定义 1.9 (等概完备事件组) 若有限个事件 $A_1, A_2, \cdots, A_n$ 满足以下条件，

（1） 完备性：$\bigcup_{n=1}^{N} A_n = \Omega$；

（2） 不相容性：$\forall i, j \in \{1, 2, \cdots, N\}, i \neq j, A_i \cap A_j = \phi$；

（3） 等可能性：$A_1, A_2, \cdots, A_N$ 可能性相同。

则称 $A_1, A_2, \cdots, A_N$ 为等概完备事件组。

定义一个等概完备事件组来划分样本空间是古典概型中计数的前提。

“基本结果”、“基本事件”是相对的概念，根据分析问题的需要，对随机现象的所有可能结果有不同的划分方法，因而给于“基本结果”不同的含义。

例 1.9 掷一个骰子，点数为 $\{1\}\{2\}\{3\}\{4\}\{5\}\{6\}$ 构成一个等概完备事件组，而“出现奇数点”与“出现偶数点”也构成一个等概完备事件组。

例 1.10 无穷次掷硬币，求下列事件的概率：

（1） A：第 k 次出现正面；

（2） B：第一次正面出现在第 k 次。

解： 设基本结果为 $(x_1, x_2, \cdots)|x_i \in \{0, 1\}, i \in \mathbb{N}$,

（1） 按照第 k 次投掷的结果来定义等概完备事件组划分 Ω,
$\omega_0 = \{(x_1, x_2, \cdots, x_k = 0, \cdots)|x_i \in \{0, 1\}, i \neq k\}$ 表示第 k 次为 *0* 的结果；
$\omega_1 = \{(x_1, x_2, \cdots, x_k = 1, \cdots)|x_i \in \{0, 1\}, i \neq k\}$ 表示第 k 次为 *1* 的结果；
等概完备事件组有两个事件，A 为 ω_1, 所以 $P(A) = \frac{1}{2}$。

（2） 按照前 k 次投掷的结果来定义等概完备事件组划分 Ω,
对于任意 $x_1, x_2. \cdots, x_k \in \{0, 1\}$，事件 $\omega_{x_1, x_2, \cdots, x_k} = \{(y_1 = x_1, y_2 = x_2, \cdots, y_k = x_k, y_{k+1}, \cdots)|y_i \in \{0, 1\}, i > k\}$ 表示前 k 次投掷结果为 $x_1, x_2. \cdots, x_k$ 的事件，B 为事件 $\omega_{x_1=0, x_2=0, \cdots, x_{k-1}=0, x_k=1}$，等概完备事件组有 2^k 个事件，所以 $P(B) = \frac{1}{2^k}$。

1.3 概率的公理化定义

1.3.1 古典概型

利用古典概型来定义概率是将事件中样本点数与样本空间样本点数的比率定义为概率，这种定义的优点是简单，直观，然而也存在很多局限性：

（1） 古典概型基于等可能性假设，但随机事件的结果可能不具备等可能性。例如扔一颗质地不均匀的骰子，如果靠近 1 点这一边较重，那么投掷时 1 点朝下的概率较大，1 点出现概率比较小，而 1 点对面出现的概率就比较大。

（2） 当样本空间样本点个数为无穷多时，古典概型往往不适用。

（3） 不是建立在实验基础上，“等可能性”全凭感觉，无法证明比率是否真的代表了事件发生可能性大小，缺乏实证性。

1.3.2 概率的频率定义

在随机试验中，与事件 A 有关的随机现象可重复进行，考虑随机试验，有了频率的定义。

定义 1.10 在相同条件下，重复做 n 次随机试验，设 $n(A)$ 是事件 A 出现的次数，$\frac{n(A)}{n}$ 称为事件 A 出现的频率。

注： 在随机试验中，充分扰动可以视为相同条件，例如充分混匀重复做 n 次盒中抽球。

人们的长期实践表明，随着试验重复次数 n 的增加，频率会稳定在某一常数附近，我们称这个常数为频率的稳定值，将其定义为概率。这一定义面临的问题:

（1） 在等可能性假设的条件下，当试验次数越来越大时，频率 $\frac{n(A)}{n}$ 会不会愈来愈接近比率 $\frac{|A|}{|\Omega|}$？

