§1.5　概率空间

一　走向概率论公理化体系

在概率论的发展的历史上，曾有过概率的古典定义、概率的几何定义、概率的频率定义、概率的主观定义．在现实世界里有一些随机现象是不能重复的或不大量重复的，这时有关事件的概率是如何确定的？

统计界的贝叶斯派认为：一个事件的概率是人们根据经验对该事件发生的可能性所给出的个人信念，这样给出的概率称为主观概率．这种利用经验确定随机事件发生可能性大小的例子是很多的，人们也常根据这些主观概率来行事．

例如：在气象预报中，往往说:“明天下雨的概率为80%”，一个外科医生根据自己多年的临床经验和患者的病情，认为“手术成功的可能性为90%”. 主观概率不是主观臆造，主观概率要求当事人对所考察的事件有透彻的了解和丰富的经验. 能对历史信息和当前信息进行仔细分析，如此确定的主观概率是可信的，从某种意义上说，不利用这些丰富的经验也是一种浪费. 对于无法大量重复出现的随机现象，用主观概率方法去做决策和判断是适合的，主观方法至少是频率方法的一种补充.

概率的上述四种定义，各适用于特定的随机现象. 那么如何给出适合一切随机现象的概率的最一

般定义呢？ **1900**年数学家希尔伯特（**1862-1943**）提出要建立概率的公理化定义以解决这个问题，即以最少的几条本质特性出发去刻画概率的概念. **1933** 年, 前苏联数学家科尔莫哥洛夫首次提出概率的公理化定义，这个定义概括了历史上几种概率定义的共同特性，又避免了各自的局限性和含糊之处，不管什么随机现象，只有满足定义中的三条公理，才能说它是概率. 这一公理化体系是概率论发展的一个里程碑.

二　事件域

前面我们知道, 事件都是样本空间的某个子集. 在几何概型中, 不能把不可测集看作为是事件，那么样本空间的那些子集可以看作是事件？

我们把全体事件构成的集合记为$\mathbb{F}$，要求$\mathbb{F}$满足

(1)　$\Omega \in \mathbb{F}$;

(2)　若$A \in F$,则$\overline{A} \in \mathbb{F}$;

(3)　若$A_i \in \mathbb{F}, i = 1, 2, \cdots$,则$\bigcup_{i=1}^{\infty} A_i \in \mathbb{F}$.

满足上述3条规定的子集类叫做σ域或σ代数. 在概率论中我们称为事件域，σ域具有很多没有写出的一些性质，这些性质足以保证它能满足我们研究概率的需要. 例如：

(4)　$\phi = \overline{\Omega} \in \mathbb{F}$;

(5)　若$A_n \in \mathbb{F}$，$n = 1, 2, \cdots$.　　则

$$\bigcap_{n=1}^{\infty} A_n = \overline{\bigcup_{n=1}^{\infty} \overline{A}_n} \in \mathbb{F};$$

$$\bigcup_{n=1}^{k} A_n = A_1 \cup A_2 \cup \cdots \cup A_k \cup \phi \cup \phi \cup \cdots \in \mathbb{F}$$

$$\bigcap_{n=1}^{k} A_n = A_1 \cap A_2 \cap \cdots \cap A_k \cap \Omega \cap \Omega \cap \cdots \in \mathbb{F}$$

$$A_1 - A_2 = A_1 \overline{A}_2 \in \mathbb{F}$$

一些常见的σ域：

$$\mathbb{F} = \{\Omega,\ \phi\}; \qquad \mathbb{F} = \{\Omega,\ \phi,\ A,\ \overline{A}\}; \qquad \mathbb{F} = 2^{\Omega};$$

$$\mathbb{F} = \{\Omega,\ \phi,\ A,\ \overline{A},\ B,\ \overline{B},\ AB,\ A\overline{B},\ \overline{A}B,\ \overline{A}\,\overline{B},$$

$$A \cup B, \overline{A} \cup B,\ A \cup \overline{B},\ \overline{A} \cup \overline{B},\ AB \cup \overline{A}\,\overline{B},\ A\overline{B} \cup \overline{A}B\}.$$

一维博雷尔σ域：以后我们将以R^1记实数全体，并称由一切形如$[a,b)$的有界左闭右开的开区间构成的集类所产生的σ域为一维博雷尔域，该σ域中的任一元素为一维博雷尔点集.

若x,y为任意实数，由于

$$\{x\}=\bigcap_{n=1}^{\infty}[x,x+\frac{1}{n});\quad (x,y)=[x,y)-\{x\};$$

$$[x,y]=[x,y)+\{y\};\quad (x,y]=[x,y)-\{x\}+\{y\}.$$

因此，一维博雷尔σ域包含一切开区间，闭区间，单个实数，可列个实数，以及由它们经过可列次逆、并、交运算而得到的集合.这是一个相当大的集类.

n维博雷尔σ域：以R^n记 n 维欧几里得空间，称由一切 n 维矩形产生的域为 n 维博雷尔σ域.

三 概率及其性质

假定Ω是一个样本空间，已经在上面给出了一个事件σ域$\mathbb{F}$, 于是我们有了可测空间(Ω, $\mathbb{F}$). 我们只需对$\mathbb{F}$中的集合（事件）定义概率. 由于概率是描述事件发生可能性大小的一个量，是定义在事件域上的集函数，它必须满足如下三条性质：

（1）非负性　　对任何事件A，　有　$P(A)\geq 0$.

（2）规范性　　$P(\Omega)=1$.

（3）可列可加性　若$A_1,A_2,\cdots$两两互不相容，则

$$P(\sum_{n=1}^{+\infty}A_n)=\sum_{n=1}^{+\infty}P(A_n).$$

利用概率的上述基本性质，我们可得到概率的另外一些重要性质.

