表示至少有 2 人的生日相同, 则对立事件 $\overline{A}$ 则表示任意两人的生日均不相同. k 个人的生日共有 365^k 种可能, 而 k 个人的生日两两互不相同的有 $(365)_k = k!\binom{365}{k}$ 种可能, 因此

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{(365)_k}{365^k}.$$

□

易知当 $k = 30$ 时, $P(A) = 70.6\%$; 当 $k = 40$时, $P(A) = 89.1\%$; 当 $k = 50$时, $P(A) = 97\%$; 当 $k = 60$时, $P(A) = 99.4\%$; 当 $k = 100$时, $P(A) = 99.99\%$.

例1.8. *设有 N 件产品, 其中有 M 件次品, 现从 N 件产品中任选 n 件, 求其中恰有 k 件次品的概率. 考虑两种情况: i) 不放回抽样, ii) 有放回抽样.*

解. 对于不放回抽样, 从 N 件产品中任选 n 件, 有 $\binom{N}{n}$ 种不同的方法; 在 n 件产品中恰有 k 件次品, 则有 $\binom{M}{k}\binom{N-M}{n-k}$ 种不同的方法, 因此在不放回抽样的情况下, 任选 n 件其中恰有 k 件次品的概率为

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}.$$

对有放回抽样, 每次抽到一件非次品的概率为 $\frac{N-M}{N}$, 每次抽到一件次品的概率为 $\frac{M}{N}$, 因此 n 件中恰有 k 件次品的概率是

$$\binom{n}{k}\left(\frac{M}{N}\right)^k\left(\frac{N-M}{N}\right)^{n-k}.$$

□

例1.9. *将 n 个男生和 m 个女人随机排成一列 $(m < n)$, 任意两女生不相邻的概率是多少.*

解. 先将 n 个男生排成一列, 有 $n!$ 种排法; 然后将 m 个女生排入 n 个男生之间, 使得女生互不相邻, 有 $(n+1)n\cdots(n-m+2) = (n+1)_m$ 种排法. 因此任何两女生不相邻的概率为

$$\frac{n!(n+1)_m}{(m+n)!}.$$

□

课题练习. *若排列成一圈, 首尾相接, 则任意两女生不相邻的概率是多少.*

例1.10. *从 $\{1, 2, \ldots, 9\}$ 数中有放回取 n 个, 试求取出 n 个数的乘积被 10 整除的概率.*

解. 令 $A = \{取出n个整数的乘积能被10整除\}$, $B = \{取出的n个数中有偶数\}$, $C = \{取出的n个数中至少有一个5\}$. 于是有 $A = BC$, 进一步有

$$P(A) = P(BC) = 1 - P(\overline{BC}) = 1 - P(\overline{B} \cup \overline{C}) = 1 - P(\overline{B}) - P(\overline{C}) + P(\overline{B} \cap \overline{C}) = 1 - \frac{5^n}{9^n} - \frac{8^n}{9^n} + \frac{4^n}{9^n}.$$

□

1.3.2 几何概型

古典概型只能考虑有限的样本空间, 即有限个等可能的基本事件, 限制了其使用范围. 本节介绍一种新的概型, 具有如下特征:

- **无限性** 样本空间包含无限多个样本点, 是一个测度有限的区域 (如 1 维线段, 2 维平面区域等), 相应的几何测度有限非零;
- **等可能性** 每个基本事件发生的可能性相等; 每个事件发生的概率与该事件的几何测度相关, 与具体位置无关.

称之为**几何概型**. 其形式化定义如下:

定义1.3. 在一个测度有限的区域 Ω 内等可能性投点, 落入 Ω 内的任意子区域 A 的可能性与 A 的测度成正比, 与 A 的位置与形状无关, 这样的概率模型称之为**几何概型**. 事件 A 发生的概率为

$$P(A) = \frac{A \text{ 的测度}}{\Omega \text{ 的测度}} = \frac{\mu(A)}{\mu(\Omega)}.$$

例1.11. 二人约定中午 $12:00-13:00$ 到某地会面, 两人到达时间随机, 且约定先到者等另一人 15 分钟后离开, 求两人见面的概率.

解. 用 x, y 分别表示两人的到达时间 (分钟), 则样本空间 $\Omega = \{(x,y)\big|0 \le x, y \le 60\}$. 用 A 表示两人见面的事件, 则 $A = \{(x,y)\big||x-y| \le 15\} = \{(x,y)\big|x-y \le 15 \text{ 且 } x-y \ge -15\}$. 于是有

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{60^2-45^2}{60^2} = \frac{7}{16}.$$

□

例1.12. 在区间 $[0,1]$ 内随机取两数, 求两数之积小于 $1/4$ 的概率.

