第 1 章　随机事件与概率

概率论是研究随机现象的统计规律的学科, 是统计学的理论基础. 事件和概率是概率论中最基本的两个概念. 本章重点介绍概率论的两个最基本概念：随机事件及其概率. 主要内容包括：随机试验、样本空间、随机事件的频率与概率、条件概率以及事件的独立性等概率论中的最基本概念. 介绍古典概型和几何概型、条件概率、乘法公式、全概率公式与贝叶斯公式及二项概率公式等计算概率的基本公式. 这些内容是进一步学习概率论的基础.

1.1　随 机 事 件

1.1.1　随机现象

在自然界和人类社会生活中普遍存在着两类现象, 一类是在一定条件下必然发生的现象, 称为**确定性现象**. 例如, 在没有外力作用下, 作匀速直线运动的物体必然继续作匀速直线运动; 人从地面向上抛起的石块经过一段时间必然落到地面; 在标准大气压下, 水加热到 100℃ 必然会沸腾; 等等.

另一类现象是在一定条件下可能发生也可能不发生的现象, 称为**随机现象**. 例如, 投掷一枚硬币, 我们不能事先预知将出现正面还是反面; 投掷一枚骰子出现的点数也不能预知; 从一大批产品中任取一个产品, 这个产品可能是正品, 也可能是废品, 其结果带有偶然性. 但人们通过长期实践并深入研究之后, 发现这类现象在大量重复试验或观察下, 它的结果却呈现出某种规律性. 概率论与数理统计就是研究和揭示随机现象这种规律性的一门数学学科. 其理论和方法被广泛地应用于自然科学、社会科学、工程技术和经济管理等诸多领域.

1.1.2　随机试验与随机事件

为了研究和揭示随机现象的统计规律性, 我们要对随机现象进行大量重复观察. 我们把观察的过程称为试验, 满足下列条件的试验称为**随机试验**, 本书以下简称为**试验**. 一般地, 一个随机试验要求满足下列条件：

(1) 可重复性：试验可以在相同条件下重复进行多次, 甚至进行无限多次;

(2) 可观测性：每次试验的结果具有多种可能性, 并能事先明确试验的所有可能结果;

(3) 随机性：试验之前不能准确预言该次试验将出现哪一种结果.

我们用 E 表示一个试验, 用 ω 表示随机试验 E 的最基本的结果, 称为**样本点**, 用 $\Omega=\{\omega\}$ 表示随机试验 E 的最基本结果的集合, 称为**样本空间**或**基本空间**.

例 1.1.1 在投掷一枚硬币观察其出现正面还是出现反面的试验中, 有两个样本点：正面、反面. 样本空间为 $\Omega=\{$正面, 反面$\}$. 记 $\omega_1=$“正面”, $\omega_2=$“反面”, 则样本空间可表示为

$$\Omega=\{\omega_1,\omega_2\}.$$

例 1.1.2 投掷一枚骰子, 观察出现的点数, 则基本结果是“出现 i 点”, 分别记为 $\omega_i(i=1,2,3,4,5,6)$, 则试验的样本空间为

$$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6\}.$$

例 1.1.3 将一枚硬币抛掷两次, 观察正面出现的次数, 记为 $\omega_i(i=0,1,2)$, 则样本空间为

$$\Omega=\{\omega_0,\omega_1,\omega_2\}.$$

也可简记为 $\Omega=\{0,1,2\}$.

例 1.1.4 记录某电话台在一分钟内接到的呼叫次数, 则样本空间为

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

例 1.1.5 测量某一零件的长度, 考察其观测结果与真正长度的误差, 样本空间可取作 $[-M,M]$, 其中 M 为最大误差. 如果无法确定这一最大值, 将 Ω 取作 $(-\infty,+\infty)$ 也可.

对于随机现象, 我们关心的通常是在随机试验中某一结果是否会出现, 或会出现什么结果, 这些结果称为**随机事件**. 习惯上, 用大写字母 A,B,C 等表示.

例如在例 1.1.2 中, 样本空间 $\Omega=\{1,2,3,4,5,6\}$, 如果以 A 表示“得到的为偶数点”, 则显然 $A=\{2,4,6\}$, B 表示“得到的点数大于等于 3”, 则 $B=\{3,4,5,6\}$. 这里 A,B 均为随机事件, 它们都由 Ω 中的若干个样本点所构成.

由此可见, 准确地讲, **随机事件**是由若干个样本点组成的集合, 或者说是样本空间的子集. 称某个事件 A 发生, 当且仅当该事件所包含的某个样本点出现.

由一个样本点组成的事件, 称为**基本事件**. 样本空间 Ω 本身也是 Ω 的子集, 它包含 Ω 的所有样本点, 在每次试验中 Ω 必然发生, 称为**必然事件**. 空集 $\varnothing$ 也是 Ω 的子集, 它不包含任何样本点, 在每次试验中都不可能发生, 称为**不可能事件**.

例 1.1.6 在抛掷一枚骰子的实验中, 分别记“点数是 1”为 A, “点数是偶数”为 B, “点数大于等于 2”为 C, “点数大于 6”为 D, “点数小于等于 6”为 F, 则 $A=\{1\},B=\{2,4,6\},C=\{2,3,4,5,6\},D=\varnothing,F=\Omega=\{1,2,3,4,5,6\}$.

在一个样本空间中, 如果只有有限个样本点, 则称它为**有限样本空间**; 如果有无限多个样本点, 则称它为**无限样本空间**. 例 1.1.1 $\sim$ 例 1.1.3 中的样本空间都是有限样本空间, 在例 1.1.4 和例 1.1.5 中的样本空间都是无限样本空间.

1.1.3 随机事件的关系和运算

在一个随机试验中, 一般有很多随机事件. 为了通过对简单事件的研究来掌握比较复杂的事件的规律, 需要研究事件的关系及事件的运算. 由于事件是样本空间的子集, 因此事件的关系及运算与集合的关系及运算是相互对应的. 关键是要理解事件的关系及运算的概率含义.

