第一章 古典概型与概率测度的公理化

§1 古典概型

1．随机事件和概率

随机事件是在试验中可能出现也可能不出现的结局，而概率就是用数量来刻划随机事件发生的可能性的大小。事件 A 发生的概率记为 $P(A)$，习惯上有 $0\leqslant P(A)\leqslant 1$。对于空集，$P(\emptyset)=0$.

基本事件，是不能再细分的事件。事件可以表示为某些基本事件组成的集合。例如，投骰子，A_i={出现 i}，i =1，2，3，4，5，6 就是基本事件。

样本，就是一次观察试验可能出现的结果。样本空间，为所有样本点全体所组成的集合，记为 Ω。显然的我们有 $P(\Omega)=1$。以下我们将随机事件定义为样本点的某个集合，事件的运算就是集合的运算。

对于概率 P，应该满足以下几个条件：

(1)$P(A+B)=P(A)+P(B)$;

(2)$P(A^c)=1-P(A)$;

(3)$A\subseteq B\Longrightarrow P(A)\leqslant P(B)$;

2．古典概型

最简单的样本空间是元素有限的集合。而对每一个基本事件，最简单的情形就是它们发生的概率都相等。于是，我们得到了古典概型的数学描述：

(1)$\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$;

(2)$P(\omega_k)=1/n, k=1,2,\cdots,n$;

古典概型是离散样本空间的最常见的概率模型。

例 1.1 投掷一枚均匀的硬币，事件 $A=\{$ 国徽朝上 $\}$，求 $P(A)$。

解 该试验的结果有两个：国徽朝上和国徽朝下，它们是等可能的。于是，

$$P(A)=1/2.$$

例 1.2 检验 100 件产品，其中有 5 件是次品。任取 50 件，求其中恰有

两件是次品的概率。

解 从中取出 50 件共有 C_{100}^{50} 种方法，其中恰有 2 件次品的取法共有 $C_{95}^{48}C_5^2$ 种，于是，取出 50 件恰有两件是次品的概率为

$$\frac{C_{100}^{50}}{C_{95}^{48}C_5^2}=\frac{8}{25}.$$

例 1.3（生日问题）一个班上有 n 个学生，$A=\{$ 至少有两个人生日相同 $\}$，求事件 A 发生的概率。

解 直接求 A 发生的概率并不容易，所以我们先考虑 A 的补集 A^c 发生的概率。

$$P(A)=1-P(A^c)=1-\frac{A_{365}^n}{365^n}.$$

当 n=40，45，50，55 时，以上概率分别等于 0.89，0.94，0.97，0.99。这一结果与直观的印象并不相符。

古典概型中有一类问题，就是摸球问题。简单的说，就是一个容器中有 a 个红球，b 个黑球，每次从容器中摸出一个球，求 n 次中恰好有 k 个红球的概率 p_k。

对此，应该分为两种情况：有放回和无放回。

有放回时，每次的情况都相同，n 次中有 k 次摸到红球，$n-k$ 次摸到黑球。于是有

$$p_k=C_n^k\frac{a^kb^{n-k}}{(a+b)^n}=C_n^k\Big(\frac{a}{a+b}\Big)^k\Big(\frac{b}{a+b}\Big)^{n-k}.$$

无放回时，a 个红球中取出 k 个，b 个黑球中取出 $n-k$ 个，于是

$$p_k=\frac{C_a^kC_b^{n-k}}{C_{a+b}^n}.$$

上面的例 1.2 就属于无放回的摸球问题。

§2 几何概型

古典概型利用等概率，成功的计算了样本空间为有限的一些问题。但是当样本空间为无限时，古典概型就不能解决了。于是，我们引进几何概型，作为古典概型在无限情形的推广。

先看下面的一个例子。

例 2.1（约会问题）两人相约某天 5 点至 6 点在某地会面，先到者等候另一人 20 分钟，过时就离去。求这两个人能会面的概率。

解 以 x, y 分别表示两人到达的时刻，则会面的充要条件是 $|x-y|\leqslant 20$。记

$$G=\{(x,y)|0\leqslant x\leqslant 60, 0\leqslant y\leqslant 60\},$$

$$g=\{(x,y)||x-y|\leqslant 20, (x,y)\in G\}.$$

参见图 1-1，可知两人能会面的概率为

$$P(A)=\frac{S(g)}{S(G)}=\frac{60^2-40^2}{60^2}=\frac{5}{9}.$$

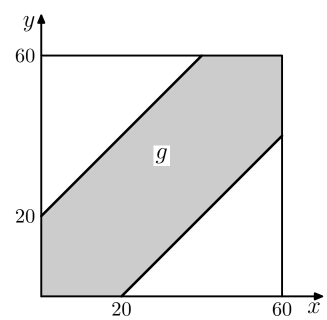


图 1.1

在这里，我们实际上假定了两人到达的时间在 5 点至 6 点之间的机会均等，而且互不影响。对于一般的情形，我们都可以用图形表示出，若全体试验结果的图形表示为 G，且任何两点的机会均等，而有利于事件 A 的试验结果的图形表示为 g，则事件 A 发生的概率为

$$P(A)=\frac{S(g)}{S(G)}.$$

这里的 $S(\cdot)$ 可以表示图形的面积，也可以表示长度，体积等。

下面再给出一个例子。

例 2.2（$Buffon$ 问题）平面上划着一些平行线，它们之间的距离都等于 a，向此平面上任意的投掷一长度为 $l(l\leqslant a)$ 的针，求此针与某一平行线相交的概率。

