第一章 随机事件与概率

第一章 随机事件与概率

§2. 频率与概率

- 在相同条件下，进行了n次试验，其中事件A发生n_A次，称为事件A发生的频数，比值

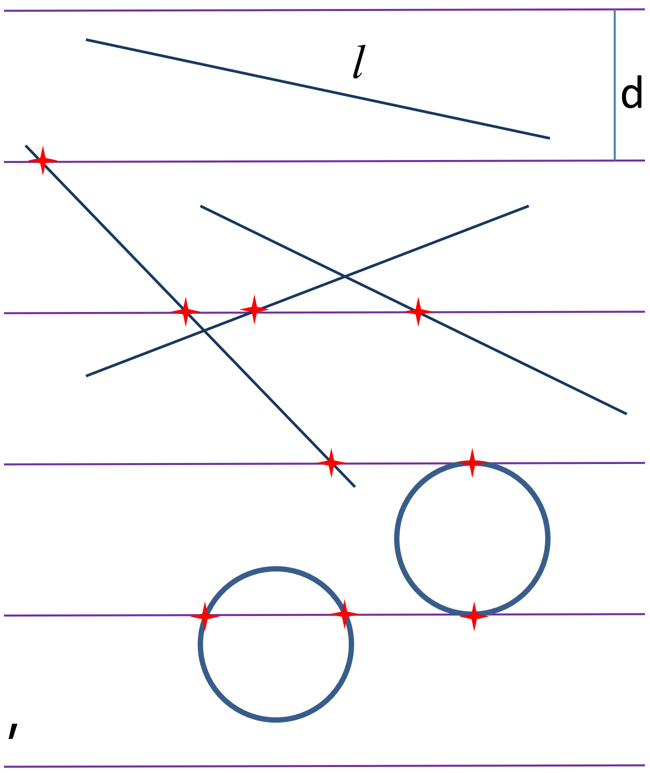
$$n_A/n$$

称为事件A发生的频率，并记为$f_n(A)$

- 基本性质
    - $0 \le f_n(A) \le 1$
    - $f_n(\Omega) = 1$
    - 若$\mathrm{A}_1, \mathrm{A}_2, \dots, \mathrm{A}_\mathrm{k}$是两两互不相容的事件，则 $f_n(A_1 \cup A_2 \cup \cdots \cup A_k) = f_n(A_1) + f_n(A_2) + \cdots + f_n(A_k)$

蒲丰投针——通过随机交点估算π

- d -- 平行线的间距，n – 随机投下的次数
- m – 交点总数，l – 针的长度
- 如果圆的直径为d，投下n次，会产生$2n$个交点
- 设想把圆圈拉直，变成一条长为πd的铁丝段，随机扔下，每次可能与平行线产生0, 1, 2, 3，4个交点，考虑机会均等，总的交点数大致不变
- 长度为l 的针，投下后与平行线相交的交点总数$m \propto l$，即$m = kl$
- 考虑 $l = \pi d$ 的特殊情形，此时有m=2n，于是求得$k = \frac{2n}{\pi d}$，于是：$m = \frac{2l \cdot n}{\pi d}$，从而$\pi = \frac{2l \cdot n}{m \cdot d}$



概率

- 定义：设E是随机试验，Ω是其样本空间，对于E的每一个事件A赋予一个实数P(A) ，如果集合函数P(.)满足下列条件，则称P(A)为事件A的概率。
  - 非负性：P(A)>=0;
  - 规范性：P(Ω)=1;
  - 可列可加性：若$\mathrm{A}_1, \mathrm{A}_2, \ldots$，是两两互不相容的事件，即对于$\mathrm{A}_i\mathrm{A}_j = \emptyset, i \neq j, i, j = 1,2, \ldots$，有
    $$\mathrm{P}(A_1 \cup A_2 \cup \cdots) = \mathrm{P}(A_1) + \mathrm{P}(A_2) + \cdots$$

基本性质

- 性质1：（空集的概率）$\mathrm{P}(\emptyset)=0$
- 性质2：（有限可加性）若$\mathrm{A}_1, \mathrm{A}_2, \ldots, \mathrm{A}_\mathrm{n}$ 是两两互不相容的事件，有

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

- 性质3：（差事件的概率）设A, B是两个事件，若$\mathrm{A} \subset \mathrm{B}$，则有

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

- 性质4：$\mathrm{P}(\mathrm{A}) \leq 1$

基本性质

- 性质5：（逆事件的概率）对任一事件A，有

$$\mathrm{P}(\bar{A}) = 1 - \mathrm{P}(\mathrm{A})$$

- 性质6：（加法公式）对任意两事件A, B，有

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

- 一般推广：

$$\mathrm{P}(A_1 \cup A_2 \cup \cdots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \le i < j \le n} P(A_i A_j) + \sum_{1 \le i < j < k \le n} P(A_i A_j A_k) + \cdots + (-1)^{n-1} P(A_1 A_2 \ldots A_n)$$

