概率理论与数理统计

基本概念、古典概率论

概率论：基本概念

随机实验

- 可以在相同条件下重复进行。
- 试验可能结果不止一个，但知道所有可能结果.
- 无法预知哪一个结果出现。

样本空间/ 随机空间

- 随机试验的所有可能结果的集合，通常用Ω表示。Ω中可以包含无穷多个样本.
- 通常用ω来表示实验的某一个结果, 也称为一个样本.

随机事件

一个随机事件A总是表示为满足某个条件的试验样本ω集合:

$$A = \{\omega : \text{实验样本}\omega\text{可以实现}A\}$$

随机试验

- 掷一次骰子, $\Omega = \{1, 2, 3, 4, 5, 6\}$, ω 为1到6之间的整数,
  随机事件A $=$\{点数为偶数\}.
- 在$[0, 1]$中随机取一个实数, $\Omega = [0, 1]$. $0 \leq \omega \leq 1$.
  随机事件A $=\{\omega \in [0, \frac{1}{2}]\}$.
- 连续掷硬币无穷多次, $\Omega = \{H$与T的无穷序列$\}$,
  $\omega = HTTHH\cdots$,
  随机事件A $=$\{前两次为头像\}.

事件的关系与运算

- 必然会发生的样本的集合即为Ω,
- 不可能发生样本集合为$\emptyset$,
- 与事件A要求相反的样本，即“非A”，记为A的补集，A^c或者$\bar{A}$. 显然$\Omega^c = \emptyset$.
- 事件A包含于B, 即A发生必导致B发生，若B不满足的话则A也不满足，$A \subset B$. 若$A \subset B$与$B \subset A$ 同时成立, 则称A与B相等.
- 事件A与B中之一发生, 即两者取并集，称为和事件，$A \cup B$. 若A与B不相容，可写为$A + B$.
- 两个事件A与B同时发生，即两者取交集，称为积事件，$A \cap B$.
- 事件$A - B = A \cap B^c$, 即A发生但B不发生.
- 两个事件A_1与A_2不相容，或者互斥，即$A_1 \cap A_2 = \emptyset$.
- 上极限事件，下极限事件。

概率公理化：事件域

事件的运算

- 交换律：$A \cup B = B \cup A$, $A \cap B = B \cap A$.
- 结合律: $(A \cup B) \cup C = A \cup (B \cup C)$,
- 分配率: $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$,
- De Morgan 公式: $(A \cup B)^c = A^c \cap B^c, (A \cap B)^c = A^c \cup B^c$.

事件域

$\mathcal{F}$为在Ω上的事件域，也称σ-代数，σ-域，即它是Ω的子集构成的集合，满足

1. Ω, $\emptyset \in \mathcal{F}$;
2. 若$A \in \mathcal{F}$, 则$A^c \in \mathcal{F}$;
3. 若$A_i \in \mathcal{F}$, $i \in \mathbb{N}$，则$\cup_i A_i \in \mathcal{F}$, 且$\cap_i A_i \in \mathcal{F}$.

设$\mathcal{G}$ 为$\mathcal{F}$的子集, 包含$\mathcal{G}$且满足定义要求的最小集族, 称为由$\mathcal{G}$生成的σ-代数$\sigma(\mathcal{G})$.

Borel域

在一个拓扑空间Ω上，由其上的所有开集生成的最小σ-域，成为Borel-域，记为$\mathcal{B}(\Omega)$，其中的集合称为Borel-集.

$$\begin{aligned}\mathcal{B}(\mathbb{R}) &= \sigma((a,b),\ a,b\in\mathbb{R}) = \sigma([a,b],\ a,b\in\mathbb{R}) \\ &= \sigma((a,b],\ a,b\in\mathbb{R}) \\ &= \sigma((-\infty,a),\ a\in\mathbb{R}) = \sigma((-\infty,a],\ a\in\mathbb{R})\end{aligned}$$

例子

- $[0,1]$区间上由闭区间生成的σ-域
- 无穷次独立抛掷硬币的空间: $\mathcal{F}_0=\{\varnothing,\Omega\}$,
  $\mathcal{F}_1=\{\varnothing,\Omega,A_H,A_T\}$,
  $\mathcal{F}_2=\{\varnothing,\Omega,A_H,A_T,A_{HH},A_{HT},A_{TH},$
  $A_{TT},A_{HH}^c,A_{HT}^c,A_{TH}^c,A_{TT}^c,A_{HH}\cup A_{TH},A_{HT}\cup A_{TH},$
  $A_{HT}\cup A_{TT},A_{HH}\cup A_{TT}\}$,

概率公理化：定义

频率

在相同条件下，进行n次实验，事件A发生的次数n_A称为频数，比值$\frac{n_A}{n}$称为A发生的频率，常被记为$f_n(A)$.