在以下的讨论中, 假定 Ω 是试验 E 的样本空间, 所论及的事件都是同一试验 E 的事件.

1. 事件的包含

如果事件 A 发生必然导致事件 B 发生, 则称事件 B**包含**事件 A, 或称事件 A**包含于**事件 B, 或称事件 A 是事件 B 的子事件, 记作

$$B \supset A \quad 或 \quad A \subset B.$$

易知, 对任意事件 A, 有

$$\varnothing \subset A \subset \Omega.$$

2. 事件的相等

如果事件 A 包含事件 B, 事件 B 也包含事件 A, 则称事件 A 与事件 B**相等**, 记作 $A = B$. 事件 A 与事件 B 相等, 表明 A 和 B 是样本空间的同一子集.

3. 事件的并 (或和)

如果事件 A 和事件 B 至少有一个发生, 则称这样的一个事件为事件 A 与事件 B 的**并**或**和**, 记作 $A \cup B$, 即

$$A \cup B = \{A\ 发生或 B\ 发生\} = \{\omega | \omega \in A\ 或\ \omega \in B\}.$$

在例 1.1.6 中, $A \cup B = \{1, 2, 4, 6\}$.

事件 A 和事件 B 作为样本空间 Ω 的子集, 事件 $A \cup B$ 就是子集 A 与 B 的**并集**.

4. 事件的交 (或积)

如果事件 A 和事件 B 同时发生, 则称这样的一个事件为事件 A 与事件 B 的**交**或**积**, 记作 $A \cap B$ 或 AB, 即

$$A \cap B = \{A\ 发生且\ B 发生\} = \{\omega | \omega \in A\ 且\ \omega \in B\}.$$

在例 1.1.6 中, $B \cap C = \{2, 4, 6\}$.

事件 A 和事件 B 作为样本空间 Ω 的子集, 事件 $A\cap B$ 就是子集 A 与 B 的**交集**.

5. 事件的差

如果事件 A 发生而事件 B 不发生, 则算这样的一个事件为事件 A 与 B 的**差**事件, 记作 $A-B$, 即

$$A-B=\{A\text{ 发生但 }B\text{不发生}\}=\{\omega|\omega\in A\text{ 但 }\omega\notin B\}.$$

在例 1.1.6 中, $C-B=\{3,5\}$.

6. 互不相容事件

如果事件 A 和事件 B 在同一次试验中不能同时发生, 则称事件 A 与事件 B 是**互不相容**的, 或称事件 A 与事件 B**互斥**.

在例 1.1.6 中, 事件 A 和事件 B 是互不相容的.

7. 对立事件 (或互逆事件)

如果在每一次试验中事件 A 和事件 B 都有一个且仅有一个发生, 则称事件 A 与事件 B 是**对立**的或**互逆**的, 其中的一个事件是另一个事件的对立事件, 记作 $\overline{A}=B$, 或 $\overline{B}=A$. 显然 $\overline{\overline{A}}=A$.

在例 1.1.6 中, 事件 A 与事件 C 是对立的.

8. 有限个或可数无穷多个事件的并与交

设有 n 个事件 $A_1,A_2,\cdots,A_n$, 则称 "$A_1,A_2,\cdots,A_n$ 至少有一个发生" 这一事件为事件 $A_1,A_2,\cdots,A_n$ 的**并**, 记作 $A_1\cup A_2\cup\cdots\cup A_n$ 或 $\bigcup\limits_{i=1}^{n}A_i$. 称 "$A_1,A_2,\cdots,A_n$ 同时发生" 这一事件为事件 $A_1,A_2,\cdots,A_n$ 的**交**, 记作 $A_1A_2\cdots A_n$ 或 $\bigcap\limits_{i=1}^{n}A_i$.

设有可数个事件 $A_1,A_2,\cdots,A_n,\cdots$, 则称 "$A_1,A_2,\cdots,A_n,\cdots$ 至少有一个发生" 这一事件为事件 $A_1,A_2,\cdots,A_n,\cdots$ 的**并**, 记作 $\bigcup\limits_{i=1}^{\infty}A_i$. 称 "$A_1,A_2,\cdots,A_n,\cdots$ 同时发生" 这一事件为事件 $A_1,A_2,\cdots,A_n,\cdots$ 的**交**, 记作 $\bigcap\limits_{i=1}^{\infty}A_i$.

9. 完备事件组

设 $A_1,A_2,\cdots,A_n,\cdots$ 是有限个或可数无穷多个事件, 如果满足:

(1) $A_iA_j=\varnothing,i\neq j,i,j=1,2,\cdots$;

(2) $\bigcup\limits_{i}A_i=\Omega$.

则称 $A_1,A_2,\cdots,A_n,\cdots$ 为样本空间 Ω 的一个**完备事件组**或 (**有限**)**分割**.

10. 事件的关系与运算的文氏图

上述关于事件的各种关系与运算可直观地用图形 (文氏图) 来表示 (见图 1.1).

类似集合的运算, 事件的运算有以下运算法则:

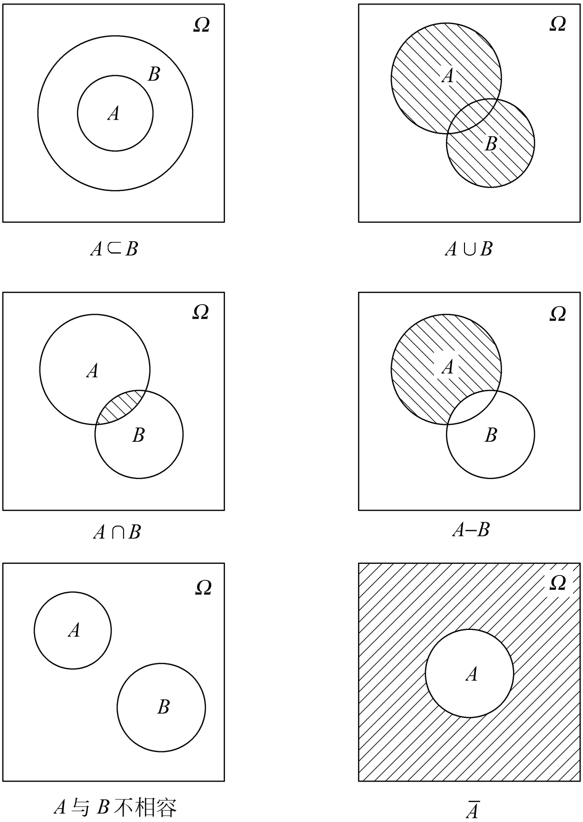

图 1.1 事件的关系与运算的文氏图

(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：$(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$,

$(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;

(4) 德摩根 (De Morgan) 定律 (对偶律)：

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \overline{A_i}, \quad \overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \overline{A_i}.$$

对于多个随机事件，以上的运算法则也是成立的.

