4、利用事件间的关系及概率的性质：(逆事件公式、加法公式、全概率公式、Bayes 公式等)。

例：生日问题。

例㉕(匹配问题). 某人一次写了n封信，又写了n个信封。如果他任意地将n张信纸装入n个信封中。求：

(1) 至少有一封信的信纸和信封是一致的概率；

(2) n封信均未配对的概率；

(3) 恰有m封信配对的概率。

第一章　事件与概率

第5节　概率空间

李 娟

Email: juanlisdu@163.com

山东大学(威海)　　数学与统计学院

授课要求：1. 掌握概率的定义与性质；
2. 掌握概率空间的定义。
3. 理解一维（n维）Borel域的定义。

一、走向概率论公理化结构（略）

1933年，前苏联数学家Kolmogorov提出了概率论公理化结构，使概率论成为严谨的数学分支。

二、事件域

分析：一般并不把Ω的一切子集都作为事件，因为这将对给定概率带来困难，譬如在几何概率中，若把不可测集也作为事件，将带来不可克服的困难。

另一方面，又必须把问题中感兴趣的事件都包括进来，例如，若A是事件，则应要求$\bar{A}$也是事件；若A与B是事件，则$A\cup B$及AB也应是事件。当样本空间Ω 由无限多个点构成时–在几何概率中就是如此–显然还得考虑可列个事件的并与交。此外，把Ω及ϕ作为事件有很大方便。

1、σ域（σ-代数）

记$\mathcal{F}$是由Ω的一些子集构成的集类，且满足：

（i）$\Omega \in \mathcal{F}$;

（ii）若$A \in \mathcal{F}$，则$\bar{A} \in \mathcal{F}$;

（iii）若$A_n \in \mathcal{F}$，$n = 1, 2, \ldots$，则$\cup_{n=1}^{\infty} A_n \in \mathcal{F}$。

一般地，称空间Ω上满足上述三个要求的集类为σ域，亦称σ代数。

注1: 若$\mathcal{F}$为σ域，则由（i）及（ii）可得：$\phi \in \mathcal{F}$;

$= \bar{\Omega}$

注2: 若$A_n \in \mathcal{F}$，$n = 1, 2, \ldots$，则$\cap_{n=1}^{\infty} A_n$，$\cup_{n=1}^{k} A_n$，$\cap_{n=1}^{k} A_n \in \mathcal{F}$.

证明：

$\bigcap_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \bar{A}_n \in \mathcal{F}$

$\Downarrow$

$A_1 A_2 \cdots A_k \, \Omega \, \Omega \, \Omega, \ldots$

$A_1 A_2 \cdots A_k \, \phi \, \phi \, \phi \ldots$

二、事件域

2、定义1.5.1 若$\mathcal{F}$是由样本空间Ω的一些子集构成的一个σ域，则称它为事件域（event field），$\mathcal{F}$中的元素称为事件，Ω称为必然事件，ϕ称为不可能事件。

注: 值得指出，按照这种定义，样本点并不一定是事件。

下面我们来举一些事件域的例子。

例1. $\mathcal{F}=\{\phi,\Omega\}$，不难验证$\mathcal{F}$是一个$\sigma$域，这时只有必然事件$\Omega$与不可能事件$\phi$是事件。 平凡的 σ 域

例2. $\mathcal{F}=\{\phi,\Omega,A,\bar{A}\}$，这时$\mathcal{F}$也是一个$\sigma$域，$\Omega$，$\phi$，$A$，$\bar{A}$是事件。

例3. $\Omega=\{\omega_1,\ldots,\omega_n\}$，$\mathcal{F}$由$\Omega$的一切子集构成，因此有：

$$C_n^0+C_n^1+\cdots+C_n^{n-1}+C_n^n=2^n$$

个元素，不难验证，$\mathcal{F}$是一个σ域。

例4. 对于一般的Ω，若$\mathcal{F}$由Ω的一切子集构成，可以验证$\mathcal{F}$是一个σ域。

二、事件域

从上面几个例子中看到，事件域可以很简单，也可以选得十分复杂，这就需要我们根据问题的不同要求来选择适当的事件域。

表面看，当Ω确定后，把事件域$\mathcal{F}$选得越大，能处理的事件越多，就越方便。但是概率论最关心的毕竟是概率，过大的事件域对概率的给定带来困难，并不可取。不过，如果定义概率没有困难，那么，事件域当然可以尽量选大。因此对有限样本空间和离散样本空间，我们将看到，通常都取Ω的一切子集作为事件域。