定义

给定一个概率空间Ω及其上的一个事件域$\mathcal{F}$，**概率$\mathbf{P}$**是一个定义在集合$A \in \mathcal{F}$上的映射，满足

1. $P(A) \geq 0$,
2. $P(\Omega) = 1$,
3. P具有可列可加性：若(A_i)为一列可数个两两不相容/互斥集合，则

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

三元组$(\Omega, \mathcal{F}, \mathbf{P})$称为概率空间。

概率性质

1. $\mathbf{P}(\emptyset) = 0$.
2. 有限可加性: 若$A_1, \cdots, A_n \in \mathcal{F}$互斥, 则

$$\mathbf{P}(\cup_{i=1}^{n} A_i) = \mathbf{P}(\sum_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mathbf{P}(A_i).$$

3. 设若$\mathcal{F}$中的A, B，满足$A \subset B$, 则

$$\mathbf{P}(B - A) = \mathbf{P}(B) - \mathbf{P}(A),$$

$$\mathbf{P}(B) \geq \mathbf{P}(A) \Rightarrow \mathbf{P}(A) \leq 1.$$

4. 任意$A \in \mathcal{F}$, $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$.
5. 加法公式：$\mathcal{F}$中任意A_1, A_2，有

$$\mathbf{P}(A_1 \cup A_2) = \mathbf{P}(A_1) + \mathbf{P}(A_2) - \mathbf{P}(A_1 \cap A_2)$$

推广到n个事件$A_1, \cdots, A_n \in \mathcal{F}$,

$$\begin{aligned}\mathbf{P}(\cup_{i=1}^{n} A_i) = \quad & \textstyle\sum_{i=1}^{n} \mathbf{P}(A_i) - \sum_{1 \leq i \leq j \leq n}^{n} \mathbf{P}(A_i \cap A_j) \\ & + \textstyle\sum_{1 \leq i \leq j \leq k \leq n}^{n} \mathbf{P}(A_i \cap A_j \cap A_k) + \cdots \\ & + (-1)^{n-1} \mathbf{P}(\cap_{i=1}^{n} A_i)\end{aligned}$$

概率性质：连续性

- 次可加性：对于$\mathcal{F}$中的任意可数事件组(A_i)，

$$\mathbf{P}(\cup_i A_i) \le \sum_i \mathbf{P}(A_i).$$

- 连续性：
  1. 若$A_n \downarrow \emptyset$，则$\mathbf{P}(A_n) \downarrow 0$;
  2. 若$A_n \downarrow A$，则$\mathbf{P}(A_n) \downarrow \mathbf{P}(A)$;
  3. 若$A_n \uparrow \Omega$，则$\mathbf{P}(A_n) \uparrow 1$;
  4. 若$A_n \uparrow A$，则$\mathbf{P}(A_n) \uparrow \mathbf{P}(A)$;

定理

假设$\mathbf{P}$满足1,2,及有限可加性，那么以上4条与可列可加性等价.

离散事件概率的计算

概率空间Ω是有限的，或可数的，可以通过为每一个样本ω_i赋概率值来定义概率P，记$\mathbf{P}(\omega_i) = p_i$，则要满足

1. $0 \le \mathbf{P}(\omega_i) = p_j \le 1$;
2. $\sum_{\omega_i \in \Omega} \mathbf{P}(\omega_i) = \sum_i p_i = 1$.

