

随机事件与概率

基本概念

随机事件

随机试验是对随机现象的一次观测，满足三个条件：可重复（相同条件下试验可重复进行），多样性（每次试验有多种结果，试验前知道所有可能结果）以及随机性（每次试验结果随机）

样本空间是随机试验中所有可能结果的集合，记为 Ω ，样本空间的每个元素称为样本点

随机事件是样本空间的子集，由一个样本点组成的集合称为基本事件，基本事件不可再分

样本空间 Ω 称为必然事件，而空集 \emptyset 称为不可能事件

事件的关系与运算

事件本质上是集合。事件的关系有

- 包含关系，如 $A \subset B$ 表示 A 发生则 B 一定发生
- 互不相容关系，即两个事件不可能同时发生，用集合可表示为 $A \cap B = \emptyset$
- 相等关系，即 $A \subset B, B \subset A$

事件的运算也同集合一样，有

- 并：记作 $A \cup B$ ，多个事件的并可表示为 $\bigcup_{i=1}^n A_i$ ，含义为至少发生一个
- 交：记为 $A \cap B$ （有时省略 \cap ），多个事件的交可表示为 $\bigcap_{i=1}^n A_i$ ，含义为同时发生
- 差：事件 A 与事件 B 的差为 $A - B$ ，意为 A 发生且 B 不发生
- 对立：不发生 A 称为 A 的对立事件，记为 \bar{A} ，有 $\bar{\bar{A}} = A, A \cup \bar{A} = \Omega$

事件的运算有许多性质在证明中十分重要，以事件 A, B 为例

$$A = (A - B) \cup AB$$

事件 A 可表示为两个不相容事件的并

$$A \cup B = (A - B) \cup AB \cup (B - A)$$

$A \cup B$ 可表示为三个不相容事件的并

$$A - B = A - AB = A\bar{B}$$

事件的差可表示为事件的交

事件的运算满足交换律，结合律，分配律和**德摩根律**

事件的概率

频率与概率

频率：设随机事件 A 在 n 次重复试验中出现 n_A 次，则 A 发生的频率为

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

频率满足三个性质

- 非负性： $0 \leq f_n(A) \leq 1$
- 规范性： $f_n(\Omega) = 1$
- 有限可加性：若事件 A_1, A_2, \dots, A_n 互不相容，则

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

随试验次数的增加，频率逐渐稳定在某个数值 p ，称其为**统计概率**

概率的公理化定义：设 E 为随机试验， Ω 为其样本空间，对于每个随机事件 A 对应一个实数 $P(A)$ ，若函数 $P(\cdot)$ 满足

1. 非负性： $P(A) \geq 0$
2. 规范性： $P(\Omega) = 1$
3. 可列可加性：若事件 $A_1, A_2, \dots, A_n, \dots$ 两两不相容，则

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

则称 $P(A)$ 为事件 A 的概率

概率的性质

性质 1（可由可列可加性证明）

$$P(\emptyset) = 0$$

性质 2, 有限可加性 (由性质 1 和可列可加性证明), 设事件 A_1, A_2, \dots, A_n 互不相容, 则

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

性质 3 (可由性质 2 证明, $A - B$ 与 AB 不相容)

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

性质 4, 加法定理 (可由性质 2 证明)

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

加法定理的推广即容斥原理

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j) + \dots + (-1)^{n-1} P(A_1 A_2 \dots A_n)$$

性质 5

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

等可能概型

等可能概型是一类随机试验, 满足

- 样本空间 Ω 的样本点个数为有限个
- 样本空间中每个样本点发生的可能性相同

设随机试验 E 为等可能概型, 其样本空间 Ω 含有 n 个样本点, 而事件 A 含有 m 个样本点, 则事件 A 的概率为

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

N 件产品, 其中有 M 件次品, 抽取 n 件, 问抽取到 k 件次品的概率

不放回抽样 (without replacement) 的超几何分布:

$$P = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

有放回抽样 (with replacement), 独立重复试验:

$$P = C_n^k \left(\frac{M}{N}\right)^k \left(1 - \frac{M}{N}\right)^{n-k}$$

小概率事件： $P \leq 0.01$

小概率原理：试验中一般不会发生小概率事件，若发生，试验本身有问题

条件概率

条件概率与乘法公式

B 发生下 A 发生的概率，称为条件概率，记为 $P(A|B)$

本质上是事件 B 的发生削减了 A 的样本空间

设 A, B 为两个随机事件，且 $P(B) > 0$ ，称

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

为在事件 B 发生的条件下，事件 A 发生的概率。条件概率满足

1. 非负性： $P(A|B) \geq 0$
2. 规范性： $P(\Omega|B) = 1$
3. 可列可加性：若事件 A_1, A_2, \dots, A_n 互不相容，则

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

由条件概率公式，当 $P(B) > 0$ 时

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

称为乘法公式。乘法公式可以推广到多个事件

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

全概率公式与贝叶斯公式

划分：若 n 个事件 A_1, A_2, \dots, A_n 满足

- 两两互不相容
- $A_1 \cup A_2 \cup \dots \cup A_n = \Omega$

则称其为样本空间 Ω 上的一个划分或完备事件组

全概率公式：设 A_1, A_2, \dots, A_n 为完备事件组，则对事件 B 有

$$P(B) = \sum_{k=1}^n P(B|A_k)P(A_k)$$

$P(A_k)$ 可以看作是 B 发生的原因, 称为先验概率, 而 $P(A_k|B)$, 即 B 发生时, 原因为 A_k 的概率, 称为后验概率, 可使用贝叶斯公式求得

贝叶斯公式:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{k=1}^n P(B|A_k)P(A_k)}$$

独立性

独立性的定义

两个事件 A, B 若 B 的发生对于 A 没有影响, 即 $P(A|B) = P(A)$, 则称事件 A, B 相互独立

设 A, B 是两个随机事件, 若满足

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

则称其相互独立

若随机事件 A, B 相互独立, 则 $A, \bar{B}, \bar{A}, B, \bar{A}, \bar{B}$ 也相互独立

显然, “事件 A, B 相互独立”与“事件 A, B 互不相容”不能同时发生

独立性也可以推广到多个事件。 n 个事件相互独立需要其中任意 $k(2 \leq k \leq n)$ 个事件相互独立, 即需要 $C_n^2 + C_n^3 + \dots + C_n^n = 2^n - n - 1$ 个条件

分组独立性定理: 设 A_1, A_2, \dots, A_n 相互独立, 将其分成没有公共事件的 k 个组, 每个组中任意运算得到新事件, 则这 k 个新事件也互相独立

相互独立的事件的加法原理可以得到很大的简化

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(\bar{A}_1)P(\bar{A}_2) \dots P(\bar{A}_n)$$

可靠性分析

可靠性分析主要讨论按一定方式连接的多个元件或子系统正常工作的概率。设 n 个**独立**工作的元件构成系统, 第 i 个元件正常工作的事件为 $A_i, P(A_i) = p_i$, 系统正常工作的事件为 B

串联系统: 串联系统正常工作的充要条件是 n 个元件全都正常工作

$$P(B) = P(A_1, A_2, \dots, A_n) = \prod_{i=1}^n p_i$$

并联系统: 并联系统正常工作的充要条件是至少有一个元件正常工作

$$P(B) = P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - P(\overline{A_1} \overline{A_2} \dots \overline{A_n}) = 1 - \prod_{i=1}^n (1 - p_i)$$

混联系统可归化成子系统分别计算

独立重复试验概型

有一类独立重复试验概型，具有如下特点

- 每次试验只有两种结果， A, \overline{A} ，且 $P(A) = p$
- 试验进行 n 次，每次试验结果相互独立

这样的试验称为 n 重伯努利试验，其中 A 发生 k 次的概率记为 $P_n(k)$

$$P_n(k) = C_n^k p^k (1 - p)^{n-k}$$

有

$$\lim_{n \rightarrow \infty} C_n^k p^k (1 - p)^{n-k} = \frac{\lambda^k}{k!} e^{-\lambda}$$

其中 $\lambda = np$ ，在 n 很大而 p 很小时可用此式近似 n 重伯努利试验