第一章、随机事件与概率

§1.1 随机事件及其概率

- **随机事件**: 在一定条件下/做一次(随机)试验, 可能发生也可能不发生的事件.
- 例1.1. 掷分币/抛(公平)硬币.

$A=$ "正面朝上" / "正面"; $B=$ "正面朝下" / "反面".

- 例1.2. 掷两枚分币/抛两枚硬币/抛两次硬币.

$A=$ "都是正面"; $B=$ "都是反面"; $C=$ "一正一反".

- 例1.3. 10件同类产品中有8个正品, 2个次品. 任意抽取3个.

$A=$ "都是正品", $B=$ "至少一个次品".

$V=$ "都是次品" (不可能事件), $U=$ "至少一个正品" (必然事件).

概率

- 事件是否发生无法预知, 但是其可能性大小可以定量描述.
- 如, 抛硬币, “正面”和“反面”的可能性大小相同.
- 又如, 抛两枚硬币,
  “都是正面”和“都是背面”的可能性大小相同;
  “一正一反”的可能性比“都是正面”的可能性大.
- **随机事件的概率**: 定量描述它发生可能性大小.
  A 的概率记为$P(A)$.
- 概率: 频率、置信度、公理化定义.

- 一定条件(记为条件组S)大量重复实现时，
  事件A 发生的次数(称为**频数**)与总试验次数之比：

$$A \text{ 发生的频率} = \frac{A\text{的频数}}{\text{总试验次数}}.$$

- 频率为长期经验积累所得的，趋于稳定值p，这是客观事实.
- 例. 抛硬币，“正面朝上”的频率趋于1/2. (见P3表格).
- 例. 100 个球. 其中黑球50 个，白球50 个. 任取一个. 则“取到黑球”的频率趋于1/2. 若条件改为黑球60 个，白球40 个，则“取到黑球”的频率趋于0.6.
- 定义1.1. 称上述稳定值p为A (在条件组S 下) 发生的概率，记作$P(A) = p$.

- 频率具有稳定性的事件称为随机事件, 简称**事件**.
  其频率的稳定值称为**该事件的概率**.
- 注: 实际中遇到的事件一般都是随机事件.
- 频率取值于$[0,1]$. 因此$0 \leqslant P(A) \leqslant 1$.
- 不可能事件V 的概率: $P(V)=0$;
  必然事件U 的概率: $P(U)=1$.
- 频率是测量值, 是概率的近似值. 不要怀疑概率的存在性.

概率的主观含义

- 不能重复或不能大量重复的事件A, 如何定义其概率?
- 人们根据已有的知识和经验, 对A 发生可能性给出个人信念, 用$[0,1]$ 中的数q 来表示. 可能性大的事件A 对应较大的数q, (单调性).
- 定义1.2. 称上述个人信念q 为主观概率. 这是主观定义.
- 例. 企业家对产品畅销可能性的**预测**;

  医生对某特定病人手术成功的**预测**.
- 主观概率q: 对事件作了详细考察(与条件组S 吻合), 充分利用已有经验(与频率吻合, 与稳定值p 近似).

§1.2 古典概型

- 某些问题本身具有“对称性”. 可直接计算其概率.
- 这是用**数学模型**求解概率的方法.
- 如. 抛硬币, **认为**“正面”和“反面”概率相等(对称性), 故各为0.5.
- 注: “正面”和“反面”的频率之和为1.

例2.1. 盒中5个球, 3白2黑. 从中任取一个. 问: 取到白球的概率?

- 直观看为$3/5$.
- 建模: 把5个球编号, 1, 2, 3号为白球, 4, 5号为黑球.
- 分析: 取到每个球的概率相同(对称性). 事件互相排斥. 频率之和为1. 故概率各为$1/5$.
- 把3个白球的概率加起来即可. 注: 频率加起来即可.
- 注: 白球的概率与白球的比例吻合.

例2.2. 盒中5个球, 3白2黑. 从中任取两个. 问: 都是白球的概率?

- 注: 这时不能直观得出概率.
- 建模: 把5个球编号, 1—3号为白球, 4—5号为黑球.
- 所有可能结果有$C_5^2 = 10$ 个: $1+2$ $1+3$ $1+4$ $1+5$ $2+3$ $2+4$ $2+5$ $3+4$ $3+5$ $4+5$.
- 每个结果发生的机会相同, 互斥. 故每个结果的概率为1/10.
- 其中, $1+2, 1+3, 2+3$ 这3个结果满足“都是白球”, 其他结果则不满足. 故所求为3/10.

等概完备事件组

- 定义2.1 设$A_1, A_2, \ldots, A_n$ 为事件, 若
  (1) $A_1, A_2, \ldots, A_n$ 发生的机会相同(等可能性),
  (注: 源于对称性);
  (2) $A_1, A_2, \ldots, A_n$ 至少有一个发生(完备性),
  (注: 所有可能结果, 总频率为1);
  (3) $A_1, A_2, \ldots, A_n$ 至多有一个发生(互不相容性).
  (注: 互相排斥, 频率可加起来).

  则称$A_1, A_2, \ldots, A_n$ 为**等概完备事件组/等概基本事件组**.
  称每个A_i 为**基本事件**.
- 如, 例1.1, 抛硬币, 两个基本事件: “正面”和“反面”.

- 若$A_1, A_2, \ldots, A_n$是一个等概基本事件组, 事件B由其中的m个基本事件所构成, 则

$$P(B) = \frac{m}{n}. \tag{2.1}$$

- **古典概型**: 用等概基本事件组和(2.1)来计算事件的概率.
- 注: 建模, 不重不漏地算出n 与m.

例2.2(续). 5个球, 3白2黑. 从中任取两个. 问: 都是白球的概率?