§3. 古典概型与几何概型

- 古典概型的两个特点：
  - 样本空间只包含有限个元素，样本空间
  $$\Omega = \{e_1, e_2, \dots, e_n\}$$
  - 试验中每个基本事件发生的可能性相同
  $$P(e_1) = P(e_2) = \cdots = P(e_n)$$
- 古典概型中的概率
  - 设事件A中所含样本点个数为N(A)，以N(Ω)记样本空间Ω中样本点总数，则有
  $$P(A) = \frac{N(A)}{N(\Omega)}$$
  - 基本性质
    - $0 \le P(A) \le 1$
    - $P(\Omega) = 1, P(\emptyset) = 0$
    - $AB = \emptyset$，则$P(A \cup B) = P(A) + P(B)$

一些古典概型问题——超几何分布

例：N件产品中有D件是次品，从中任取n件，问其中恰有k(k≤D)件次品的概率

解：总的取法(可能空间)$\binom{N}{n}$，实际可以分为取出k件次品和(n-k)件合格品两步，于是，可能取法数为：

$$\binom{D}{k}\binom{N-D}{n-k}$$

所求概率为：

$$p=\binom{D}{k}\binom{N-D}{n-k}/\binom{N}{n}$$

动物种群调查

- 鱼塘中捕捉1000尾鱼，做上标记后放回鱼塘，经过一段时间，再次捕捉2000尾鱼，发现其中有100尾有标记，能否对鱼塘中鱼的数量做出些结论

解：1000尾~次品数量D

100尾~恰有k件次品

2000尾~任取n件

池塘中鱼的数量~N件产品

满足条件的抽样概率为：

$$p_N = \binom{1000}{100}\binom{N-1000}{2000-100}/\binom{N}{2000}$$

假设极端情况：若N=2900

$$p_{2900} = \frac{\binom{1000}{100}\binom{1900}{1900}}{\binom{2900}{2000}} \approx 3.34 * 10^{-639}$$

$$f(N)=\frac{p_N}{P_{N-1}}=\frac{\dfrac{\binom{1000}{100}\binom{N-1000}{1900}}{\binom{N}{2000}}}{\dfrac{\binom{1000}{100}\binom{N-1-1000}{1900}}{\binom{N-1}{2000}}}$$

$$=\frac{\dfrac{(N-1000)!\,(N-2000)!\,2000!}{1900!\,(N-1900-1000)!\,N!}}{\dfrac{(N-1-1000)!\,(N-1-2000)!\,2000!}{1900!\,(N-1-1900-1000)!\,(N-1)!}}$$

$$=\frac{(N-1000)(N-2000)}{(N-1900-1000)N}=\frac{(N-1000)(N-2000)}{(N-2900)N}$$

满足条件的解：

$$f(N) \to 1, or\ f(N) - 1 \to 0$$

可以写成：

$$\min_N\{g(N) = |f(N) - 1|\}, or\ \min_N\{g(N) = (f(N) - 1)^2\}$$

于是

$$g'(N) = \frac{(N - 1000)(N - 2000)}{(N - 2900)N} - 1 = 0$$

解得：

$$N = 20000$$

一些古典概型问题——二项分布

考虑有放回抽样(二项分布模型)

从含有D件有缺陷的N件产品中有放回地抽取n件产品进行排列，可能的排列数为N^n个，将每一排列看作基本事件，总数为N^n 。

而在N件产品中取n件，其中恰有k件次品的取法共有$C_n^k D^k (N-D)^{n-k}$

于是所求的概率为

$$p = \frac{C_n^k D^k (N-D)^{n-k}}{N^n} = C_n^k \left(\frac{D}{N}\right)^k \left(1 - \frac{D}{N}\right)^{n-k}$$

此式即为二项分布的概率公式

一些古典概型问题

例：抛一枚硬币三次，设事件A1为至少出现一次正面，求P(A1)

解：样本空间：

$\{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}, \mathrm{THH}, \mathrm{THT}, \mathrm{TTH}, \mathrm{TTT}\}$

$\underline{A1} = \{\mathrm{HHH}, \mathrm{HHT}, \mathrm{HTH}, \mathrm{HTT}, \mathrm{THH}, \mathrm{THT}, \mathrm{TTH}\}$

