第 1 周讲稿

第一章　概率与概率空间

§1 引言

1. 随机数学的研究对象
2. 发展简史

概率论发展简史：

16 世纪 欧洲学者开始研究赌博中的一些简单问题（掷骰子、扑克等）；

17 世纪中叶 Pascal，Fermat 等基于排列组合（组合数学）研究一些较为复杂赌博问题（得分问题，破产问题等）；

1713 年第一个分水岭，J.Bernoulli（大数定律）$\overset{1716年}{\Rightarrow}$ $De\ Moivre$（中心极限定理）（在简单情形下，用他导出的 $n! \sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n}$ 首次说明了正态分布的重要性，其一般结论后由 Laplace 给出）

1812 年第二个分水岭，P.S. Laplace《概率的分析理论》（给出了概率的古典定义，运用了 Laplace 变换，母函数及差分方差的方法，完成了组合计算到分析方法的过渡）；其后的概率主流，是推广和发展大数定律和中心极限定理；Chebyshev(1866 年)$\Rightarrow$Liapunov（1901 年（特征函数方法））

20 世纪，Khintchine, Kolmogorov, Levy, Wiener 等，特别是 1933 年的 Kolmogorov 的概率公理化体系的提出。

随机过程发展简史：

1900 年（1905 年）Bachelier （Einstein）的 Brown 运动的提出；

1907 年 Markov 的 Markov 链（1931 年 Kolmogorov 的 Markov 过程）；

1923 年 Wiener 的 Brown 运动的数学定义；

1934 年 Khintchine 的平稳过程；

1938 年 Levy 的独立增量过程；

1951 年 Ito 的随机积分，1953 年 Doob 的随机过程一般理论和鞅论。

随机过程的研究方法：概率方法（轨道性质、停时、SDE 等）；分析方法（测度论、半群、函数论、微分方程等）

§2 随机事件及其概率

1. 一些基本概念：样本空间 Ω，事件，事件的运算与法则
   - ☆ 能正确写出恰当描述随机试验的样本空间；
   - ☆ 样本点和样本空间的选取并不是唯一的（但不管选取哪个，确定事件的概率是唯一的），要选择容易计算概率的那一个样本空间；
   - ☆ 同一样本空间可以表示不同的随机试验.

附：事件的四种关系，三种运算及运算法则：

☆ $\omega \in A \Leftrightarrow$ 事件 ***A*** 发生

☆ **运算法则：着重注意对偶律（De Morgan 律）**

（1） 事件之间的四种关系

关系	符号	概率论	集合论
包含关系	$A \subset B$	事件 ***A*** 发生则事件 ***B*** 必发生	***A*** 是 ***B*** 的子集
等价关系	$A = B$	事件 ***A*** 与事件 ***B*** 相等	***A*** 与 ***B*** 相等
对立关系	A^c	事件 ***A*** 的对立事件（或逆事件）	***A*** 的余集
互斥关系	$AB = \phi$	事件 ***A*** 与事件 ***B*** 不能同时发生(互不相容)	***A*** 与 ***B*** 无公共元素

（2） 事件之间的三种运算

运算	符号	概率论	集合论
事件的和（并）	$A \cup B$（或 $A+B$）	事件 ***A*** 与事件 ***B*** 至少有一个发生	***A*** 与 ***B*** 的并集
	$\bigcup_{i=1}^{n} A_i$	事件 $A_1, \cdots, A_n$ 至少有一个发生	$A_1, \cdots, A_n$ 的并集
事件的积（交）	$A \cap B$（或 AB）	事件 ***A*** 与事件 ***B*** 同时发生	***A*** 与 ***B*** 的交集
	$\bigcap_{i=1}^{n} A_i$	事件 $A_1, \cdots, A_n$ 同时发生	$A_1, \cdots, A_n$ 的交集
事件的差	$A - B$（或 $A \setminus B$）	事件 ***A*** 发生而事件 ***B*** 不发生	***A*** 与 ***B*** 的差集