- 共有$n = C_5^2 = 10$ 种不同取法, 每种取法对应一个基本事件.
- 其中, 满足“都是白球”的基本事件有$m = C_3^2 = 3$ 个.
- 故所求为$m/n = 3/10$.
- 误解: 三个基本事件: “两白”, “两黑”, “一黑一白”.
- 解决: 建模时务必将所有对象(即, 球)编号并加以区分.

例2.3. 100件产品, 有5件次品. 任取50件. 求: “无次品”的概率.

- 共有$n = C_{100}^{50}$ 个结果, 每个结果对应一个基本事件.
- 符合事件$B =$ “无次品”的基本事件: 从95个正品中取出50件. 故$m = C_{95}^{50}$.
- 故所求为

$$
\begin{aligned}
P(B) &= \frac{C_{95}^{50}}{C_{100}^{50}} = \frac{95!/(50!45!)}{100!/(50!50!)} \\
&= \frac{50!/45!}{100!/95!} = \frac{50 \cdot 49 \cdot 48 \cdot 47 \cdot 46}{100 \cdot 99 \cdot 98 \cdot 97 \cdot 96} \\
&= \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{47}{99} \cdot \frac{46}{97} = \frac{1081}{38412} \approx 2.8\%.
\end{aligned}
$$

例2.4. 100件产品, 有5件次品. 任取50件.

记$A=$ “恰好有2件次品”. 求: $P(A)$.

- $n=C_{100}^{50}$.
- 符合事件A 的基本事件: 从5个次品中任取2个, 从95个正品中任取48个. 故$m=C_5^2C_{95}^{48}$.
- 故所求为

$$P(A)=\frac{C_5^2C_{95}^{48}}{C_{100}^{50}}=\frac{\frac{5!}{2!3!}\cdot\frac{95!}{47!48!}}{\frac{100!}{50!50!}}\approx 0.32.$$

例2.5. 设一批产品共N个，其中次品共M个. 从中任取n个. 问：恰好出现m个次品的概率？

- 其中，$0 \leqslant m \leqslant n,\ \ m \leqslant M,\ \ n-m \leqslant N-M.$
- 所求为

$$P(\text{恰好出现}m\text{个次品}) = \frac{C_{N-M}^{n-m} C_M^m}{C_N^n}. \tag{2.2}$$

- 当$k < 0$ 或$k > n$ 时，约定$C_n^k = 0$.

定理2.1. 设有N 个对象, 分成k 类, 其中第i 类有N_i 个.
其中, $N_i \geqslant 0$, $i = 1, 2, \ldots, k$; 且$N_1 + N_2 + \cdots + N_k = N$.
从这N 个对象中任取n 个. 记

$$A = \text{“恰有} m_i \text{ 个属于第} i \text{ 类}, i = 1, \cdots, k\text{”}.$$

其中, $0 \leqslant m_i \leqslant N_i$, $i = 1, 2, \ldots, k$; 且$n = m_1 + m_2 + \cdots + m_k$.
则

$$P(A) = \frac{C_{N_1}^{m_1} C_{N_2}^{m_2} \ldots C_{N_k}^{m_k}}{C_N^n}. \tag{2.3}$$

§1.3 事件的运算及概率的加法公式 &

§1.4 集合与事件、概率的公理化定义

- 定义4.1. 集合是若干不同的元素的全体.
- 集合的符号: $A, B, C, \ldots$, 元素的符号: $a, b, c, \ldots$.
- $a \in A$: 元素a 属于集合A;

  $a \notin A$: 元素a 不属于集合A.
- **空集**, $\varnothing$: 不含任何元素.
- 例4.1. 全体正整数的集合, $\{1, 2, 3, \cdots\}$.
- 例4.2. 不大于10的正整数的集合, $\{1, 2, \cdots, 10\}$.
- 例4.5. 红、黄、白三个球, 有放回地抽取三次的(所有)结果组成的集合, 共有$3^3 = 27$ 个元素.

- 包含/包含于: A 的元素都是B 的元素(即, $a \in A \Longrightarrow a \in B$), 记为$A \subseteq B$ (称为A 包含于B) 或$B \supseteq A$ (称为B 包含A). 也称A 是B 的**子集**.
- 例4.3. $\mathbb{R}$: 实数集. $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$. 二维单位圆, $A = \{(x, y) : x, y \in \mathbb{R}; x^2 + y^2 < 1\}$ (开), $B = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leqslant 1\}$ (闭). 则$A \subseteq B \subseteq \mathbb{R}^2$.
- 相等: A 的元素与B 的元素完全相同, 记作$A = B$.
- $A = B \Longleftrightarrow A \subseteq B$ 且$B \supseteq A$.

集合的运算

- 并集, $A \cup B$: 所有“属于A 或属于B ”的元素组成.
- 交集, $A \cap B$, AB: 所有“属于A 且属于B”的元素组成.
- 若只讨论某非空集合Ω 的子集之间的关系, 则称Ω 为**全集**.
  A 的余集/补集: $A^c = \{x \in \Omega : x \notin A\}$. 注: $(A^c)^c = A$.
- 差集: $A \setminus B := A \cap B^c$. 例, $A^c = \Omega \setminus A$. (注: 维恩图).
- 交换律: $A \cup B = B \cup A$, $A \cap B = B \cap A$.
- 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$, 交集类似.
- 分配律: $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$, $\cap$ 与$\cup$ 可互换.
- 对偶律: $(A \cup B)^c = A^c \cap B^c$, $(A \cap B)^c = A^c \cup B^c$.
- $A \cup A = A$, $A \cup A^c = \Omega$, $A \cup \Omega = \Omega$, $A \cup \varnothing = A$.
  $A \cap A = A$, $A \cap A^c = \varnothing$, $A \cap \Omega = A$, $A \cap \varnothing = \varnothing$.