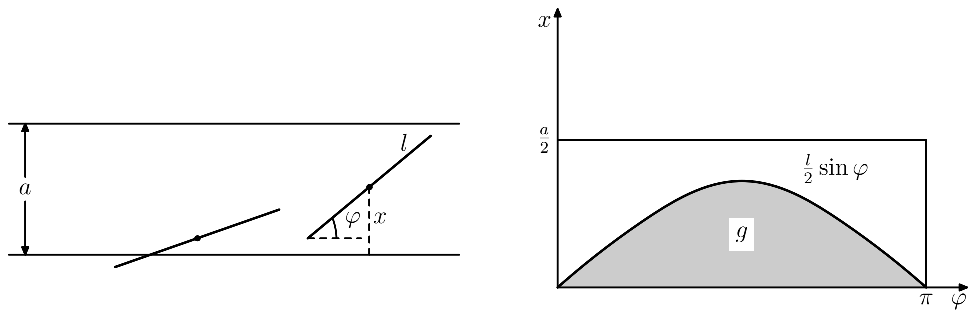


图 1.2

解 以 x 表示针的中点到最近的一条平行线的距离，φ 表示该平行线与针的夹角（见图 1-2），显然有

$$0 \leqslant x \leqslant \frac{a}{2}, 0 \leqslant \varphi \leqslant \pi,$$

因此，

$$G = \{(\varphi, x) | 0 \leqslant x \leqslant \frac{a}{2}, 0 \leqslant \varphi \leqslant \pi\},$$

而当针与平行线相交时，有 $x \leqslant \frac{l}{2}\sin\varphi$，所以有

$$g = \{(\varphi, x) | x \leqslant \frac{l}{2}\sin\varphi, (\varphi, x) \in G\}.$$

这样就可以得到

$$P(A) = \frac{S(g)}{S(G)} = \frac{\int_0^\pi \frac{l}{2}\sin\varphi d\varphi}{\frac{a}{2}\pi} = \frac{2}{a}\frac{l}{\pi}.$$

在看接下来的这个问题。

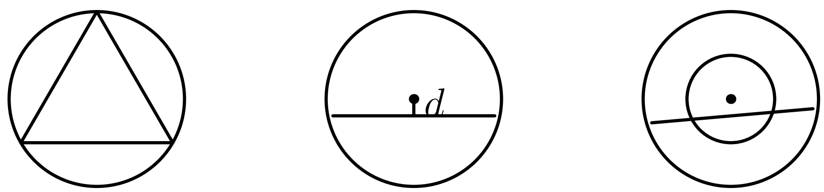
例 3.3（$Berfrand$ 悖论）在半径为 1 的圆内随机的取一条弦，问其长度超过该圆内接等边三角形的边长 $\sqrt{3}$ 的概率是多少？

这是一个几何概型的问题，但是对“随机地”一词的含义的不同解释，可以得到不同的答案。下面给出其中的三种。

解 1 任何弦都和圆周交于两点。不失一般性，先固定其中一点在圆周上，以此点为一顶点作圆的内接等边三角形，将圆周分为三段等长的弧。显然，只有当弦的另一顶点落在与该顶点不相邻的一段弧上的时候，弦长才会超过 $\sqrt{3}$。故所求的概率为 $\frac{1}{3}$。见图 1-3。

解 2 弦长只跟它的弦心距有关，与方向无关，于是可以假设它垂直于某一条直径。对于这样的弦，其长度大于 $\sqrt{3}$ 当且仅当它的弦心距小于 $\frac{1}{2}$。因此所求的概率为 $\frac{1}{2}$。

解 3 弦被其中点唯一决定，故当且仅当其中点属于半径为 $\frac{1}{2}$ 的同心圆时，弦长才大于 $\sqrt{3}$。而此小圆面积为大圆面积的 $\frac{1}{4}$，故所求的概率为 $\frac{1}{4}$。



(a) (b) (c)

图 1.3

同一问题得到三种不同的答案，原因是采用了不同的等可能性。在解 1 中，假定一端固定而另一端在圆周上均匀分布；解 2 中假设弦的中点在一固定直径上均匀分布；而解 3 则假定弦的中点在圆内均匀分布。因此，三种不同的答案是针对三种不同的随机试验（尽管看起来相同），对于各自的随机试验而言，它们都是正确的。