对一个试验，当Ω给定后，总有些子集必须作为事件处理，但它们未必能满足σ域的要求，如何处理？

二、事件域

3、定理 若给定Ω的一个非空集类$\mathcal{G}$，必存在唯一的一个Ω上的σ域$m(\mathcal{G})$，具有如下性质：

（1）包含$\mathcal{G}$；

（2）若有其他σ域包含$\mathcal{G}$，则必包含$m(\mathcal{G})$。

这个$m(\mathcal{G})$称为包含$\mathcal{G}$的最小σ域，亦称由$\mathcal{G}$产生的σ域。

证明：

设 $\mathcal{A} = \{F \mid F$ 为σ域，且 $F \supset \mathcal{G}\}$ (非空)

$m(\mathcal{G}) = H = \bigcap_{F\in\mathcal{A}} F$ 为σ域，满足(1)(2)

因此从必须作为事件处理的子集出发，通过添加其它子集，必能得到Ω上的σ域。不过，上述最小σ域因便于给定概率而受重视。

二、事件域

例① 按照这种观点，例1是只把不可能事件ϕ及必然事件Ω看作事件的平凡事件域；而例2是由事件A产生的事件域。

$g = A_0$ $m(g)$

例② 当Ω有限或可列，如果要求每一个样本点都是事件，则包含它的最小σ域就是Ω的一切子集（如例3）。因此，在这两种场合，事件域的选取实际上没有困难。

例③ 真正要关心的是样本空间为一维或n维Euclid（欧几里德）空间的场合。这里的许多结果是由Borel（1871-1956，博雷尔）建立的。

二、事件域

至于如何在函数空间上定义事件域则超出本书的讨论范围。下面介绍两个非常有用的σ域。

1、一维Borel点集

$\mathcal{G}_0 = \{[a,b) \mid \forall a,b \in \mathbb{R}\}$

以$\mathbb{R}^1$记数直线或实数全体，并称由一切形为$[a,b)$的有界左闭右开区间构成的集类所产生的σ域为一维Borelσ域，记为$\mathcal{B}_1$，称$\mathcal{B}_1$中的集为一维Borel点集。$\mathcal{B}_1 = m(\mathcal{G})$

注1: $\mathcal{B}_1$中包含一切开区间，闭区间，单个实数，可列个实数，以及由它们经可列次逆、交、并运算而得出的集合。这是相当大的一个集类，足够把实际问题中感兴趣的点集都包括在内。

注2: 显然，若不从左闭右开区间$[a,b)$出发，而从(a,b)或$(a,b]$，或$[a,b]$，甚至$(-\infty,x)$出发，都将产生同一个σ域。

$\forall a$, $\{a\} = \bigcap_{n=1}^{\infty} [a, a+\frac{1}{n})$

$\forall$ $(a,b) = [a,b) - \{a\}$

2、n维Borel点集

以$\mathbb{R}^n$记n维欧式空间，可以类似地定义n维Borel点集，它们是由一切n维矩形产生的n维Borelσ域$\mathcal{B}_n$中的集合，也可以把$\mathbb{R}^n$中我们感兴趣的点集都包括在内。

三、概率

（一）定义

1. 集合函数: 在公理化结构中，概率是针对事件定义的，即对应于事件域$\mathcal{F}$中的每一个元素A有一个实数$P(A)$与之对应。一般把这种从集合到实数的映射称为集合函数。

注: 因此，概率是定义在事件域$\mathcal{F}$上的一个集合函数。

此外，在公理化结构中只规定概率应满足的性质，而不具体给出它的计算公式或计算方法。概率应有什么性质呢?

