随机试验 (random experiment)

满足下列 3 个条件的对随机现象的观察、记录或实验被称为随机试验 (简称为试验):

(1) 在相同的条件下可重复进行;

(2) 试验前事先知道所有可能出现的结果;

(3) 每次试验会出现哪一结果是无法预知的.

例如: 抛一枚质量均匀的硬币, 观察正反面; 对听课人数进行一次登记.

1.1.2 概率的统计定义

在相同的条件下重复做 N 次试验, 各次试验互不影响. 考察事件 A 出现的次数 (频数) n, 称

$$F_N(A) = \frac{n}{N}$$

为 A 在 N 次试验中出现的**频率** (frequency).

频率的稳定性: 人们发现, 当 N 很大时, 频率会呈现某种稳定性, 即在某常数附近摆动. ($N \to \infty$ 时, $F_N(A) \to c$)

注 “频率的稳定性”最早由瑞士数学家伯努利 (Bernoulli) 正式指出并证明 (见第四章).

频率的性质:

(1) 非负性: $F_N(A) \geqslant 0$;

(2) 规范性: 对必然事件 Ω, $F_N(\Omega) = 1$;

(3) 可加性: 若 A 与 B 是两个不会同时发生的事件, 以 $A + B$ 记 “A 与 B 至少有一个发生” 这一事件, 则

$$F_N(A + B) = F_N(A) + F_N(B).$$

(可推广到任意有限个事件)

注 由概率的统计定义, 可知 (统计) 概率也具有非负性、规范性和可加性.

1.2.1 样本空间与样本点

样本点 (sample point): 随机试验的每一可能结果, 记为 ω. 也称样本点为**基本事件**.

样本空间 (sample space): 样本点的全体, 即随机试验的所有可能结果, 记为 Ω.

注 即使对同一试验, 若试验目的不同, 样本空间也可能不同. 讨论问题前需事先取定样本空间.

例 3 杭州今天发生交通事故的次数.

$$\omega_1 = \{0\},\ \ \omega_2 = \{1\},\ \ \omega_3 = \{2\},\ \ \cdots$$

$$\Omega = \{0, 1, 2, \cdots\}.$$

例 4 发射一枚炮弹, 考察落地点与目标之间的距离.

$$\Omega = \{\omega : 0 \leqslant \omega \leqslant a\} = [0, a].$$

1.2.2 古典概型 (classical probability model)

1. 定义及计算

定义（古典概型/拉普拉斯试验）

满足下列两个条件的试验模型被称为古典概型,

(1) 样本空间是有限的, 即 $\Omega = \{\omega_1, \omega_2, \cdots, \omega_n\}$;　(有限性)

(2) 各样本点的出现是等可能的.　(等可能性)

注　一项试验的全部结果是否具有等可能性是无法确切证明的, 这个概念本身就不具备一种可供证明 (或证伪) 的清晰的含义, 它更多是从感觉上去理解.

古典概率 (classical probability) 的定义:

定义（古典概率） 古典概型的条件

设一试验有 n 个等可能发生的样本点，而事件 A 包含其中的 m 个样本点，则事件 A 发生的概率 $P(A)$ 定义为

$$P(A) = \frac{m}{n} = \frac{A\text{ 包含的样本点个数}}{\text{样本空间中的样本点总数}} =: \frac{\#A}{\#\Omega}.$$

古典概率的性质:

(1) 非负性: $\mathsf{P}(A) \geqslant 0$;

(2) 规范性: 对必然事件 Ω, $\mathsf{P}(\Omega) = 1$;

(3) 可加性: 若 A 与 B 是两个不会同时发生的事件, 以 $A + B$ 记 "A 与 B 至少有一个发生" 这一事件, 则

$$\mathsf{P}(A + B) = \mathsf{P}(A) + \mathsf{P}(B).$$

(可推广到任意有限个事件)

注 如果 A 和 B 两个事件有且只有一个发生 (即 A 与 B 不会同时发生, 但至少发生一个), 则 $\mathsf{P}(A) + \mathsf{P}(B) = 1$.

古典概型计算中常用的计数方法:

- 从 n 个不同的元素中有放回地每次抽取一个, 依次抽取 m 个排成一列, 可以得到 n^m 个不同的排列. 当随机抽取时, 得到的不同排列是等可能发生的.
- 从 n 个不同的元素中 (不放回地) 抽取 m 个元素排成一列时, 可以得到 $A_n^m = \frac{n!}{(n-m)!}$ 个不同的排列. 当随机地抽取和排列时, 得到的不同排列是等可能发生的.
- 从 n 个不同的元素中 (不放回地) 抽取 m 个元素, 不论次序地组成一组, 可以得到 $C_n^m = \frac{n!}{m!(n-m)!}$ 个不同的组合. 当随机抽取时, 得到的不同组合是等可能发生的.
- 将 n 个不同的元素分成有次序的 k 组, 不考虑每组中元素的次序, 第 i $(1 \leqslant i \leqslant k)$ 组恰有 n_i $(1 \leqslant i \leqslant k)$ 个不同的元素的结果数是

$$C_n^{n_1} C_{n-n_1}^{n_2} \cdots C_{n-n_1-\cdots-n_{k-1}}^{n_k} = \frac{n!}{n_1! n_2! \cdots n_k!}.$$

注意到

$$\frac{A_N^n}{N^n} = \frac{N(N-1)\cdots(N-n+1)}{N^n} = \left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{n-1}{N}\right),$$

$\ln(1-x) = -x + O(x^2)$, $x \to 0$. 因此,

$$\begin{aligned}
&\ln\left[\left(1-\frac{1}{N}\right)\left(1-\frac{2}{N}\right)\cdots\left(1-\frac{n-1}{N}\right)\right] \\
=&\sum_{k=1}^{n-1}\ln\left(1-\frac{k}{N}\right) = -\sum_{k=1}^{n-1}\frac{k}{N} + O\left(\sum_{k=1}^{n-1}\frac{k^2}{N^2}\right) \\
=&-\frac{n(n-1)}{2N} + O\left(\frac{n^3}{N^2}\right).
\end{aligned}$$

若 N^2 比 n^3 大得多, 则有近似计算公式:

$$\mathsf{P}(B) = \frac{A_N^n}{N^n} \approx \exp\left\{-\frac{n(n-1)}{2N}\right\}.$$

3. 古典概型的推广

古典概型要求样本空间是有限的, 且各样本点的出现是等可能的.

