# 目 录

# 第一章　事件与概率

**教学目的:**

1) 掌握随机事件的概念和相关运算.

2) 了解概率的不同定义，掌握古典概型的基本计算.

3) 掌握条件概率的概念，熟练运用全概率公式和Bayes公式.

4) 掌握事件独立的概念和有关运算.

## §1.1　概率论发展简史

概率论起源于17世纪, 现在公认是1654年Pascal与Fermat就赌博中的数学问题所展开的讨论, 在讨论中提出了一些基本概念, 最典型的例子是如何分赌本的问题. 两个赌徒相约赌若干局, 谁先赢s局就算谁赢. 由此提出期望的概念. 之后几个数学大家Huygens, Bernoulli, J, De Moivre 等研究了这个问题, Bernoulli 对频率与概率接近这一事实给予了理论上的阐述. 1812年Laplace 在《分析概率论》中最早叙述了概率论的几个基本定理, 给出了古典概率的明确定义. 1814年在《概率的哲学探讨》一书中, 记载了一个有趣的统计故事, 根据伦敦、彼得堡、柏林和全法国的统计资料, 得出几乎一致的男婴和女婴出生的比例为22:21, 即男婴比例为51.16%, 或男婴与女婴的比值为104.76:100, 可是统计1745-1784年整整40年巴黎男婴的出生率时, 得到的比例为25:24 (104.17:100), 调查研究后发现巴黎人有遗弃男婴的陋习. 1900年Hilbert 在第二届世界数学家大会上提出了23个有名的问题, 主体是对新世纪数学发展方向的探讨. 关于建立概率论的公理体系是他所提的第六个问题“借助公理来研究那些在其中数学起重要作用的物理科学; 首先是概率和力学”. 随后Poincare, Borel等都对概率论公理体系的建立做出了努力, 1933年苏联的大数学家Kolmogorov(1903-1987)正式提出了概率论的公理体系. 概率论从此得到迅速的发展, 在此基础上, 数理统计也得到了迅速的发展.

## §1.2　概率论的几个基本概念

### §1.2.1　随机试验和随机事件

随机现象: 自然界中的客观现象, 当人们观测它时, 所得结果不能预先确定, 而仅仅是多种可能结果之一.

举例说明随机现象.

随机试验: 随机现象的实现和对它某个特征的观测.

随机试验中要求试验的结果至少2个，每次试验或观测得到其中的一个结果，在试验或观测之前不能预知是哪个结果发生。此外，一般还要求试验能够重复。

如观测把硬币抛4次后正面向上的次数; 观测某地的温度变化; 某电话总机单位时间内转接的电话次数.

**定义 1.2.1.** *基本事件: 随机试验中的每个单一结果, 它犹如分子中的原子, 在化学反应中不能再分, 所以有"基本"两字.*

如把硬币抛3次后有8种可能结果: 正正正、正正反、正反正、反正正、正反反、反正反、反反正、反反反. 这8种可能结果的每一个都是基本事件.

**定义 1.2.2.** *随机事件: 简称事件(Event), 在随机试验中我们所关心的可能出现的各种结果, 它由一个或若干个基本事件组成.*

随机事件常用大写英文字母$A, B, C, D$等表示. 如果用语言表达, 则要用花括号括起来.

**定义 1.2.3.** *样本空间(Sample Space): 随机试验中所有基本事件所构成的集合, 通常用$\Omega$或$S$表示. 样本空间中的元素,称为样本点, 通常用$\omega$等表示.*

**例 1.2.1.** *掷一枚骰子, 观察出现的点数. 则 $\Omega = \{1, 2, 3, 4, 5, 6\}$.*

**例 1.2.2.** *考察某一地区的年降雨量, 则 $\Omega = \{x|0 \le x < T\}$, 这里 $T$ 表示某个常数, 表示降雨量不会超过 $T$.*

样本空间的元素应该是相互不同的, 根据试验的不同目的,样本空间应该予以不同的选择. 但是总的原则是样本空间应该尽可能详细, 即尽可能包含所有可能的结果. 看下面的例子

**例 1.2.3.** *(1). 将一枚硬币抛三次，考察正反面出现的情况;*
*(2). 将一枚硬币抛三次，考察正面出现的次数。*

这两个试验的目的不同，因此样本空间的选取也不同。

**定义 1.2.4.** *必然事件($\Omega$): 在试验中一定会发生的事件;*

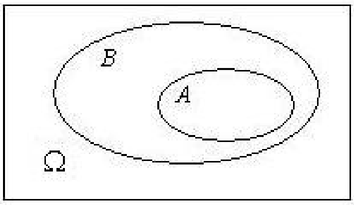
*不可能事件($\phi$): 在试验中不可能发生的事件.*

因此,我们可以说**样本空间的子集称为随机事件**.