$\overline{A1}$={TTT}

$P(A1) = 7/8$

或者

$$P(A1) = 1 - P(\overline{A1}) = 1 - 1/8 = 7/8$$

一些古典概型问题

例：n个球，随机放入N（N≥n）个盒子中，试求每个盒子至多有一个球的概率(假设盒子最多可以装n个球)

解：样本空间大小：每个球可有N个选择，共有N^n中可能。满足条件的空间大小：N(N-1)(N-2)...(N-n+1)

$$p=\frac{N(N-1)(N-2)\ldots(N-n+1)}{N^n}=\frac{N!}{N^n(N-n)!}$$

- 类似问题：n(≤365)个同学生日不同的概率(假定每年365天)

$$p=\frac{365!}{365^n(365-n)!}$$

n	20	30	40	50	100
p	0.589	0.294	0.109	0.030	0.0000003

- 一般地，把n个球随机地分成m组(n>m),要求第 *i* 组恰有 n_i 个球(*i*=1,…, *m*)，共有分法数量。

$$\binom{n}{n_1}\binom{n-n_1}{n_2}\binom{n-n_1-n_2}{n_3}\binom{n-\sum_{i=1}^{m-1} n_i}{n_m}$$

$$=\frac{n!}{(n-n_1)!\,n_1!}\cdot\frac{(n-n_1)!}{(n-n_1-n_2)!\,n_2!}\cdot\cdots\cdot\frac{\left(n-\sum_{i=1}^{m-1} n_i\right)!}{0!\,n_m!}$$

$$=\frac{n!}{n_1!\,n_2!\cdot\cdots\cdot n_m!}$$

例：袋中有a个白球，b个红球，$k(\leq a+b)$个人抽，分别做放回和不放回抽样，求第$i(=1,2,\dots,k)$个人抽到白球的概率

解：(1)放回抽样显然为：

$$p=\frac{a}{a+b}$$

(2) 不放回抽样，空间大小：

$$(a+b)(a+b-1)\dots(a+b-k+1)$$

满足条件的事件可以理解为第i个人优先取一个白球，于是有a种取法，剩下$(a-1+b)$个球供其他$k-1$个人抽取，于是事件空间大小为：

$$a(a+b-1)(a+b-1-1)\dots(a+b-1-(k-1)+1)$$

概率为：

$$p=\frac{a}{a+b}$$

从工作记录分析工作状态

某工作应该每天进行观察，每天可能等概率发生异常，并将异常情况加以记录，从记录本上看到20条异常记录，均发生在周一、二和五，你认为记录者是在每天观察吗？

分析：20次异常均发生在三天中的概率$\frac{3^{20}}{7^{20}} \approx 4.37 * 10^{-8}$

可以推断记录不可信

例：30名学生中有3名运动员，将这30名学生平均分成3组，求：
（1）每组有一名运动员的概率；
（2）3名运动员集中在一个组的概率。

解: 设事件A: 每组有一名运动员;
事件B: 3名运动员集中在一组

$$N(\Omega) = \binom{30}{10}\binom{20}{10}\binom{10}{10} = \frac{30!}{10!\,10!\,10!}$$

$$P(\mathrm{A}) = \frac{3!\,\frac{27!}{9!\,9!\,9!}}{N(\Omega)} = \frac{6000}{30 * 29 * 28} \approx 0.246$$

$$P(\mathrm{B}) = \frac{3 \times \binom{27}{7}\binom{20}{10}\binom{10}{10}}{N(\Omega)} = \frac{3 * 27!}{7!\,\frac{30!}{10!}} \approx 0.0887$$

随机取数问题

从1到2000这2000个自然数中任取一个数,

(1) 求取到的数能被6整除的概率(事件A)

(2) 求取到的数能被8整除的概率(事件B)

(3) 求取到的数既能被6整除也能被8整除的概率(事件C)

(4) 求取到的数既不能被6整除也不能被8整除的概率(事件D)

解：$\mathrm{N}(\Omega) = 2000,$

$$\mathrm{N}(\mathrm{A}) = \lfloor 2000/6 \rfloor = 333, P(A) = \frac{N(A)}{N(\Omega)} = \frac{333}{2000}$$

$$\mathrm{N}(\mathrm{B}) = \lfloor 2000/8 \rfloor = 250, P(B) = \frac{N(B)}{N(\Omega)} = \frac{1}{8}$$

$$\mathrm{N}(\mathrm{C} = \mathrm{AB}) = \lfloor 2000/24 \rfloor = 83, P(C) = \frac{N(C)}{N(\Omega)} = \frac{83}{2000}$$

$$\mathrm{N}(\mathrm{D}) = \mathrm{N}(\Omega) - \mathrm{N}(\mathrm{A}) - \mathrm{N}(\mathrm{B}) + \mathrm{N}(AB) \ = 1500, P(D) = \frac{N(D)}{N(\Omega)} = \frac{3}{4}$$