一般地, 如果样本空间 $\Omega = \{\omega_1, \omega_2, \cdots\}$ 含有可列个样本点, ω_i 出现的可能性为 $p(\omega_i)$, 其中 $p(\omega_i) \geqslant 0$, $\sum_{i=1}^{\infty} p(\omega_i) = 1$, 则事件 A 发生的概率定义为

$$\mathsf{P}(A) = \sum_{\omega_i \in A} p(\omega_i).$$

注 这样定义的 $\mathsf{P}(\cdot)$ 仍满足非负性、规范性和可加性.

1.2.3 几何概型

若样本空间中有不可列个样本点, 则不能利用古典概型求概率. 但类似的算法可以进行推广.

若样本空间是一个包含无限个点的区域 Ω (一维, 二维, $\cdots$, n 维), 样本点是区域中的一个点, 此时用点数度量样本点的多少就毫无意义.

"等可能性" 可理解为: 对任意两个区域, 当它们的测度 (长度, 面积, 体积, $\cdots$) 相等时, 样本点落在这两区域上的可能性相等, 而与形状和位置都无关.

几何概率 (geometric probability) 的定义:

定义（几何概率）

假设上述“等可能性”成立. 记

$$A_g = \{\omega : \omega \in g \subset \Omega\},$$

则定义 A_g 发生的概率为

$$P(A_g) = \frac{g \text{ 的测度}}{\Omega \text{ 的测度}} =: \frac{m(g)}{m(\Omega)}.$$

几何概率的性质:

(1) 非负性: $\mathsf{P}(A_g) \geqslant 0$;

(2) 规范性: 对必然事件 A_Ω, $\mathsf{P}(A_\Omega) = 1$;

(3) 可加性: 若 A_{g_1} 与 A_{g_2} 是两个不会同时发生的事件, 以 $A_{g_1} + A_{g_2}$ 记 "A_{g_1} 与 A_{g_2} 至少有一个发生" 这一事件, 则

$$\mathsf{P}(A_{g_1} + A_{g_2}) = \mathsf{P}(A_{g_1}) + \mathsf{P}(A_{g_2}).$$

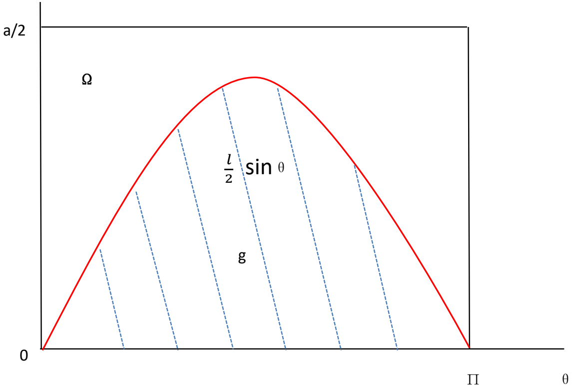
(可推广到任意有限个事件)

样本空间为

$$\Omega = \{(\theta, x) : 0 \leqslant \theta \leqslant \pi,\ 0 \leqslant x \leqslant a/2\}.$$

针与平行线相交的充要条件是 $x \leqslant \frac{l}{2}\sin\theta$. 记 $E = \{$针与平行线相交$\}$, 则 E 在样本空间中所对应的区域是

$$g = \left\{(\theta, x) : x \leqslant \frac{l}{2}\sin\theta\right\}.$$



概率的统计定义属于描述性定义. 概率的古典定义和几何定义都建立在"等可能性"基础之上, 只适用于某些特定的概率模型; 同时, 对于"等可能性"的不同理解可能会导致不同的答案.

例如: 有一家锯木厂, 它会把木头切成不同的木方, 木方的截面都是正方形, 边长会在 $1 \sim 3$ 尺之间随机浮动. 那么根据几何概率, 该锯木厂生产出边长在 $1 \sim 2$ 尺之间的木方的概率有多大? 如果只认同长度上的等可能性假设, 那么所求概率为 $\frac{2-1}{3-1} = \frac{1}{2}$; 如果只认同面积上的等可能性假设, 那么所求概率为 $\frac{2^2-1^2}{3^2-1^2} = \frac{3}{8}$; 如果只认同体积上的等可能性假设, 那么所求概率为 $\frac{2^3-1^3}{3^3-1^3} = \frac{7}{26}$.

历史上有个著名的类似悖论, 叫贝特朗 (Bertrand) 悖论, 它促使数学家们开始认真思考/反思概率的定义.

1.3.1 事件

在随机试验中, 人们除了对基本结果 (样本点) 感兴趣外, 往往还对某些复杂的结果感兴趣. 例如, 口袋中装有 10 个球: 3 个红球, 3 个白球以及 4 个黑球, 对它们进行编号 $1, 2, \cdots, 10$, 任取一球. 除了关心 10 个基本结果 $\omega_1, \omega_2, \cdots, \omega_{10}$ 外, 往往还关心类似下列的结果:

- $A = \{$取得红球或白球$\} = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\}$;
- $B = \{$取得球的号码小于 $5\} = \{\omega_1, \omega_2, \omega_3, \omega_4\}$;
- $C = \{$没有取得红球$\} = \{\omega_4, \omega_5, \omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}\}$.

A, B, C 都是样本空间的子集.

人们把样本空间 Ω 的子集定义为事件, 用大写英文字母 A, B, C 等表示.
优点: 可以用集合论的方法来探讨事件.

如果在一次试验中某样本点 ω 出现, 而 $\omega \in A$, 则称事件 A 发生.

全集 Ω 以及空集 ϕ 都是 Ω 的子集, 因此它们都是事件.

因为在每次试验中必然出现 Ω 中的一个样本点, 即 Ω 必然发生, 所以称 Ω 为**必然事件**.

因为在每次试验中 ϕ 必定不发生, 因此称它为**不可能事件**.

事件的关系:

1. **包含**: 若事件 A 的发生能导致事件 B 的发生 (或者说, 若 $\omega \in A$ 则 $\omega \in B$), 则称事件 A 包含于事件 B (或者说 B 包含 A), 记作 $A \subset B$ (或 $B \supset A$).

2. **相等**: 若事件 A 的发生能导致事件 B 的发生, 且事件 B 的发生能导致事件 A 的发生 (即 $A \subset B$ 且 $B \subset A$), 则称事件 A 和事件 B 相等, 记作 $A = B$.

3. **和 (并) 事件**: 称 A 与 B 至少有一个发生 (或者说: A 发生或 B 发生) 这一事件为 A 与 B 的和 (并) 事件, 记作 $A \cup B$.