- (可列)无穷个集合的并集和交集. 设$A_1, A_2, \ldots$ 是一列集合.
- 并集, $\bigcup_{k=1}^{\infty} A_k$, $A_1 \cup A_2 \cup \ldots$:
  a 属于$A_1, A_2, \ldots$ 中的至少一个; 存在 k 使得a 属于A_k.

$$\bigcup_{k=1}^{\infty} A_k := \{a : \exists\ k \in \{1, 2, \ldots\} \text{ 使得 } a \in A_k\}.$$

- 交集: $\bigcap_{k=1}^{\infty} A_k$, $A_1 \cap A_2 \cap \ldots$, $A_1 A_2 \ldots$:
  a 属于所有的 $A_1, A_2, \ldots$; a 属于任意一个A_k.

$$\bigcap_{k=1}^{\infty} A_k := \{a : a \in A_k, \forall\ k \in \{1, 2, \ldots\}\}.$$

- 例, $\bigcup_{k=1}^{\infty} [\frac{1}{k}, 1] = (0, 1]$, $\bigcap_{k=1}^{\infty} (-\frac{1}{k}, \frac{1}{k}) = \{0\}$.

- 条件组S 下(即, 某个随机试验中)所有可能不同结果的集合记作Ω (视为全集). 事件则对应于Ω 的子集.
- Ω 是必然事件, $\varnothing$ 是不可能事件.
- $\omega \in A$: (ω 使得) A 发生.
- 事件的关系与运算: 包含/包含于, 相等, 交事件, 并事件, $\cdots$
- **对立事件**/补事件, A^c, 也记为$\bar{A}$. 含义: A 不发生.
- **差事件**, $A \backslash B$. 含义: 事件A 发生但事件B 不发生.
- A, B **互不相容/不相交/互斥**: $A \cap B = \varnothing$.
  含义: 不能同时发生. (注: $A \subseteq B^c$, $B \subseteq A^c$.)
- 多个事件互不相容: 两两互不相容. 例, 基本事件互不相容.

例4.6. 抛两枚硬币.

- $A =$ "两个都是正面", $B =$ "恰好一个正面",
  $C =$ "至少一个正面", $D =$ "一正一反".
- 则

$$A \subseteq C, \quad B = D \subseteq C, \quad \text{因此 } AC = A, \quad BC = B.$$

$$A \cup B = C, \quad AB = \varnothing.$$

- 数学语言: 用H(Head) 表示正面, T(Tail)表示反面.
  譬如, HT: 第一枚硬币正面朝上, 第二枚硬币反面朝上. 则

$$\Omega = \{\mathrm{HH}, \mathrm{HT}, \mathrm{TH}, \mathrm{TT}\}, \quad A = \{\mathrm{HH}\},$$

$$B = D = \{\mathrm{HT}, \mathrm{TH}\}, \quad C = \{\mathrm{HT}, \mathrm{TH}, \mathrm{HH}\}.$$

例3.2 一射手向某目标连续射击.

- $A_1 =$ “第一次射击, 命中”,
- $A_k =$ “前$k-1$次射击都未中, 第k次射击命中”, 其中, $k = 2, 3, \ldots$.
- 则$A_1, A_2, \ldots$ 互不相容, 且$\bigcup_{k=1}^{\infty} A_k =$ “终于命中”.
- 数学语言: 用H(Head) 表示“命中”, T(Tail)表示“未中”. 譬如, HTH$\cdots$: 第一次命中, 第二次未中, 第三次命中, $\cdots$

$$
\begin{aligned}
\Omega &= \text{所有无穷长的H-T字符串}, \\
A_1 &= \text{所有形如H} *** \text{的字符串}, \\
A_2 &= \text{所有形如TH} *** \text{的字符串}, \cdots
\end{aligned}
$$

$$\bigcup_{k=1}^{\infty} A_k = \Omega \backslash \{\mathrm{TT} \cdots \mathrm{T} \cdots\}.$$

- 频率的稳定值是直观的/易接受，主观概率不易被接受. 二者的数学严密性不足.
- 柯尔莫戈罗夫(Kolmogorov A. N., 1903-1987), 1933年: 用集合论、测度论严格定义概率，需要作公理化假设.
- 设Ω 为一个非空集合(全集), 称为**基本事件空间/样本空间**.
- 设$\mathscr{F}$ 是Ω的一些子集(事件)组成的集合. 若
  (1) $\Omega \in \mathscr{F}$;
  (2) 若$A \in \mathscr{F}$, 则$A^c = \Omega \backslash A \in \mathscr{F}$;
  (3) 若$A_1, A_2, \cdots \in \mathscr{F}$, 则$\bigcup_{n=1}^{\infty} A_n \in \mathscr{F}$,

  则称$\mathscr{F}$ 为σ **代数**.
- $\mathscr{F}$ 中的元素是事件, 有限个或可列个集合的运算封闭.
- 例, 若Ω 为有限集或可数集, 则通常$\mathscr{F}$ = "所有子集" .

- 设$P = P(\cdot)$ 是$\mathscr{F}$ 上的函数, 若

$$(1)\ P(A) \geqslant 0, \forall A \in \mathscr{F}, \tag{4.17}$$

$$(2)\ P(\Omega) = 1; \tag{4.18}$$

(3) 完全可加性: 若$A_1, A_2, \cdots \in \mathscr{F}$, 且它们两两不相交, 则

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n). \tag{4.19}$$

则称P 为**概率测度**, 简称**概率**.

- 赋有$\mathscr{F}, P$ 的Ω 叫做**概率空间**.
- 注: 完全可加性/可列可加性, 由经验/理性得出.
- 注: 有些情况下, 无法在Ω 的所有子集上合理定义概率.
- 注: 以下, 所有的事件均假设在$\mathscr{F}$ 中.

1 $P(\varnothing) = 0$.

2 有限可加性: 若$A_1, \ldots, A_n$ 两两不相交, 则

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i). \tag{4.20}$$

3 $P(A^c) = 1 - P(A)$. (注: 可利用$P(A^c)$ 计算/估算$P(A)$).

