

概率论与数理统计

任课教师: 胡家信

清华大学数学科学系

2022 年 9 月 27 日

上周内容

- σ -代数：设 X 是一个集合， \mathcal{F} 是 X 的子集构成的集合，如果 \mathcal{F} 关于补运算和可数并运算是封闭的：

$$1) \text{若 } A \in \mathcal{F}, \text{ 则 } A^c \in \mathcal{F}; \quad 2) \text{若每个 } A_i \in \mathcal{F}, \text{ 则 } \bigcup_{i=1}^{\infty} A_i \in \mathcal{F},$$

则称 \mathcal{F} 是 X 上的一个 σ 代数。

- 概率的定义：概率是一个集函数（定义在集合上的函数），满足：非负性、正规化、可数可加性。

概率的定义

概率的定义： 设 Ω 是一个样本空间(可数或者不可数)， \mathcal{F} 是 Ω 上的一个 σ -代数，**概率** P 是一个**集函数**，满足下列三个条件：

- $P(A) \geq 0$ ，对每个 $A \in \mathcal{F}$ ；
- $P(\Omega) = 1$ ；
- 对每个两两互不相交的 $A_n \in \mathcal{F}$ ，均有

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n).$$

(可数可加性)

概率的定义（续）

总之，概率 P 是一个集函数，其定义域 \mathcal{F} 一般为：

- 有限样本空间 Ω : \mathcal{F} 是所有 Ω 的子集构成，包含空集和本身。
- 可数样本空间 Ω : \mathcal{F} 是所有 Ω 的子集构成，包含空集和本身。
- 不可数样本空间 Ω : \mathcal{F} 是一个 σ -代数，即 \mathcal{F} 关于补运算和可数并运算是封闭。

此时概率理解为一个测度，满足

$$P(\Omega) = 1.$$

概率的定义（续）

概率和函数的区别:

- 概率 $P(A)$
 - 集合 A
 - 定义域为集合的集合
 - 值域为 $[0, 1]$
- ♡ 函数 $f(x)$
 - ♡ 实数 x
 - ♡ 定义域为某个实数集
 - ♡ 值域不定

概率的定义（续）

例：设样本空间 Ω 如下：

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

定义概率 P 如下：对任意 $1 \leq k \leq 6$ ，令 $P(\{k\}) = \frac{1}{6}$ 。

计算：

$$P(\{1, 2, 3\}) = ?$$

$$P(\{1, 2, 3, 4, 5\}) = ?$$

若集合只含一个点，有时写成 $P(k) := P(\{k\})$ ，与函数符号一样！

概率的定义（续）

例：设 $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ 是一个无限集合。令 $\{p_i\}_{i=1}^{\infty}$ 是一个数列，满足每个 $p_i \geq 0$ ，且

$$\sum_{i=1}^{\infty} p_i = 1.$$

构造集函数 P 如下：

$$P(\omega_i) = p_i, \quad i \geq 1;$$

$$P(A) = \sum_{\omega \in A} P(\omega).$$

这样构成的集函数 P 是 Ω 上的一个概率。

独立事件

定义：一个事件就是概率 P 定义域 \mathcal{F} 中的一个点，即样本空间 Ω 的一个可测子集 A ，使得 $P(A)$ 有意义。

定义：两个事件 A 和 B 独立，若 $P(A \cap B) = P(A)P(B)$ 。

更一般，若干事件 $\{A_m\}_{m=1}^n$ 独立，如果下列成立：

$$P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k})$$

对任意 $\{1, 2, \cdots, n\}$ 一组数 $\{i_1, i_2, \cdots, i_k\}$ 。

独立事件（续）

例 $n = 3$: 三个事件 A_1, A_2, A_3 独立, 当且仅当下列均成立:

$$P(A_1 \cap A_2) = P(A_1) P(A_2)$$

$$P(A_1 \cap A_3) = P(A_1) P(A_3)$$

$$P(A_2 \cap A_3) = P(A_2) P(A_3)$$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2) P(A_3).$$

注意: 前三个式子都成立不一定意味第四个式子成立; 反之, 第四个式子成立不一定意味前三个式子都成立。

给出反例 (课后思考题)。

例子1 (Bernoulli公式)

