Ch 1-2 频率与概率

回顾前一次课

- 概率与统计关系
- 随机现象：二重性
- 随机试验：三特点
- 样本空间、样本点
- 随机事件：基本事件、不可能事件、必然事件
- 事件关系：$\subset$、$=$、$\cup$、 $-$ 、$\cap$ 、$\bar{A}$ 互斥与对立事件的关系
- 事件运算：幂等、交换、结合、分配、德摩根律

可测空间

用2^{Ω}表示样本空间Ω所有子集所构成的集合，称为Ω的幂集

- 对有限或可列的样本空间Ω，可考虑事件的集合为2^{Ω}
- 对无限样本Ω，一般不将样本空间的一切子集都看作事件
  例如，后面的几何概型中不可测空间，无法计算概率

定义：设Ω是一个样本空间且$\Sigma \subseteq 2^{\Omega}$，若$\Sigma$满足以下三个条件：

- 必然事件$\Omega \in \Sigma$;
- 若任意$A \in \Sigma$，则有补集$\bar{A} \in \Sigma$;
- 若任意$A_i \in \Sigma\ (i = 1,2,\cdots)$，则有$\bigcup_{i=1}^{+\infty} A_i \in \Sigma$,

则称Σ是样本空间Ω的***σ*代数**，又称***σ*域**，称Σ中的元素（一个子集）是**可测集**，以及(Ω, Σ)是一个**可测空间**

可测空间

设Ω是一个样本空间且$\Sigma \subseteq 2^{\Omega}$，若$\Sigma$满足以下三个条件：

- 必然事件$\Omega \in \Sigma$；
- 若任意$A \in \Sigma$，则有补集$\bar{A} \in \Sigma$；
- 若任意$A_i \in \Sigma$ $(i = 1,2,\cdots)$，则有$\bigcup_{i=1}^{+\infty} A_i \in \Sigma$，

则称Σ是样本空间Ω的**$\boldsymbol{\sigma}$代数**，又称**$\boldsymbol{\sigma}$域**

σ−代数Σ本质上是一个集合，每一个元素也是集合

根据定义可知：

$$\emptyset \in \Sigma \qquad \bigcup_{i=1}^{n} A_i \in \Sigma \qquad \bigcap_{i=1}^{n} A_i \in \Sigma \qquad \bigcap_{i=1}^{+\infty} A_i \in \Sigma$$

给定样本空间Ω，最小的σ−代数为$\Sigma = \{\emptyset, \Omega\}$

最大的σ−代数为$\Sigma = 2^{\Omega}$

最小σ−代数

当非空事件集合$\mathcal{F} \subset 2^{\Omega}$并不构成$\sigma$代数时，可以构造包含$\mathcal{F}$的最小$\sigma$代数

定义：给定样本空间Ω和事件集合$\mathcal{F} \subset 2^{\Omega}$，用$\sigma(\mathcal{F})$表示包含事件集合$\mathcal{F}$的最小$\sigma$−代数，即若$\Sigma$是一个代数且$\mathcal{F} \subset \Sigma$，则有$\sigma(\mathcal{F}) \subset \Sigma$

$\mathcal{F} = \{A\}$，仅包含单一事件A，最小σ−代数$\sigma(\mathcal{F}) = \{\emptyset, A, \bar{A}, \Omega\}$

对于一般事件集合$\mathcal{F} \subset 2^{\Omega}$，最小$\sigma$−代数为

$$\sigma(\mathcal{F}) = \bigcap_{\Sigma \supset \mathcal{F}} \Sigma$$

常用的σ代数

对有限或可列的样本空间Ω，常见的σ代数为2^{Ω}

在实数集$\mathbb{R}$中，最重要的σ−代数是**博雷尔$\boldsymbol{\sigma}$ 代数**，即

$$\begin{aligned}\sigma(\{(a,b):\ a<b\}) &= \sigma\left(\{[a,b]:\ a<b\}\right) = \sigma\left(\{(-\infty,b)\}\right)\\ &= \sigma(\ \{(-\infty,b]\}\ = \sigma\left(\{(a,+\infty)\}\right) = \sigma(\{[a,+\infty)\})\end{aligned}$$

频率

随机事件在一次试验中可发生也可不发生, 通常关心随机事件发生可能性有多大, 为此引入频率, 描述随机事件发生的频繁程度

定义 在相同的条件下，进行了n次试验，n次试验中事件A发生的次数为n_A，称n_A为事件A发生的频数，事件A发生的频率为

$$f_n(A)=\frac{n_A}{n}$$

事件的频率在一定程度上反应了事件发生的可能性. 频率的性质包括

- $0\le f_n(A)\le 1$
- $f_n(\Omega)=1$
- 若$A_1, A_2, \dots A_k$两两互不相容，则$f_n(A_1 \cup A_2 \cup \cdots \cup A_k)=\sum_{i=1}^{k} f_n(A_i)$

频率的稳定性

频率在试验中表现出**随机性**

在大量重复试验中，事件频率通常在一个常数p附近摆动，随着试验次数的增大摆幅越来越小，将这种规律称为 **频率的稳定性**

例如，历史上多人对重复投掷硬币的试验，下面是试验统计结果：

实验者	n	n_H	$f_n(H)$
德·摩根	2048	1061	0.5181
蒲 丰	4040	2048	0.5069
K·皮尔逊	12000	6019	0.5016
K·皮尔逊	24000	12012	0.5005

频率的稳定性-计算机模拟

利用计算机随机数对投币试验进行仿真，下图为试验结果

当试验次数不多时频率呈现波动性

当试验次数充分大时，频率具有稳定性

概率的统计定义

频率的稳定性即通常所说的统计规律性，是**随机事件本身所固有的客观属性**，可用于度量事件发生的可能性大小.

