目录

1 事件与概率

1.1 事件与样本空间

在自然界与人类社会中存在着一类特殊的现象，这类现象的结果是不确定的，但是我们可以确定的是这类现象的结果只能是某些特定的结果，而不会是其他结果。我们把这类现象称为**随机现象**，把这类现象的结果称为**随机事件**。例如，抛硬币、掷骰子、抽奖等都是随机现象。

定义 1.1.1 事件域
定义事件域类型.

定义 1.1.2 事件 (Event)
定义**事件**为事件域上的集合.

当我们多次观察自然现象和社会现象后，会发现许多事情在一定的条件下必然会发生，另外一些事情在一定条件下一定不可能发生。对此我们定义：

定义 1.1.3 决定性事件
在一定条件下，必然会发生的事情称为必然事件；反之，一定条件下，必然不会发生的事情就称为不可能事件。必然事件与不可能事件统称为**决定性事件**。

定义 1.1.4 频率
在相同条件下，某一事件发生的次数与试验次数的比值称为**频率**。
对于随机事件 A，在相同条件下，进行 n 次试验，事件 A 发生的次数为 m，则事件 A 的频率为

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

一个随机事件出现的频率常在某个固定的常数附近摆动，这种规律性我们称之为统计规律性，频率的稳定性说明随机事件发生的可能性大小是随机事件本身固有的、不随人们意志而改变的一种客观属性，因此可以对它进行度量。我们把这种度量称为**概率**。

定义 1.1.5 概率
概率是随机事件发生的可能性大小的度量，记作 $P(A)$，其中 A 为随机事件。

稍后会严谨地定义。

定义 1.1.6 试验
为了研究某个随机现象，为了观察这个现象的规律，我们进行的事情称为**试验**。并且总假定试验可以在相同条件下重复进行。

定义 1.1.7 样本点与样本空间
为了研究随机试验，首先需要知道这个试验可能出现的结果，这些结果称为**样本点**，一般用 ω 表示，样本点全体构成**样本空间 (samplespace)**，指一个随机现象所有可能结果的集合，用 Ω 表示.

在具体问题中，给定样本空间是描述随机现象的第一步.

性质 1.1.7.1
事件 A 的概率等于其中包含的所有样本点的概率之和。

定义 1.1.8　事件的包含关系
若 A 中的每一个样本点都包含在 B 中，则记为 $A \subset B$ 或 $B \supset A$，并称 A 是 B 的特款，亦称事件 A 包含于事件 B（或事件 B 包含了事件 A），这时事件 A 发生必然使得事件 B 发生。

定义 1.1.9　事件相等
若 $A \subset B$ 且 $B \subset A$，则称事件 A 与事件 B 相等，记作 $A = B$。相等的事件可以视作一样的。

定义 1.1.10　事件的交与并
用 $A \cap B$ 或 AB 表示所有同时属于 A 及 B 的样本点的集合，称它为 A 与 B 的交，事件 $A \cap B$ 表示事件 A 与事件 B 同时发生；
用 $A \cup B$ 表示至少属于 A 或 B 中的一个的所有样本点的集合，称它为 A 与 B 的并，事件 $A \cup B$ 即表示事件 A 与事件 B 至少发生一个。

定义 1.1.11　互斥
设 A, B 是事件, 定义 A 与 B **互斥**当且仅当: $A \cap B = \varnothing$.

定义 1.1.12　和事件
设 A, B 是事件, A 与 B 互斥, 定义 A 与 B 的**和事件**为 $A \cup B$, 记为 $A + B$.

定义 1.1.13　事件的差
用 $A - B$ 表示所有属于 A 但不属于 B 的样本点的集合，称它为 A 与 B 的差，事件 $A - B$ 表示事件 A 发生而事件 B 不发生。

性质 1.1.13.1

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

特别地，若 A 与 B 互斥，则

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

定义 1.1.14　对立事件
对于事件 A，由所有不包含在 A 中的样本点所组成的事件称为 A 的**对立事件**，或称 A 的**逆事件**，记为 $\bar{A}$，表示 A 不发生。即：

$$\bar{A} = \Omega - A.$$

性质 1.1.14.1
事件 A 与事件 B 互为对立事件，当且仅当 $A \cup B = \Omega$ 且 $A \cap B = \varnothing$。自然语言解释为，对立事件是指在一次试验中，只能发生其中一个事件的事件。

性质 1.1.14.2 事件的运算定律

1. 交换律：$A \cap B = B \cap A$，$A \cup B = B \cup A$；
2. 结合律：$(A \cap B) \cap C = A \cap (B \cap C)$，$(A \cup B) \cup C = A \cup (B \cup C)$；
3. 分配律：$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$，$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$；