4. **积 (交) 事件**: 称 A 与 B 同时发生 (或者说: A 发生且 B 发生) 这一事件为 A 与 B 的积 (交) 事件, 记作 $A \cap B$ (或者 AB).

5. **差事件**: 称 A 发生而 (且) B 不发生这一事件为 A 与 B 的差事件, 记作 $A\backslash B$. 如果 $B \subset A$, 则 $A\backslash B$ 可写作 $A - B$.

6. **互不相容事件**: 若 A 与 B 不能同时发生, 即 $AB = \phi$, 则称 A 与 B 为互不相容事件. 这时, $A \cup B$ 可写作 $A + B$.

7. **逆 (对立, 余) 事件**: 若 A 与 B 有且只有一个发生, 即 $AB = \phi$ 且 $A \cup B = \Omega$, 则称 A 与 B 互为逆 (对立, 余) 事件, 记作 $B = \overline{A} = A^c$, $A = \overline{B} = B^c$. $(A\backslash B = A\overline{B} = A - AB)$

事件的运算:

1. **交换律**: $A \cup B = B \cup A$, $A \cap B = B \cap A$.

2. **结合律**: $(A \cup B) \cup C = A \cup (B \cup C)$, $(AB)C = A(BC)$.

3. **分配律**: $(A \cup B) \cap C = AC \cup BC$,
   $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$.

4. **德摩根 (De Morgan) 律**: $\overline{A \cup B} = \overline{A} \cap \overline{B}$, $\overline{A \cap B} = \overline{A} \cup \overline{B}$.

要求: 集合论的写法和事件的语言能互相翻译, 能把复杂事件分解成简单事件.

例 1 若 A, B, C 是 3 个事件, 则

$\{A$ 与 B 都发生而 C 不发生$\}=AB\overline{C} = AB \backslash C = AB - ABC$.

$\{A, B, C$ 都发生$\}=ABC$.

$\{$三事件恰好发生一个$\}=A\,\overline{B}\,\overline{C} + \overline{A}\,B\,\overline{C} + \overline{A}\,\overline{B}\,C$.

$\{$三事件恰好发生两个$\}=AB\overline{C} + A\overline{B}C + \overline{A}BC$.

$\{$三事件至少发生一个$\}=A \cup B \cup C$.

1.3.2 概率空间

概率空间由三个要素组成:

(1) 样本空间 Ω;

(2) 事件域 $\mathscr{F}$ (σ-域, σ-代数);

(3) 概率 (测度) P.

在几何概型中, 若区域 g 的测度不存在, 则无法定义事件 A_g 的概率. 因此, 我们并不是对所有的事件感兴趣.

事件域 $\mathscr{F}$ 是指 Ω 中满足下列条件的子集所组成的集类:

(1) $\Omega \in \mathscr{F}$;

(2) 若 $A \in \mathscr{F}$, 则 $\overline{A} \in \mathscr{F}$;

(3) 若 $A_1, A_2, \cdots, A_n, \cdots \in \mathscr{F}$, 则 $\bigcup_{n=1}^{\infty} A_n \in \mathscr{F}$.

事件域的性质:

(1) $\phi \in \mathscr{F}$;

则 $\bar{A}_1, \bar{A}_2, \dots, \bar{A}_n, \dots \in F$，则 $\bar{A}_1 \cup \dots \cup \bar{A}_n \cup \dots \in F$

则 $\overline{\bigcup_{n=1}^{\infty} \bar{A}_n} \in F$

(2) 若 $A_1, A_2, \cdots, A_n, \cdots \in \mathscr{F}$, 则 $\bigcap_{n=1}^{\infty} A_n \in \mathscr{F}$;

(因为 $\bigcap_{n=1}^{\infty} A_n = \overline{\overline{\bigcap_{n=1}^{\infty} A_n}} = \overline{\bigcup_{n=1}^{\infty} \overline{A}_n}$)

$\overline{\bigcap_{n=1}^{\infty} A_n} = \bigcup_{n=1}^{\infty} \bar{A}_n$

(3) 若 $A_1, A_2, \cdots, A_n \in \mathscr{F}$, 则 $\bigcup_{k=1}^{n} A_k \in \mathscr{F}$, $\bigcap_{k=1}^{n} A_k \in \mathscr{F}$.

令 $A_{n+1} = A_{n+2} = \dots = \phi$

因此 $\bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{\infty} A_k \in \mathscr{F}$

事件域的选取:

1. 若只对 Ω 的一个子集 A 感兴趣, 则通常取

$$\mathscr{F} = \{A, \overline{A}, \phi, \Omega\}.$$

2. 若样本空间由有限个或者可列个样本点组成, 则常取 Ω 的一切子集所构成的集类作为 $\mathscr{F}$.

3. 若 $\Omega = \mathbb{R} = \mathbb{R}^1$ (一维实数全体), 则取一切左开右闭有界区间和它们的 (有限和可列) 并运算、(有限和可列) 交运算、补运算所构成的集类作为 $\mathscr{F}$ (一维博雷尔 (Borel) σ-域).

注 若 x, y 表示任意实数, 由于

$$\{x\} = \bigcap_{n=1}^{\infty} (x - \frac{1}{n}, x],$$

$$(x, y) = (x, y] - \{y\},$$

$$[x, y] = (x, y] + \{x\},$$

$$[x, y) = \{x\} + (x, y),$$

因此 $\mathscr{F}$ 包含一切开区间、闭区间、半开半闭区间、单个实数、可列个实数, 以及由它们经 (有限或可列) 并、(有限或可列) 交、补运算而得到的集合. 这是足够大的一个集类, 能把实际问题中感兴趣的点集都包含在内.

4. 若 $\Omega=\mathbb{R}^n$ (n 维实数全体), 则取一切左开右闭有界 n 维矩形和它们的 (有限和可列) 并运算、(有限和可列) 交运算、补运算所构成的集类作为 $\mathscr{F}$ (n 维博雷尔 (Borel) σ-域).

定义（概率, Kolmogorov, 1933）

概率 (测度) 是定义在 $\mathscr{F}$ 上的实值集函数: $A\ (\in \mathscr{F}) \mapsto P(A)$, 且满足下列三个条件:

(1) 非负性: 对任一 $A \in \mathscr{F}$, $P(A) \geqslant 0$;

(2) 规范性: $P(\Omega) = 1$;

(3) 可列可加性: 若 $\{A_n, n \geqslant 1\}$ 是 $\mathscr{F}$ 中两两互不相容事件序列, 则

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

注 若无特别声明, 今后遇到的事件都假定来自事件域 $\mathscr{F}$, 使得它有概率.