几何概型

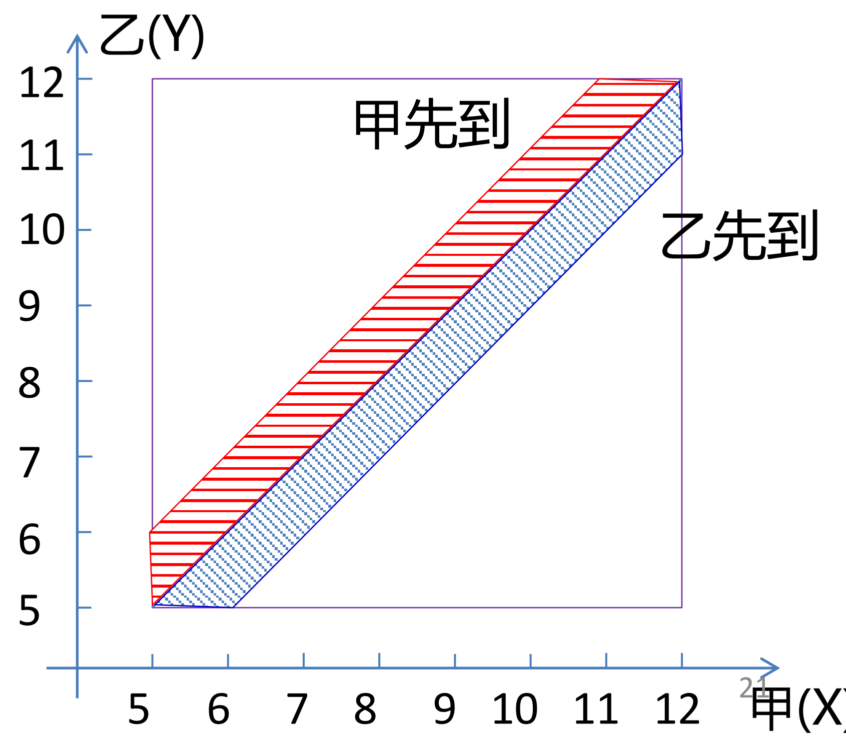
- 古典概型只考虑了有限可能的随机试验
- 几何概型考虑的是有无穷多个等可能结果的随机试验

例 (会面问题)　甲、乙二人约定在 5 点到 12 点之间在某地会面，先到者等一个小时后即离去。设二人在这段时间内的各时刻到达是等可能的，且二人互不影响。求二人能会面的概率。