4 若$A \subseteq B$, 则$P(A) \leqslant P(B)$ 且$P(B \backslash A) = P(B) - P(A)$.

5 极限事件的概率: 设$A_1 \subseteq A_2 \subseteq \cdots$, $B_1 \supseteq B_2 \supseteq \cdots$, 则

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} P(A_n), \quad P\left(\bigcap_{n=1}^{\infty} B_n\right) = \lim_{n\to\infty} P(B_n).$$

例(若当公式).

$$P(A \cup B) = P(A) + P(B) - P(AB). \tag{3.6}$$

- 证: A 与$B\backslash A$ 互不相交, 且并集为$A \cup B$. 故

$$P(A \cup B) = P(A) + P(B\backslash A). \tag{3.7}$$

- 类似地, $P(B) = P(B\backslash A) + P(AB)$. 将$\star$ 代入上式即可.
- 若当公式: 对任意$n \geqslant 2$, $P(\bigcup_{k=1}^n A_k) = \sum_k P(A_k) - \sum_{k<i} P(A_k A_i) + \sum_{k<i<j} P(A_k A_i A_j) + \cdots$.

例3.1. 袋中有红、黄、白球各一个. 每次任取一个, 有放回地抽取三次. 记A = “抽到的三个球中没有红球或没有黄球”. 求$P(A)$.

- 解: 记G = “三个球都不是红球”, H = “三个球都不是黄球”. 则$A = G \cup H$. (注: G 和H 不是互不相容).
- $P(G) = \frac{2^3}{3^3} = \frac{8}{27}$. 同理, $P(H) = \frac{8}{27}$.
- GH = “三个球都是白球”, 故$P(GH) = \frac{1^3}{3^3} = \frac{1}{27}$.
- 根据若当公式, 所求为

$$P(G \cup H) = P(G) + P(H) - P(GH) = \frac{15}{27} = \frac{5}{9}.$$

§1.5 条件概率、乘法公式、独立性

- 在条件S下/在某随机试验中，事件B 的概率记为$P(B)$.
- 在条件S的基础上，再附加/追加条件A.
- 讨论在附加条件A之后，事件B 的概率，记为$P(B|A)$.
- 称$P(B|A)$ 为条件概率.

例5.1. 16个球, 6个玻璃球(2红4蓝), 10个木球(3红7蓝).
从16个球中任取一个.

- $A=$ “取到蓝球”,
  $P(A)=\frac{11}{16}$.
- $B=$ “取到玻璃球”,
  $P(B)=\frac{6}{16}$.

	玻璃	木质	
红	2	3	5
蓝	4	7	11
	6	10	16

- 现随机取出一球, 看出是蓝球. 问: 该球是玻璃球的概率?
- 在已知事件A 发生的(附加前提)条件下, 事件B 发生的概率, 记为$P(B|A)$.
- 继承古典概型的对称性,

$$P(B|A)=\frac{4}{11}=\frac{4/16}{11/16}=\frac{P(AB)}{P(A)}\approx\frac{AB\text{的频数}}{A\text{的频数}}.$$

- 古典概型: 设条件组S 下, 有n 个基本事件, A 由其中m 个组成, (B 由其中l 个组成), AB 由其中k 个组成. 则

$$
\begin{aligned}
P(B|A) =& \frac{\text{在}A\text{发生的前提下}B\text{中包含的基本事件数}}{\text{在}A\text{发生的前提下的基本事件总数}} \\
=& \frac{k}{m} = \frac{k/n}{m/n} = \frac{P(AB)}{P(A)}.
\end{aligned}
$$

- 定义5.1. 设$P(A) \neq 0$. 则

$$
P(B|A) = \frac{P(AB)}{P(A)}. \tag{5.1}
$$

为已知A 发生的前提条件下, B 发生的(条件)概率.

例5.2. 5个乒乓球, 3新2旧. 每次取一个, 无放回取两次.
A= "第一次取到新球"; B = "第二次取到新球".
求: $P(A)$, $P(B)$, $P(B|A)$.

- 建模: 把5个球编号, 1 — 3 为新球, 4 — 5 为旧球.
- $P(A) = \frac{3\cdot 4}{5\cdot 4} = \frac{3}{5}$, $P(AB) = \frac{3\cdot 2}{5\cdot 4} = \frac{3}{5}\cdot\frac{1}{2}$. 故$P(B|A) = \frac{1}{2}$.
- 直观: 若A 发生, 则还剩2新2旧(追加条件后, 产生新的随机试验). 于是第二次取到新球的概率为$\frac{2}{4} = \frac{1}{2}$, 即$P(B|A) = \frac{1}{2}$.
- $P(B)$: (i,j) 表示第一次取到i 号球, 第二次取到j 号球.
  $B = \{(i,j) : i \leqslant 5, j \leqslant 3, i \neq j\}$, 含$3 \times 4$ 个基本事件.
  故$P(B) = \frac{3\cdot 4}{5\cdot 4} = \frac{3}{5}$.
- 抽签是公平的: $A = \{(i,j) : i \leqslant 3, j \leqslant 5, i \neq j\}$.

- 条件概率的定义:

$$P(B|A) = \frac{P(AB)}{P(A)}. \tag{5.1}$$

- 乘法公式: 将(5.1)改写为

$$P(AB) = P(A)P(B|A). \tag{5.1'}$$

- 注: 以上两个公式的应用.

  (5.1): 已知$P(A)$ 和$P(AB)$, 求$P(B|A)$.

  (5.1'): 已知$P(A)$ 和$P(B|A)$ (在新的随机试验直接得到), 求$P(AB)$.

例5.3. 5个乒乓球, 3新2旧, 每次取1个, 有放回取2次.