例 1.1.7 某工人加工了三个零件, 设 A_i 为事件“加工的第 i 个零件是合格品”($i = 1, 2, 3$), 试用 A_1, A_2, A_3 表示下列事件：

(1) 只有第一个零件是合格品;

(2) 只有一个零件是合格品;

(3) 至少有一个零件是合格品;

(4) 最多有一个零件是合格品.

解 用 A, B, C, D 分别表示 (1)、(2)、(3)、(4) 所述事件, 则有

(1) 事件 A 发生, 即"只有第一个零件是合格品"这一事件发生, 即事件 A_1 发生并且 A_2, A_3 都不发生, 故

$$A = A_1\overline{A_2}\ \overline{A_3}.$$

(2) 事件 B 发生, 意味着在三个零件中有一个是合格品, 并且另外两个是不合格品, 因此

$$B = A_1\overline{A_2}\ \overline{A_3} \cup \overline{A_1}A_2\overline{A_3} \cup \overline{A_1}\ \overline{A_2}A_3.$$

(3) 事件 C 发生, 即在三个零件中至少有一个是合格品, 故

$$C = A_1 \cup A_2 \cup A_3.$$

(4) 事件 D 发生, 也就是事件"在三个零件中任意两个都不能同时为合格品"发生, 从而

$$D = \overline{A_1}\ \overline{A_2} \cup \overline{A_1}\ \overline{A_3} \cup \overline{A_2}\ \overline{A_3}.$$

也可以表示成

$$D = \overline{A_1}\ \overline{A_2}\ \overline{A_3} \cup A_1\overline{A_2}\ \overline{A_3} \cup \overline{A_1}A_2\overline{A_3} \cup \overline{A_1}\ \overline{A_2}A_3.$$

例 1.1.8 一名射手连续向某个目标射击三次, 事件 A_i 表示该射手第 i 次射击时击中目标 $(i = 1, 2, 3)$. 试用文字叙述下列事件:

$$A_1 \cup A_2; \quad A_1 \cup A_2 \cup A_3; \quad A_1A_2A_3; \quad A_3 - A_2; \quad A_3\overline{A_2}; \quad \overline{A_1 \cup A_2};$$

$$\overline{A_1}\ \overline{A_2}; \quad \overline{A_2} \cup \overline{A_3}; \quad \overline{A_2A_3}; \quad A_1A_2 \cup A_1A_3 \cup A_2A_3.$$

解 $A_1 \cup A_2$：前两次中至少有一次击中目标;

$A_1 \cup A_2 \cup A_3$：三次射击中至少有一次击中目标;

$A_1A_2A_3$：三次射击都击中了目标;

$A_3 - A_2 = A_3\overline{A_2}$：第三次击中但第二次未击中;

$\overline{A_1 \cup A_2} = \overline{A_1}\ \overline{A_2}$：前两次均未击中目标;

$\overline{A_2} \cup \overline{A_3} = \overline{A_2A_3}$：后两次中至少有一次未击中目标;

$A_1A_2 \cup A_1A_3 \cup A_2A_3$：三次射击中至少有两次击中目标.

例 1.1.9 化简事件 $\overline{(\overline{A}\ \overline{B} \cup C)\overline{AC}}$.

解 原式 $= \overline{\overline{A}\ \overline{B} \cup C} \cup AC = \overline{\overline{A}\ \overline{B}}\ \overline{C} \cup AC = (A \cup B)\overline{C} \cup AC$

$$= A\overline{C} \cup B\overline{C} \cup AC = A(\overline{C} \cup C) \cup B\overline{C} = A \cup B\overline{C}.$$

1.2　随机事件的频率与概率

1.2.1　频率

概率是概率论中最基本的概念. 在引入此概念之前, 需要介绍频率的概念.

定义 1.2.1　设在相同条件下进行的 n 次试验中, 事件 A 发生了 n_A 次, 则称 n_A 为事件 A 发生的**频数**, 称 $\dfrac{n_A}{n}$ 为事件 A 发生的**频率**, 记作 $f_n(A)$, 即

$$f_n(A)=\frac{n_A}{n}.$$

由频率的定义易见其具有下列基本性质：

(1)　$0\leqslant f_n(A)\leqslant 1$;

(2)　$f_n(\Omega)=1$;

(3)　若 $A_1,A_2,\cdots,A_m$ 是两两互不相容的事件, 则

$$f_n(A_1\cup A_2\cup\cdots\cup A_m)=f_n(A_1)+f_n(A_2)+\cdots+f_n(A_m).$$

事件 A 的频率反映了事件 A 发生的频繁程度. 频率越大, 事件 A 发生越频繁, 这意味着 A 在一次试验中发生的可能性越大.

通过实践, 人们发现, 当试验次数 n 很大时, 事件 A 发生的频率总会在某个确定的数值附近摆动, 随机事件的频率的这一特性称为频率的稳定性.

历史上有不少人通过投掷硬币的试验研究过频率的稳定性, 表 1.1 列出了他们的试验记录.