解. 用 $x, y \in [0,1]$ 表示随机取出的两数, 样本空间 $\Omega = \{(x,y)\big|0 \le x, y \le 1\}$. 用 A 表示事件两数之积小于 $1/4$, 则 $A = \{(x,y)\big|xy \le \frac{1}{4}\}$. 样本空间的测度为 $\mu(\Omega) = 1$, 而事件 A 的测度为

$$\mu(A) = \frac{1}{4} + \int_{\frac{1}{4}}^{1} \frac{1}{4x}\,dx = \frac{1}{4} + \frac{1}{4}\ln 2.$$

事件 A 的概率为

$$P(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{1+\ln 2}{4}.$$

□

很多的几何概型可通过计算机模拟近似计算求解, 即 **统计模拟法** 或 **蒙特卡洛(Monte Carlo)法**. 先构造相应的概率模型, 在进行计算机模拟试验, 用统计的方法计算其估计值, 作为所求问题的近似值. 例如, 利用蒙特卡洛法求解例 1.12, 伪代码如下:

- -

$n_A \leftarrow 0$
For $i = 1:N$
  $x \leftarrow \mathrm{Random}(0,1)$
  $y \leftarrow \mathrm{Random}(0,1)$
  If $xy < 1/4$ then
    $n_A \leftarrow n_A + 1$
  Endif
Endfor
Return n_A/N

- -

1.4 组合计数: 十二路

在古典概型中, 概率的计算往往与组合计数相关, 考虑到组合技术在人工智能、计算机等相关学科具有重要的应用. 本节将系统介绍组合组合计数十二路 (The twelvefold way), 该问题由著名组合学家 Gian-Carlo Rota (1932-1999) 提出. 最初问题研究两个集合的函数映射 $f:[n]\to[m]$, 考虑无任何约束, 1-1 单射, 满射三个条件下函数映射的个数. 为问题简介期间, 我们采用 Knuth 的简化问题: 将 n 个球放入 m 个箱子, 如果考虑没有任何限制、每个箱子的球数 ≤ 1 或 ≥ 1 三种情况下有多少种不同的方法.

将 n 个球放入	m 个箱子	每个箱子的球数不限	每个箱子的球数 ≤ 1	每个箱子的球数 ≥ 1
不同	不同	m^n	$(m)_n$	$m!S(n,k)$
相同	不同	$\binom{n+m-1}{n}$	$\binom{m}{n}$	$\binom{n-1}{m-1}$
不同	相同	$\sum_{k=1}^{m} S(n,k)$	$\begin{cases}1 & n\le m\\ 0 & n>m\end{cases}$	$S(n,m)$
相同	相同	$\sum_{k=1}^{m} p(n,k)$	$\begin{cases}1 & n\le m\\ 0 & n>m\end{cases}$	$p(n,m)$

1.4.1 排列, 组合与多重组合

排列: 从 n 个不同的元素中取出 r 个元素进行排列, 有 $(n)_r=n(n-1)\cdots(n-r+1)$ 种不同的方法.

组合: 从 n 个不同的元素中取出 r 个元素, 取出的元素之间无顺序关系, 共有 $\binom{n}{r}$ 种方法, 其中

$$\binom{n}{r}=\frac{n!}{r!(n-r)!}=\frac{(n)_r}{r!}.$$

下面将组合的概念进行推广到多重组合.

多重组合: 将 n 个不同的元素分成 k 组, 每组依次有 $n_1,n_2,\cdots,n_k$ 个元素, 即 $n=n_1+n_2+\cdots+n_k$, 则共有

$$\binom{n}{n_1,n_2,\ldots,n_k}=\binom{n}{n_1}\binom{n-n_1}{n_2}\cdots\binom{n-n_1-n_2-\cdots-n_{k-1}}{n_k}=\frac{n!}{n_1!n_2!\cdots n_k!}$$

种不同的分组方法, 称 $\binom{n}{n_1,n_2,\ldots,n_k}$ 为多重组合, 组合本质上也属于多重组合, 即 $\binom{n}{r}=\binom{n}{r,n-r}$.

多重集: 集合中的元素是可以重复的, 且重复的元素之间不可分辨. 例如 $A=\{1,1,1,2,2,2,3,3,4\}$.

假设多重集 A 有 k 类不同, 每类元素的个数分别为 $n_1,n_2,\ldots,n_k$ 且有 $n=n_1+n_2+\cdots+n_k$, 将多重集 A 中的元素进行排列, 则有

$$\binom{n}{n_1,n_2,\ldots,n_k}$$

中不同的排列方法, 因此多重组合可用于多重集合的排列.