因此在使用“随机”、“任取”、“等可能”、“均匀”时，应该明确其含义。

§3 概率空间与概率测度公理化

前两节给出了概率直观上的定义，接下来将更精确的描述概率测度和概率空间。

前面已经讲过，事件就是某些样本点组成的集合，事件之间的运算也就是集合运算。但是，并没有对集合进行限制。对于事件，一个很明显的要求，就是所有事件组成的集合对于并 ($A \cup B$)、交 ($A \cap B$ 或记为 AB)、余 (A^c) 等三种运算封闭。前苏联学者柯尔莫哥洛夫于 1933 年在《概率论基础概念》一书中，用公理化的方法与集合论的观点成功地解决了这个问题，提出了概率空间的概念。

1 概率空间及其三要素

(1) 样本空间 Ω

Ω 是一非空集合，为样本空间。

(2) $\mathcal{F}$ 与可测空间

$\mathcal{F}$ 是由 Ω 的某些子集所构成的一个 σ 代数，即它满足：

(1) $\Omega \in \mathcal{F}$;

(2) 若 $A \in \mathcal{F}$，则 $A^c \in \mathcal{F}$;

(3) 若 $A_n \in \mathcal{F}$，$n =1$，2，$\cdots$，则 $\bigcup\limits_{n=1}^{\infty} A_n \in \mathcal{F}$;

称 $(\Omega$，$\mathcal{F})$ 为可测空间。

(3) P 与概率空间

概率 P 是定义在 $\mathcal{F}$ 上的函数，即它是一个从 $\mathcal{F}$ 到 $[0,1]$ 的映射：$P:\ \mathcal{F} \longrightarrow [0,1]$，且它满足：

(1) $\forall A \in \mathcal{F}$，$P(A) \geqslant 0$;

(2) $P(\Omega) = 1$;

(3) 若 $A_n \in \mathcal{F}$，$n =1$，2，$\cdots$，且 A_n 互不相交，则

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n);$$

称这样的 P 为可测空间 $(\Omega, \mathcal{F})$ 上的一个概率测度（简称为概率），称 $(\Omega, \mathcal{F}, P)$ 为一概率空间。

可以推出，概率测度 P 有以下性质：

(1) $P(\emptyset) = 0$;

(2) 若 $A_k \in \mathcal{F}$，$k =1$，2，$\cdots$，n，且 A_k 互不相交，则

$$P(\bigcup_{k=1}^{n} A_k) = \sum_{k=1}^{n} P(A_k);$$

(3) 若 $P(A^c) = 1 - P(A)$;

(4) $P(A - B) = P(A) - P(B)$;

(5) 若 $A \in B$，则 $P(A) \leqslant P(B)$;

(6) $P(A \cup B) \leqslant P(A) + P(B)$;

(7) $P(\bigcup_{k=1}^{n} A_k) \leqslant \sum_{k=1}^{n} P(A_k)$;

(8) 若 $A_1 \in A_2 \in \cdots$, 则

$$P(\bigcup_{n=1}^{\infty}) = \lim_{n\to\infty} P(A_n);$$

2 加法公式

由以上概率的性质可以得到加法公式的简单情形：对任意两个事件 A, B, 有

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

简要证明如下：由 $A \cup B = A \cup BA^c$, $B = BA \cup BA^c$, 且 A 与 BA^c 不相交, BA 与 BA^c 不相交, 有

$$P(A \cup B) = P(A) + P(BA^c);$$

$$P(BA^c) = P(B) - P(BA);$$

两式相加即得所要的公式。

对于多个集合的情形，我们有下面的 $Jordan$ 公式：

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1\leqslant i_1 < i_2 < \cdots < i_k \leqslant n} P(A_{i_1} A_{i_2} \cdots A_{i_k}).$$

用容斥原理或者数学归纳法均可给出证明。

下面给出一个例子。

例 3.1 （配对问题）n 对夫妇随机入座坐在长桌的两侧，先生全部坐在其中的一侧，问至少有一对夫妇面对面坐的概率是多少？

解 记

$B =$ 至少有一对夫妇面对面坐,

A_i= 第 i 对夫妇面对面坐, i =1, 2, ···, n。

易知 $B = \bigcup_{i=1}^{n} A_i$。对任意的 $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$, 有

$$P(A_{i_1} A i_2) \cdots A_{i_k} = \frac{(n-k)!}{n!},$$

$$\sum_{1\leqslant i_1 < i_2 < \cdots < i_k \leqslant n} P(A_{i_1} A_{i_2} \cdots A_{i_k}) = C_n^k \frac{(n-k)!}{n!} = \frac{1}{k!},$$

于是，由 $Jordan$ 公式可知

$$P(B) = \sum_{k=1}^{n} (-1)^{k+1} \frac{1}{k!}.$$

当 $n \to \infty$ 时，$P(B) = \sum_{k=1}^{n} (-1)^{k+1} \frac{1}{k!} \to 1 - e^{-1}$。