表　1.1

试验者	抛掷次数	正面朝上的次数	正面朝上的频率
德摩根	2048	1061	0.5181
蒲　丰	4040	2048	0.5069
费歇尔	10000	4979	0.4979
皮尔逊	12000	6019	0.5016
皮尔逊	24000	12012	0.5005

从表 1.1 可以看出频率有随机波动性. 当抛掷次数 n 较小时, 频率随机波动的幅度较大. 但是, 随着抛掷次数 n 的增大, 频率呈现出稳定性, 大致在 0.5 附近摆动, 且逐渐稳定于常数 0.5. 而且研究任何随机事件都有这样一个客观存在的常数与之对应. 这种“频率稳定性”即通常所说的统计规律性, 它是概率这一概念的经验基础. 而所谓事件发生的可能性大小, 就是这个“频率的稳定值”.

定义 1.2.2 在相同条件下重复进行 n 次试验, 如果当 n 增大时, 事件 A 的频率 $\dfrac{n_A}{n}$ 稳定地在某一常数 p 附近摆动, 则称常数 p 为事件 A 的**概率**, 记为 $P(A) = p$.

概率的这个定义, 称为概率的统计定义. 根据这一定义, 可以把大量重复试验所得到的事件的频率作为事件概率的近似值.

1.2.2 概率

任何一个数学概念都是对现实世界的抽象, 这种抽象使得其具有广泛的适应性, 并成为进一步数学推理的基础. 1933 年, 苏联数学家**柯尔莫戈洛夫**综合已有的大量成果, 提出了概率的公理化结构, 明确定义了概率等基本概念, 使得概率论成为严谨的数学分支, 推动了概率论的发展.

定义 1.2.3 设随机试验 E 的样本空间为 Ω, 如果对于 Ω 中的每一个事件 A, 有惟一的实数 $P(A)$ 和它对应, 并且这一事件的函数 $P(A)$ 满足以下条件:

(1) 非负性：对于任一事件 A, 有 $P(A) \geqslant 0$;

(2) 规范性：对于必然事件 Ω, 有 $P(\Omega) = 1$;

(3) 可列可加性：对于两两互不相容的事件 $A_1, A_2, \cdots$(即当 $i \neq j$ 时, 有 $A_iA_j = \varnothing, i, j = 1, 2, \cdots$), 有

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

则称 $P(A)$ 为事件 A 的**概率**.

由概率的公理化定义, 可以推导出概率的一些性质.

性质 1.2.1 对于不可能事件 $\varnothing$, 有 $P(\varnothing) = 0$.

证明 取 $A_i = \varnothing (i = 1, 2, \cdots)$, 显然这是一列两两不相容的事件, 且 $\bigcup\limits_{i=1}^{\infty} A_i = \varnothing$, 由定义 1.2.3 的 (3), 有

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

因此

$$P(\varnothing) = 0.$$

□

性质 1.2.2 对于两两互不相容的事件 $A_1, A_2, \cdots, A_n$, 有

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

这一性质称为概率的**有限可加性**.

证明　令 $A_i=\varnothing(i=n+1,n+2,\cdots)$, 根据概率的可列可加性及性质 1.2.1, 有

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

□

性质 1.2.3　对任一事件 A, 有

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

证明　因为 $A\cup\overline{A}=\Omega$, 且 $A\overline{A}=\varnothing$, 由性质 1.2.2 得

$$P(A)+P(\overline{A})=1,$$

即

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

□

性质 1.2.4　如果 $A\subset B$, 则有

$$P(B-A)=P(B)-P(A),\quad P(A)\leqslant P(B).$$

证明　因为 $A\subset B$, 从而有 $B=A\cup(B-A)$, 且 $A(B-A)=\varnothing$, 由性质 1.2.2 得

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

所以

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

由于 $P(B-A)\geqslant 0$, 因此

$$P(A)\leqslant P(B).$$

□

性质 1.2.5　对任一事件 A, 有 $P(A)\leqslant 1$.

证明　因为 $A\subset\Omega$, 由性质 1.2.4 及概率的规范性, 可得

$$P(A)\leqslant P(\Omega)=1.$$

□

性质 1.2.6　对于任意两个事件 A,B, 有

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

证明 由于 $B-A=B-AB$, 而 $AB\subset B$, 根据性质 1.2.4, 可得

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

□

性质 1.2.6 称为概率的**减法公式**.

性质 1.2.7 对于任意两个事件 A,B, 有

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

$$P(A\cup B)\leqslant P(A)+P(B).$$

证明 因为 $A\cup B=A\cup(B-AB)$, 且 $A(B-AB)=\varnothing, AB\subset B$, 由性质 1.2.2 及性质 1.2.4, 可得

$$\begin{aligned}P(A\cup B)&=P(A)+P(B-AB)\\&=P(A)+P(B)-P(AB).\end{aligned}$$

由于 $P(AB)\geqslant 0$, 因此

$$P(A\cup B)\leqslant P(A)+P(B).$$

□

性质 1.2.7 中的第一个公式称为概率的**加法公式**. 加法公式可以推广到任意有限个事件的情形：设 $A_1,A_2,\cdots,A_n$ 是 n 个随机事件, 则有

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

这个公式称为概率的**一般加法公式**.

例 1.2.1 已知 $P(\overline{A})=0.5$, $P(\overline{A}B)=0.2$, $P(B)=0.4$, 求：(1)$P(AB)$; (2)$P(A-B)$; (3)$P(A\cup B)$; (4)$P(\overline{A}\ \overline{B})$.

解 (1) 因为 $AB\cup\overline{A}B=B$, 且 AB 与 $\overline{A}B$ 互不相容, 故有

$$P(AB)+P(\overline{A}B)=P(B),$$

于是

$$P(AB)=P(B)-P(\overline{A}B)=0.4-0.2=0.2;$$

(2) $P(A)=1-P(\overline{A})=1-0.5=0.5,$

$P(A-B)=P(A)-P(AB)=0.5-0.2=0.3;$

(3)　$P(A\cup B)=P(A)+P(B)-P(AB)=0.5+0.4-0.2=0.7$;

(4)　$P(\overline{A}\ \overline{B})=P(\overline{A\cup B})=1-P(A\cup B)=1-0.7=0.3$.