概率空间的构造:

1. 从一批产品里任取一个产品, 若只关心它是否是正品, 则记 $A=$ {产品为正品}. 构造 $\Omega=\{A,\overline{A}\}$, $\mathscr{F}=\{A,\overline{A},\phi,\Omega\}$. 再赋予 $\mathscr{F}$ 中各事件以概率:

$$\mathsf{P}(A)=p\in[0,1],\quad \mathsf{P}(\overline{A})=1-p,\quad \mathsf{P}(\phi)=0,\quad \mathsf{P}(\Omega)=1.$$

则 P 为概率测度, $(\Omega,\mathscr{F},\mathsf{P})$ 为概率空间.

2. 令 $\Omega=\{\omega_1,\omega_2,\cdots\}$, $\mathscr{F}$ 为 Ω 的所有子集构成的集类. 赋予概率: 令 $\mathsf{P}(\{\omega_i\})=p_i\in(0,1)$, 它满足 $\sum_{i=1}^{\infty}p_i=1$. 再对来自 $\mathscr{F}$ 的事件 A, 定义

$$\mathsf{P}(A)=\sum_{i:\omega_i\in A}p_i.$$

则 P 为概率测度, $(\Omega,\mathscr{F},\mathsf{P})$ 为概率空间. 特别地, 若 Ω 是有限集, 且各 p_i 全相等, 则 $(\Omega,\mathscr{F},\mathsf{P})$ 为古典概率空间.

3. 令 $\Omega=[0,1]$, $\mathscr{F}=\mathscr{B}_{[0,1]}$ ($[0,1]$ 上的博雷尔 σ-域), m 为勒贝格 (Lebesgue) 测度 (它定义 $m((a,b])=b-a$, 且要求 m 满足非负性和可列可加性, 其中 $(a,b]\subset[0,1]$, $a<b$). 则 m 是概率测度, $(\Omega,\mathscr{F},m)$ 是概率空间, 它其实是 $[0,1]$ 上的几何概率空间.

概率的性质:

1. $\mathsf{P}(\phi) = 0.$

证明 因为

$$\mathsf{P}(\Omega) = \mathsf{P}(\Omega) + \mathsf{P}(\phi) + \mathsf{P}(\phi) + \cdots,$$

所以两边消去 $\mathsf{P}(\Omega)$ 得 $\mathsf{P}(\phi) = 0.$

2. 有限可加性: 若 $A_iA_j = \phi$, $i, j = 1, 2, \cdots, n; i \neq j$, 则

$$\mathsf{P}(\sum_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathsf{P}(A_i).$$

证明 在公理化定义的可列可加性中令 $A_i = \phi, i > n$, 再结合性质 1 可得有限可加性.

3. $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$.

证明 利用有限可加性以及 $\Omega = A \cup \overline{A}$ 可证此性质.

4. 若 $B \subset A$, 则 $\mathsf{P}(A \backslash B) = \mathsf{P}(A - B) = \mathsf{P}(A) - \mathsf{P}(B)$. (可减性)

证明 因为 $A = B \cup (A - B)$ 且 B 与 $A - B$ 互不相容, 所以

$$\mathsf{P}(A) = \mathsf{P}(B) + \mathsf{P}(A - B).$$

推论 1: 若 $B \subset A$, 则 $\mathsf{P}(B) \leqslant \mathsf{P}(A)$. (单调性)

推论 2: 对任何 $B \in \mathscr{F}$, 有 $\mathsf{P}(B) \leqslant 1$. (有界性)

5. $\mathsf{P}(A \backslash B) = \mathsf{P}(A) - \mathsf{P}(AB)$.

证明 因为 $A \backslash B = A \backslash (AB)$ 且 $AB \subset A$, 再利用性质 4, 可得此性质.

6. 加法公式: $\mathsf{P}(A\cup B)=\mathsf{P}(A)+\mathsf{P}(B)-\mathsf{P}(AB)$.

证明　因为 $A\cup B=A\cup(B\backslash A)$, A 与 $B\backslash A$ 互不相容, 所以

$$\mathsf{P}(A\cup B)=\mathsf{P}(A)+\mathsf{P}(B\backslash A)=\mathsf{P}(A)+\mathsf{P}(B)-\mathsf{P}(AB).$$

7. 多还少补 (exclusion-inclusion) 公式/定理:

$$\begin{aligned}\mathsf{P}\left(\bigcup_{i=1}^{n}A_i\right)=&\sum_{i=1}^{n}\mathsf{P}(A_i)-\sum_{1\leqslant i<j\leqslant n}\mathsf{P}(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}\mathsf{P}(A_iA_jA_k)\\&-\cdots+(-1)^{n-1}\mathsf{P}(A_1A_2\cdots A_n).\end{aligned}$$

证明　数学归纳法 (略). (若记 $p_k=\sum_{1\leqslant i_1<i_2<\cdots<i_k\leqslant n}\mathsf{P}(A_{i_1}A_{i_2}\cdots A_{i_k})$, 则上述公式可写成 $\mathsf{P}(\bigcup_{i=1}^{n}A_i)=\sum_{k=1}^{n}(-1)^{k-1}p_k$, 称为 Jordan 公式)

8. 次可加性: $\mathrm{P}(\bigcup_{i=1}^{\infty} A_i) \leqslant \sum_{i=1}^{\infty} \mathrm{P}(A_i)$. ($\mathrm{P}(\bigcup_{i=1}^{n} A_i) \leqslant \sum_{i=1}^{n} \mathrm{P}(A_i)$ 也成立)

证明 令

$$B_1 = A_1, \quad B_2 = \overline{A}_1 A_2, \quad \cdots, \quad B_n = \overline{A}_1 \cdots \overline{A}_{n-1} A_n, \quad \cdots,$$

$= A_2 \setminus A_1$ $\quad$ $= A_n \setminus (A_1 \cup \dots \cup A_{n-1})$

则 $\{B_i, i \geqslant 1\}$ 两两互不相容，且

$$B_i \subset A_i, \quad \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i.$$

所以由可列可加性得

单调性

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

利用概率的定义和性质求“复杂”事件的概率:

例 3 某城有 N 辆卡车, 车牌号为 $1, 2, \cdots, N$. 有一个外地人到该城去, 把遇到的 n 辆车子的牌号抄下 (可能重复抄到某些车牌号), 求抄到的最大号码正好为 k 的概率.