- $A =$ "第一次取到新球".
  $B =$ "第二次取到新球"
- 直观: B 的(条件)概率与A 是否发生无关.
- 数学表达:

$$P(AB) = P(A)P(B|A) = P(A)P(B).$$

- 定义5.2. 若下式成立, 则称A, B **(相互)独立**.

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

- 定义中不要求$P(A) > 0$ 或$P(B) > 0$.
- 在$P(A) > 0$ (或, $P(B) \neq 0$) 时, 独立等价于

$$P(B|A) = P(B), \quad (\text{或}, P(A|B) = P(A)).$$

即, 条件概率等于无条件的原始概率.

- 独立性的直观解释:

A 是否发生不影响B 的发生概率;

B 是否发生不影响A 的发生概率.

定理5.1. 在四对事件A, B; $A, \bar{B}$; $\bar{A}, B$; $\bar{A}, \bar{B}$ 中, 若有一对相互独立, 则另外三对也相互独立.

- 即, 这四对事件或者都相互独立, 或者都不相互独立.
- 证: A, B 独立$\Longrightarrow$ $A, \bar{B}$, 这是因为

$$
\begin{aligned}
P(A\bar{B}) =& P(A) - P(AB) = P(A) - P(A)P(B) \\
=& P(A)[1 - P(B)] = P(A)P(\bar{B}).
\end{aligned}
$$

- 类似地, $\bar{A}, B$ 独立.
  或, 进一步, A, B 独立$/B, A$ 独立$\Longrightarrow$ $B, \bar{A}$ 独立/ $\bar{A}, B$ 独立.
- 再进一步, $\bar{A}, B$ 独立$\Longrightarrow$ $\bar{A}, \bar{B}$ 独立.
- 由$A = \bar{A}^c$ 知, 命题成立.

例5.4. 甲、乙同时向一敌机炮击. 甲击中概率0.6; 乙击中概率0.5. 求: 敌机被击中的概率.

- 注: 独立性作为(默认且合理的)假设.
- 解: $A =$ “甲击中”, $B =$ “乙击中”; $C =$ “敌机被击中”.
- $P(C) = P(A \cup B) = P(A) + P(B) - P(AB)$. (若当公式).
- $P(AB) = P(A) \times P(B) = 0.6 \times 0.5 = 0.3$,
  故$P(C) = 0.6 + 0.5 - 0.3 = 0.8$.
- 另解: $\bar{C} = \overline{A \cup B} = \bar{A} \cap \bar{B}$, (对偶律).
  故$P(\bar{C}) = P(\bar{A} \cap \bar{B}) = P(\bar{A})P(\bar{B}) = (1-0.6)(1-0.5) = 0.2$.
  从而$P(C) = 1 - P(\bar{C}) = 0.8$.
- 注: 两种方法将“并事件”的概率转化为“交事件”的概率.

- 定义5.3. 若以下4 个等式均成立, 则称A, B, C **相互独立**.

$$\begin{aligned} P(AB) &= P(A)P(B), \\ P(AC) &= P(A)P(C), \\ P(BC) &= P(B)P(C), \\ P(ABC) &= P(A)P(B)P(C). \end{aligned} \tag{5.3}$$

- 定义5.4. 若对任意满足$2 \leqslant k \leqslant n$ 的整数k, 以及从$1, 2, \ldots, n$ 中任意取出的k 个不同的$i_1, i_2, \ldots, i_k$,

$$P(A_{i_1} A_{i_2} \ldots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \ldots P(A_{i_k}), \tag{5.4}$$

则称$A_1, A_2, \ldots, A_n$ **(相互)独立**.

- 必要但非充分的要求:

$$P(A_1 A_2 \ldots A_n) = P(A_1)P(A_2) \ldots P(A_n).$$

例5.5. 某型号高射炮单发击中敌机的概率为0.6. 若干门同时发射单发, 欲以99%概率击中敌机. 问: 至少需要多少门高射炮?

- 注: 独立性作为(默认且合理的)假设.
- 解: 设需要n门, A_i = "第i门高炮击中敌机".
- A = "敌机被击中". $A = A_1 \cup A_2 \cup \cdots \cup A_n$. 则

$$\begin{aligned}
P(A) =& P(A_1 \cup A_2 \cup \cdots \cup A_n) = 1 - P(\bar{A}_1 \cap \bar{A}_2 \cap \cdots \cap \bar{A}_n) \\
=& 1 - P(\bar{A}_1)P(\bar{A}_2)\ldots P(\bar{A}_n) \qquad (\text{独立性}) \\
=& 1 - (1-0.6)^n = 1 - 0.4^n.
\end{aligned}$$

- 目标是$1 - 0.4^n \geqslant 0.99$. 即

$$n \geqslant \frac{\ln 0.01}{\ln 0.4} \approx 5.026.$$

- 因此, 至少需要6门高射炮.

例5.6(反例). 均匀正四面体, 四面分别涂成(1)红色、(2)黄色、(3)蓝色、(4)红黄蓝混杂. 投掷一次, 考察底面出现的颜色. 记

$A=$ "红色出现", $B=$ "黄色出现", $C=$ "蓝色出现".

试分析A,B,C 之间的独立性.

- 基本事件: $A_i=$ "第i面在底面", $i=1,2,3,4$.
- $A=A_1\cup A_4$, $B=A_2\cup A_4$, $C=A_3\cup A_4$.
- $P(A)=P(B)=P(C)=\frac{1}{2}$.
- $AB=AC=BC=A_4$, $P(AB)=P(AC)=P(BC)=\frac{1}{4}$.
  故, A,B; A,C; B,C 这三对都是相互独立的.
- 若$A_1,\cdots,A_n$ 中的任意两个相互独立, 则称它们**两两独立**.
- $ABC=A_4$, $P(ABC)=\frac{1}{4}\neq P(A)P(B)P(C)$.
  故, A,B,C 不是相互独立的.
- 注: 本例中的A,B,C 两两独立, 但不是相互独立的.