例 1.2.2　设 A 和 B 是同一试验 E 的两个随机事件, 证明

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

证明　因为 $AB\subset A\subset (A\cup B)$, 所以

$$P(AB)\leqslant P(A\cup B).$$

由概率的性质 1.2.7 及事件运算的对偶律, 可得

$$P(\overline{A})+P(\overline{B})\geqslant P(\overline{A}\cup\overline{B})=P(\overline{AB})=1-P(AB),$$

因此

$$1-P(\overline{A})-P(\overline{B})\leqslant P(AB).$$ □

1.2.3　古典概型

如果随机试验具有以下两个特点:

(1) 随机试验只有有限个可能结果;

(2) 每一个可能结果发生的可能性相同.

则称这种试验为**等可能概型**或**古典概型**. 古典概型曾经是概率发展初期的主要研究对象.

设试验 E 是古典概型, 样本空间为 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$, 则基本事件 $\{\omega_1\}$, $\{\omega_2\}$,$\cdots$, $\{\omega_n\}$ 两两互不相容, 且

$$\Omega=\{\omega_1\}\cup\{\omega_2\}\cup\cdots\cup\{\omega_n\}.$$

由于 $P(\Omega)=1$ 及 $P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n)$, 因此

$$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n)=\frac{1}{n}.$$

如果事件 A 包含 r 个基本事件: $A=\{\omega_{i_1}\}\cup\{\omega_{i_2}\}\cup\cdots\cup\{\omega_{i_r}\}$, 其中 $i_1,i_2,\cdots,i_r$ 是 $1,2,\cdots,n$ 中某 r 个不同的数, 则有

$$P(A)=P(\omega_{i_1})+P(\omega_{i_2})+\cdots+P(\omega_{i_r})=\frac{r}{n}.$$

即

$$P(A)=\frac{A\text{ 包含的基本事件的个数}}{\Omega\text{ 包含的基本事件总数}}=\frac{r}{n}.$$

这样定义的概率称为**古典概率**. 古典概型问题的计算大致可分为三类, 下面分类举例.

1. 摸球问题 (产品的随机抽样问题)

例 1.2.3 袋中有 5 个红球, 3 个黄球, 从中一次随机地摸出两个球, 求摸出的两个球都是红球的概率.

解 E: 从 $(5+3)$ 个球中等可能地任取两球, 观察颜色.

Ω 含有 $n=\mathrm{C}_{5+3}^2=\mathrm{C}_8^2$ 个基本事件.

设 $A=\{$所取的二球全红$\}$. 则 A 含有 C_5^2 个基本事件, 即 $r=\mathrm{C}_5^2$, 所以

$$P(A)=\frac{r}{n}=\frac{\mathrm{C}_5^2}{\mathrm{C}_8^2}=\frac{5}{14}.$$

例 1.2.4 某人有 5 把钥匙, 其中有 2 把房门钥匙, 但忘记了开房门的是哪 2 把, 只好逐把试开, 问此人在 3 次内能打开房门的概率是多少?

解 E: 从 5 把钥匙中任选 3 把 (每次 1 把) 逐把试开房门 (试后不放回).

Ω 含有 $n=\mathrm{C}_5^1\times\mathrm{C}_4^1\times\mathrm{C}_3^1=5\times4\times3=60$ 个基本事件.

设 $A=\{3$ 次内打开房门$\}$, 则 A 含有 $r=60-6=54$ 个基本事件 (从 3 次开房门的所有开法中, 除去打不开房门的种数). 所以

$$P(A)=\frac{r}{n}=\frac{54}{60}=0.9.$$

例 1.2.5 袋中有 a 个白球, b 个黑球, 从中任意地连续一个一个地摸出 $k+1$ 个球 $(k+1\leqslant a+b)$, 每次摸出的球不放回袋中, 试求最后一次摸到白球的概率.

解法 1 E: 从 $a+b$ 个球中不放回地一个个地任意摸出 $k+1$ 个球进行排列 (与顺序有关).

Ω 含有 A_{a+b}^{k+1} 个基本事件.

设 A=$\{$在摸出的 $k+1$ 个球的排列中, 最后一个是白球$\}$. 考察 A: 第一步, 从 a 个白球中任取一个排到最后一个位置上, 有 A_a^1 种取法; 第二步, 从剩下的 $a+b-1$ 个球中任取 k 个排到前面的 k 个位置上, 有 A_{a+b-1}^k 种取法, 由乘法原理得出 A 含有 $\mathrm{A}_a^1\times\mathrm{A}_{a+b-1}^k$ 个基本事件. 所以

$$\begin{aligned}P(A)&=\frac{\mathrm{A}_a^1\times\mathrm{A}_{a+b-1}^k}{\mathrm{A}_{a+b}^{k+1}}\\&=a\times\frac{(a+b-1)!}{(a+b-1-k)!}\times\frac{(a+b-1-k)!}{(a+b)!}\\&=\frac{a}{a+b}.\end{aligned}$$

解法 2　$E: a+b$ 个球中不放回地一次任意摸出 $k+1$ 个球进行排列 (取时不考虑顺序, 但取出后考虑顺序).

Ω 含有 $\mathrm{C}_{a+b}^{k+1}\times(k+1)!$ 个基本事件 (第一步, 从 $a+b$ 个球中任取 $k+1$ 个球, 有 C_{a+b}^{k+1} 种取法; 第二步, 将取出的球进行全排列, 共有 $(k+1)!$ 种排法, 由乘法原理得基本事件总数).

设 A={取出的 $k+1$ 个球, 最后一个是白球}. 考察 A: 从 a 个球任取一白球, 再从 $a+b-1$ 个球中任取 k 个球, 共有 $\mathrm{C}_a^1\times\mathrm{C}_{a+b-1}^k$ 种取法, 将取出的一个白球固定在最后一个位置上, 其余的 k 个球在其余的位置上作全排列, 有 $k!$ 种排法, 由乘法原理得 A 含有 $\mathrm{C}_a^1\times\mathrm{C}_{a+b-1}^k\times k!$ 个基本事件. 所以

$$P(A)=\frac{\mathrm{C}_a^1\times\mathrm{C}_{a+b-1}^k\times k!}{\mathrm{C}_{a+b}^{k+1}\times(k+1)!}=\frac{a}{a+b}.$$

注　例 1.2.5 中所有事件的概率与 k 无关, 即每一次摸到白球的概率是一样的, 这是抽签问题 (也叫抓阄问题) 的模型, 即抽签时各人机会均等, 与抽签的先后顺序无关.