对于任一个事件A，由概率的可加性

$$\mathbf{P}(A) = \mathbf{P}(\cup_{\omega_i \in A}\{\omega_i\}) = \sum_{\omega_i \in A} \mathbf{P}(\omega_i) = \sum_{\omega_i \in A} p_i.$$

等概率

若概率空间中的每个样本等概率取到，i.e. 对于$\forall \omega_i, \omega_j \in \Omega$，$p_i = p_j$，则

$$\mathbf{P}(A) = \frac{A\text{中基本事件的个数}}{\text{所有基本事件的总数}}.$$

原理

1. 乘法原理：完成需要k个步骤，步骤i有m_i个选择，那么一共有$\prod m_i$种方法.
2. 加法原理：完成需要k个途径，途径i有m_i个选择，那么一共有$\sum m_i$种方法.

排列与组合

1. 排列：从n个不同元素中取$r(r \leq n)$个元素排成一排, 共有$\frac{n!}{(n-r)!}$种方法,
2. 重复排列，从n个不同元素中每次取一个，记录后放回后再取下一个, 连续r次得到的排列，共有n^r种方法,
3. 组合: 从n个不同元素中取$r(r \leq n)$个并成一组，共有$C_n^r = \frac{n!}{r!(n-r)!}$种方法,
4. 重复组合：从n个不同元素中每次取一个，记录后放回后再取下一个, 连续r次得到的组合，共有$C_{n+r-1}^r = \frac{(n+r-1)!}{r!(n-1)!}$ 种方法.

几何概率

- 在不可数集合上定义概率比之前的情况难很多,
- 想法：概率常与区间长度有关$\rightarrow$ 概率就和测度论结合，
- 概率可以看做测度的特例，即要求全空间的测度需为1.

几何概率

随机试验的样本空间为$\mathbb{R}^d$的一个区域$|\Omega|$(不可数多个点), 区域中的点视为实验样本ω. 概率事件A为Ω的子集$\{\omega\in A\}$(需要一定可测性). 在等概率假设下，A的概率为

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

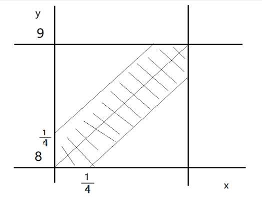
这里$|A|$, $|\Omega|$表示对应集合的面积。

约会问题：

甲乙两人约定8点到9点见面，并约定先到的的人只等15分钟就离开，那么两人见面的概率是多少？

解：假定两人在约定时间段内的任何时间到达的可能性相同。用x,y记甲乙分别到达的时间点，

$$\Omega = \{(x, y) : 8 \leq x, y \leq 9\} \;\; A = \{(x, y) : |x - y| \leq \frac{1}{4}\}.$$

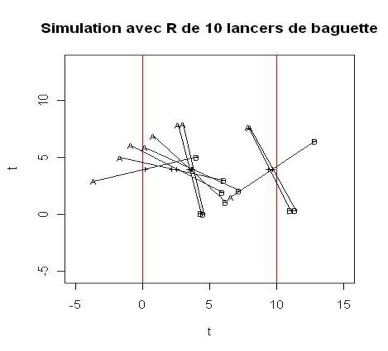


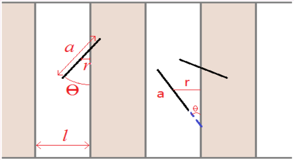
于是$P(A) = \frac{7}{16}$.

布丰掷针

Georges Louis Leclerc, count of Buffon–1733年

向地板上间隔为l 的平行线上掷针，这些针的长度为$a \leq l$, 并观测是否有针与所铺的地板的缝相交(即它是否碰到两块不同的木板)。





解：设A={针与地板缝相交}. 记θ为针与平行线的夹角，r为针的中点到平行线的距离。$\Omega = \{(r,\theta) : 0 \le r \le \frac{l}{2}, 0 \le \theta \le \frac{\pi}{2}\}$. 显然，$(0,0) \in A$，当$\theta \neq 0$ 时，

$$\begin{aligned}\text{针与平行线相交} \quad &\Leftrightarrow \quad \frac{r}{\sin\theta} \le \frac{a}{2} \\ &\Leftrightarrow \quad r \le \frac{a}{2}\sin\theta.\end{aligned}$$

于是

$$\mathbf{P}(A) = \frac{\int_0^{\frac{\pi}{2}} \frac{a}{2}\sin\theta d\theta}{\frac{l}{2}\cdot\frac{\pi}{2}} = \frac{2a}{\pi l}.$$

最初的蒙特卡罗方法

以实验的方法计算π,

$$\pi = \frac{2a}{l} \times \frac{1}{\mathbf{P}(A)}.$$

几何面积⇒ Lebesgue测度

问题.