§1.6 全概公式与逆概公式

例6.1. 5个乒乓球, 3新2旧. 无放回取两次.

求: 第二次取到新球的概率.

- 注: 抽取是公平的(参见例5.2), 所求为$\frac{3}{5}$.
- 另解: $A =$ "第一次取到新球", $B =$ "第二次取到新球".

$$B = BA \cup B\bar{A}, \tag{6.1}$$

$$\begin{aligned} \text{故}, P(B) &= P(BA) + P(B\bar{A}) \\ &= P(A)P(B|A) + P(\bar{A})P(B|\bar{A}) \quad (\text{乘法公式}) \\ &= \frac{3}{5}\cdot\frac{2}{4} + \frac{2}{5}\cdot\frac{3}{4} = \frac{3}{5}. \end{aligned}$$

- 注: (6.1)将复杂的事件(情况)分解为简单的事件(情况).

定理6.1(全概公式). 设$A_1, A_2, \ldots, A_n$ 满足:

(1) $A_1, A_2, \ldots, A_n$ 互不相容, 且$P(A_i) > 0$, $(i = 1, 2, \ldots, n)$.

(2) 完备性: $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$,

则对任一事件B,

$$P(B) = \sum_{i=1}^{n} P(A_i)P(B|A_i). \tag{6.2}$$

- 注: 称$A_1, \ldots, A_n$ 为完备事件组, 它们是Ω 的**划分/分割**.
- 证: $B = BA_1 \cup BA_2 \cup \cdots \cup BA_n$, 故

$$\begin{aligned} P(B) =& P(BA_1) + \cdots + P(BA_n) \\ =& P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n). \end{aligned}$$

- 注: 完备事件组还可以包含可列个事件.
- 应用时的关键点: 找出完备事件组.

例6.2. 甲、乙、丙三人射击敌机. 击中概率:

甲: 0.4,　乙: 0.5,　丙: 0.7.

若只有一人击中, 敌机坠毁概率:

只有一人击中: 0.2,　恰好二人击中: 0.6,　三人全中: 1.

求: 敌机坠毁概率.

- 解: 记$B=$"飞机坠毁", $A_0=$"三人都不中",
  A_1, A_2, A_3 按颜色如上定义.
- A_0, A_1, A_2, A_3构成完备事件组.
  且　$P(B|A_0)=0$, $P(B|A_1)=0.2$, $P(B|A_2)=0.6$, $P(B|A_3)=1$.
- 记$H_1=$"甲中", $H_2=$"乙中", $H_3=$"丙中". 则

$$P(A_0)=P(H_1^cH_2^cH_3^c)=(1-0.4)(1-0.5)(1-0.7)=0.09.$$

- 类似地,

$$
\begin{aligned}
P(A_1) =& P(H_1 H_2^c H_3^c) + P(H_1^c H_2 H_3^c) + P(H_1^c H_2^c H_3) \\
=& 0.4(1-0.5)(1-0.7) + *** + *** = 0.36. \\
P(A_2) =& P(H_1^c H_2 H_3) + P(H_1 H_2^c H_3) + P(H_1 H_2 H_3^c) \\
=& (1-0.4)0.5 \times 0.7 + *** + *** = 0.41. \\
P(A_3) =& P(H_1 H_2 H_3) = 0.4 \times 0.5 \times 0.7 = 0.14.
\end{aligned}
$$

- 所求为

$$
\begin{aligned}
P(B) =& \sum_{i=0}^{3} P(A_i) P(B|A_i) \\
=& 0.09 \times 0 + 0.36 \times 0.2 + 0.41 \times 0.6 + 0.14 \times 1 = 0.458.
\end{aligned}
$$

例6.3(赌徒输光问题). 设甲有赌本M元, 乙有赌本N元. 每一局输赢为1元, 没有和局. 每局甲胜概率为p. 求: “甲输光”的概率.

- 其中, M, N 是正整数, $0 < p < 1$.
- 解: 记$L = M + N$, 则$L \geqslant 2$.
- 若$L = 2$, 即$M = N = 1$, 则所求为$1 - p$. 下设$L \geqslant 3$.
- 扩充问题: 若甲、乙共有赌本L元, 甲有赌本i 元, 乙有赌本$L - i$元, 则甲输光的概率p_i 是多少? (原问题所求为p_M.)
- 固定L. i 为模型参数, 初始假设条件, $0 \leqslant i \leqslant L$.
- 记A_i = 在参数为i 的模型中“甲最后输光”这一事件.
  B = “甲赢了第一局”. (任意参数i 的模型均适用.)
- $p_i = P(A_i)$. $p_0 = 1$, $p_L = 0$, (边界条件). 直观: $p_i \searrow$.

- 首步(首局)分析法:

  当$1 \leqslant i \leqslant L-1$ 时, 在参数为i 的模型中, 由全概公式,

$$\begin{aligned} p_i =& P(B)P(A_i|B) + P(\bar{B})P(A_i|\bar{B}) \\ =& p \cdot p_{i+1} + q \cdot p_{i-1}. \qquad (q = 1-p.) \end{aligned} \tag{6.3}$$

边界条件: $p_0 = 1$, $p_L = 0$.