例 1.2.6　将一枚匀称的骰子抛掷两次, 求两次出现的点数之和等于 7 的概率.

解　E: 抛掷两次骰子所有可能出现的结果.

Ω 含有 $6\times6=36$ 个基本事件.

设 A={两次出现的点数之和等于 7}, 则 $r=6$, 所以

$$P(A)=\frac{r}{n}=\frac{6}{36}=\frac{1}{6}.$$

2. 分房问题

例 1.2.7　将 n 个人等可能地分配到 $N(n\leqslant N)$ 间房中的每一间去, 试求下列事件的概率:

(1) 某指定的 n 间房中各有 1 人;

(2) 恰有 n 间房各有 1 人.

解　E: 将 n 个人等可能地分配到 N 间房中去.

Ω 含有 N^n 个基本事件 (将每一个人分配到 N 间房中去都有 N 种分法, 因为没有限制每间房住多少人).

(1) 设 $A=\{$ 某指定的 n 间房中各有 1 人 $\}$. 考察 A: n 个人要分到指定的 n 间房中去, 保证每间房各有 1 人, 第一个人有 n 种分法, 分走一间之后, 第二个人有 $n-1$ 种分法, $\cdots\cdots$, 最后一间分配给第 n 个人, 故共有 $r=n(n-1)\cdot\cdots\cdot3\cdot2\cdot1=n!$ 种分法, 即 A 含有 $n!$ 个基本事件. 所以

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

(2) 设 $B=\{$ 恰有 n 间房各有 1 人 $\}$. 考察 $B: n$ 个人要分到 n 间房中去, 保证每间房各有 1 人, 有 $n!$ 种分法, 而 n 间房未指定, 故可以从 N 间房中任意选取, 有 C_N^n 种取法. 因此 B 含有 $r=\mathrm{C}_N^n\times n!$ 个基本事件. 所以

$$P(B)=\frac{\mathrm{C}_N^n\times n!}{N^n}.$$

例 1.2.8 某年级有 10 名大学生是 1985 年出生的, 试求下列事件的概率:

(1) 至少有两人同年同月同日生;

(2) 至少有一人在 10 月 1 日过生日.

解 E: 考察 10 人的生日是一年中的哪一天 (将 10 人的生日分配到一年的 365 天中去).

显然 Ω 含有 $n=365^{10}$ 个基本事件.

(1) 设 $A=\{$ 至少有两人的生日是同一天 $\}$.$\overline{A}$ 含有 A_{365}^{10} 个基本事件 (第一个人的生日放到 365 天中去, 有 365 种放法; 第二个人的生日只能放到剩下的 364 天中去, 有 364 种放法; 依此类推, 第 10 个人的生日只能放到 $365-9$ 天中去, 有 $365-9$ 种放法, 故共有 A_{365}^{10} 种放法). 所以

$$P(\overline{A})=\frac{\mathrm{A}_{365}^{10}}{365^{10}},$$

$$P(A)=1-P(\overline{A})\approx 0.1233.$$

(2) 设 $B=\{$ 至少有一人的生日是 10 月 1 日 $\}$,$\overline{B}=\{$ 没有人的生日是 10 月 1 日 $\}$.$\overline{B}$ 含有 364^{10} 个基本事件 (将 10 人的生日分配到除 10 月 1 日以外的 364 天中去).

$$P(\overline{B})=\frac{364^{10}}{365^{10}},$$

所以

$$P(B)=1-P(\overline{B})=1-\frac{364^{10}}{365^{10}}\approx 0.0271.$$

3. 随机取数问题

例 1.2.9 在 $0\sim 9$ 这 10 个整数中无重复地任意取 4 个数字, 试求所取到的 4 个数字能组成四位偶数的概率.

解 E: 从 10 个数字中任取 4 个进行排列.

Ω 含有 A_{10}^4 个基本事件.

设 $A=\{$ 排成的是四位偶数 $\}$. 考察 A: 先从 $0,2,4,6,8$ 等 5 个偶数中任取一个排在个位上, 有 A_5^1 种排法, 然后从剩下的 9 个数字中任取 3 个排到剩下的 3 个位置上, 有 A_9^3 种排法, 故个位上是偶数的排法共 $\mathrm{A}_5^1\times\mathrm{A}_9^3$ 种. 但是在这

种 4 个数字的排列中, 包含了 “0” 排在千位上的情况，故应除去这种情况的排列数：$\mathrm{A}_1^1\times\mathrm{A}_4^1\times\mathrm{A}_8^2$(“0” 排在千位上, 剩下的 4 个偶数任选一个排在个位上, 剩下 8 个数字中任选两个排在中间两位上), 故 A 含有 $(\mathrm{A}_5^1\times\mathrm{A}_9^3-\mathrm{A}_1^1\times\mathrm{A}_4^1\times\mathrm{A}_8^2)=56\times41$ 个基本事件. 所以

$$P(A)=\frac{56\times41}{\mathrm{A}_{10}^4}=\frac{41}{90}\approx0.4556.$$

例 1.2.10　从 $1\sim100$ 的 100 个整数中任取一个, 试求取到的整数能被 6 或 8 整除的概率.

解　E：从 $1,2,3,\cdots,100$ 中任取一数.

Ω 显然含有 100 个基本事件.

设 $A=\{$ 取到的整数能被 6 整除 $\}$; $B=\{$ 取到的整数能被 8 整除 $\}$; $C=\{$ 取到的整数能被 6 或 8 整除 $\}$. 显然 $C=A\cup B$.