解：　以X, Y 分别表示甲乙二人到达的时刻，于是5≤X, Y≤12。

即两人到达时间M(X,Y) 落在图中的矩形部分，有无穷多个可能。

由于每人在任一时刻到达都是等可能的，所以落在正方形内各点是等可能的。

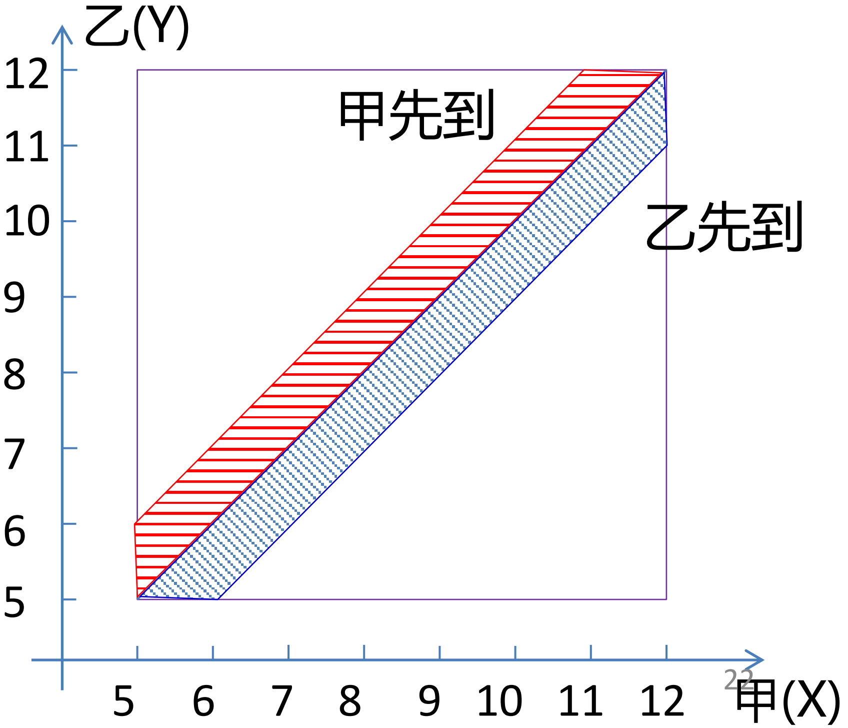


几何概型

二人会面的条件是： |X-Y|≤1

于是：

$$p=\frac{\text{阴影面积}}{\text{正方形的面积}}=\frac{49-2\times\frac{6^2}{2}}{49}=\frac{13}{49}$$



几何概型

一般地，利用某个可度量区域D中的度量m(如线段长度，平面区域面积，空间区域的体积等)。如果随机实验 E 相当于在区域内任意地取点，且取到每一点都是等可能的，则称此类试验为几何概型。

如果试验 E 是向区域内任意取点，事件 A 对应于点落在 D 内的某区域 A，则

$$P(A) = \frac{m_A}{m_D}$$

蒲丰投针的几何概型解释

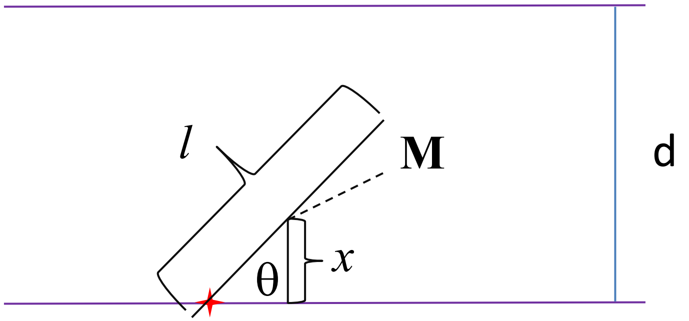
平面上有一族平行线。其中任何相邻的两线距离都是d（d>0）。向平面任意投一长为 l (l<d) 的针，试求针与一条平行线相交的概率。

解：设 x 是针的中点 M 到最近的平行线的距离，θ是针与此平行线的交角，投针问题就相当于向平面区域 D 取点的几何概型。

可能的区域$D=\{(\theta,x),|0\le\theta\le\pi,0\le x\le\frac{d}{2}\}$

满足相交的条件的区域，

$A=\{(\theta,x),|0\le\theta\le\pi,0\le x\le\frac{l}{2}\sin\theta\}$



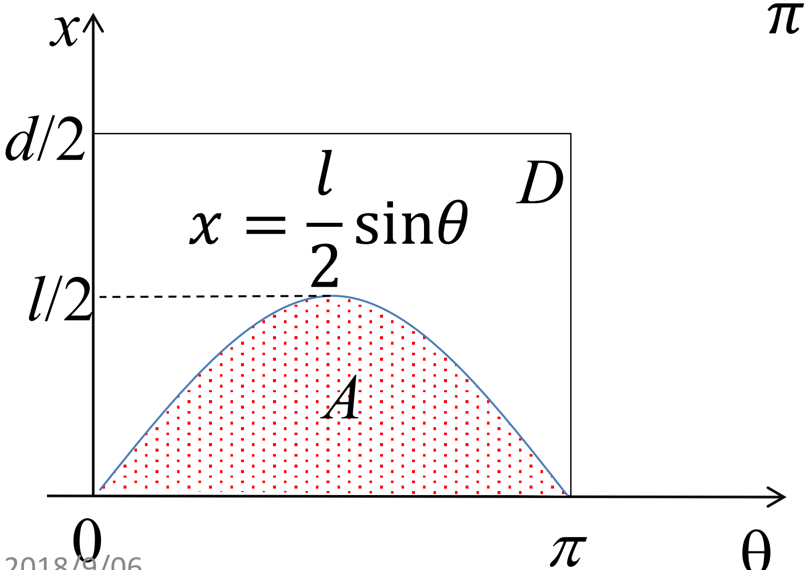
蒲丰投针的几何概型解释

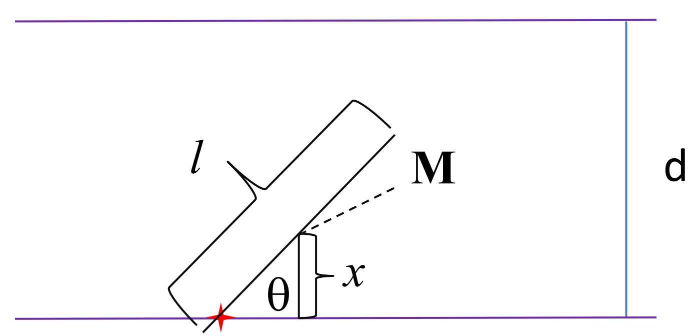
可能的区域$D=\{(\theta,x),|0\le\theta\le\pi,0\le x\le\frac{d}{2}\}$