定理 1.1.15 德摩根 (De Morgan) 定律
设 A 与 B 是两个事件，则有：

1. $\overline{A \cup B} = \bar{A} \cap \bar{B}$；
2. $\overline{A \cap B} = \bar{A} \cup \bar{B}$。

可以注意到，事件之间的运算与集合运算和布尔代数运算完全相同。

定义 1.1.16 离散样本空间
若样本空间 Ω 是有限、可列的，则称样本空间 Ω 是**离散样本空间**。

1.2 古典概型

定义 1.2.1 古典概型
若随机试验的样本空间是有限的，且每个样本点发生的可能性相同，则称这种试验为**古典概率模型**，简称为**古典概型**。

由古典概型出发可以得到一些组合计数中的重要分析公式。

例 1.2.2
从 n 个不同元素（有放回地）中依次取出 m 个元素并排成一列，这样的排列有 n^m 种。随机抽取时，得到的各种排列是等可能的。

例 1.2.3 排列
从 n 个不同元素（无放回地）中依次取出 m 个元素并排成一列，称为从 n 个元素中取出 m 个元素的排列，这样的排列的种数记作 A_n^m，有

$$A_n^m = n(n-1)(n-2)\cdots(n-m+1) = \frac{n!}{(n-m)!}.$$

随机抽取时，得到的各种排列是等可能的。
另外定义 $m < n$ 时的排列为**选排列**，$m = n$ 时的排列为**全排列**。此时有

$$A_m^m = m!$$

另外规定 $0! = 1$。

例 1.2.4 组合
从 n 个不同元素（无放回地）中依次取出 m 个元素，不考虑元素的排列顺序，称为从 n 个元素中取出 m 个元素的组合，这样的组合的种数记作 C_n^m，有

$$C_n^m = \binom{n}{m} = \frac{A_n^m}{m!} = \frac{n!}{m!(n-m)!}.$$

随机抽取时，得到的各种组合是等可能的。

例 1.2.5
将 n 个不同的元素分成 r 组，每组的元素个数分别为 $n_1, n_2, \cdots, n_r$，且 $n_1 + n_2 + \cdots + n_r = n$，这样的分组的种数记作 $P(n_1, n_2, \cdots, n_r)$，有

$$\binom{n}{n_1, n_2, \cdots, n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}.$$

1.3 几何概型与概率空间

定义 1.3.1 几何概型
设样本空间 Ω 的体积 $m(\Omega)$ 是正数，样本点等可能地分布在 Ω 中，且事件 A 是 Ω 的一个子集，且 A 的体积是 $m(A)$，则称这种试验为**几何概率模型**，简称为**几何概型**。事件 A 发生的概率为

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

那么问题来了，该如何定义这种抽象的“体积”?

定义 1.3.2 事件域、可测空间
设 Ω 是一个集合，$\mathcal{F}$ 是 Ω 的一个子集族，若 $\mathcal{F}$ 满足：

1. $\Omega \in \mathcal{F}$;
2. 若 $A \in \mathcal{F}$，则 $\bar{A} \in \mathcal{F}$;
3. 若 $A_1, A_2, \cdots \in \mathcal{F}$，则 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$。

则称 $\mathcal{F}$ 是 Ω 的一个**事件域**，称 $\mathcal{F}$ 中的元素为事件，称 $(\Omega, \mathcal{F})$ 为**可测空间**。

我们经常研究的有如下的事件域：

定义 1.3.3 Borel 事件域
设 Ω 是 n 维的实数空间 $\mathbb{R}^n$，$\mathcal{B}$ 是 $\mathbb{R}^n$ 的一个子集族，若 $\mathcal{B}$ 满足：

1. $\mathbb{R}^n \in \mathcal{B}$;
2. 若 $A \in \mathcal{B}$，则 $\bar{A} \in \mathcal{B}$;
3. 若 $A_1, A_2, \cdots \in \mathcal{B}$，则 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{B}$;

4. 若 $A_1, A_2, \cdots \in \mathcal{B}$，则 $\bigcap_{i=1}^{\infty} A_i \in \mathcal{B}$。

就称 $\mathcal{F}$ 是 $\mathbb{R}$ 的一个 **Borel 事件域**。

定义 1.3.4　测度
称 $\mathcal{F}$ 上的函数 m 是 Ω 上的一个**测度**，若 m 满足：

1. 非负性：对于任意 $A \in \mathcal{F}$，有 $m(A) \geq 0$；
2. 空间的可加性：若 $A_1, A_2, \cdots \in \mathcal{F}$ 两两互不相容，则有