解 看成是古典概型问题: 设每辆卡车被遇到的可能性相等, 把抄到的 n 个号码排成一列, 看成样本点. 记 A_k 为抄到的最大号码为 k 这一事件, B_k 为抄到的最大号码不超过 k 这一事件. 则

$$B_{k-1} \subset B_k, \quad A_k = B_k - B_{k-1}.$$

所以, $\mathsf{P}(B_k) = k^n / N^n$,

$$\mathsf{P}(A_k) = \mathsf{P}(B_k) - \mathsf{P}(B_{k-1}) = \frac{k^n - (k-1)^n}{N^n}, \quad k = 1, 2, \cdots, N.$$

例 4 (匹配问题) 某班有 n 个士兵, 每人各有一支枪, 这些枪外形完全一致. 在一次夜里紧急集合中, 每人随机取一支枪. 求至少有一人拿到自己的枪的概率.

解 记 $A_i = \{$第 i 个人拿到自己的枪$\}$, $i = 1, 2, \cdots, n$. 则所求概率为

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

因为

$$= \frac{1 \times (n-2)!}{n!} = \frac{1}{n(n-1)}$$

$$\mathsf{P}(A_i) = \frac{1}{n}, \quad \mathsf{P}(A_iA_j) = \frac{1}{n(n-1)}, \quad \cdots, \quad \mathsf{P}(A_1A_2\cdots A_n) = \frac{1}{n!},$$

共 C_n^2 项 $(i<j)$

所以由多还少补公式得

$$
\begin{aligned}
p_n :=& \mathsf{P}\left(\bigcup_{i=1}^{n} A_i\right) \\
=& n \cdot \frac{1}{n} - C_n^2 \cdot \frac{1}{n(n-1)} \\
& + C_n^3 \cdot \frac{1}{n(n-1)(n-2)} - \cdots + (-1)^{n-1} \cdot \frac{1}{n!} \\
=& 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1} \frac{1}{n!}.
\end{aligned}
$$

可以看出

$$\lim_{n\to\infty} p_n = 1 - e^{-1}.$$

例 5　证明下面的概率不等式:

$$\mathsf{P}\left(\bigcap_{i=1}^{n} A_i\right) \geqslant \sum_{i=1}^{n} \mathsf{P}(A_i) - (n-1).$$

证明　利用概率的次可加性,

$$\begin{aligned}
\mathsf{P}\left(\bigcap_{i=1}^{n} A_i\right) &= 1 - \mathsf{P}\left(\bigcup_{i=1}^{n} \overline{A}_i\right) \geqslant 1 - \sum_{i=1}^{n} \mathsf{P}(\overline{A}_i) \\
&= 1 - \left[n - \sum_{i=1}^{n} \mathsf{P}(A_i)\right] \\
&= \sum_{i=1}^{n} \mathsf{P}(A_i) - (n-1).
\end{aligned}$$

例 6 (握手问题) 某次大型会议, 其中参会者 (超过 5 人) 每人都至少与比例为 p 的人群握过手 (不包括自己), 其中 $p > 2/3$. 证明: 一定能找到其中的 4 个人, 他们互相都握过手.

证明 把参会者编号为 $1, 2, \cdots, n$, 其中 n 为参会者总数. 考虑以 $\Omega = \{1, 2, \cdots, n\}$ 为样本空间的古典概型. 记 A_i 为与第 i 号参会者握过手的参会者的编号集合, 则

$$\mathsf{P}(A_i) \geqslant p > 2/3, \quad \forall i \geqslant 1.$$

取 $a \in A_1$, 根据例 5 的概率不等式, 有

$$\mathsf{P}(A_1 A_a) \geqslant 2p - 1 > 0,$$

所以存在 $b \in A_1 A_a$. 再次利用例 5 的概率不等式, 有

$$\mathsf{P}(A_1 A_a A_b) \geqslant 3p - 2 > 0,$$

所以存在 $c \in A_1 A_a A_b$. 可以看出: 编号为 $1, a, b, c$ 的四位参会者, 他们互相都握过手.

1.3.3 概率测度的连续性

给定一概率空间 $(\Omega, \mathscr{F}, \mathsf{P})$, 假定 $\{A_n \in \mathscr{F}, n \geqslant 1\}$ 是一列单调增加的事件序列, 即

$$A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots,$$

记 $A = \bigcup_{n=1}^{\infty} A_n$, 称 A 为 A_n 的极限.

若 $\{A_n \in \mathscr{F}, n \geqslant 1\}$ 是一列单调减少的事件序列, 即

$$A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots,$$

记 $A = \bigcap_{n=1}^{\infty} A_n$, 称 A 为 A_n 的极限.

从事件域的定义可以看出, A 仍然是一来自 $\mathscr{F}$ 的事件, 如何求 A 的概率?

定理

如果 $\{A_n \in \mathscr{F}, n \geqslant 1\}$ 是一列单调增加 (或减少) 的事件序列, 具有极限 A, 那么

$$P(A) = \lim_{n\to\infty} P(A_n).$$

证明 只证单调增加情形, 单调减少情形类似可证. 令 $A_0 = \phi$, 且

$$B_k = A_k - A_{k-1}, \quad k \geqslant 1.$$

则

$$A = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n.$$

由可列可加性,

$$\mathsf{P}(A) = \mathsf{P}(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mathsf{P}(B_n) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathsf{P}(B_k).$$

又因为 $\mathsf{P}(B_k) = \mathsf{P}(A_k) - \mathsf{P}(A_{k-1})$, 所以

$$\mathsf{P}(A) = \lim_{n\to\infty} \sum_{k=1}^{n} [\mathsf{P}(A_k) - \mathsf{P}(A_{k-1})] = \lim_{n\to\infty} \mathsf{P}(A_n).$$

证毕.

对单调增加的事件序列 $\{A_n \in \mathscr{F}, n \geqslant 1\}$, 记

$$\lim_{n\to\infty} A_n = A = \bigcup_{n=1}^{\infty} A_n,$$

那么上述定理说明

$$\mathsf{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mathsf{P}(A_n),$$

称之为概率 (测度) 的下连续性.

对单调减少的事件序列 $\{A_n \in \mathscr{F}, n \geqslant 1\}$, 记

$$\lim_{n\to\infty} A_n = A = \bigcap_{n=1}^{\infty} A_n,$$

那么上述定理说明

$$\mathsf{P}(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} \mathsf{P}(A_n),$$

称之为概率 (测度) 的上连续性.