分析：统计概率、古典概型；几何概型。在一般场合，处理可列个事件之和是完全必要的，因此保留这种可列可加性要求看来是合理的。综上所述，我们如下定义概率。

三、概率

2. 定义1.5.2. 定义在事件域$\mathcal{F}$上的一个集合函数P称为概率，若它满足如下三个要求：

（i）$P(A) \geq 0$，对一切$A \in \mathcal{F}$；

（ii）$P(\Omega) = 1$；

（iii）若$A_i \in \mathcal{F}, i = 1, 2, \cdots$，且两两互不相容，则

$$P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i). \tag{1.5.1}$$

注：性质（i）称为非负性；性质（ii）称为规范性；性质（iii）称为可列可加性或完全可加性。

利用概率的基本性质（i),（ii),（iii）可以推出概率的另外一些重要性质。

三、概率

（二）性质

$\Omega = \Omega + \phi + \phi + \cdots$

$\Rightarrow P(\Omega) = 1 = P(\Omega) + \sum P(\phi)$

性质1. 不可能事件的概率为0，i.e., $P(\phi) = 0$. $\Rightarrow P(\phi) = 0$

注: 不可能事件的概率为0，但逆命题不一定成立。

性质2. 概率具有有限可加性，i.e. ,若$A_iA_j = \phi$, $i \neq j$，则

∞

$$P(\sum_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i). \quad + \phi + \phi + \cdots \qquad (1.5.2)$$

注: 这一性质说明，由概率的完全可加性可以推得概率的有限可加性，但反之不一定成立。

三、概率

性质3. 对任何事件A有$P(\bar{A}) = 1 - P(A)$.（1.5.3）

$\forall A \in \mathcal{F}$

性质4. 若$A \supset B$，则$P(A - B) = P(A) - P(B)$.（1.5.4）

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

推论1.（单调性） 若$A \supset B$，则$P(A) \geq P(B)$.

$P(A-B) = P(A\bar{B}) = P(A) - P(AB)$

注1: 由此即知，对任意事件A，有$P(A) \leq 1$，再注意到概率的非负性，所以成立$0 \leq P(A) \leq 1$.（1.5.5）

注2: 任意的A, $B \in \mathcal{F}$，$P(A - B) = P(A) - P(AB)$.

三、概率

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

性质5. $P(A\cup B)=P(A)+P(B)-P(AB).$ （1.5.6）

注: 公式（1.5.6）称为概率的加法公式。

推论2.（布尔不等式） $P(A\cup B)\leq P(A)+P(B).$ （1.5.7）

推论3.（Bonferroni不等式） $P(AB)\geq P(A)+P(B)-1.$ （1.5.8）

注1: 利用归纳法不难把这两个不等式推广到n个事件的场合。

①$P(A_1\cup A_2\cup\cdots\cup A_n)\leq P(A_1)+P(A_2)+\cdots+P(A_n);$

②$P(A_1A_2\cdots A_n)\geq P(A_1)+P(A_2)+\cdots+P(A_n)-(n-1).$

注2: $P(A_1\cup A_2\cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1A_2)$
$-P(A_1A_3)-P(A_2A_3)+P(A_1A_2A_3).$

一般地有：

性质6（一般加法公式）若$A_1,\cdots,A_n$为n个事件，则：

$$P(A_1\cup A_2\cup\cdots\cup A_n)=\sum_{i=1,\cdots,n}P(A_i)-\sum_{\substack{i<j\\ i,j=1,\cdots,n}}P(A_iA_j)+\sum_{\substack{i<j<k\\ i,j,k=1,\cdots,n}}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n). \qquad (1.5.9)$$

公式（1.5.9）可以用数学归纳法证明。

（三）概率计算的公式法

$A=\{\cdots=k\}$

利用上面推导的公式来作概率计算，常能使解题思路清晰，计算便捷。

$B_k=\{\cdots\le k\}$ $P(A)=P(B_k-B_{k-1})$ $B_{k-1}\subset B_k$ $\Longrightarrow$ $P(B_k)-P(B_{k-1})$

例5（最大车牌号）某城有N辆卡车，车牌号从1到N，有一个外地人到该城去，把遇到的n辆车子的牌号抄下（可能重复抄到某些车牌号），求抄到的最大号码正好为k 的概率（$1\le k\le N$）.