我们还可以得到组合、多重组合与多项式系数有如下关系:

$$(x+y)^n = \sum_{i=0}^{n}\binom{n}{i}x^i y^{n-i}$$

$$(x_1+x_2+\cdots+x_k)^n = \sum_{n=n_1+n_2+\cdots+n_k}\binom{n}{n_1,n_2,\cdots,n_k}x_1^{n_1}x_2^{n_2}\cdots x_k^{n_k}$$

根据排列组合, 有如下结论:

将 n 个球放入 m 个箱子	每个箱子的球数不限	每个箱子的球数 ≤ 1	每个箱子的球数 ≥ 1
不同　　不同	m^n	$(m)_n$	
相同　　不同		$\binom{m}{n}$	

1.4.2 整数的有序分解

研究方程 $x_1+x_2+\cdots+x_k=n$ 有多少个非负整数解, 该问题等价于将 n 个相同的球放入 k 个箱子中, 共有多少种不同的放法. 有如下定理:

性质1.8. 方程 $x_1+x_2+\cdots+x_k=n$ 有 $\binom{n+k-1}{n}=\binom{n+k-1}{k-1}$ 种不同的非负整数解.

Proof. 考虑对应关系: 将 n 个星号 '$*$' 和 $k-1$ 条竖线 '$|$' 排列成一排, 最后再加入一条竖线 '$|$'. 如下例所示:

$$\underbrace{*\ *\ *\ *}_{x_1}\ |\ |\ \cdots\ \underbrace{*\ *\ *\ *\ *}_{x_i}|\ \cdots\ \underbrace{*\ *}_{x_k}|$$

第 i 条竖线 '$|$' 与第 $i-1$ 条竖线 '$|$' 之间星号 '$*$' 的个数表示 x_i 的值, 而第 1 条竖线 '$|$' 之前的星号的个数表示 x_1 的值. 可发现方程 $x_1+x_2+\cdots+x_k=n$ 与这种排列之间存在一一对应关系, 因此方程 $x_1+x_2+\cdots+x_k=n$ 有

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

种非负整数解. □

例如: 方程 $x_1+x_2+x_3=10$ 有 $\binom{12}{2}$ 个不同的非负整数解.

推论1.3. 方程 $x_1+x_2+\cdots+x_k=n$ $(k\le n)$ 有 $\binom{n-1}{k-1}=\binom{n-1}{n-k}$ 个不同的正整数解.

解. 令 $x_1'=x_1-1$, $x_2'=x_2-1$, $\ldots$, $x_k'=x_k-1$, 方程 $x_1+x_2+\cdots+x_k=n$ $(k\le n)$ 有多少个不同的正整数解转化为方程

$$x_1'+x_2'+\cdots+x_k'=n-k$$

有多少个不同的非负整数解. 根据性质 1.8 上述方程有

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

个不同的正整数解. □

课题练习. 求不等式 $x_1 + x_2 + \cdots + x_k \leq n$ 有多少种不同的非负整数解.

课题练习. 求不等式 $x_1 + x_2 + \cdots + x_k \leq n$ $(k \leq n)$ 有多少种不同的正整数解.

根据整数的有序分解, 有如下结论:

将 n 个球放入 m 个箱子	每个箱子的球数不限	每个箱子的球数 ≤ 1	每个箱子的球数 ≥ 1
相同　　　不同	$\binom{n+m-1}{n}$		$\binom{n-1}{m-1}$

1.4.3 第二类 Stirling 数 (The Stirling Number of the Second Kind)

定义1.4. *将 n 个不同的元素分成 k 个非空的块 (Block), 不同的划分方法数称为* ***第二类 Stirling 数****, 记为 $S(n,k)$.*

这里以 $n = 3$ 为例, 假设集合 $A = \{1,2,3\}$, 则有

- 若分成 $k = 1$ 个非空的块, 则有 $\{1,2,3\}$, 因此 $S(3,1) = 1$;
- 若分成 $k = 2$ 个非空的块, 则有 $\{\{1\},\{2,3\}\}$, $\{\{2\},\{1,3\}\}$, $\{\{3\},\{1,2\}\}$, 因此 $S(3,2) = 3$;
- 若分成 $k = 3$ 个非空的块, 则有 $\{\{1\},\{2\},\{3\}\}$, 因此 $S(3,3) = 1$;

针对更一般的情况, 我们有如下性质:

性质1.9. *对 $n \geq 1$ 有 $S(n,n) = S(n,1) = 1$, 以及递推关系*

$$S(n,k) = kS(n-1,k) + S(n-1,k-1) \quad (1 < k < n).$$

Proof. 根据块的定义可知

$$S(n,n) = S(n,1) = 1.$$

对任意 $1 < k < n$, 将 n 个不同的元素集合 $A = [n]$ 分成 k 个非空的块, 有 $S(n,k)$ 种不同的划分数, 这些划分可分如下两种情况:

- 若最后一个元素 n 单独为一个块 $\{n\}$, 则其他元素构成 $k-1$ 块, 有 $S(n-1,k)$ 种不同的划分数;
- 若最后一个元素 n 没有单独成一个块, 则其他剩余元素构成 k 块, 有 $S(n-1,k)$ 中不同的划分数, 再将第 n 个元素放入 k 个块中的一个, 有 $kS(n-1,k)$ 种不同的划分数.

从而完成证明. □

根据上述性质有

推论1.4. 第二类 *Stirling* 数的一般表达式为:

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

根据第二类 Stirling 数, 有

将 n 个球放入 m 个箱子	每个箱子的球数不限	每个箱子的球数 ≤ 1	每个箱子的球数 ≥ 1
不同 不同			$m!S(n,m)$
不同 相同	$\sum_{k=1}^{m} S(n,k)$	$\begin{cases} 1 & n\leq m \\ 0 & n>m \end{cases}$	$S(n,m)$

1.4.4 整数的无序分拆 (Partition)

整数的无序分拆: 将正整数 n 划分(partition)成 k 个部分, 每个部分都是正数, 且这 k 个部分之间无序. 整数 n 划分(partition)成 k 个部分, 有多少种不同的划分数, 记为 $p(n,k)$.

这里以 $n=7$ 为例, 考虑其无序拆分:

$k=1$	$\{7\}$	$p(7,1)=1$
$k=2$	$\{1,6\},\{2,5\},\{3,4\}$	$p(7,2)=3$
$k=3$	$\{1,1,5\},\{1,2,4\},\{1,3,3\},\{2,2,3\}$	$p(7,3)=4$
$k=4$	$\{1,1,1,4\},\{1,1,2,3\},\{1,2,2,2\}$	$p(7,4)=3$
$k=5$	$\{1,1,1,1,3\},\{1,1,1,2,2\}$	$p(7,5)=2$
$k=6$	$\{1,1,1,1,1,2\}$	$p(7,6)=1$
$k=7$	$\{1,1,1,1,1,1,1\}$	$p(7,7)=1$

针对更一般的情况, 我们有如下性质:

性质1.10. 对 $n\geq 1$ 有 $p(n,n)=p(n,1)=1$, 以及递推关系

$$p(n,k)=p(n-1,k-1)+p(n-k,k)\quad (1<k<n).$$

Proof. 根据整数拆分的定义可知

$$p(n,n)=p(n,1)=1.$$

对任意 $1<k<n$, 将正整数 n 划分成 k 个无序的部分, 该问题可转化为方程

$$x_1+x_2+\cdots+x_k=n \quad \text{s.t.} \quad x_1\geq x_2\geq\cdots\geq x_k\geq 1.$$

该方程有 $p(n,k)$ 种不同的划分数, 考虑最后一个划分 x_k 的取值, 可分如下两种情况:

- 若最小部分 $x_k=1$, 则 $(x_1,x_2,\cdots,x_{k-1})$ 是整数 $n-1$ 的 $k-1$ 部分的无序划分, 有 $p(n-1,k-1)$ 种不同的划分数;

- 若最小部分 $x_k > 1$, 则 $(x_1 - 1, x_2 - 1, \cdots, x_k - 1)$ 是整数 $n - k$ 的 k 部分的无序划分, 有 $p(n-k, k)$ 种不同的划分数.

从而完成证明. $\square$

进一步有以下性质:

性质1.11. 对任意正整数 $n > 0$, $k \in [n]$, 有

$$\frac{\binom{n-1}{k-1}}{k!} \leq p(n,k) \leq \frac{\binom{n-1+k(k-1)/2}{k-1}}{k!}.$$

性质1.12. 对任意给定正整数 $k > 0$, 当 n 非常大或 $n \to \infty$ 有

$$p(n,k) \approx \frac{n^{k-1}}{k!(k-1)!}.$$

根据正整数分拆, 有

将 n 个球放入 m 个箱子		每个箱子的球数不限	每个箱子的球数 ≤ 1	每个箱子的球数 ≥ 1
相同	相同	$\sum_{k=1}^{m} p(n,k)$	$\begin{cases} 1 & n \leq m \\ 0 & n > m \end{cases}$	$p(n,m)$