考察 A：设 100 个整数中有 x 个数能被 6 整除, 则

$$6x\leqslant100,$$

所以

$$x=16.$$

即 A 含有 16 个基本事件.

考察 B：设 100 个整数中有 y 个数能被 8 整除, 则

$$8y\leqslant100,$$

所以

$$y=12.$$

即 B 含有 12 个基本事件.

考察 AB：能被 6 整除又能被 8 整除的数就是能被 24 整除的数, 设共有 z 个数, 则

$$24z\leqslant100,$$

所以

$$z=4.$$

即 AB 含有 4 个基本事件.

所以

$$\begin{aligned}P(C)&=P(A\cup B)=P(A)+P(B)-P(AB)\\&=\frac{16}{100}+\frac{12}{100}-\frac{4}{100}=\frac{24}{100}=0.24.\end{aligned}$$

1.2.4 几何概型

古典概型是关于存在有限等可能结果的随机试验的概率模型. 人们希望把这种做法推广到有无限多个基本事件, 而这些基本事件又有某种等可能性的情形.

如果一个随机试验相当于从直线、平面或空间的某一区域 Ω 任取一点, 而所取的点落在 Ω 中任意两个度量 (长度、面积、体积) 相等的子区域内的可能性是一样的, 则称此试验模型为**几何概型**. 对于任意有度量的子区域 $A \subset \Omega$, 定义事件 "任取一点落在区域 A 内" 的概率为

$$P(A) = \frac{A \text{ 的度量}}{\Omega \text{ 的度量}}.$$

这样定义的概率称为**几何概率**.

例 1.2.11 (会面问题) 甲、乙两人约定在 0 到 T(单位: h) 这段时间内在某处会面, 先到者等候另一个人 t ($t \leqslant T$) 后即可离去. 如果每一个人可在指定的这段时间内的任一时刻到达并且彼此独立, 求两人能会面的概率.

解 以 x 和 y 分别表示甲、乙两人到达约会地点的时刻, 则两人能会面的充分必要条件是

$$|x-y| \leqslant t.$$

在平面上建立直角坐标系如图 1.2 所示, 则 (x, y) 的所有可能结果是边长为 T 的正方形里的点, 能会面的点的区域用阴影标出. 根据几何概率的定义, 所求概率为

$$p = \frac{\text{阴影区域的面积}}{\text{正方形的面积}} = \frac{T^2-(T-t)^2}{T^2}.$$

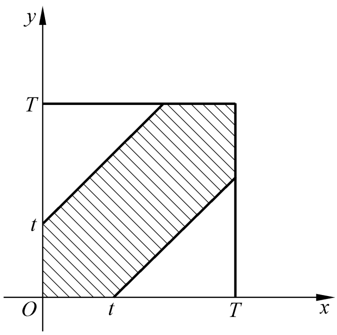

图 1.2

例 1.2.12 (Buffon 投针问题) 在平面上画有等距离的平行线, 平行线间的距离为 $2a(a>0)$, 向平面任意投掷一枚长为 $2l(l<a)$ 的圆柱形针, 试求此针与任一平行线相交的概率.

解 以 M 表示针的中点, 以 x 表示针投在平面上时点 M 到最近的一条平行线的距离, 以 φ 表示针与此直线的交角 (见图 1.3). 易知有

$$0 \leqslant \varphi \leqslant \pi, \quad 0 \leqslant x \leqslant a.$$

由这两式确定出 φOx 平面上的一个矩形 Ω. 针与最近的一条平行线相交的充分

必要条件是

$$x \leqslant l\sin\varphi.$$

由这个不等式表示的区域 A 是图 1.4 中的阴影部分, 所求概率为

$$p=\frac{A\text{ 的面积}}{\Omega\text{ 的面积}}=\frac{\displaystyle\int_0^{\pi} l\sin\varphi\mathrm{d}\varphi}{\pi a}=\frac{2l}{\pi a}.$$

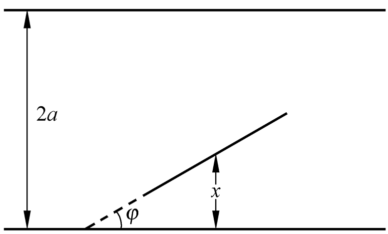

图 1.3

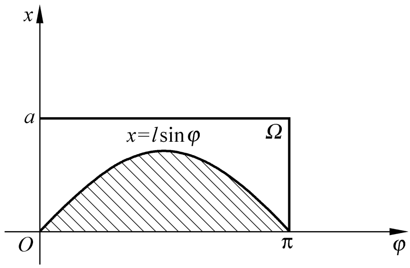

图 1.4

注 如果 l 和 a 为已知, 则以 π 值代入上式就可以算得 p. 反之, 也可以利用上式求 π. 如果投针 N 次, 其中针与平行线相交 n 次, 以频率值 $\dfrac{n}{N}$ 作为概率 p 的近似值, 代入上式可得

$$\pi\approx\frac{2lN}{an}.$$

历史上有一些学者曾做过这个实验. 例如, Wolf 在 1850 年投掷 5000 次, 得到 π 的近似值 3.1596; Smith 在 1855 年投掷 3204 次, 得到 π 的近似值 3.1554; Lazzerini 在 1901 年投掷 3408 次, 得到 π 的近似值 3.1415929; 等等.

例 1.2.13 设有任意两数 x 和 y 满足 $0<x<1, 0<y<1$, 求 $xy<\dfrac{1}{3}$ 的概率.

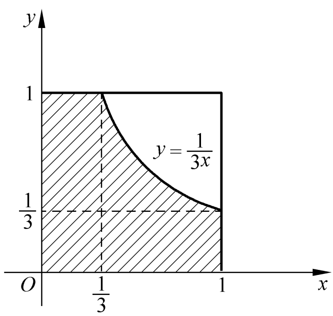

图 1.5

解 试验的样本空间为区域 $\Omega=\{(x,y)|0<x<1, 0<y<1\}$, 设所求事件为 A, 则 A 即为区域 $\{(x,y)|(x,y)\in\Omega$ 且 $xy<\dfrac{1}{3}\}$.

