定义2.6. 若事件 A 在一次试验中发生的概率非常小, 但经过多次独立地重复试验, 事件 A 的发生是必然的, 称之为 **小概率原理**.

小概率原理可根据严格的数学推理得到: 若事件 $A_1, A_2, \cdots, A_n, \cdots$ 独立且每事件发生的概率 $P(A_i) = p > 0$ 非常小, 则有

$$P(A_1A_2\cdots A_n) = 1 - P(\bar{A}_1\bar{A}_2\cdots\bar{A}_n) = 1-(1-p)^n \to 1 \quad 当 \quad n\to\infty,$$

即独立重复多次的小概率事件亦可成立必然事件.

若独立事件 $A_1, A_2, \cdots, A_n$ 发生的概率 $P(A_i) = p\ (i\in[n])$, 则 n 个事件中恰有 k 个事件发生的概率为 $\binom{n}{k}p^k(1-p)^{n-k}$.

例2.14. 冷战时期美国的导弹精度 99%, 苏联的导弹精度 60%, 但苏联的导弹数量特别多, 导弹的数量能否弥补精度的不足?

解. 假设每次独立发射 n 枚导弹, 用事件 A_i 表示第 i 枚导弹命中目标, 则 n 枚导弹击中目标的概率为

$$P(A_1\cup A_2\cup\cdots\cup A_n) = 1-(1-0.6)^n \geq 0.99 \quad\Rightarrow\quad n\geq 5,$$

因此每次独立发射 5 枚导弹, 击中目标的概率高于 99%. □

在上例中, 若美国的导弹精度为 90%, 苏联的导弹精度为 70%, 则苏联每次只需独立发射两枚导弹即可达到 91%.

例2.15. 一串电路图: A, B, C, D, E, F, G 是电路元件, 电路元件各自下方的数字表示正常工作的概率. 若各电路元件之间相互独立. 求电路正常工作的概率.

解. 用事件 W 表示电路正常工作, 则有 $W = A\cap B\cap(C\cup D\cup E)\cap(F\cup G)\cap H$. 根据独立性假设有

$$P(W) = P(A)P(B)P(C\cup D\cup E)P(F\cup G)P(H).$$

根据 $P(C\cup D\cup E) = 1-P(\bar{C})P(\bar{D})P(\bar{E}) = 1-(2/3)^3 = 19/27$ 和 $P(F\cup G) = 1-P(\bar{E})P(\bar{G}) = 7/16$, 可得 $P(W) = 133/1800$. □

2.3 案例分析

2.3.1 利用独立性验证大矩阵乘法是否相等

本节研究的问题: 给定矩阵 $A, B, C\in\{0,1\}^{n\times n}$ (n 非常大, 如 $n\geq 10000000$), 验证 $AB = C$ 是否成立? 若直接执行矩阵乘法运算、并验证等式是否成立, 计算复杂度为 $O(n^3)$; 若采用分治法, 计算复杂度为 $O(n^{\log_2^7})$, 目前最好的计算复杂度为 $O(n^{2.37})$. 为进一步降低计算复杂度, 可利用独立性验证 $AB = C$ 是否成立?

独立随机产生一个向量 $r\in\{0,1\}^n$, 判断

$$A(Br) = Cr?$$

计算 $A(Br)$ 和 Cr 的复杂度均为 $O(n^2)$. 若 $A(Br) \neq Cr$ 则直接可得 $AB \neq C$; 若 $A(Br) = Cr$ 并不能得出 $AB = C$. 将上述过程独立进行 K 次, 可以以证明以较大的概率有 $AB = C$ 成立, 该过程被称为 Freivalds 算法.

- -

Freivalds 算法

- -

Input: A, B, C

Output: Yes/No

- -

```
For i = 1 : K
    Select a random vector r = (r_1, r_2, ··· , r_n) with P(r_j = 0) = P(r_j = 1) = 1/2 (j ∈ [n])
    Compute p = A × (Br) − Cr
    If p ≠ 0 then
        Return ‘No’.
    EndIf
EndFor
Return ‘Yes’.
```

- -

首先发现该算法的计算复杂度为 $O(Kn^2)$, 若 K 比较小则显著降低了计算复杂度. 进一步研究算法的有效性, 若返回 ‘No’, 则必然有 $AB \neq C$; 若返回 ‘Yes’, 然而并不一定有 $AB = C$ 成立, 下面研究当算法返回 ‘Yes’ 时 $AB = C$ 成立的概率.

设 $D = AB - C \neq 0$, 则 D 中必存在一些元素不为 0, 不妨令 $d_{11} \neq 0$. 对任意一轮循环, 不妨设随机向量 $r = \{r_1, r_2, \cdots, r_n\}$, 根据返回 ‘Yes’ 可知 $Dr = 0$, 进一步可得向量 Dr 的第一个元素等于 0, 即

$$\sum_{j=1}^{n} d_{1j} r_j = 0 \quad \Longrightarrow \quad r_1 = -\frac{1}{d_{11}} \sum_{j=2}^{n} d_{1j} r_j$$

无论 $r_2, \ldots, r_n$ 取何值, 等式 $\sum_{j=1}^{n} d_{1j} r_j = 0$ 是否成立由 r_1 的值决定. 根据 $P(r_1 = 0) = P(r_1 = 1) = 1/2$ 可知 $\sum_{j=1}^{n} d_{1j} r_j = 0$ 成立的概率不超过 1/2. 因此在 K 轮独立的循环中, 等式 $\sum_{j=1}^{n} d_{1j} r_j = 0$ 成立的概率不超过 $1/2^K$.