问：一枚硬币连续掷 n 次，正面出现 j 次的概率是多少？

课堂讨论：连续抛硬币4次，正面出现2次的概率是多少？

例子1 (Bernoulli公式)

解: $n = 4, j = 2$ (即连续抛硬币4次, 正面出现2次):

- 正正反反, 正反正反, 正反反正;
(第1次出现正面的情况)
- 反正正反, 反正反正;
(第1次出现反面、同时第2次出现正面的情况)
- 反反正正。
(第1次出现反面、同时第2次出现反面的情况)
- 剩余情况。无

故所求概率为

$$P(A) = 6 \times \left(\frac{1}{2}\right)^4 = \frac{3}{8}.$$

例子1 (Bernoulli公式)

问：一枚硬币连续掷 n 次，正面出现 j 次的概率是多少？

解：掷 n 次，正面出现 j 次，共有

$$\binom{n}{j}$$

种不同情况，每种情况出现的概率为

$$\left(\frac{1}{2}\right)^j \left(\frac{1}{2}\right)^{n-j} = \left(\frac{1}{2}\right)^n. \quad (\text{利用独立性})$$

所求概率为

$$P(A) = \binom{n}{j} \left(\frac{1}{2}\right)^n.$$

课堂思考题

问：一枚不均匀硬币，正面出现的概率为 p ，反面出现的概率为 $q = 1 - p$ 。连续掷 n 次，正面出现 j 次的概率是多少？

课堂思考题

解：掷 n 次，正面出现 j 次，共有

$$\binom{n}{j}$$

不同种情况，每种情况出现的概率为

$$p^j (1-p)^{n-j}. \quad (\text{利用独立性})$$

所求概率为

$$P(A) = \binom{n}{j} p^j (1-p)^{n-j}.$$

例子3

例子3: 已知 $\Omega = \{1, 2, \dots, 120\}$,

$$A = \{\omega \in \Omega : \omega \text{ 是 } 3 \text{ 的倍数}\},$$

$$B = \{\omega \in \Omega : \omega \text{ 是 } 4 \text{ 的倍数}\},$$

求 $P(A \cup B) = ?$ 事件 A 和 B 是独立吗?

例子3

解:

$$P(A) = \frac{40}{120} = \frac{1}{3}, \quad P(B) = \frac{30}{120} = \frac{1}{4}.$$

事件 $AB = \{\omega \in \Omega : \omega \text{ 是 } 12 \text{ 的倍数}\}$, 故

$$P(AB) = \frac{10}{120} = \frac{1}{12}.$$

所求概率为

$$P(A \cup B) = P(A) + P(B) - P(AB) = \frac{1}{3} + \frac{1}{4} - \frac{1}{12} = \frac{1}{2}.$$

事件 A 和 B 是独立吗?

例子3

又知 $C = \{\omega \in \Omega : \omega \text{ 是6的倍数}\}$ 。求 $P(BC) = ?$

解： $P(C) = \frac{20}{120} = \frac{1}{6}$ 。事件

$$BC = \{\omega \in \Omega : \omega \text{ 是12的倍数}\}.$$

故 $P(BC) = \frac{10}{120} = \frac{1}{12}$ 。

事件 B 和 C 是独立吗??

概率的计算

例子4：一家餐馆菜谱如下：

- 1 汤，果汁，开胃酒。
- 2 牛肉，火腿，炸鸡，狮子头。
- 3 土豆泥，兰花菜，豌豆。
- 4 冰激凌，可乐。
- 5 咖啡，茶，牛奶。

每样只选一种，有多少选择？

概率的计算

解：基本原则：做多种选择，

第一个选择有 m_1 种可能，

第二个选择有 m_2 种可能，

第三个选择有 m_3 种可能，

\dots ,

第 n 个选择有 m_n 种可能。

则总共有

$$m_1 \times m_2 \times m_3 \times \dots \times m_n$$

种可能。

例子: $3 \cdot 4 \cdot 3 \cdot 2 \cdot 3 = 216$ 种菜谱。

概率的计算

常见问题：

- 抽样问题。
- 分配问题。

抽样问题

例子5：一只黑箱装标号为1至 m 的球。现抽取 n 个球，有多少种抽法（或结果）？

分下列四情况：

- 抽取后将球放回黑箱，并考虑顺序；
- 抽取后球不放回黑箱，并考虑顺序；
- 抽取后球不放回黑箱，不考虑顺序；
- 抽取后将球放回黑箱，不考虑顺序。

例如： $m = 3$ ， $n = 2$ 。

抽样问题

情形1: 假如抽取后将球放回黑箱，并考虑顺序，有多少个抽法(或结果)?