这种方法曾在第二次世界大战中被盟军用来估计敌方的军火生产能力，从被击毁的战车上的出厂号码推测其生产批量，得到相当准确的有用情况。

例6（匹配问题）某人写好n封信，又写好n只信封，然后在黑暗中把每封信放入一只信封中，试求至少有一封信放对的概率。

四、可列可加性与连续性

下面我们对可列可加性作进一步讨论，从性质2知道，由可列可加性可以推出有限可加性，但是一般来讲，由有限可加性并不能推出可列可加性。

分析：为了具有可列可加性，还需要下式成立：

$$\lim_{n\to\infty} P(\sum_{i=1}^{n} A_i) = P(\sum_{i=1}^{\infty} A_i). \tag{1.5.12}$$

或者写成更富有启发性的等式：

$$\lim_{n\to\infty} P(\sum_{i=1}^{n} A_i) = P(\lim_{n\to\infty} \sum_{i=1}^{n} A_i). \tag{1.5.13}$$

即要允许把极限号移到概率号里面去，这就提出了一个新要求。

四、可列可加性与连续性

现在就来考察这个新要求。若记$S_n = \sum_{i=1}^{n} A_i$，则$S_n \in \mathcal{F}, n = 1, 2, \cdots$ 而且$S_n \subset S_{n+1}$，i.e.，S_n是$\mathcal{F}$中一个单调不减的集序列，这时可改写（1.5.13）式成为下式：

$$\lim_{n\to\infty} P(S_n) = P(\lim_{n\to\infty} S_n). \tag{1.5.14}$$

一般，对于$\mathcal{F}$上的集合函数，若它对$\mathcal{F}$中任何一个单调不减的集序列$\{S_n\}$均成立（1.5.14）式，则我们称它是下连续的。因此，前面的推导表明，为保证概率的可列可加性成立，除要求它具有有限可加性外，还要求它是下连续的。

四、可列可加性与连续性

定义 对于$\mathcal{F}$上的取有限值的集合函数Φ，若对$\mathcal{F}$中的任一单调不减的集合序列$\{A_n\}$，有$\lim\limits_{n\to\infty}\Phi(A_n)=\Phi(\lim\limits_{n\to\infty}A_n)$，则称集合函数$\Phi$在$\mathcal{F}$上是下连续的，其中$\lim\limits_{n\to\infty}A_n=\bigcup\limits_{n=1}^{\infty}A_n$.

若对$\mathcal{F}$中的任一单调不增的集合序列$\{A_n\}$，有$\lim\limits_{n\to\infty}\Phi(A_n)=\Phi(\lim\limits_{n\to\infty}A_n)$，则称集合函数$\Phi$在$\mathcal{F}$上是上连续的，其中$\lim\limits_{n\to\infty}A_n=\bigcap\limits_{n=1}^{\infty}A_n$。

四、可列可加性与连续性

下面的定理明确地阐明了这三个概念之间的关系。

Th1.5.1 若P是$\mathcal{F}$上满足$P(\Omega)=1$的非负集合函数，则它具有可列可加性的充要条件为：

（i）它是有限可加的；（ii）它是下连续的。

系1 概率是下连续的。

系2 概率是上连续的，i.e.，若$B_i \in \mathcal{F}$，而且$B_i \supset B_{i+1}, i=1,2,\cdots$，则$\lim\limits_{n\to\infty} P(B_n) = P(\lim\limits_{n\to\infty} B_n)$. （1.5.15）

四、可列可加性与连续性

系3 $P(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} P(A_i)$. 次可列可加性

系4 若P为$\mathcal{F}$上满足$P(\Omega)=1$的非负集合函数，则它具有可列可加性的充要条件为：

（i）它是有限可加的；（ii）它是次可列可加的。

$\pi R \cdot 2\pi x = 4\pi R^2$

$x = \frac{4\pi R^2}{2\pi^2 R} = \frac{2}{\pi} R$

五、概率空间$(\Omega, \mathcal{F}, P)$

在科尔莫戈罗夫的概率论公理化结构中，称三元总体$(\Omega, \mathcal{F}, P)$为**概率空间**，其中Ω是样本空间，$\mathcal{F}$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种种问题。至于问题中，如何选定Ω，怎样构造$\mathcal{F}$，怎样给定P，则要视具体情况而定。