（3） 事件的运算法则

交换律：$AB = BA; A \cup B = B \cup A$	结合律：$(A \cup B) \cup C = A \cup (B \cup C)$； $(AB)C = A(BC)$
分配律：$(A \cup B)C = AC \cup BC$ $(A \cap B) \cup C = (A \cup C)(B \cup C)$	**对偶律(De Morgan 律):** $(A \cup B)^c = A^c \cap B^c; (A \cap B)^c = A^c \cup B^c$ $(\bigcup_{i=1}^{n} A_i)^c = \bigcap_{i=1}^{n} A_i^c$； $(\bigcap_{i=1}^{n} A_i)^c = \bigcup_{i=1}^{n} A_i^c$
补元律：$AA^c = \phi; A \cup A^c = \Omega$	还原律：$(A^c)^c = A$
蕴涵律：若 $AB = \phi$，则 $A \subset B^c, B \subset A^c$	分解律：若 $A \subset B$，则 $B = A \cup A^c B$
差积转换律：$A - B = AB^c = A - AB$	吸收律：若 $A \subset B$，则 $AB = A; A \cup B = B$
矛盾律：$AA \;\; = \phi$	排中律：$A \cup A^c = \Omega$

2. 两类**等可能概型**：古典概型与几何概型

★ 古典概型的前提； **定义：** $P(A)=\frac{\#A}{\#\Omega}$, $A\in\mathscr{F}$;

工具：排列和组合数数，要注意分子分母数数时的一致性

★ 几何概型的前提； **定义：设** Ω **为可测区域，** $A\in\mathscr{F}$**且可测，** $P(A)=\frac{L(A)}{L(\Omega)}$,

工具：微积分求区域面积、体积等

例 1（摸球问题、抽签问题）袋中装有 α 个白球及 β 个黑球，

（1）从袋中任取 $a+b$ 个球，试求所取的球恰含 a 个白球和 b 个黑球的概率（$a\le\alpha, b\le\beta$）.

【 $\frac{C_\alpha^a C_\beta^b}{C_{\alpha+\beta}^{a+b}}$ 】

（2）从袋中任意地接连取出 $k+1(k+1\le\alpha+\beta)$ 个球，如果每球被取后不放回，试求最后取出的球是白球的概率。 【 $\frac{\alpha}{\alpha+\beta}$ 】

例 2（约会问题） 两人相约于晚 7 点到 8 点间在某地会面，先到者等足 20 分钟便立即离去. 设两人的到达时刻在 7 点到 8 点间都是随机且等可能的. 求两人能会面的概率 p. 【 $\frac{5}{9}$ 】

例 3（Buffon 问题） 平面上画有一族相距为 a 的平行线. 向此平面投一长为 $l\,(<a)$的针. 求针与平行线相交的概率 p. 【 $\frac{2l}{a\pi}$ 】

存在的问题：

☆Bertrand 悖论 （在圆内任意作一弦，求其长超过圆内接正三角形边长的概率？）

问题的提法不确定，这里的“任意”至少有 3 种解释，相对于各自的解释，每种解法都正确。原因：当随机试验有无穷多个可能结果时，有时很难规定“等可能”这一概念。

☆抛硬币之例（书中例子介绍例 1.3；例 1.4，例 1.5）;

☆有限样本空间的非古典概型例子。

§3 概率空间及概率的计算

1. 事件域的引入

定义：事件族(Ω 的子集族)$\mathcal{F}$ 称为 σ **-域**(也称为 σ **-代数**或**事件体**), 如果它满足下列条件：

i) $\Omega\in\mathcal{F}$;

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

iii) 若 $A_1, A_2, \cdots \in \mathcal{F}$，则 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$，

几个特殊的 σ -域介绍（例子）。

2. 概率的公理化定义

定义：概率（也称为**概率测度**）P 为 $\mathcal{F}$ 上的非负值函数，即对每一事件 $A \in \mathcal{F}$，都可定义一个数 $P(A)$，满足下列条件：

(1) 非负性： 对一切 $A \in \mathcal{F}$，有 $P(A) \geq 0$ (1.2)

(2) 规范性： $P(\Omega) = 1.$ (1.3)

(3) 可数可加性： 若 $A_1, A_2, \cdots \in \mathcal{F}$ 为一列两两互不相容的事件，则

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \tag{1.4}$$

则称 $P(A)$为事件 A 的概率。试验的样本空间 Ω、事件 σ -域 $\mathcal{F}$ 及定义在 $\mathcal{F}$ 上的概率 P 所构成的三元组(Ω, $\mathcal{F}$, P), 称为描述该随机试验的概率空间.

