第 15 章　概　　率

在一定条件下，事先就能断定发生或不发生某种结果，这种现象就是**确定性现象**. 在一定条件下，某种结果可能发生，也可能不发生，事先不能断定出现哪种结果，这种现象就是**随机现象**. 在自然界和人类社会的生产与生活中，存在着大量的确定性现象和随机现象.

对于某个现象，如果能让其条件实现 1 次，那么就是进行了 1 次**试验**(trial). 而试验的每一种可能的结果都蕴含着某种统计规律.

概率论的主要任务是研究随机现象的统计规律. 自然地，我们首先应当运用某种数学语言来进行表述.

为了清楚地叙述，现以抛掷一颗骰子的试验为例进行说明.

我们把"抛掷一颗骰子，结果向上的点数是 k"记为 ω_k ($k=1, 2, 3, 4, 5, 6$).

像上述 ω_1, ω_2, ω_3, ω_4, ω_5, ω_6 这样不可能再细分的结果称为**样本点**(sample point)；所有样本点组成的集合称为**样本空间**(sample space)，记为 Ω；样本空间的子集称为**随机事件**(random event)，也简称**事件**(event). 事件一般用 A, B, C 等大写英文字母表示. 当一个事件仅包含单一样本点时，称该事件为**基本事件**(elementary event).

显然，Ω(全集)是必然事件，$\varnothing$(空集)是不可能事件.

不难发现 A,B 两者之间的关系为 $B\subsetneqq A$，因此“事件 B 发生必导致事件 A 发生”. 这时，我们称 B 是 A 的子事件.

不难发现 A,B,C 三者之间的关系为 $C=A\cup B$，因此“事件 A 与 B 至少有一个发生即为事件 C 发生”. 这时，我们称 C 是 A 与 B 的并，也称 C 是 A 与 B 的和，并记作 $C=A+B$.

不难发现 A,B,C 三者之间的关系为 $C=A\cap B$，因此“事件 A 与 B 同时发生即为事件 C 发生”. 这时，我们称 C 是 A 与 B 的交，也称 C 是 A 与 B 的积，并记作 $C=AB$.

一般地,对于给定的随机事件 A,在相同条件下,随着试验次数的增加,事件 A 发生的频率会在某个常数附近摆动并趋于稳定,我们可以用这个常数来刻画随机事件 A 发生的可能性大小,并把这个常数称为随机事件 A 的**概率**(probability),记作 $P(A)$.

Ω 和$\varnothing$分别表示必然事件和不可能事件

$$P(\Omega)=1,\ P(\varnothing)=0.$$

$$0\leqslant P(A)\leqslant 1.$$

(1) 基本事件只有有限个；

(2) 每个基本事件的发生都是等可能的.

我们将满足上述条件的随机试验的概率模型称为**古典概型**

如果一次试验的等可能基本事件共有 n 个，即样本空间有 n 个样本点，那么每一个等可能基本事件发生的概率都是$\frac{1}{n}$. 如果事件 A 由其中 m 个等可能基本事件组合而成，即 A 中包含 m 个样本点，那么事件 A 发生的概率为

$$P(A)=\frac{m}{n}.$$

事件 A 与 B 不可能同时发生. 即 $AB=\varnothing$，称 A,B 为互斥事件.

互斥事件 A 和事件 C 中必有一个发生. 即

$$AC=\varnothing，且\ A+C=\Omega,$$

称 A，C 为对立事件 记作 $C=\bar{A}$或 $A=\bar{C}$.

显然，对立事件必为互斥事件，但反之不然. 对立事件是必有一个发生的互斥事件.

如果事件 A，B 互斥，那么事件 $A+B$ 发生的概率，等于事件 A，B 分别发生的概率的和，即

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

互斥事件可以推广到 n 个事件的情形（$n\in\mathbf{N}$，$n>2$）：如果事件 A_1，A_2，…，A_n 中任何两个事件都是互斥事件，那么称事件 A_1，A_2，…，A_n 两两互斥. 如果事件 A_1，A_2，…，A_n 两两互斥，那么

$$P(A_1+A_2+\cdots+A_n)=P(A_1)+P(A_2)+\cdots+P(A_n).$$

$$P(A+\bar{A})=P(A)+P(\bar{A})$$

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

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

一般地，如果事件 A 是否发生不影响事件 B 发生的概率，那么称 A, B 为**相互独立事件**(independent events).

对于相互独立事件，有下面的结论：

A, B 相互独立 $\Leftrightarrow P(AB) = P(A)P(B)$.

独立事件可以推广到 n 个事件的情形 $(n \in \mathbf{N}, n > 2)$. 一般地，如果事件 $A_1, A_2, \cdots, A_n$ 相互独立，那么

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

一般地，设 A，B 为两个事件，$P(A)>0$，我们称$\frac{P(AB)}{P(A)}$为事件 A 发生的条件下事件 B 发生的条件概率(conditional probability)，记为 $P(B \mid A)$，读作"A 发生的条件下 B 发生的概率"，即

$$P(B \mid A)=\frac{P(AB)}{P(A)}\ (P(A)>0).$$

由上述公式可知

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

通常将此公式称为概率的乘法公式.

条件概率有如下性质：

(1) $P(\Omega \mid A)=1$；

(2) $P(\varnothing \mid A)=0$；

(3) 若 B_1，B_2 互斥，则 $P((B_1+B_2) \mid A)=P(B_1 \mid A)+P(B_2 \mid A)$.

若事件 $A_1, A_2, \cdots, A_n$ 两两互斥，且它们的和 $\sum_{i=1}^{n} A_i = \Omega$，$P(A_i) > 0$，$i = 1, 2, \cdots, n$，则对于 Ω 中的任意事件 B，有

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

这个公式称为**全概率公式**(total probability formula).

一般地，若事件 $A_1, A_2, \cdots, A_n$ 两两互斥，且 $A_1 \cup A_2 \cup \cdots \cup A_n = \Omega$, $P(A_i) > 0$, $i = 1, 2, \cdots, n$，则对于 Ω 中的任意事件 B, $P(B)>0$，有

$$P(A_i \mid B)P(B) = P(B \mid A_i)P(A_i).$$

因此

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(B)}.$$

再由全概率公式得

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum_{j=1}^{n} P(A_j)P(B \mid A_j)}.$$

这个公式称为**贝叶斯公式**.