下面讨论几个具体例子。

例7（有限概率空间） 设Ω中只有n个点。在这种场合，一般可取$\mathcal{F}$为Ω的所有子集全体，这仍是一个有限的集合，元素总数为2^n，它满足事件域的三个要求，而且样本点（看作一个单点集）是事件。至于概率，只要对样本点$\omega_i, i = 1, 2, \cdots, n$，给定满足：

例7（有限概率空间）（续）

$p(\omega_i) \geq 0,\ i = 1, 2, \cdots, n$；$p(\omega_1) + p(\omega_2) + \cdots + p(\omega_n) = 1$的一组数 $p(\omega_1), p(\omega_2), \cdots, p(\omega_n)$，则若$A$是$\mathcal{F}$中元素，包含样本点 $\omega_{i_1}, \cdots, \omega_{i_k}$，则由概率的可加性，自然应令：

$$P(A) = p(\omega_{i_1}) + \cdots + p(\omega_{i_k}).$$

这就给定了事件A的概率，从而构成了概率空间$(\Omega, \mathcal{F}, P)$，这时显然有：$P(\{\omega_i\}) = p(\omega_i), i = 1, 2, \cdots, n.$

从这个例子中看到下面两点。

五、概率空间（$\Omega, \mathcal{F}, P$）

（1）选定了$(\Omega, \mathcal{F})$之后，对于事件概率的给定还有相当大的灵活性，这表现在$p(\omega_i)$的选取上。因为只有这样，才能用概率空间来描述不同的随机现象。例如，在投一次硬币的试验中，Ω总是由出正面(ω_1)及出反面(ω_2)两个样本点构成。对于均匀的硬币，可以假定它出正面及反面的概率均为$\frac{1}{2}$；但对于很不均匀的硬币，例如出正面可能性大得多的硬币，则必须给定另外的概率，而这只须适当给定$p(\omega_1)$就可以了。

（2）一旦诸$p(\omega_i)$给定后，事件A的概率并不能任意给定，i.e.，在事件域中，各事件的概率有一定关系，给定概率时必须满足这些关系。

例8（离散概率空间） Ω由可列个点构成：$\Omega = \{\omega_1, \omega_2, \cdots\}$，这时$\mathcal{F}$还是可以选为$\Omega$的子集全体，它满足事件域的三个要求。这时样本点也是事件。为给定概率，可选择可列个非负的数$p_i, i = 1, 2, \cdots$，满足$\sum_{i=1}^{\infty} p_i = 1$，分别作为样本点$\omega_i$的概率，而一般事件$A \in \mathcal{F}$的概率，则必须取为它所含的样本点的概率之和。

例9 若$\Omega = \mathbb{R}^1$，即样本空间由全体实数构成，这时，$\mathcal{F}$ 不能取为Ω的一切子集，因为这个集类太大，无法在其上定义概率。这时通常取$\mathcal{F}$为直线上Borel点集全体$\mathcal{B}_1$，这是相当大的一个集类，可以把实际问题中所有感兴趣的点集都包括在内。另一方面在Borel σ域上定义概率相当方便，这只要对左闭右开区间给定概率即可。这些将在第三章中做深入讨论。

顺便指出，若Ω不是$\mathbb{R}^1$，而是它的一部分，也可以类似处理。例如Ω为一个区间，这时$\mathcal{F}$可取为该区间上的Borel点集全体：它们通过直线上的Borel点集与该区间之交而得到。

例10 若$\Omega = \mathbb{R}^n$（或$\mathbb{R}^n$的一部分），这时可类似于一维场合取n维Euclid空间中的Borel点集全体$\mathcal{B}_n$作为事件域$\mathcal{F}$，在第三章第二节中将对这种场合进行深入讨论。