注 1) 如果 Ω 只包含 n 个点，则每个单点集 $\{\omega_j\}$ $j=1,2,\cdots,n$ 是一个基本事件，取 $\mathcal{F}$ 为 Ω 的所有子集的全体，则对每个 $\{\omega_j\}$ 指定概率就足够了。因为对任意 $A \in \mathcal{F}$, 我们有 $P(A) = \sum_{\omega \in A} P(\omega)$。当基本事件都是等可能的情况下，我们进一步可推得

$$P(\{\omega_j\}) = \frac{1}{n} \qquad j=1,2,\cdots.n,$$

此时， $P(A) = \frac{\# A}{n}$ （$A \in \mathcal{F}$）。由此可见古典概型仅仅是 Kolmogorov 模型中的一个非常小的子模型.

例（有限概率空间，但不等可能）从 $1,2,\cdots,100$ 中选取一数，取到不超过 50 的数的概率为 p，取到不超过 50 的数的概率为 3p，求取一数它为平方数的概率？ 【$\frac{16}{200}$】

2) 如果 Ω 包含可数个点，我们就不能对基本事件做等可能的假设，但仍然可以通过对每个基本事件 $\{\omega\}$ 指定概率，而得到概率 P。因为对任意 $A \in \mathcal{F}$ (其中 $\mathcal{F}$ 为 Ω 的所有子集的全体)，令 $P(A) = \sum_{\omega \in A} P(\omega)$，就得到满足概率公理的概率测度 P 。

3) 如果 Ω 包含不可数多个点，每个单点集是一基本事件, 虽然这时各单点集可能完全对称,它们出现的可能性也相同，但是这时我们不能简单地指定每个基本事件的概率，因为这个值为零。更深层的原因是：此时 $\mathcal{F}$ 一般不能取 Ω 的所有子集的全体，这相当于实分析中“难测度问题”。

3. 概率的简单性质

性质 1（求逆公式） 如果 $A \in \mathscr{F}$，则 $P(A^c) = 1 - P(A)$.

性质 2（减法公式） 如果 $A, B \in \mathscr{F}$, 则 $P(A-B) = P(A) - P(AB)$；

特别地，当 $A \supset B$ 时，有 $P(A-B) = P(A) - P(B)$，从而 $P(A) \geq P(B)$ **（单调性）**.

性质 3（一般的加法公式） 如果 $A, B \in \mathscr{F}$，则

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

一般地，若 $A_1, A_2, \cdots, A_n \in \mathscr{F}$，则

$$P(\bigcup_{i=1}^{n} A_i) = \sum_i P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k) - \cdots\cdots + (-1)^{n+1} P(\bigcap_{i=1}^{n} A_i)$$

性质 4（有限可加性） 若 $A_1, A_2, \cdots, A_n \in \mathscr{F}$ 为一列两两互不相容的事件，则

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

注：可数可加性 $\Rightarrow$ 有限可加性；可数可加性与加法公式区别.

性质 5（概率的下（上）连续性） 设 $\{A_n\}$ 是 $\mathscr{F}$ 中的非减（或非增）事件序列(即 $A_n \in \mathscr{F}$, 并且 $A_n \subset A_{n+1}$ (或 $A_n \supset A_{n+1}$), $n = 1, 2, \cdots$)，则

$$\lim_{n\to\infty} P(A_n) = P(\bigcup_{n=1}^{\infty} A_n) \quad \text{（或} \lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n) \text{）}$$

注：可数可加性 $\Leftrightarrow$ 有限可加性 + 概率的连续性（思考）

例 从 $(0,1)$ 中任取一数，它为 $\frac{1}{3}$ 的概率等于 0.

提示：记 A_n ={ $(0,1)$ 中任取一数，它的前 n 位小数为 $\underbrace{33\cdots3}_{n}$ }，$A_n \downarrow A = \{\frac{1}{3}\}$，$P(A_n) = \frac{1}{10^n}$

补充：事件序列的极限

设 $\{A_n\}$ 为样本空间 Ω 中的事件序列，

定义 $\{A_n\}$ 的上极限为：$\overline{\lim_{n\to\infty}} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$；(当且仅当有无穷个 A_n 发生)

$\{A_n\}$ 的下极限为：$\underline{\lim_{n\to\infty}} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$。(当且仅当至多有有限个 A_n 不发生)

如果 $\overline{\lim_{n\to\infty}} A_n = \underline{\lim_{n\to\infty}} A_n$，则称事件序列 $\{A_n\}$ 的极限存在，记为 $\lim_{n\to\infty} A_n$。