往区间$[0,1]$中随机抛掷一个质点,

1. 质点落在区间中点的概率;
2. 质点落在区间中的有理数点的概率;
3. 质点落在区间$(0,1)$中的概率。

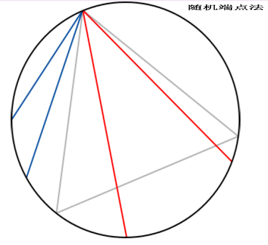
解: $\Omega = [0,1]$, 用E表示事件,

1. $E = \{\frac{1}{2}\}$, 于是$P(E) = \frac{|E|}{|\Omega|} = 0$.
2. $E = \{\text{有理数}\}$, 于是$P(E) = \frac{|E|}{|\Omega|} = 0$.
3. $E = (0,1)$, 于是$P(E) = \frac{|E|}{|\Omega|} = 1$.

Bertrand悖论

考虑圆中的一个等边三角形，在圆中随机选择一条弦，其长度大于比三角形的一条边长的长度的概率是多少？

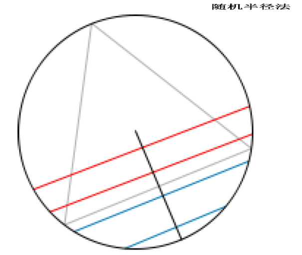
Bertrand 给出了三种方法得到三种结果：

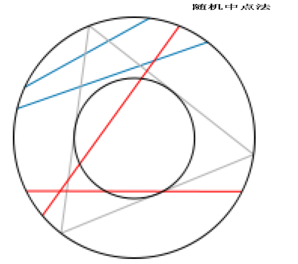


- **随机端点法**：在圆周$(=\Omega)$上随机选两个点，得到一条弦。将三角形旋转到其一个角与弦的某端点重合。$A=\{$ 弦比等边三角形的边长长$\}=\{$弦的另一个端点在这个角对着的那段弧上$\}$.

  于是$P(A)=\frac{|A|}{|\Omega|}=\frac{\text{弧长}}{\text{圆周}}=\frac{1}{3}$.

Bertrand悖论





- **随机半径法**：选择圆的一个半径(=Ω)OB，在半径上随机取一点A，通过A做一条垂直于这条半径的弦。将三角形旋转使其边同垂直于OB。
  $A=\{$弦的长度大于三角形的边长的$\}=\{$ A点比等边三角形边与这条弦的交点C 更靠近圆心O$\}$.
  于是$P(A)=\frac{|A|}{|\Omega|}=\frac{\text{OC长度}}{\text{半径长度}}=\frac{1}{2}$.
- **随机中点法**：在圆(=Ω)中随机选取一个点，做一条弦以此点为中点.
  $A=\{$弦长比三角形边长长$\}=\{$ 这个点在半径$\frac{1}{2}$的同心圆内$\}$。
  于是$P(A)=\frac{|A|}{|\Omega|}=\frac{\text{小圆面积}}{\text{大圆面积}}=\frac{1}{4}$.

实验数据的真伪

掷硬币实验

掷100次硬币，将实验结果记录下来：1表示头像，0表示数字。

- 甲的记录：
  10011101001001101101100011010010111011011110001100
  10110110001011000101100100011011101000101110101010
- 乙的记录：
  10010101001010011010011001010101100100100111001010
  01010101100011000100100100101010101101011001110110
- 丙的记录:
  11001000101001000101000000110100111110001111010000
  00100110000001111101101000101010100011001011101011

问题：三人是真正做过实验，还是伪造数据？

游程Excursion

连续的0称作0游程，连续的1称作1游程。
每个0或1的个数称为游程长度。

简单统计：

- 甲的记录：52个1，48个0，连续1最多4个，连续0至多3个，0-游程30个，
- 乙的记录：47个2，53个0，连续1最多4个，连续0至多3个，0-游程36个，
- 丙的记录：46个1，54个0，连续1最多6个，连续0至多6个，0-游程25个。