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

定义 1.3.5　体积
设有 r 维向量空间 $\mathbb{F}^r = \{(x_1, x_2, \cdots, x_i, \cdots, x_r) | x_i \in \mathbb{F}\}$，对于 $\mathbb{F}^r$ 的任意子集 A，若存在一个测度 m，使得 $m(A)$ 满足：

$$m(A) = \int_A \mathrm{d}x_1 \mathrm{d}x_2 \cdots \mathrm{d}x_r$$

则称这个测度为体积。

至此，我们可以对概率进行严谨的定义。

定义 1.3.6　概率
设 $\mathcal{F}$ 上的函数 P 是 Ω 上的一个**测度**，若它满足：

1. 非负性：对于任意 $A \in \mathcal{F}$，有 $P(A) \geq 0$；
2. 规范性：对于必然事件 Ω，有 $P(\Omega) = 1$；
3. 空间的可加性：若 $A_1, A_2, \cdots \in \mathcal{F}$ 两两互不相容，则有

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

就称 P 是 Ω 上的一个**概率测度**，简称**概率**，称 $(\Omega, \mathcal{F}, P)$ 为**概率空间**。

定义 1.3.7　几乎必然
若事件 A 的概率 $P(A) = 1$，则称事件 A 是**几乎必然**的，也称事件 A 是**概率 1 发生**或**几乎处处发生**的。
值得注意的是，几乎必然发生的事件在任意一次试验中也可能不会发生。

定理 1.3.8　概率的 Jordan 公式

设 $A_1, A_2, \cdots, A_n$ 是事件，设

$$p_k = \sum_{1 \le j_1 \le \cdots \le j_k} P(A_i A_2 \cdots A_k)$$

，则有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} p_k.$$

证明:
我们使用数学归纳法证明。

性质 1.3.8.1　概率的性质

1. $P(\varnothing) = 0$；
2. $P(A \cup B) = P(A) + P(B) - P(AB)$；
3. $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$；
4. $P(\bar{A}) = 1 - P(A)$；

现在这些规律都可以被轻松地证明。

反例 1.3.9　Bertrand 概率问题

定义 1.3.10　概率的连续性
设 $A_1, A_2, \cdots$ 是一列单调递增的事件，即 $A_1 \subset A_2 \subset \cdots$，则有

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

概率的这种性质称为**连续性**。

定理 1.3.11　概率的可列可加性定理
设 P 是 $(\Omega, \mathcal{F})$ 上的一个概率，则它具备可列可加性的充要条件为：它是连续且有限可加的。

证明:

1.4 条件概率

定义 1.4.1 条件概率
设概率空间中的事件 A, B，且 $P(B) > 0$，则称

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

为在事件 B 发生的条件下事件 A 发生的**条件概率**。

性质 1.4.1.1 概率的乘法原理
设 A, B 是概率空间中的事件，则有

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

定理 1.4.2 全概率公式
设事件 $B_1, B_2, \cdots, B_k$ 是一组互不相容的事件，且 $\bigcup_{i=1}^{k} B_i = \Omega$，则对于任意事件 A，有

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

证明:
由于 B_i 两两互不相容，所以

$$A = \sum_{i=1}^{k} B_i A$$

由可列可加性可得

$$P(A) = \sum_{i=1}^{k} P(B_i A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

定理 1.4.3 贝叶斯公式
设事件 $B_1, B_2, \cdots, B_k$ 是一组互不相容的事件，且 $\bigcup_{i=1}^{k} B_i = \Omega$，则对于任意事件 A，有

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{k} P(A|B_j)P(B_j)}.$$

证明:
使用全概率公式即可证明。

1.5 事件的独立性

定义 1.5.1 事件的独立性
设事件 A, B，若

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

，则称事件 A 与事件 B 是**相互独立**的，也称**独立**的。

定义 1.5.2 独立事件列
设事件 $A_1, A_2, \cdots, A_n$，若对于任意的 $1 \le i_1 \le i_2 \le \cdots \le i_k$，有

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

，则称事件 $A_1, A_2, \cdots, A_n$ 是**独立事件列**。

性质 1.5.2.1
设 $A_1, A_2, \cdots, A_n$ 是独立事件列，用 B_i 表示 A_i 或 $\bar{A}_i$，则 $B_1, B_2, \cdots, B_n$ 也是独立事件列。

性质 1.5.2.2

2 随机变量与概率分布

2.1 随机变量

定义 2.1.1 随机变量
设 $(\Omega, \mathcal{F}, P)$ 是一个概率空间，若对于每一个事件 $A \in \mathcal{F}$，都有一个实数 $X(A)$ 与之对应，且满足：