Ω 为边长是 1 的正方形, 其面积为 1(见图 1.5). A 的面积为

$$\frac{1}{3}+\int_{\frac{1}{3}}^{1}\frac{1}{3x}\mathrm{d}x=\frac{1}{3}+\frac{1}{3}\ln 3,$$

所以

$$P(A) = \frac{A\text{ 的面积}}{\Omega\text{ 的面积}} = \frac{\frac{1}{3} + \frac{1}{3}\ln 3}{1}$$

$$= \frac{1}{3} + \frac{1}{3}\ln 3.$$

1.3 条件概率

1.3.1 条件概率与乘法公式

在讨论事件发生的概率时, 经常会遇到这样的情况：已经知道某个事件 A 发生, 要求事件 B 发生的概率.

例 1.3.1 两台车床加工同一种机械零件, 数据如表 1.2 所示. 从这 100 个零件中任取一个零件, 令 $B=\{$取到合格品$\}$, 则

$$P(B) = \frac{85}{100} = 0.85.$$

表 1.2

加工的产品	合格品数	次品数	总计
第一台车床加工零件数	35	5	40
第二台车床加工零件数	50	10	60
总计	85	15	100

若从 100 个零件中任取一个零件, 当已知取到的是第一台车床加工的零件时, 问它是合格品的概率是多少. 我们令 $A=\{$取到零件是第一台车床加工的$\}$, 于是所求概率是事件 A 发生的条件下事件 B 发生的概率, 所以称它为 A 发生的条件下 B 发生的条件概率, 并记为 $P(B|A)$.

$P(B|A)$ 也可以用古典概型计算. 因为取到的是第一台车床加工的, 又知道第一台车床加工 40 个零件, 其中 35 个是合格品, 所以

$$P(B|A) = \frac{35}{40} = 0.875.$$

由于 AB 表示事件“取到的是第一台车床加工的并且是合格品”, 而在 100 件产品中是第一台车床加工的又是合格品的有 35 件, 所以 $P(AB) = \frac{35}{100}$. 从而

$$P(B|A) = \frac{35}{40} = \frac{\frac{35}{100}}{\frac{40}{100}} = \frac{P(AB)}{P(A)}.$$

一般地, 有下面的定义:

定义 1.3.1 设 A 和 B 是试验 E 的两个事件, 且 $P(A)>0$, 称 $\dfrac{P(AB)}{P(A)}$ 为在事件 A 已经发生的条件下, 事件 B 发生的**条件概率**, 记为 $P(B|A)$, 即

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

由这个定义可知, 对于任意两个事件 A 及 B, 如果 $P(A)>0$, 则有

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

称上式为概率的**乘法公式**.

同样可以在 $P(B)>0$ 的条件下, 定义在事件 B 已经发生的条件下, 事件 A 发生的条件概率为

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

在 $P(A)>0, P(B)>0$ 的条件下, 有

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

条件概率具有如下性质:

(1) 非负性: 对任意事件 B, 有 $P(B|A)\geqslant 0$;

(2) 规范性: 对于必然事件 Ω, 有 $P(\Omega|A)=1$;

(3) 可列可加性: 对于两两互不相容的事件 $B_1, B_2, \cdots, B_n, \cdots$, 有

$$P\left(\bigcup_{i=1}^{\infty} B_i|A\right)=\sum_{i=1}^{\infty} P(B_i|A).$$

由条件概率的三个基本性质可以推导出其他一些性质. 如

$$P(\varnothing|A)=0;$$

$$P(\overline{B}|A)=1-P(B|A);$$

$$P(B_1\cup B_2|A)=P(B_1|A)+P(B_2|A)-P(B_1B_2|A).$$

可以把乘法公式推广到任意多个事件的交的情况: 设 $A_1, A_2, \cdots, A_n$ 是同一试验的事件, 且 $P(A_1A_2\cdots A_{n-1})>0$, 则有

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

例 1.3.2 一个盒中有 6 只好晶体管, 4 只坏晶体管, 任取两次, 每次取一只. 考虑两种取产品的方式:

(i) 第一次取出一只晶体管, 观察好坏后放入盒中, 搅匀后再取一只. 这种取产品的方式叫放回抽样;

(ii) 第一次取出一只晶体管后不放回盒中, 第二次从剩余的晶体管中取出一只. 这种取产品的方式叫不放回抽样.

试分别按上述两种取晶体管的方式, 求:

(1) 第二次取到的是好的概率;

(2) 在第一次取到的是好的条件下, 第二次取到的是好的概率.

解 设 A 及 B 分别表示事件 “第一次取到的是好的” 及 “第二次取到的是好的”.

(i) 放回抽样

试验的基本事件总数 $n = 10 \times 10 = 100$.

(1) 事件 B 包含的基本事件数 $n_B = 10 \times 6 = 60$, 因此

$$P(B) = \frac{n_B}{n} = \frac{60}{100} = \frac{3}{5}.$$

(2) 事件 AB 包含的基本事件数 $n_{AB} = 6 \times 6 = 36$, 事件 A 包含的基本事件数 $n_A = 6 \times 10 = 60$, 所以

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{36}{100}}{\frac{60}{100}} = \frac{3}{5}.$$

(ii) 不放回抽样

试验的基本事件总数 $n = 10 \times 9 = 90$.

(1) 事件 B 包含的基本事件数 $n_B = 6 \times 5 + 4 \times 6 = 54$, 因此

$$P(B) = \frac{n_B}{n} = \frac{54}{90} = \frac{3}{5}.$$

(2) 事件 AB 包含的基本事件数 $n_{AB} = 6 \times 5 = 30$, 事件 A 包含的基本事件数 $n_A = 6 \times 9 = 54$, 因此

$$P(AB) = \frac{n_{AB}}{n} = \frac{30}{90} = \frac{1}{3},$$

$$P(A) = \frac{n_A}{n} = \frac{54}{90} = \frac{3}{5}.$$

所以

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{\frac{1}{3}}{\frac{3}{5}} = \frac{5}{9}.$$