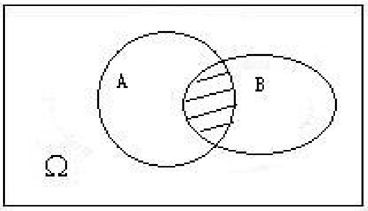
### §1.2.2 事件的运算

可以证明, 把样本空间中的基本事件与空间中的点相对应, 则事件与集合相对应, 因此事件运算与集合运算可以建立一一对应关系.

1. 子事件$A \subset B$: 事件$A$发生蕴含事件$B$一定发生, 则事件$A$称为事件$B$的子事件, 记为$A \subset B$. 若$A \subset B$, 且$B \subset A$, 则称事件$A$与事件$B$相等, 记为$A = B$.

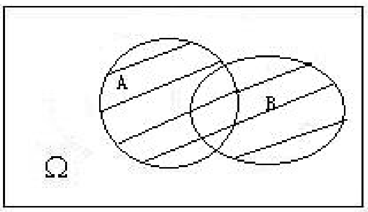



2. 事件的和$(A \cup B)$ : 事件$A$和事件$B$中至少有一个发生的这一事件称为事件$A$和事件$B$的和, 记为$A \cup B$.

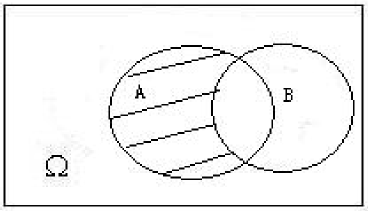



3. 事件的积$(A \cap B)$ : 事件$A$和事件$B$同时发生这一事件称为事件$A$和事件$B$的积, 记为$A \cap B$.

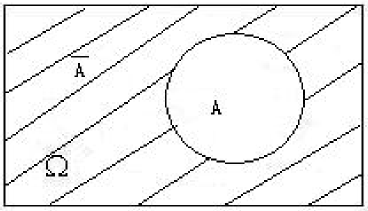
   如果$A \cap B = \phi$, 则称$A$和$B$不相容, 即事件$A$和$B$不能同时发生.




4. 对立事件$A^c$(或$\bar{A}$): $A$不发生这一事件称为事件$A$的对立事件(或余事件) .




5. 事件$A$和事件$B$的差$A-B$: 事件$A$发生而事件$B$不发生这一事件称为事件$A$和事件$B$的差, 记为$A-B$, 或等价的, $AB^c$.




De Morgan对偶法则:

$$\overline{A \cup B} = \bar{A} \cap \bar{B},$$

$$\overline{A \cap B} = \bar{A} \cup \bar{B},$$

上面公式可以推广到$n$个事件:

$$\overline{\bigcup_{i=1}^{n} A_i} = \bigcap_{i=1}^{n} \bar{A}_i$$
$$\overline{\bigcap_{i=1}^{n} A_i} = \bigcup_{i=1}^{n} \bar{A}_i$$

**例 1.2.4.** 设$A, B, C$是三个事件，试表示下列事件

1. 事件$A, B$发生而$C$不发生；*($AB\bar{C}$)*
2. 事件$A, B, C$不同时发生; *($\bar{A} + \bar{B} + \bar{C}$)*
3. 事件$A, B, C$中至多有一个发生；*($A^cB^c + A^cC^c + B^cC^c$)*
4. 事件$A, B, C$中至少发生两个；*($AB + AC + BC$)*
5. 事件$A, B, C$中恰好发生两个；*($AB\bar{C} + A\bar{B}C + \bar{A}BC$)*

### §1.2.3 概率的定义及性质

**1. 概率的定义**

什么叫概率? 直观地讲, 概率是随机事件发生可能性大小的数字表征, 其值在0和1之间, 换句话说, 概率是事件的函数. 如何求出事件$A$的概率(记为$P(A)$)?

(1) 古典概型: 有两个条件,

第一, (有限性) 试验结果只有有限个(记为$n$) ,

第二, (等可能性) 每个基本事件发生的可能性相同.

为计算事件$A$的概率, 设$A$中包含$m$个基本事件, 则定义事件$A$的概率为

$$P(A) = \frac{m}{n}$$

记号: 为方便起见，以$\#(B)$记事件$B$中基本事件的个数，因此，

$$P(A) = \frac{\#(A)}{\#(\Omega)}$$

**(2) 概率的统计定义**

古典概型的两个条件往往不能满足, 此时如何定义概率? 常用的一种方法是把含有事件$A$的随机试验独立重复做$n$次($Bernouli$试验) , 设事件$A$发生了$n_A$次, 称比值$\frac{n_A}{n}$为事件$A$发生的频率, 当$n$越来越大时, 频率会在某个值$p$附近波动, 且波动越来越小, 这个值$p$就定义为事件$A$的概率.