1. 对于任意的实数 x，有 $\{\omega | X(\omega) \le x\} \in \mathcal{F}$；
2. 对于任意的实数 x，有 $P(X \le x) = P(\{\omega | X(\omega) \le x\})$。

则称 X 是 $(\Omega, \mathcal{F}, P)$ 上的一个**随机变量**。

定义 2.1.2 离散型随机变量
若随机变量 X 的取值只能是有限个或可列个，则称 X 是**离散型随机变量**。

定义 2.1.3 连续型随机变量
若随机变量 X 的取值是一个区间，则称 X 是**连续型随机变量**。

定义 2.1.4 分布函数
设 X 是一个随机变量，对于任意的实数 x，定义

$$F(x) = P(X \le x)$$

为 X 的**分布函数**。

性质 2.1.4.1 分布函数的性质

1. 单调性：$F(x)$ 是单调不减的；
2. 有界性：$0 \le F(x) \le 1$；
3. 左连续性：$F(x)$ 是左连续的：$F(x-0) = F(x)$；
4. 极限性：$\lim_{x\to-\infty} F(x) = 0$，$\lim_{x\to\infty} F(x) = 1$。

这些性质正好与概率的三条性质对应起来。

定义 2.1.5　离散型随机变量的概率分布
设 $\{x_i\}$ 是离散型随机变量 ξ 的所有可能取值，p_i 是其取到 x_i 的概率。则称

$$P(\xi = x_i) = p_i, i = 1, 2, 3, \cdots$$

为该离散型随机变量的**概率分布**。

2.2　几种经典的概率分布

定义 2.2.1　Bernoulli 试验
若随机试验只有两个可能的结果，且这两个结果发生的概率分别为 $P(A) = p$ 和 $P(\bar{A}) = 1 - p$，则称这种试验为 **Bernoulli 试验**。

定义 2.2.2　试验的独立性
类似于事件的独立性，我们定义：对 $1 \le i \le n$，我们进行试验 S_i，结果为事件 A_i，事件所属的事件域分别是 $\mathcal{A}_i$，若

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

就称这些试验 S_i 是相互独立的。

定义 2.2.3　n 重 Bernoulli 试验
重复进行 n 次相互独立的 Bernoulli 试验（所谓重复，就是上述的 $P(A) = p$ 的值是不随试验次数而改变的情况），这种试验称为 **n 重 Bernoulli 试验**。

性质 2.2.3.1
n 重 Bernoulli 试验的概率空间为：

$$(\omega_1, \cdots, \omega_i, \cdots, \omega_n)$$

其中的 ω_i 表示 A 或 $\bar{A}$，表示在第 i 次试验中 A 是否发生。
由此可以看出，一共有 2^n 个样本点，是一个有限样本空间。

定义 2.2.4　Bernoulli 两点分布

定义 2.2.5 二项分布
称 n 重 Bernoulli 试验中事件 A 发生的次数为 X，则称 X 服从**二项分布**，记作 $X \sim B(n,p)$，其中 n 为试验次数，p 为事件 A 发生的概率。
并且简记 $P(X=k)$ 为 $B(k;n,p)$.

定理 2.2.6
若随机变量 X 服从二项分布 $B(n,p)$，则

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}.$$

性质 2.2.6.1 二项分布的性质

定理 2.2.7 Poisson 近似
在独立重复试验中，随机变量 X 服从二项分布 $B(n,p_n)$,，p_n 是某次试验中事件 A 发生的概率，且 $np_n \to \lambda$. 则当 $n \to \infty$ 时，有

$$B(k;n,p_n) \to \frac{\lambda^k}{k!}\exp(-\lambda)$$

证明:

$$\begin{aligned} B(k;n,p_n) &= \binom{n}{k}p_n^k(1-p_n)^{n-k} \\ &= \frac{n(n-1)\cdots(n-k+1)}{k!}\cdot\frac{np_n}{n}(1-p_n)^{n-k} \\ &= \frac{(np_n)^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)\left(1-\frac{np_n}{n}\right)^{n-k} \\ &\lim_{n\to\infty} np_n=\lambda, \quad \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^{n-k}=\exp(-\lambda) \end{aligned}$$

取极限，得

$$\lim_{n\to\infty} B(k;n,p_n)=\frac{\lambda^k}{k!}\exp(-\lambda)$$

定义 2.2.8 Poisson 分布
如果随机变量 X 服从如下的分布

$$Ps(k;\lambda)=P(X=k)=\frac{\lambda^k}{k!}\exp(-\lambda)$$

就称其服从 **Poisson 分布**。记作 $X \sim Ps(\lambda)$.

定义 2.2.9 Poisson 过程

定义 2.2.10　几何分布

n 重 Bernoulli 试验中，只会出现“成功”和“失败”两种结果。