解: 每次有 m 中可能，故共有 m^n 种结果。
实际上，可考虑集合

$$A = \{(a_1, a_2, \dots, a_n) : \text{每个 } 1 \leq a_i \leq m\},$$

求 A 中元素的个数。

概率的计算

情形2: 假如抽取后球不放回黑箱，并考虑顺序，有多少个抽法(或结果)?

抽样问题

解：第一次有 m 中可能；第2次有 $m - 1$ 中可能；第3次有 $m - 2$ 中可能；

...

第 n 次有 $m - n + 1$ 中可能。

故共有

$$m \cdot (m - 1) \cdots (m - n + 1)$$

种结果。

抽样问题

情形3: 假如抽取后球不放回黑箱, 不考虑顺序, 有多少个抽法(或结果)?

例子: $m = 5$, $n = 3$ 。出现结果为 $A = \{235\}$, 即标号为2, 3, 5的球各出现一次的结果, 共有下列几种情形:

$(235), (253), (352), (325), (523), (532)$.

若不讲顺序, 上面情形一样, 看成一个结果。

抽样问题

解：现 n 次抽取，球不放回黑箱，考虑顺序，这 n 个球排列时共有 $n!$ 种情形，不考虑顺序时，这些情形视为一种结果。现设有 x 种不同的结果，则

$$n!x = m(m-1) \cdots (m-n+1)$$

上述右端为考虑顺序下的所有情况，即 $n!x$ 等于讲顺序的所有情形。解得，

$$x = \frac{m(m-1) \cdots (m-n+1)}{n!} = \binom{m}{n},$$

故抽取后球不放回黑箱，不考虑顺序，有 $\binom{m}{n}$ 个抽法。

抽样问题

等价：从 m 个球中，一次抽取 n 个球，有多少个结果？

$$\binom{m}{n}.$$

抽样问题

情形4 (较复杂): 假如抽取后将球放回黑箱, 不考虑顺序, 有多少个抽法(或结果)?

课堂讨论: 设 $m = n = 3$, 即考虑箱里有3个编号球, 抽3次, 球放回, 不考虑顺序, 有多少结果?

抽样问题

解：设 $m = n = 3$ ，即考虑箱里有3个编号球，抽3次，球放回，不考虑顺序。那么，共有下列10种结果：

- 111("1"出现3次);
- 112, 113("1"出现2次);
- 123, 122, 133("1"出现1次);
- 222, 223, 332, 333("1"出现0次)。

所以，答案为10种抽法(或结果)。

例如"112"的情形可以这样看：

1	2	3
√√	√	

或者√√|√|。或者问题等价于：现有 $n = 3$ 个“√”和 $m - 1 = 2$ 个“|”排成一排，有多少种不同的排列方式？

答案：共有10种不同的情况。

抽样问题

关键想法：相当将 $m_1 = 3 = n$ 个"√", $m_2 = 2 = m - 1$ 个"|" 怎样放。

即有

$$\frac{5!}{3!2!} = \frac{(m_1 + m_2)!}{m_2!m_1!} = \frac{(m + n - 1)!}{(m - 1)!n!} = \binom{n + m - 1}{n}.$$

或者，问题相当从含 $m_1 = n$ 个黑球和 $m_2 = m - 1$ 个白球的黑箱中一一抽取 n 个球，共有多少情况。

答案： $\frac{(m_1 + m_2)!}{m_2!m_1!} = \binom{n + m - 1}{n} = \binom{n + m - 1}{m - 1}.$

分配问题

例6：有多少种方法将4人分为2组，每组2人？

课堂讨论一下

- ① 区分小组（小组编号，如第一组，第二组）。
- ② 不区分小组（不编号，如玩扑克）。

分配问题

答案:

- ① 区分小组（小组编号，如第一组，第二组）:

$(12)(34)$, $(13)(24)$, $(14)(23)$, $(23)(14)$, $(24)(13)$, $(34)(12)$

- ② 不区分小组（不编号，如玩扑克）:

$(12)(34)$, $(13)(24)$, $(14)(23)$.