取 $K = \log_2 n$, 则算法 Freivalds 计算复杂度为 $O(n^2 \log n)$, 若算法返回 ‘No’, 则 $AB \neq C$; 若返回 ‘Yes’, 则有

$$P(AB = C) > 1 - 1/n,$$

即至少以 $1 - 1/n$ 的概率有 $AB = C$ 成立.

3 离散型随机变量

有些随机试验的结果本身就是数值, 例如, 抛一枚骰子的点数分别为 $1, 2, \cdots, 6$; 国家一年出生的婴儿数分别为 $1, 2, \cdots, n, \cdots$. 有些试验结果可能与数值无关, 但结果可以用数值进行表示, 例如, 抛一枚硬币, 正面朝上用 0 表示, 正面朝下用 1 表示; 流星坠落地球的落脚点用坐标纬度表示. 当试验结果用数值表示时, 可以引入一个变量来表示随机事件, 由此产生随机变量的概念.

将样本空间 Ω 中每个样本点 ω 与一个实数 $X(\omega)$ 相对应, $X(\omega)$ 是 ω 的实值函数, 称实值函数 $X(\omega) : \Omega \to \mathbb{R}$ 为随机变量 (random variable), 简写为 r.v., 一般用大写字母 X, Y, Z 表示. $X(\omega)$ 随样本点 ω 的不同而取不同的值, 例如:

- 抛一枚骰子, 用随机变量 X 表示出现的点数, 则随机变量 $X \in [6]$. 出现的点数不超过 4 的事件可表示为 $\{X \leq 4\}$; 出现偶数点的事件可表示为 $\{X$ 为偶数$\}$.
- 用随机变量 X 表示一盏电灯的寿命, 其取值为$[0, +\infty)$, 电灯寿命不超过 500 的事件可表示为 $\{X \leq 500\}$.

通过随机变量可形式化描述随机现象或随机事件, 从而利用数学工具来研究概率, 例如 $\{X \leq -\infty\}$ 表示不可能事件, 以及 $\{X \leq +\infty\}$ 表示必然事件.

根据随机变量的取值, 可分为离散型随机变量和连续型随机变量. 若随机变量 X 的取值是有限的、或无限可列的, 则称 X 为 **离散型随机变量**; 若随机变量 X 的取值是无限不可列的, 则称 X 为 **非离散型随机变量**. 本章主要研究离散型随机变量.

3.1 离散型随机变量及分布列

离散型随机变量 X 的取值是有限或无限可列的, 不妨假设其取值为 $x_1, x_2, \cdots, x_n, \cdots$, 事件 $\{X = x_k\}$ 的概率记为

$$p_k = P(X = x_k), \quad k = 1, 2, \cdots,$$

称之为随机变量 X 的 **分布列**. 分布列包含了随机变量的取值和概率, 从而完全刻画了离散随机变量的概率属性. 分布列也可以用表格表示

X	x_1	x_2	$\cdots$	x_n	$\cdots$
P	p_1	p_2	$\cdots$	p_n	$\cdots$

根据概率的非负性和完备性有

性质3.1. *设随机变量 X 的分布列 $p_k = P(X = x_k)$ $(k \geq 1)$, 有 $p_k \geq 0$ 和 $\sum_k p_k = 1$.*

下面来看看一些离散随机变量的例子:

例3.1. *设随机变量 X 的分布列 $P(X = k) = c/4^k$ $(k = 0, 1, 2, \cdots)$, 求概率 $P(X = 1)$.*

解. 根据概率的完备性有

$$1=\sum_{k=0}^{\infty}P(X=k)=\sum_{k=0}^{\infty}\frac{c}{4^k}=\frac{4}{3}c,$$

求解得到 $c=3/4$, 进一步有 $P(X=1)=3/16$. □

例3.2. 给定常数 $\lambda>0$, 随机变量 X 的分布列 $p_i=c\lambda^i/i!\ (i\geq 0)$, 求 $P(X>2)$.

解. 根据概率的完备性有

$$1=\sum_{i=0}^{\infty}p_i=c\sum_{i=0}^{\infty}\frac{\lambda^i}{i!}=c\cdot e^{\lambda}$$

从而得到 $c=e^{-\lambda}$, 进一步得到

$$P(X>2)=1-P(X\leq 2)=1-p_0-p_1-p_2=1-e^{-\lambda}(1+\lambda+\lambda^2/2).$$

□

例3.3. 从 $\{1,2,\ldots,10\}$ 中不放回随机任意取 5 个数, 令随机变量 X 表示所取 5 个数中的最大值, 求 X 的分布列.

解. 由题意可知 X 的取值为 $5,6,7,8,9,10$, 且

$$P(X=k)=\binom{k-1}{4}\Big/\binom{10}{5}\quad(k=5,6,\ldots,10).$$

由此可得 X 的分布列表格为

X	5	6	7	8	9	10
P	1/252	5/252	15/252	35/252	70/252	126/252

□

3.2 离散型随机变量的期望和方差

随机变量的取值具有一定的随机性, 我们希望研究随机变量的一些不变量, 用以刻画随机变量的特征, 最常见的特征是期望与方差.

3.2.1 期望

定义3.1. *设离散型随机变量 X 的分布列为 $P(X=x_k)=p_k\ (k=1,2,\cdots)$, 若级数 $\sum_{k=1}^{\infty}p_kx_k$ 绝对收敛, 称级数和为随机变量 X 的期望 (expectation), 记为 $E(X)$, 即*

$$E(X)=\sum_{k=1}^{\infty}p_kx_k,$$

又被称为均值 (mean) 或加权平均 (weighted average).