17 世纪的瑞士学者 Jacob Bernoulli 在历史上第一个企图对此论断给予严格意义和数学证明，在假设基本结果的等可能性时，贝努利证明了这个论断，被称为“大数定律”，让古典概型下的比率有了实证意义。

（2） 当随机事件的结果不具备等可能性时应该如何证明？

频率稳定性本身就是一个假设，根据经验和等可能性假设下的大数定律，我们相信这一点。然而不进行无限次的条件不变的重复试验，就无法完全肯定频率的稳定性。例如事件 A 在 10 次重复试验中频率为 0.3,100 次为 0.28,1000 次为 0.295,10000 次为 0.2922，那么什么情况下可以称为稳定？应当以哪个值作为稳定值？这些都有含糊不清的缺陷。

在概率定义中，若运用古典概型，就要假设等可能性（此时，概率定义与频率定义是相洽的）。若不具备等可能性，就要假设频率的稳定性。那么有没有比等可能性、频率稳定性更基础的，对度量事件发生可能性大小更本质的前提或假设呢？通过还原历史的分析，我们可以知道概率论研究需要前提——假设。数学中的公理无法证明，源于现实，是人类观察到的经验规律的高度总结，几何学中就有不加证明而接受的公理。因此，寻找概率论中本质的前提假设，是更深入地定义概率所要解决主要问题。

1.3.3 概率的公理化定义

概率的公理化定义由 Kolmogorov 在 1933 年的著作《概率论基本概念》中提出，这一定义把事件概率的存在作为一个不需要证明的事实接受下来。

定义 1.11 设 Ω 为一样本空间，$\mathscr{F}$ 为 Ω 上的一个事件域，称定义在 $\mathscr{F}$ 上的一个实值函数 $P(\cdot):\mathscr{F}\to\mathbb{R}$ 为 $\mathscr{F}$ 上的概率测度（简称概率），若满足

（*1*） 非负性：$\forall A\in\mathscr{F}, P(A)\geqslant 0$；

（*2*） 正则性：$P(\Omega)=1$；

（3）　可列可加性：若 $A_n \in \mathscr{F}(n=1,2,\cdots)$ 两两互不相交（$1 \leqslant i,j \leqslant n, i \neq j, A_i \cap A_j = \phi$），则 $P(\bigcup_{n=1}^{+\infty} A_n) = \sum_{n=1}^{+\infty} P(A_n)$

称 $P(A)$ 为事件 A 的概率，称三元组 $\{\Omega, \mathscr{F}, P\}$ 为概率空间。

注 1：　概率的公理化定义没有直接回答“概率”是什么，而是把概率应具备的几条基本性质（经验事实）概括出来，把具有这几条性质的量叫做概率，在此基础上展开概率论的研究。在此基础上，可以证明：（1）当 $\frac{|A|}{|\Omega|}$ 有意义时是概率；（2）频率稳定于概率。

注 2：　概率的公理化定义刻画了概率论的本质，但没有告诉人们如何去确定概率，而古典概型和频率定义可视为一定场合下，确定概率的实际方法。例如对样本空间做等可能划分时，可以用古典概型确定概率；样本不具备等可能性时，可以用频率作为概率的估计值，从观测出发得到概率，这就是统计学。

注 3：　概率空间 $(\Omega, \mathscr{F}, P)$ 是研究概率的基础和出发点。

注 4：　设 Ω 为一样本空间，称 Ω 的某些子集所组成的集合类 $\mathscr{F}$ 为 Ω 上的一个事件域（σ 代数），满足

（1）　$\Omega \in \mathscr{F}$；

（2）　若 $A \in \mathscr{F}$，则 $\bar{A} \in \mathscr{F}$；

（3）　若 $A_n \in \mathscr{F}, n=1,2,\cdots$，则 $\bigcup_{n=1}^{+\infty} \in \mathscr{F}$。

若 $A \in \mathscr{F}$，则称 A 为（概率）事件，事件域中的元素称为事件，事件域的直观含义为：可以合理地定义概率的事件的全体。对于事件域有以下讨论：