*一般地, 对一列事件 $\{A_n \in \mathscr{F}, n \geqslant 1\}$, 记

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m, \quad \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m.$$

容易验证 $\liminf\limits_{n\to\infty} A_n \subset \limsup\limits_{n\to\infty} A_n$. 如果 $\liminf\limits_{n\to\infty} A_n = \limsup\limits_{n\to\infty} A_n$, 则称 A_n 的极限存在, 并记作 $\lim\limits_{n\to\infty} A_n$. 若极限 $\lim\limits_{n\to\infty} A_n$ 存在, 那么同样有

$$\mathsf{P}\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} \mathsf{P}(A_n).$$

例 7 独立投掷一枚均匀硬币无穷多次，证明：一次正面都没出现的概率为 0.

证明 令 A_n 表示前 n 次投掷中至少出现一次正面，则

$$A_n \subset A_{n+1}, \quad n \geqslant 1.$$

记

$$A = \bigcup_{n=1}^{\infty} A_n,$$

则

$$A = \{\text{正面最终会出现}\}.$$

所求概率为 $\mathsf{P}(\overline{A})$. 注意到

$$\mathsf{P}(A) = \lim_{n\to\infty} \mathsf{P}(A_n) = \lim_{n\to\infty} \frac{2^n - 1}{2^n} = 1,$$

所证结果成立.

1.4.1 条件概率 (conditional probability)

取定概率空间后, A 发生的概率为 $\mathsf{P}(A)$. 若事件 B 发生, 在这一前提下 A 发生的概率就不一定是 $\mathsf{P}(A)$了, 因为样本空间发生了变化. 其实这时样本空间已经被压缩 (变小) 了, 只由 B 中的样本点所组成. (在新的试验条件下, B 成为样本空间)

记 "在 B 发生的条件下 A 发生的概率" 为 $\mathsf{P}(A|B)$.

$\mathsf{P}(A)$ 和 $\mathsf{P}(A|B)$ 是在两个不同的概率空间里对事件 A 的概率度量, 它们通常不相等.

事实上, 在古典概型下, 当 $\mathsf{P}(B) \neq 0$ 时,

$$\begin{aligned}
\mathsf{P}(A|B) &= \frac{\text{在 } B \text{ 发生的条件下 } A \text{ 包含的样本点数}}{\text{在 } B \text{ 发生的条件下样本点总数}} \\
&= \frac{AB \text{ 包含的样本点数}}{B \text{ 包含的样本点数}} \\
&= \frac{AB \text{ 发包含的样本点数}/\#\Omega}{B \text{ 包含的样本点数}/\#\Omega} \\
&= \frac{\mathsf{P}(AB)}{\mathsf{P}(B)}.
\end{aligned}$$

注　容易看出上述等式在几何概型下也成立.

定义（条件概率）

对事件 A 和 B, 若 $P(B) \neq 0$, 则 “在 B 发生的条件下 A 发生的概率” 记作 $P(A|B)$, 定义为

$$P(A|B) = \frac{P(AB)}{P(B)}.$$

容易验证: 这样定义的条件概率满足非负性、规范性和可列可加性. 即条件概率具有概率的一切性质.

概率的乘法公式:

- 若 $\mathsf{P}(B) \neq 0$, 则 $\mathsf{P}(AB) = \mathsf{P}(B)\mathsf{P}(A|B)$;
- 若 $\mathsf{P}(A) \neq 0$, 则 $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B|A)$;
- 若 $\mathsf{P}(A_1A_2\cdots A_{n-1}) \neq 0$, 则

$$\begin{aligned} &\mathsf{P}(A_1A_2\cdots A_n) \\ =&\mathsf{P}(A_1)\mathsf{P}(A_2|A_1)\mathsf{P}(A_3|A_1A_2)\cdots\mathsf{P}(A_n|A_1A_2\cdots A_{n-1}). \end{aligned}$$

1.4.2 全概率 (total probability) 公式, 贝叶斯 (Bayes) 公式

先介绍一个概念: **完备事件组** (或叫 Ω 的分割、分划、划分)

定义 (完备事件组)

若事件组 $\{A_i, i=1,2,\cdots,n\}$ $(n \leqslant \infty)$ 满足下列两个条件:

(1) $\{A_i, i=1,2,\cdots,n\}$ 两两互不相容, 且 $P(A_i) > 0$; (不交)

(2) $\sum_{i=1}^{n} A_i = \Omega$, (不漏)

则称 $\{A_i, i=1,2,\cdots,n\}$ 是 Ω 的一个完备事件组.

注 最简单的完备事件组为: $\{A, \overline{A}\}$, 其中 $\mathrm{P}(A) \in (0,1)$.

定理 (全概率公式)

若 $\{A_i, i=1,2,\cdots,n\}$ $(n \leqslant \infty)$ 是 Ω 的一个完备事件组, 则对任意事件 B,

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

证明 由 $A_iB \subset A_i$ 知 $\{A_iB, i=1,2,\cdots,n\}$ 是两两互不相容的, 所以由有限可加性或可列可加性得

$$\begin{aligned}\mathrm{P}(B) &= \mathrm{P}(B\Omega) = \mathrm{P}\left(B\sum_{i=1}^{n} A_i\right) = \mathrm{P}\left(\sum_{i=1}^{n} A_iB\right) \\ &= \sum_{i=1}^{n} \mathrm{P}(A_iB) = \sum_{i=1}^{n} \mathrm{P}(A_i)\mathrm{P}(B|A_i).\end{aligned}$$

注 若直接求 $\mathrm{P}(B)$ 比较困难, 而 B 随 $\{A_i, i=1,2,\cdots,n\}$ 伴出, 则可通过全概率公式解决问题. (化整为零, 分块计算)

例 8 (赌金分配问题)　设甲、乙两赌徒进行公平的赌博 (即两人单局获胜概率相等; 假设没有平局). 事先约定谁先赢得 6 局便算赢家, 赢家获得全部赌金. 如果在甲赢了 5 局且乙赢了 3 局时因故终止了赌博, 应如何分配赌金?