概率的统计定义 随机事件A在大量重复试验中发生的频率总是稳定地在一个常数p附近摆动，随着试验次数的增加而摆幅逐渐越小，则称常数p为事件A发生的**概率**, 记为 $P(A)=p$.

概率的性质：

- $0 \le P(A) \le 1$
- $P(\Omega)=1$
- 若$A_1, A_2, \dots A_k$两两互不相容，则 $P(A_1 \cup A_2 \cup \cdots \cup A_k)=\sum_{i=1}^{k} P(A_i)$

概率与频率

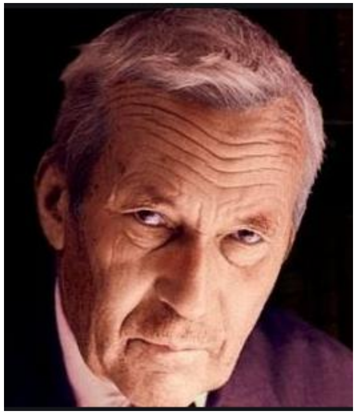
- 概率用于度量事件发生的可能性
- 频率在一定程度上反映了事件发生的可能性
- 概率是恒定的，而频率在试验中具有随机性
- 若试验次数足够多，频率与概率非常接近
- 概率可以通过频率来“测量”，频率是概率的一个近似

概率的统计定义存在数学上的不严谨性，在实际中几乎不可能每一个事件做大量重复的试验来计算频率，进而近似概率

受到频率的稳定性及其性质的启发给出严谨的概率公理化体系

概率的公理化定义

苏联数学家柯尔莫哥洛夫于1933年给出了概率的公理化定义，通过规定概率具备的基本性质来定义概率



概率公理化 在可测空间(Ω, Σ)上，若函数$P: \Sigma \to R$满足下列条件

- **非负性**：$P(A) \geq 0$
- **规范性**：$P(\Omega) = 1$
- **可列可加性**：若$A_1, A_2, \ldots$可列个两两互不相容的事件，则

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

称$P(A)$为事件$\boldsymbol{A}$**的概率**，(Ω, Σ, P)为**概率空间**

简明扼要地刻画了概率的定义，为现代概率论奠定了基础，是概率论发展历史上的一个里程碑，从此被公认为数学的一个分支

概率的性质

- 对不可能事件∅有 $\mathbf{P(\emptyset) = 0}$

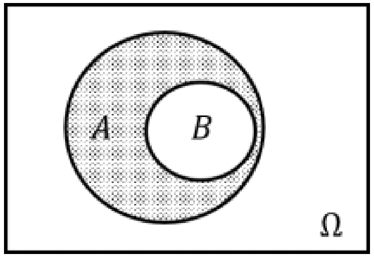
- **有限可加性** 若$A_1, A_2 \cdots A_n$是两两不相容事件, 则

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

- 对任意事件A有$P(\bar{A}) = 1 - P(A)$

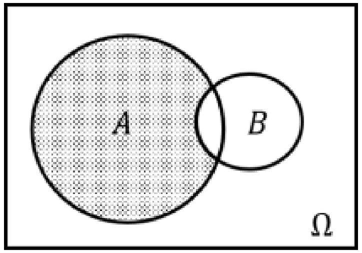
概率的性质

- 若$\mathrm{B} \subset \mathrm{A}$，则$\mathrm{P}(\mathrm{A} - B) = \mathrm{P}(\mathrm{A}) - P(B)$和$\mathrm{P}(\mathrm{B}) \leq \mathrm{P}(\mathrm{A})$



- 对任意事件A和B

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



容斥原理 (Inclusion-Exclusion Principle)

- 对任意随机事件A和B有

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

- 对三个随机事件A,B,C有

$$P(A\cup B\cup C)=P(A)+P(B)+P(C)-P(AB)-P(AC)-P(BC)+P(ABC)$$

- 对事件$A_1,A_2,\cdots,A_n$有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{i<j}P(A_iA_j)+\sum_{i<j<k}P(A_iA_jA_k)+\cdots+(-1)^{n-1}P(A_1\cdots A_n)$$

可进一步简化为

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{r=1}^{n}(-1)^{r+1}\sum_{i_1<\cdots<i_r}P(A_{i_1}\cdots A_{i_r})$$

不等式

Union bounds: 对事件 $A_1, A_2, \dots A_n$ 有

$$P(A_1 \cup A_2 \cup \cdots \cup A_n) \leqslant P(A_1) + P(A_2) + \cdots + P(A_n)$$

Bonferroni不等式： 对事件 $A_1, A_2, \dots A_n$ 有

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n} A_i\right) &\leqslant \sum_{i=1}^{n} P(A_i); \\
P\left(\bigcup_{i=1}^{n} A_i\right) &\geqslant \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i A_j); \\
P\left(\bigcup_{i=1}^{n} A_i\right) &\leqslant \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k);
\end{aligned}$$

可以依次类推

例子

设$P(A)=p,\ P(B)=q, P(AB)=r,$ 用p,q,r分别表示下述事件的概率

1) $P\left(\bar{A}\cup\bar{B}\right)$
2) $P\left(\bar{A}B\right)$
3) $P\left(\overline{\mathrm{A}}\cup B\right)$
4) $P\left(\overline{\mathrm{A}}\cap\bar{B}\right)$

例子

设三个随机事件 A, B, C 满足

$$P(A) = P(B) = P(C) = 1/4,$$

$$P(AB) = 0, P(AC) = P(BC) = 1/16,$$

求事件 A, B, C 中至少有一个事件发生的概率.