- 将p_i 改写为$p \cdot p_i + q \cdot p_i$. 推出

$$\begin{aligned} p(p_i - p_{i+1}) =& q(p_{i-1} - p_i). \quad (\text{记 } \delta_i = p_i - p_{i+1}.) \\ \delta_i =& \frac{q}{p}\delta_{i-1} = r^2\delta_{i-2} = \cdots = r^i\delta_0. \quad (\text{记 } r = \frac{q}{p}.) \end{aligned}$$

$$p_0 - p_{i+1} = \sum_{k=0}^{i} \delta_k = \sum_{k=0}^{i} r^k \delta_0. \tag{6.4}$$

- $p_0 - p_i = \sum_{k=0}^{i-1} r^k \delta_0$, $\delta_i = p_i - p_{i+1}$, $r = \frac{q}{p}$, $q = 1 - p$.
  边界条件: $p_0 = 1$, $p_L = 0$.
- $p = \frac{1}{2}$:

$$p_0 - p_i = i\delta_0 \Rightarrow p_0 - p_L = L\delta_0 \Rightarrow \delta_0 = \frac{1}{L} \Rightarrow p_i = 1 - \frac{i}{L}.$$

- $p \neq \frac{1}{2}$:

$$p_0 - p_i = \frac{1-r^i}{1-r}\delta_0 \Rightarrow p_0 - p_L = \frac{1-r^L}{1-r}\delta_0 \Rightarrow \delta_0 = \frac{1-r}{1-r^L}$$

$$\Rightarrow p_i = 1 - \frac{1-r^i}{1-r} \cdot \frac{1-r}{1-r^L} = 1 - \frac{1-r^i}{1-r^L} = \frac{r^i - r^L}{1-r^L}.$$

- 甲输光的概率: $L = M + N$, $i = M$.

$$p_M = \begin{cases} \frac{N}{M+N}, & p = \frac{1}{2}. \\ \frac{r^M - r^{M+N}}{1-r^{M+N}}, & p \neq \frac{1}{2}. \end{cases}$$

例6.4. 敏感问题的调查.

- 在问卷调查时，某些敏感问题会遭到拒绝回答或谎报. 如，调查运动员“是否曾服用过兴奋剂”.
- 解决办法: 增设无害问题. 如，
  问题A: 你的生日是否在7月1日之前(不含7月1日)?
  问题B: 你是否服用过兴奋剂?
- 在盒中放入红、白球，其中红球数量占比为α(已知).
  若抽出白球，则回答问题A；若抽出红球，则回答问题B.
- 若样本量较大(如，有$n = 200$ 位受调查者)，则可用统计来估计服用兴奋剂的比例(记为p).

- 设有n 张答卷, 其中k 张答“是”.
- 答“是”的比例/频率: $\varphi = \frac{k}{n}$.
- 全概公式:

$$\begin{aligned} P(\text{答“是”}) =& P(\text{抽到白球})P(\text{生日在7月1日前}|\text{抽到白球}) \\ & + P(\text{抽到红球})P(\text{服用过兴奋剂}|\text{抽到红球}) \\ =& (1-\alpha)\cdot 0.5 + \alpha \cdot p \approx \frac{k}{n}. \quad (\text{概率} \approx \text{频率}) \\ \Rightarrow p \approx& \frac{\frac{k}{n} - \frac{1-\alpha}{2}}{\alpha}. \end{aligned}$$

- 例如, 50个球中有30个红球, $\alpha = 0.6$.
- 某国$n = 246$ 位运动员接受调查, 答“是”者有$k = 54$ 位.
- $p \approx 0.0325$. 即, 约3.25% 的运动员服用过兴奋剂.
- 思考: 比例α 的选取有何影响?

例6.5(发报与接收). 发报台分别以概率0.6和0.4发出信号“•”和“–”. 信号可能误码. 正确接收与错误接收的概率如下表:

设收到信号“•”.

求: 发报台发出的是“•”的概率.

		接收 •	接收 –
发出	•	0.8	0.2
	–	0.1	0.9

- $A=$ “发出信号‘•’”,
  $B=$ “收到信号‘•’”.
- 所求如下. (注: 基本可以判断原来是发出的“•”.)

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(A^c)P(B|A^c)}$$
$$=\frac{0.6\times 0.8}{0.6\times 0.8+0.4\times 0.1}=0.923.$$

- 注: 两个基本情况(A, A^c), 观测到一个结果(B 发生), 逆推A 与A^c 中哪一个情况是实际情况.

定理6.2(逆概公式). 设$A_1, \ldots, A_n$ 为一完备事件组, $P(B) > 0$. 则

$$P(A_j|B) = \frac{P(A_j)P(B|A_j)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}. \tag{6.6}$$

- 证: 联合条件概率的定义与全概公式.
- 注: 完备事件组还可以包含可列个事件.
- 逆概公式也称为贝叶斯(Bayes)公式.

例6.6(艾滋病检测).

- 美国的艾滋病人比例保守估计1000 分之一.
- 是否应该对新婚夫妇进行艾滋病毒血液检测?
- “血液检测法”各种结果及其概率:

		检验结果	
		报告阳性	报告阴性
实际情况	AIDS	真阳性(0.95)	假阴性(0.05)
	非AIDS	假阳性(0.01)	真阴性(0.99)

- 设某人报告阳性. 求: “真是患病者”的概率.
- $A =$“受试者带有艾滋病毒”, $T =$“检测结果呈阳性”.
- $P(A) = p \approx 0.001$, $P(T|A) = 0.95$, $P(T|A^c) = 0.01$.
  求$P(A|T)$.

- $P(A) = p \approx 0.001$, $P(T|A) = 0.95$, $P(T|A^c) = 0.01$. 求$P(A|T)$.
- 由逆概公式,

$$P(A|T) = \frac{P(A)P(T|A)}{P(A)P(T|A) + P(A^c)P(T|A^c)}$$
$$= \frac{p \times 0.95}{p \times 0.95 + (1-p) \times 0.01} = \frac{0.95}{0.94 + \frac{0.01}{p}} = f(p).$$

- $f(\cdot) \nearrow$. $f(p) \approx f(0.001) \approx 0.087$.
- 思考: 全面的检验是否必要?

例7.1 (独立)重复抛5 次硬币. 求: “恰有两次正面”的概率.

- 解: 古典概型. 共有$2^5 = 32$个等概基本事件.
- 其中恰有两次正面朝上的个数为$C_5^2 = 10$.
- 所求为

$$\frac{C_5^2}{32} = C_5^2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3. \tag{7.1}$$

- 若改为“分币不均匀”, 每一次“正面”概率为p, 则所求为

$$C_5^2 p^2 (1-p)^3.$$

- 此时, 完备事件组中的事件不等概.