解　应根据 "在甲赢了 5 局且乙赢了 3 局的条件下甲乙成为赢家的可能性之比" 分配赌金. 记

$$\begin{aligned}&f_6(x,y)\\ =&\mathrm{P}(\text{甲先赢得 6 局}|\text{甲赢了 } x \text{ 局且乙赢了 } y \text{ 局}), \quad x,y=0,1,\cdots,6.\end{aligned}$$

通过后续的第一局的胜负情况, 建立递推公式 (应用条件概率的全概率公式: $\mathrm{P}(B|A)=\mathrm{P}(C|A)\mathrm{P}(B|AC)+\mathrm{P}(\overline{C}|A)\mathrm{P}(B|A\overline{C})$):

$$f_6(x,y)=\frac{1}{2}f_6(x+1,y)+\frac{1}{2}f_6(x,y+1), \quad x,y=0,1,\cdots,5.$$

例 8 (赌金分配问题)　设甲、乙两赌徒进行公平的赌博 (即两人单局获胜概率相等; 假设没有平局). 事先约定谁先赢得 6 局便算赢家, 赢家获得全部赌金. 如果在甲赢了 5 局且乙赢了 3 局时因故终止了赌博, 应如何分配赌金?

解　应根据 “在甲赢了 5 局且乙赢了 3 局的条件下甲乙成为赢家的可能性之比” 分配赌金. 记

$$\begin{aligned} & f_6(x,y) \\ =& \mathsf{P}(\text{甲先赢得 6 局}|\text{甲赢了 } x \text{ 局且乙赢了 } y \text{ 局}), \quad x,y=0,1,\cdots,6. \end{aligned}$$

通过后续的第一局的胜负情况, 建立递推公式 (应用条件概率的全概率公式: $\mathsf{P}(B|A)=\mathsf{P}(C|A)\mathsf{P}(B|AC)+\mathsf{P}(\overline{C}|A)\mathsf{P}(B|A\overline{C})$):

$$f_6(x,y)=\frac{1}{2}f_6(x+1,y)+\frac{1}{2}f_6(x,y+1), \quad x,y=0,1,\cdots,5.$$

注意到

$$f_6(x,6)=0, \quad f_6(6,x)=1, \quad x=0,1,\cdots,5,$$

且根据对称性有

$$f_6(x,x)=\frac{1}{2},\quad x=0,1,\cdots,5.$$

所以

$$\begin{aligned}f_6(5,3)&=\frac{1}{2}[f_6(6,3)+f_6(5,4)]\\&=\frac{1}{2}+\frac{1}{4}[f_6(6,4)+f_6(5,5)]\\&=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}=\frac{7}{8}.\end{aligned}$$

这意味着甲、乙两赌徒应按照比例 $f_6(5,3):(1-f_6(5,3))=7:1$ 分配赌金.

定理（贝叶斯公式）

若 $\{A_i, i = 1, 2, \cdots, n\}$ $(n \leqslant \infty)$ 是 Ω 的一个完备事件组, 则

$$P(A_i|B) = \frac{P(A_iB)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{k=1}^{n} P(A_k)P(B|A_k)}, \quad i = 1, 2, \cdots, n.$$

证明　容易, 略.

P(A_i) 是在没有进一步信息 (不知 B 是否发生) 的情况下人们对于 A_i 发生的可能性大小的认知, 称为**先验 (priori) 概率**; 现在有了新的信息 (知道 B 发生), 人们对 A_i 发生的可能性大小有了新的认知, 得到条件概率 P$(A_i|B)$, 称为**后验 (posteriori) 概率**.

若汽车在 1 号门后, 则

$$\mathsf{P}(B_3|A_1) = 1/2;$$

若汽车在 2 号门后, 则

$$\mathsf{P}(B_3|A_2) = 1;$$

若汽车在 3 号门后, 则

$$\mathsf{P}(B_3|A_3) = 0.$$

所以

$$\mathsf{P}(B_3) = \sum_{i=1}^{3} \mathsf{P}(A_i)\mathsf{P}(B_3|A_i) = \frac{1}{3} \times (\frac{1}{2} + 1 + 0) = \frac{1}{2},$$

$$\mathsf{P}(A_1|B_3) = \frac{\mathsf{P}(A_1)\mathsf{P}(B_3|A_1)}{\mathsf{P}(B_3)} = \frac{1}{3},$$

$$\mathsf{P}(A_2|B_3) = \frac{\mathsf{P}(A_2)\mathsf{P}(B_3|A_2)}{\mathsf{P}(B_3)} = \frac{2}{3}.$$

因此, 改选 2 号门是正确的决策.

$$\mathsf{P}(A|B) = \frac{\mathsf{P}(A)\mathsf{P}(B|A)}{\mathsf{P}(B)} = \frac{\mathsf{P}(A)\mathsf{P}(B|A)}{\mathsf{P}(A)\mathsf{P}(B|A) + \mathsf{P}(\overline{A})\mathsf{P}(B|\overline{A})}$$

$\mathsf{P}(A)$ 是先验概率, $\mathsf{P}(A|B)$ 是后验概率, 随着 B 的更新, 利用上述 Bayes 公式不断迭代求解 $\mathsf{P}(A|B)$, 最后识得 $\mathsf{P}(A)$ 的真容:

$$\mathsf{P}(A) \xrightarrow{B_1} \mathsf{P}(A|B_1) \triangleq \mathsf{P}(A) \xrightarrow{B_2} \mathsf{P}(A|B_2) \triangleq \mathsf{P}(A) \cdots .$$

哲学思想: Bayes 公式告诉我们在面对不确定性时, 可以结合先验知识和新的信息来不断更新我们的认知/信念, 做到“吃一堑长一智”.

Bayes 公式在生活中最直接的体现就是: 从经验中学习, 做到有效地反思和行动.

通过获得的信息, 利用贝叶斯公式得到后验概率, 并根据后验概率改进决策的过程非常近似于人类的学习模式. 正是由于这一特点, 贝叶斯方法在机器学习领域发挥了重要的作用, 已成为学习型人工智能的理论基础. (另一典型的应用是“垃圾邮件分类问题”)

在这个不断变化的世界中, 我们可能无法完全控制我们的起点, 但我们完全有能力掌握如何迭代和进步. 每一次的尝试/失败/微小成功, 都是我们不断进化和适应的机会. (起点不重要, 迭代/奋斗很重要)

定义（两个事件的独立性）

对于两个事件 A 和 B, 若

$$P(AB) = P(A)P(B),$$

则称 A 与 B 相互独立 (或统计独立, 或独立).

注 由独立性的定义可知, 任何事件与 ϕ 相互独立, 任何事件与 Ω 相互独立.

例 14 已知 A 与 B 相互独立, 证明: A 与 $\overline{B}$, $\overline{A}$ 与 B, $\overline{A}$ 与 $\overline{B}$ 也相互独立.