（1）　事件域对集合的运算（可列并，可列交，差，对立）有封闭性；

（2）　若 Ω 为离散样本空间（Ω 为有限集或可列集），则由 Ω 的所有子集构成的集合类 $\mathscr{F}$ 是一个事件域；

（3）　若 Ω 为连续样本空间（Ω 为不可列无限集），由 Ω 的所有子集不一定组成事件域，没有统一的办法来确定。

若 $\Omega = \mathbb{R}$，可由一个基本集合类逐步扩展成一个事件域：

- 基本集合类 $\mathscr{F} = \{$全体半直线$\}$，即 $\{(-\infty, a) | -\infty < a < +\infty\}$；
- $[a,b) = (-\infty, b) - (-\infty, a)$；
- $[a,b] = \bigcap_{n=1}^{+\infty} [a, b+\frac{1}{n})$；
- $\{b\} = [a,b] - [a,b)$；
- $(a,b] = [a,b] - \{a\}$；
- $(a,b) = [a,b) - \{a\}$；

- 上述集合经过（有限个或可列个）并运算和交运算得到的实数集的全体称为 Borel 事件域，域中的每个元素（实数集）称为 Borel 集。

1.4 概率的性质

利用概率的公理化定义，可以导出概率的一系列性质。

1.4.1 基本性质

定理 1.1 空集概率为 0，即 $P(\phi)=0$。

定理 1.2 (有限可加性) 若 $A_i(i=1,2,\cdots,n)$ 互不相容，则 $P(\bigcup_{i=1}^n A_i)=\sum_{i=1}^n P(A_i)$。

定理 1.3 $P(\bar{A})=1-P(A)$。

在很多情况下，若利用此定理考虑对立事件会相对比较简单。

例 1.11 抛一枚硬币 5 次，求既出现正面，又出现反面的概率。

解： 样本空间 $\Omega=\{(x_1,x_2,\cdots,x_5)|x_i\in\{0,1\},1\leqslant i\leqslant 5\}$，$|\Omega|=2^5$，

$$\begin{aligned}P(A)&=1-P(\bar{A})\\&=1-[P(\text{全为正面})+P(\text{全为反面})]\\&=1-(\frac{1}{2^5}+\frac{1}{2^5})=\frac{15}{16}\end{aligned}$$

1.4.2 单调性

概率的单调性与事件的差相关联，用来求两个事件的差事件的概率。

定理 1.4 若 $A\supset B$，则 $P(A-B)=P(A)-P(B)\geqslant 0$。

证明： 因 $A\supset B$，故 $A=B\cup(A-B)$ 且 $B\cap(A-B)=\phi$，由概率定义得 $P(A)=P(B)+P(A-B)$。

定理 1.5 *P(A-B)=P(A)-P(AB)*。

证明： 因 $A=(A-B)\cup(AB)$ 且 $(A-B)\cap(AB)=\phi$，故 $P(A)=P(A-B)+P(AB)$。

难点： 利用概率的性质求较为复杂事件的概率时，要学会通过复杂事件的转化来解决，概率的性质都足够简单，综合运用才是难点。

1.4.3 加法公式

加法公式用来求多个事件（不一定不相容）的并事件的概率。

定理 1.6 $P(A\cup B)=P(A)+P(B)-P(A\cap B)$。

证明： $P(A\cup B)=P(A-B)+P(B)=P(A)-P(AB)+P(B)$。

定理 1.7 $P(\bigcup_{i=1}^{n}A_i)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(\bigcap_{i=1}^{n}A_i)$。

一般来说，可以利用文氏图来理解概率公式，事件与概率的关系可以类比于集合与计数的关系：$P(\cdot):\mathscr{F}\rightarrow\mathbb{R}$，$P(\cdot)$ 是一个测度，$P(A)$ 是衡量集合 A 大小的实值函数。

1.4.4 概率的连续性

定义 1.12 对 $\mathscr{F}$ 中的任何一个递增（渐张）列 $E_1\subset E_2\subset\cdots$，称可列并 $\bigcup_{n=1}^{+\infty}E_n$ 为 $\{E_n\}$ 为极限事件，记为 $\lim_{n\rightarrow+\infty}E_n$；对 $\mathscr{F}$ 中的任何一个递减（渐缩）列 $E_1\supset E_2\supset\cdots$，称可列交 $\bigcap_{n=1}^{+\infty}E_n$ 为 $\{E_n\}$ 为极限事件，亦记为 $\lim_{n\rightarrow+\infty}E_n$。