例7.2 某人打靶, 命中率为0.7, 独立重复射击5次. 求: “恰好命中2 次”的概率.

- 记$p=0.7, q=1-p=0.3$.
- $P(\text{恰好命中2次})=C_5^2p^2q^3$.
- $P(\text{恰好命中3次})=C_5^3p^3q^2$.
- $P(\text{恰好命中4次})=C_5^4p^4q^1$.
- $P(\text{恰好命中5次})=C_5^5p^5q^0$.
- $P(\text{恰好命中1次})=C_5^1p^1q^4$.
- $P(\text{恰好命中0次})=C_5^0p^0q^5$.

定理. 设单次试验中, 事件A 发生(称为成功)的概率为p (称为成功概率). 则在n次(独立)重复试验中, 对$k = 0, 1, 2, \ldots, n$,

$$P(A\text{发生}k\text{次}) = C_n^k p^k q^{n-k}, \quad (q = 1 - p).$$

- 一般而言, 假设$0 < p < 1$. 否则答案是平凡的.
- 证: 记A_i = "第i 次试验中, A 发生".
- 2^n 个形如$C_1 C_2 \cdots C_n$ (其中C_i 可以是A_i 或A_i^c) 的事件构成完备事件组. (若$p \neq 1/2$, 则它们不等概.)
- 其中, 有$m = C_n^k$ 个符合"A发生k次". 记为$B_1, \cdots, B_m$.
- $P(B_1) = \cdots = P(B_m) = p^k q^{n-k}$. 故结论成立.

- 注: “重复”蕴含两重含义:
- (1) 每次试验的条件相同, 从而成功概率不变;
- (2) 各次试验的结果(是否成功)相互独立.
- 反例: 已知80个产品中有5个次品, 从中每次任取一个, 无放回地取20次, 求: “其中有2个次品”的概率.
- (1)反例中, 每次抽取的试验条件不同, 不能直接认为每次的成功(“取到次品”)概率不变(虽然这是事实);
- (2) 反例中, 各次的抽取结果(是否“取到次品”)不独立.
- 注: 若产品批量特别大, 则可用独立试验序列概型近似计算.

例7.3. 设每次射击打中目标的概率等于0.001. 设射击5000 次, 求: "至少两次打中目标" 的概率.

- $p = 0.001, q = 0.999.$

$$P(\text{至少两次打中目标}) = \sum_{k=2}^{5000} P(\text{恰有}k\text{ 次打中目标})$$

$$=1 - P(\text{都不中}) - P(\text{仅中一次})$$

$$=1 - q^{5000} - 5000 \times pq^{4999} \approx 1 - 0.006721 - 0.03364$$

$$\approx 0.9596.$$

P(成功k次) 的近似公式

- 公式一、当n 很大且p 很小的时候, 将np 视为常数λ.
  在n 次重复试验中,

$$P(\text{成功}k\text{ 次}) = C_n^k p^k q^{n-k} \approx \frac{\lambda^k}{k!}e^{-\lambda}. \tag{7.2}$$

  称为**泊松分布**近似, 见§2.2(P55).

- 如, 例7.3, $n = 5000$, $p = 0.001$, $np = 5$,

$$P(\text{都不中}) \approx e^{-5000\times 0.001} \approx 0.006738,$$

$$P(\text{仅中一次}) \approx \frac{5000 \times 0.001}{1!}e^{-5000\times 0.001} \approx 0.03369,$$

$$P(\text{至少两次打中}) \approx 1 - 0.006738 - 0.03369 \approx 0.9596.$$

- 公式二、当n 很大, 但p 不是很小时, 将p 视为常数.
  在n 次重复试验中,

$$P(\text{成功}k\text{次}) = C_n^k p^k q^{n-k} \approx \frac{1}{\sqrt{np(1-p)}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x_k^2},$$

$$\text{其中,}\ x_k = \frac{k-np}{\sqrt{np(1-p)}}.$$

- 参见§4.6(P153)的中心极限定理.
- 注: 记$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$. 则$\star\star = \phi(x_k)\Delta x_k$.

例7.4. 设每次射击打中目标概率为$p = \frac{1}{6}$. 设射击$n = 6000$ 次. 求: “击中次数在900到1100之间”的概率.

- 用公式二: $P(\text{成功}k\text{次}) \approx \phi(x_k)\Delta x_k$, 其中

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad x_k = \frac{k - np}{\sqrt{np(1-p)}}.$$

- 所求为

$$\sum_{k=900}^{1100} P(\text{成功}k\ \text{次}) = \sum_{k=900}^{1100} \phi(x_k)\Delta x_k \approx \int_a^b \phi(x)dx = \Phi(b) - \Phi(a),$$

其中$\Phi(y) = \int_{-\infty}^{y} \phi(x)dx$.

- $a = x_{900} \to x_{900-0.5}$, $b = x_{1100} \to x_{1100+0.5}$.
- 查表得到$\Phi(b) - \Phi(a) \approx 0.99950$.

例7.5(简单随机游动). 质点从原点出发, 每一步向右或左移动一个单位, 概率: p, $q = 1 - p$. 求: "质点n 步后位于K" 的概率.

- 注: $0 < p < 1$, n 是正整数, K 是非负整数, $n - K$ 为偶数.
- 设质点右移了x 步, 左移y 步, 则

$$x + y = n, \quad x - y = K \Rightarrow x = \frac{n+K}{2}.$$

- 将右移视为成功. 则所求为在n 次重复试验中,

$$P\left(\text{成功}\frac{n+K}{2}\text{ 次}\right) = C_n^{\frac{n+K}{2}} p^{\frac{n+K}{2}} q^{\frac{n-K}{2}}.$$