证明 因为 A 与 B 相互独立, 所以 $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$. 由此可得

$$\begin{aligned}
\mathsf{P}(A\overline{B}) =& \mathsf{P}(A) - \mathsf{P}(AB) \\
=& \mathsf{P}(A) - \mathsf{P}(A)\mathsf{P}(B) \\
=& \mathsf{P}(A)[1 - \mathsf{P}(B)] \\
=& \mathsf{P}(A)\mathsf{P}(\overline{B}),
\end{aligned}$$

所以 A 与 $\overline{B}$ 相互独立. 类似可证其它结论.

注　在实际问题中, 用 (1.4.1) 或 (1.4.2) 来判断两个事件的独立性往往是比较困难的. 这时, 只需通过生活常识、生活经验进行判断即可. 例如, 一个电路系统中的两个不同原件是否出现故障常常被认为是相互独立的.

定义（两个事件域相互独立）

两个事件域 $\mathscr{F}_1$ 和 $\mathscr{F}_2$, 被认为关于 P 独立, 如果对于任意的事件 $A_1 \in \mathscr{F}_1$, $A_2 \in \mathscr{F}_2$ 都有

$$P(A_1A_2) = P(A_1)P(A_2).$$

2. 多个事件的独立性

定义（三个事件的独立性）

若

$$\begin{aligned} P(AB) &= P(A)P(B), \\ P(AC) &= P(A)P(C), \\ P(BC) &= P(B)P(C), \end{aligned}$$

且

$$P(ABC) = P(A)P(B)P(C),$$

则称 A, B, C 相互独立.

两两独立推不出相互独立

例 15 一个均匀的正四面体，其第一面染成红色，第二面染成白色，第三面染成黑色，第四面红白黑三色都有．分别用 A, B, C 记投一次四面体时底面出现红、白、黑的事件．易知

$$\mathsf{P}(A) = \mathsf{P}(B) = \mathsf{P}(C) = \frac{1}{2},$$

$$\mathsf{P}(AB) = \mathsf{P}(BC) = \mathsf{P}(AC) = \mathsf{P}(ABC) = \frac{1}{4}.$$

所以，A, B, C 是两两独立的．但同时

$$\mathsf{P}(ABC) = \frac{1}{4} \neq \mathsf{P}(A)\mathsf{P}(B)\mathsf{P}(C) = \frac{1}{8}.$$

定义（n 个事件的独立性）

对于 n 个事件 $A_1, A_2, \cdots, A_n$, 若对其中的任意 k $(2 \leqslant k \leqslant n)$ 个事件: $A_{i_1}, A_{i_2}, \cdots, A_{i_k}$, 有

$$P(A_{i_1} A_{i_2} \cdots A_{i_k}) = P(A_{i_1}) P(A_{i_2}) \cdots P(A_{i_k}), \tag{1.4.3}$$

则称 n 个事件 $A_1, A_2, \cdots, A_n$ 相互独立.

注　要说明 n 个事件 $A_1, A_2, \cdots, A_n$ 是相互独立的, 需验证 (1.4.3) 中的

$$C_n^2 + \cdots + C_n^n = 2^n - n - 1$$

个等式.

3. 试验的独立性

事件的独立性: 判断 $\mathsf{P}(AB) = \mathsf{P}(A)\mathsf{P}(B)$ 是否成立, 其中 A, B 来自同一样本空间. 启发: 可以用类似的方式去定义试验的独立性.

设 $E_1, E_2, \cdots, E_n$ 为 n 个试验, E_i 的样本空间为 Ω_i. 构造复合试验

$$E = (E_1, E_2, \cdots, E_n),$$

对应的样本空间为

$$\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n := \{(\omega_1, \omega_2, \cdots, \omega_n) : \omega_i \in \Omega_i, i = 1, 2, \cdots, n\}.$$

把 E_i 的任一事件 $A^{(i)}$ 放到复合的样本空间, 成为复合事件

$$\Omega_1 \times \cdots \times \Omega_{i-1} \times A^{(i)} \times \Omega_{i+1} \times \cdots \times \Omega_n,$$

不妨仍记为 $A^{(i)}$.

若对一切的复合事件 $A^{(1)}, A^{(2)}, \cdots, A^{(n)}$ (它们来自同一样本空间了), 有

$$\mathsf{P}(A^{(1)}A^{(2)}\cdots A^{(n)}) = \mathsf{P}(A^{(1)})\mathsf{P}(A^{(2)})\cdots\mathsf{P}(A^{(n)}),$$

则称试验 $E_1, E_2, \cdots, E_n$ 相互独立, 这里的 $\mathsf{P}(\cdot)$ 是定义在由复合样本空间 $\Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$ 构造的事件域上.

4. 伯努利 (Bernoulli) 概型

若试验 E_0 只有两个试验结果: $\{A, \overline{A}\}$, 则称试验 E_0 为伯努利试验.

定义 (伯努利概型)

若试验 E 满足下列三个条件:

(1) 是 n 次的独立重复试验;

(2) 每次试验只有/关心两个结果: $\{A, \overline{A}\}$;

(3) 每次试验中, A 发生的概率是固定不变的: $P(A) = p$,

则称这种试验 E 为伯努利概型 (或叫 n 重伯努利试验).

注 在伯努利概型里, 样本点 $\omega = (\omega_1, \omega_2, \cdots, \omega_n)$, 其中每个 ω_i 为 A 或 $\overline{A}$. 所以样本点总数为 2^n, 这是有限样本空间, 但不一定是古典概型.

例 19 求 n 重伯努利试验中事件 $B_k = \{$事件 A 恰好发生 k 次$\}$ 发生的概率.

解 记 $A_i = \{$第 i 次试验, A 发生$\}$, $i = 1, 2, \cdots n$. 则

$$\begin{aligned} B_k = & A_1 A_2 \cdots A_k \overline{A}_{k+1} \overline{A}_{k+2} \cdots \overline{A}_n + A_1 A_2 \cdots \overline{A}_k A_{k+1} \overline{A}_{k+2} \cdots \overline{A}_n \\ & + \cdots + \overline{A}_1 \overline{A}_2 \cdots \overline{A}_{n-k} A_{n-k+1} A_{n-k+2} \cdots A_n. \end{aligned}$$

上述等式的右边共有 C_n^k 项, 两两互不相容, 每一项中各事件又是相互独立的, 任一项的概率都是

$$[\mathsf{P}(A)]^k [1 - \mathsf{P}(A)]^{n-k} = p^k (1-p)^{n-k}.$$