定理 1.8 对任意递增列 $\{E_n\}$，称对应概率 $P(\cdot)$ 是下连续的，有 $\lim_{n\rightarrow+\infty}P(E_n)=P(\lim_{n\rightarrow+\infty}E_n)$；对任意递减列 $\{E_n\}$，称对应概率 $P(\cdot)$ 是上连续的，有 $\lim_{n\rightarrow+\infty}P(E_n)=P(\lim_{n\rightarrow+\infty}E_n)$。

证明： 设 $\{E_n\}$ 为 $\mathscr{F}$ 中的一递增列，设 $E_0=\phi$，则 $\bigcup_{i=1}^{+\infty}E_i=\bigcup_{i=1}^{+\infty}(E_i-E_{i-1})$，诸 E_i-E_{i-1} 互不相容，

$$\begin{aligned}右边&=P(\bigcup_{i=1}^{+\infty}E_i)=\sum_{i=1}^{+\infty}P(E_i-E_{i-1})=\lim_{n\rightarrow+\infty}\sum_{i=1}^{n}P(E_i-E_{i-1})\\&=\lim_{n\rightarrow+\infty}\sum_{i=1}^{n}[P(E_i)-P(E_{i-1})]=\lim_{n\rightarrow+\infty}P(E_n)=左边\end{aligned}$$

概率的连续性，即 $P(\cdot)$ 是连续的，可以与将来要学习的随机变量分布函数的连续性相类比。

概率的下连续性与上连续性，两者等价，统称连续性。考虑递减列 $\{E_n\}$, 则 $\{\overline{E_n}\}$ 为递增列，

$$\begin{aligned}\lim_{n\rightarrow+\infty}P(\overline{E_n})&=P(\lim_{n\rightarrow+\infty}\overline{E_n})=P(\bigcup_{n=1}^{+\infty}\overline{E_n})=P(\overline{\bigcap_{n=1}^{+\infty}E_n})\\&=1-P(\bigcap_{n=1}^{+\infty}E_n)=\lim_{n\rightarrow+\infty}[1-P(E_n)]=1-\lim_{n\rightarrow+\infty}P(E_n)\end{aligned}$$

例 1.12 无限次连续抛掷一枚均匀硬币，求事件 A“正面永不出现”的概率。

解 1： 考虑补集 $\bar{A}$“至少出现一次正面”，设 E_n 表示首次正面出现在第 n 次抛掷，$n=1,2,\cdots$，则 $\bar{A}=\bigcup_{n=1}^{+\infty}E_n$，$\bar{A}$ 可分解为可列个互不相容事件的并，$P(\bar{A})=\sum_{n=1}^{+\infty}P(E_n)=\sum_{n=1}^{+\infty}\frac{1}{2^n}=1$，$P(A)=0$。由此可以看出概率为 *0* 并非为不可能事件，概率为 *1* 并非为必然事件。

解 2： A“正面永不出现”即“永远反面”，记 $A_n=$“前 n 次投掷均出现反面”,$n=1,2,\cdots$，则 $P(A_n)=\frac{1}{2^n}$，注意 $A_n\supset A_{n+1}$，构成递减事件列 $A_1\supset A_2\supset A_3\supset\cdots$。始终出现反面 A 为 $\bigcap_{n=1}^{+\infty}A_n$，由连续性 $P(A)=P(\bigcap_{n=1}^{+\infty}A_n)=\lim_{n\rightarrow+\infty}P(A_n)=\lim_{n\rightarrow+\infty}\frac{1}{2^n}=0$。

解 3： $\forall n \in \mathbb{N}$，$A_n \supset A$，$P(A_n) = \frac{1}{2^n} \geqslant P(A)$，两边取极限 $\lim_{n\to+\infty} \frac{1}{2^n} = 0 \geqslant P(A) \geqslant 0$，$P(A) = 0$。