满足相交的条件的区域$A=\left\{(\theta,x),\middle|0\le\theta\le\pi,0\le x\le\frac{l}{2}\sin\theta\right\}$

$$p=\frac{A\text{的面积}}{D\text{的面积}}=\frac{\int_0^\pi\frac{l}{2}\sin\theta d\theta}{\frac{d}{2}\pi}=\frac{2l}{\pi d}$$

$$\pi=\frac{2l}{dp}$$





§4. 条件概率

例：抛一枚硬币两次，

事件A：至少出现一次正面

事件B：两次为同一面

求：事件A发生条件下，事件B出现的概率

Ω={HH, HT, TH, TT}

A= {HH, HT, TH}

B={HH, TT}

A发生的条件下，意味着事件的空间$\Omega \rightarrow$A

而B和A的交集只有{HH} ➔ $P(B|A)=1/3$

例 一盒中混有100只新，旧乒乓球，各有黄、白两色，分布如下表。从盒中随机取出一球，若取得的是一只黄球，试求该黄球是新球的概率。

A——取出一只黄球，$n_A = 60$

B——取出一只新球，$n_{AB} = 40$

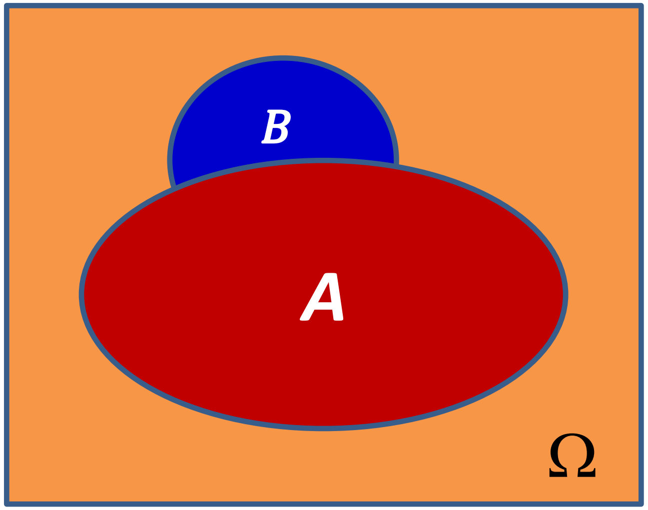
	黄	白
新	40	30
旧	20	10

B

A

$$P(B|A) = \frac{n_{AB}}{n_A} = \frac{2}{3}$$

- 条件概率是概率论中一个重要而实用的概念。它所考虑的是事件 A 已经发生的条件下事件B发生的概率。
- 设A、B是某随机试验中的两个事件，且P(A)>0，则称事件B在“事件A已发生”这一附加条件下的概率为在事件A已发生的条件下事件B的条件概率，简称为B在A之下的条件概率，记为P(B|A)



条件概率与无条件约束的概率

例　盒中有4个外形相同的球，它们的标号分别为1、2、3、4，每次从盒中取出一球，有放回地取两次．求取出的两球标号之和为 4的概率.

解: 设B={取出的两球标号之和为 4}.该试验的所有可能的结果为

$$\begin{gathered}\{(1,1),(1,2),\mathbf{(1,3)},(1,4),\\(2,1),\mathbf{(2,2)},(2,3),(2,4),\\\mathbf{(3,1)},(3,2),(3,3),(3,4),\\(4.1),(4,2),(4,3),(4,4)\}\end{gathered}$$

其中(i,j)表示第一次取i号球，第二次取j号球

于是P(B)=3/16

考虑另一个问题：已知第一次取出球的标号为 2,求取出的两球标号之和为4的概率。

设A={ 第一次取出球的标号为 2 }

即求在事件A发生的条件下，事件B发生的概率P(B|A)