（4）$P(\phi)=0$

（5）有限可加性 若事件$A_1,A_2,\cdots,A_n$两两互不相容，则　$P(A_1+A_2+\cdots+A_n)=P(A_1)+\cdots+P(A_n)$.

（6）可减性　如果$B\supset A$，则

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

（7）单调性　如果$B \supset A$，则　$P(B) \geq P(A)$;

（8）$P(\bar{A}) = 1 - P(A)$;

（9）加法公式

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

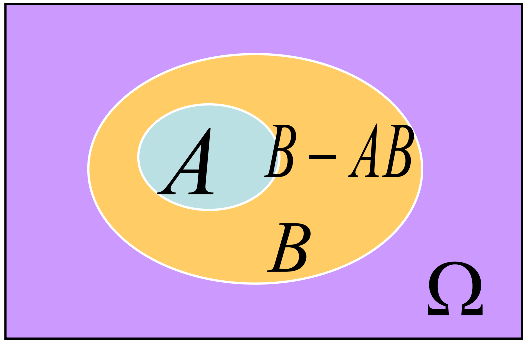
（10）一般加法公式

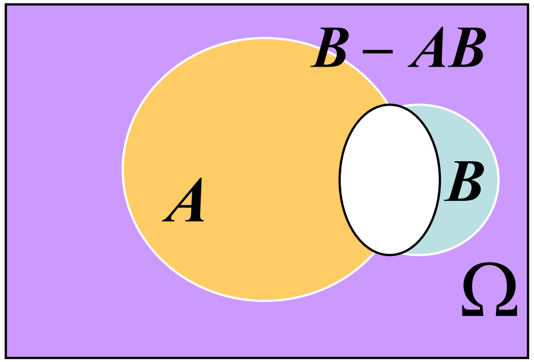
$$P(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{1 \leq i \leq n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i A_j)$$

$$+ \sum_{1 \leq i < j < k \leq n} P(A_i A_j A_k) - \cdots + (-1)^{n-1} P(A_1 A_2 \cdots A_n);$$

（11）下连续性 若$A_1 \subset A_2 \subset \cdots \subset A_n \subset \cdots$，则

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





可减性　　如果$B \supset A$, 则

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

加法公式

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

（12）上连续性 若$A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$，则

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

可列可加性与有限可加性之间的相互关系：

定理1.5.1 若P是$\mathbb{F}$上满足$P(\Omega)=1$的非负集合函数，则具有可列可加性的充要条件是

（1）它是有限可加的；

（2）它是下连续的.

证明： 必要性即是概率的性质（5），（11）.

充分性

若$A_1, A_2, \cdots$两两互不相容，由有限可加性得

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

因为级数$\sum_{n=1}^{\infty} P(A_n)$是正项级数．显然有

$$A_1 \subset A_1 + A_2 \subset \cdots \subset A_1 + A_2 + \cdots + A_n \subset \cdots$$

根据下连续性得

$$P(\sum_{n=1}^{+\infty} A_n) = \lim_{n\to+\infty} P(\sum_{i=1}^{n} A_i) = \lim_{n\to+\infty} \sum_{i=1}^{n} P(A_i) = \sum_{n=1}^{\infty} P(A_n).$$

此即为可列可加性.

$$
\begin{tikzcd}
\sum_{i=1}^{n} A_i \arrow[r] \arrow[d] & \sum_{i=1}^{\infty} A_i \arrow[d] \\
\sum_{i=1}^{n} P(A_i) \arrow[r] & \sum_{i=1}^{\infty} P(A_i)
\end{tikzcd}
$$

四　概率空间

称三元总体$(\Omega,\mathbb{F},\mathrm{P})$为概率空间. 其中$\Omega$是样本空间，$\mathbb{F}$是事件域，P是概率，它们都认为是预先给定的，并以此作为出发点讨论各种问题. 至于实际问题中, 如何选定Ω, 如何构造$\mathbb{F}$, 怎样定义P, 则要视具体情况而定.

最常见的几种概率空间

（有限概率空间）设 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$, 此时一般把$\Omega$的任意子集看成是事件，它是事件域. 每个样本点作为单点集也是事件. 在这种样本空间中引

进概率，只需要对每个样本点定义了概率就可以了，记之为$P(\omega_i)$，它是非负的，而且满足

$$P(\omega_1)+P(\omega_2)+\cdots+P(\omega_n)=1$$

至于$P(\omega_i)$的具体数值，要根据随机现象的特点而定. 例如在古典概型中，$P(\omega_i)=1/n$. 抛掷均匀的骰子是古典概型问题，但抛掷不均匀的骰子就不是古典概型问题，此时$P(\omega_i)$必须定义其他的数值. 但是，一旦所有的$P(\omega_i)$给定以后，任何事件A的概率$P(A)$已经确定了，$P(A)$是A中各样本点的概率之和.

（离散概率空间） 设$\Omega=\{\omega_1,\omega_2,\cdots,\omega_n,\cdots\}$，此时我们一般把$\Omega$的任意子集看成是事件，它是事件域.每个样本点作为单点集也是事件.在这种样本空间中引进概率，只需要对每个样本点定义了概率$P(\omega_i)$就可以了,它是非负的，而且满足

$$\sum_{i=1}^{\infty}P(\omega_i)=1$$

任何事件A的概率是A中各样本点的概率之和.

若$\Omega=R^1$，这时通常取$\mathbb{F}$为直线上的博雷尔点集的全体.

若$\Omega=R^n$(或R^n的一部分),这时通常取$\mathbb{F}$为R^n上的n维博雷尔点集的全体.