注意: 为什么不能写为$\lim_{n\to\infty}\frac{n_A}{n} = p$? 因为$\frac{n_A}{n}$不是$n$的函数.

**例 1.2.5.** 抛硬币的试验

| 试验者 | 掷硬币的次数 | 正面出现的次数 | 频率 |
|---|---|---|---|
| 蒲丰 | *4040* | *2048* | *.5069* |
| 皮尔逊 | *12000* | *6019* | *.5016* |
| 皮尔逊 | *24000* | *12012* | *.5005* |

从这个例子可以看出随着试验次数的增加，频率越来越接近1/2.

其他例子: 英文字母被使用的频率是相当稳定的; 福尔摩斯探案集第四本《跳舞的小人》, 福尔摩斯用频率破了丘比特和埃尔茜之间联络密码; 1872年英国人$Shix, W$ 把$\pi$算到707位, 1944.5-1945.3数学家法格逊认为$\pi$的小数位的数字对0到9应该是等可能的, 但核

对$Shix$的结果发现数字7太少, 故对$Shix$的结果有怀疑, 重新计算发现前527位是正确的, 后面不对了. 计算机出现后, 法国人让.盖尤计算了$\pi$的前100万位小数, 发现各个数字出现的频率相同.

**(3) 主观概率**

关于概率的统计定义, 我们可能会想到, 如果试验不能在相同的条件下独立重复很多次时该怎么办? 还有人们常谈论种种事件出现机会的大小, 如某人有80%的可能性办成某事.另一人则认为仅有50%的可能性. 即我们常常会拿一个数字去估计这类事件发生的可能性, 而心目中并不把它与频率挂钩. 这种概率称为主观概率, 这类概率有相当的生活基础. 在金融和管理等方面有大量的应用, 这一学派称为$Bayes$ 学派, 近来得到越来越多的认可. 但是当前用频率来定义概率的频率派仍是数理统计的主流. 焦点是频率派认为概率是客观存在，不可能因人而异.

(4) 概率的公理化定义: 对概率运算规定一些简单的基本法则,

$(i)$ 设$A$ 是随机事件, 则$0 \le P(A) \le 1$,

$(ii)$ 设$\Omega$为必然事件, 则$P(\Omega) = 1$,

$(iii)$ 若事件$A$和$B$不相容, 则$P(A \cup B) = P(A) + P(B)$,

为了对可数无穷个事件仍能成立, 我们要把上面公式中的两个事件推广到可数无穷个两两不相容的事件序列

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

由概率的公理化体系，可以得到有关概率的一些性质. 以下$\mathcal{F}$为由样本空间所生成的$\sigma-$代数,它表示了样本空间中所有可以计算概率的事件的集合.

1. $P(\phi) = 0$
2. (有限可加性) 若$A_k \in \mathcal{F}, k = 1, \cdots, n$且两两互斥，则

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

3. (可减性) 若$A, B \in \mathcal{F}$且$A \subset B$，则$P(B - A) = P(B) - P(A)$.
4. (单调性) 若$A, B \in \mathcal{F}$且$A \subset B$，则$P(A) \le P(B)$.
5. $P(\bar{A}) = 1 - P(A)$

6. (加法定理) 对任意的事件$A_1,\cdots,A_n\in\mathcal{F}$，有

$$\begin{aligned}P(\sum_{k=1}^{n}A_k) &= \sum_{k=1}^{n}P(A_k)-\sum_{1\le i<j\le n}P(A_iA_j)+\sum_{1\le i<j<k\le n}P(A_iA_jA_k)\\ &\quad -\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)\end{aligned}$$

7. (次可加性) 对任意的事件$A_1,\cdots,A_n,\cdots\in\mathcal{F}$，有$P(\sum\limits_{n=1}^{\infty}A_n)\le\sum\limits_{n=1}^{\infty}P(A_n)$.

8. (下连续性) 若$A_n\in\mathcal{F}$且$A_n\subset A_{n+1},n=1,2,\cdots$,则

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

9. (上连续性) 若$A_n\in\mathcal{F}$且$A_n\supset A_{n+1},n=1,2,\cdots$,则

$$P(\prod_{n=1}^{\infty}A_n)=\lim_{n}P(A_n)$$

**例 1.2.6.** 求证对任意$n$个事件$A_1,\cdots,A_n$有

$$P(\prod_{k=1}^{n}A_k)\ge\sum_{k=1}^{n}P(A_k)-n+1$$