由于事件A已经发生，则该试验的所有可能结果为

$$\{(2,1),\mathbf{(2,2)},(2,3),(2,4)\}$$

于是P(B|A)=1/4

附加”条件A已发生”的概率与不附加这个条件的概率是不同的

条件概率

一般化，

设试验的基本事件总数为n，A所包含的基本事件数为m(m>0)，AB所包含的基本事件数为k，即有

$$P(B|A)=\frac{k}{m}=\frac{k/n}{m/n}=\frac{P(AB)}{P(A)}$$

条件概率定义：设A, B是两个事件，且P(A)>0，称

$$\mathrm{P(B|A)}=\frac{P(AB)}{P(A)}$$

为在事件A发生的条件下事件B发生的条件概率。

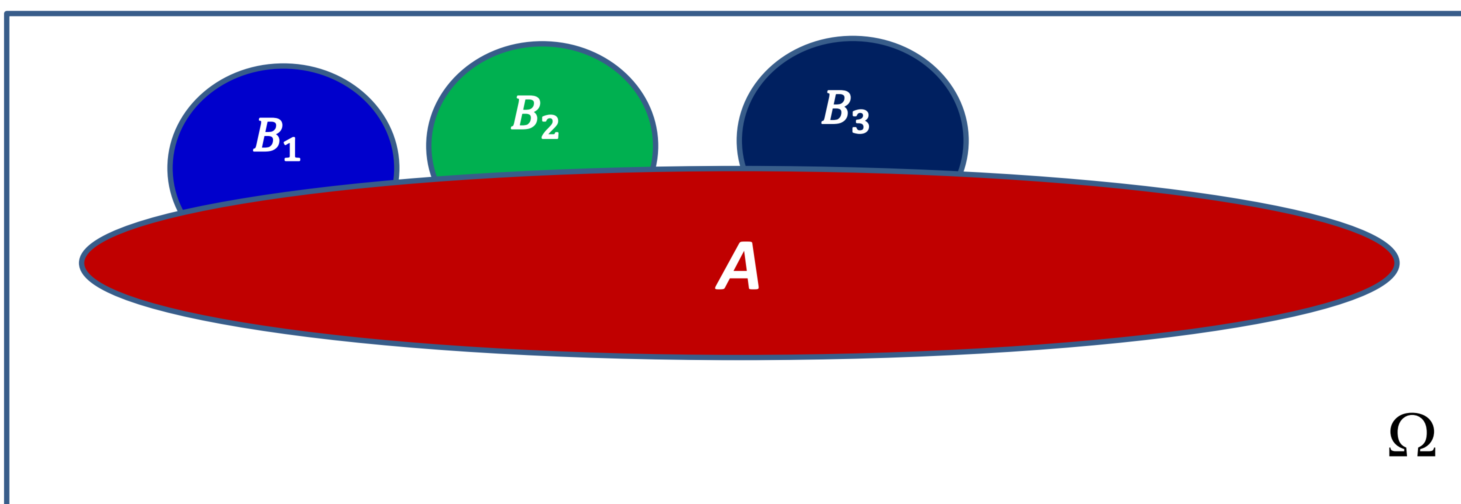
条件概率的计算(对应上述)

- 缩小样本空间法——适用于古典概型
- 公式法

“条件概率”是“概率”吗？

- 概率三性质检验
  - 非负性：对任一事件B，满足P(B|A)≥0
  - 规范性：对必然事件Ω，满足P(Ω|A)=1
  - 可列可加性：设$\mathrm{B}_1, \mathrm{B}_2, \ldots$是一系列两两互不相容的事件，则有

$$\mathrm{P}\left(\bigcup_{i=1}^{\infty} B_i | A\right) = \sum_{i=1}^{\infty} P(B_i|A),$$



“条件概率”是“概率”吗？

- 概率三性质检验
  - 非负性：对任一事件B，满足P(B|A)≥0
  - 规范性：对必然事件Ω，满足P(Ω|A)=1
  - 可列可加性：设$B_1, B_2, \ldots$是一系列两两互不相容的事件，则有

$$P\left(\bigcup_{i=1}^{\infty} B_i | A\right) = \sum_{i=1}^{\infty} P(B_i|A),$$

- 其他性质
  - (逆事件) $P(\overline{B}|A) = 1 - P(B|A)$
  - (加法公式) 对任意两事件B_1, B_2，有
    $P(B_1 \cup B_2|A) = P(B_1|A) + P(B_2|A) - P(B_1B_2|A)$

市场上的手机中，甲产品的市场占有率50%，乙产品的市场占有率30%，其他品牌的占有率为20%。甲产品的合格率是95%，乙产品的合格率是80%，其他品牌的合格率为60%。若用事件A,B,C分别表示甲、乙和其他品牌的产品，Q表示产品为合格品，相关的概率如下：

$P(A)=0.5,\ P(Q|A)=0.95$

$P(B)=0.3,\ P(Q|B)=0.8$

$P(C)=0.2,\ P(Q|C)=0.6$

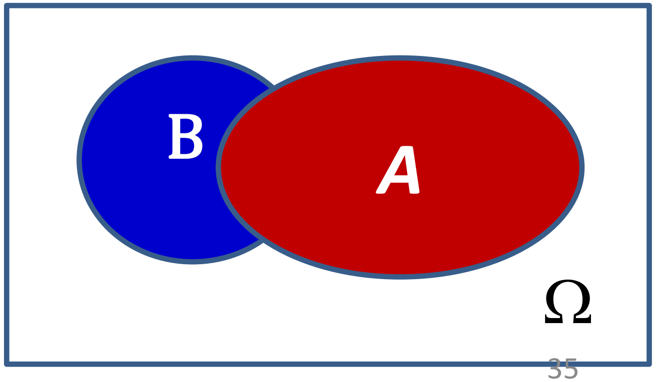
某地区连续两天出现清晨大雾的概率为0.1，根据天气情况，第一天大雾的概率为0.6，第二天大雾的概率为0.3，游客第一天遇到大雾，求第二天没有大雾的概率。

解：设事件A_1与A_2分别表示第一与第二天出现大雾

$$\mathrm{P}(\overline{A_2}|A_1) = \frac{P(\overline{A_2}A_1)}{P(A_1)} = \frac{P(A_1) - P(A_1A_2)}{P(A_1)} = \frac{0.6 - 0.1}{0.6} = \frac{5}{6}$$

- 条件概率与无条件概率之间的大小无确定关系

$$\frac{|\mathrm{B}|}{|\Omega|} \text{ vs. } \frac{|\mathrm{AB}|}{|\mathrm{A}|}$$



乘法公式

由条件概率

$$\mathrm{P(B|A)} = \frac{P(AB)}{P(A)}$$

有下面的乘法定理：

对A, B$\subset\Omega$，P(A)>0, $\mathrm{P(AB) = P(B|A)P(A)}$

推广到三事件A, B, C，假设P(AB)>0，有：

$$\mathrm{P(ABC) = P(C|AB)P(B|A)P(A)}$$

推广到更一般情况，假设$\mathrm{P(A_1A_2 \ldots A_{n-1}) > 0}$，有：

$$\mathrm{P(A_1A_2 \ldots A_n)}$$
$$\mathrm{= P(A_n|A_1A_2 \ldots A_{n-1})P(A_{n-1}|A_1A_2 \ldots A_{n-2}) \ldots P(A_2|A_1)P(A_1)}$$

波利亚罐子模型(传染模型) 盒中有r个红球，t个白球，每次从袋中任取一只，观察其颜色后放回，并再放入a只与所取之球颜色相同的球，若从盒中连续取球4次

(1) 试求第1、2次取得红球、第3、4次取得白球的概率

$A_i(i=1,2,3,4)$表示第i次取得红球，而$\bar{A}_i(i=1,2,3,4)$表示第i次取得白球

$$\mathrm{P}(A_1A_2\bar{A}_3\bar{A}_4)=\mathrm{P}(\bar{A}_4|A_1A_2\bar{A}_3)\mathrm{P}(\bar{A}_3|A_1A_2)\mathrm{P}(A_2|A_1)\mathrm{P}(A_1)$$
$$=\frac{t+a}{r+t+3a}\cdot\frac{t}{r+t+2a}\cdot\frac{r+a}{r+t+a}\cdot\frac{r}{r+t}$$

(2) 试求第1、2次取得白球、第3、4次取得红球的概率

$$\mathrm{P}(\bar{A}_1\bar{A}_2A_3A_4)=\mathrm{P}(A_4|A_3\bar{A}_2\bar{A}_1)\mathrm{P}(A_3|\bar{A}_2\bar{A}_1)\mathrm{P}(\bar{A}_2|\bar{A}_1)\mathrm{P}(\bar{A}_1)$$
$$=\frac{r+a}{r+t+3a}\cdot\frac{r}{r+t+2a}\cdot\frac{t+a}{r+t+a}\cdot\frac{t}{r+t}$$

- 一般性地，对波利亚罐子模型，先取出n_1个红球，再取出n_2个白球的概率为

$$p=\frac{r(r+a)\dots(r+n_1a-a)t(t+a)\dots(t+n_2a-a)}{(r+t)(r+t+a)\dots(r+t+(n_1+n_2)a-a)}$$

作业

- 概率论及其应用
  - p. 20, #14, #16, #17, #19
  - pp. 42-44, #9, #11, #15, #24, #38
  - pp. 107-108, #6, #7, #18, #19
- 概率论与数理统计
  - pp. 25-29, #2, #8, #12, #13, #16, #18, #19