分配问题

例7：6名登山运动员分3组向顶峰冲击，每组分别为1,2,3人，有多少种方法？

- ❶ 不区分小组，有多少情况？
- ❷ 区分小组，有多少情况？

分配问题

解:

- ① 不区分小组。 G_1, G_2, G_3 (下标分别表示“1人”、“2人”、“3人”，即 G_1 表示该组有1人， G_2 表示该组有2人， G_3 表示该组有3人)。

$$\binom{6}{1} \binom{6-1}{2} \binom{6-1-2}{3} = 60.$$

- ② 若区分小组，有多少情况？答案： $60 \times 3! = 360$ ，这是因为在区分小组情况下，必须将标签“第1组”、“第2组”、“第3组”分别贴在 G_1, G_2, G_3 的运动员背后（例如，将 G_3 的3名运动员后背都贴上同一个标签“第1组”，将 G_2 的2名运动员后背都贴上同一个标签“第2组”，将 G_1 的1名运动员后背贴上标签“第3组”），共有 $3!$ 种情况。解毕。

分配问题

例8：一个篮子装550个苹果，有2%是烂的，随机拿25个苹果有2个是烂的概率是多少？

解：烂苹果 = $550 \times 2\% = 11$ 个。

好苹果 = $550 - 11 = 539$ 个。所求概率为：

$$\frac{\binom{539}{23} \binom{11}{2}}{\binom{550}{25}}.$$

分配问题

例9：一个52张牌4个A在一起的概率是多少？

课堂讨论一下

分配问题

例9：一个52张牌4个A在一起的概率是多少？

解：52张牌有52!种放法。

第1个A放的位置有49种，而4个A的放置方法有4!种，其余48张牌有48!种放法。

所求概率为：

$$\frac{49 \cdot 4! \cdot 48!}{52!} = \frac{24}{52 \times 51 \times 50}.$$

分配问题

课堂讨论题：有15个学生均匀分配成3组，其中有3个尖子生，问每组有1个尖子生的概率是多少？尖子生都在一个组的概率是多少？

- ❶ 区分小组，答案？
- ❷ 不区分小组，答案？

分配问题

解:

(1) 区分小组 (第1组、第2组、第3组)。15名学生均匀分成3组, 共有 $\binom{15}{5} \cdot \binom{10}{5} \cdot \binom{5}{5}$ 种分法, 每组有一个尖子生的分法有

$$\binom{3}{1} \binom{12}{4} \cdot \binom{2}{1} \binom{8}{4}.$$

所以, 每组有一个尖子生的概率为:

$$\frac{\binom{3}{1} \binom{12}{4} \cdot \binom{2}{1} \binom{8}{4}}{\binom{15}{5} \cdot \binom{10}{5}} = \frac{6 \cdot 5^3}{15 \times 14 \times 13} \approx 0.274725.$$

分配问题

3个尖子生在同一个组的分法有

$$3 \cdot \binom{12}{2} \cdot \binom{10}{5} \cdot \binom{5}{5},$$

这是因为3个尖子生放在同一个小组，有3种情况。故所求概率为：

$$\frac{3 \binom{12}{2} \cdot \binom{10}{5}}{\binom{15}{5} \cdot \binom{10}{5}} = \frac{5 \cdot 4 \cdot 3^2}{15 \times 14 \times 13} \approx 0.065934.$$

(2) 若不区分小组，答案是什么？(课后思考题，参考钟书第64-65页)

生日问题

课堂练习（生日问题）：有 n 个人至少有2人生日在同一天的概率是多少？

生日问题

解： n 个人生日互不相同的天数有

$$\binom{365}{n}$$

种情况，即从365天中选取 n 天的情况。故 n 个人生日都不相同的概率为：

$$\frac{\binom{365}{n} n!}{365^n}.$$

所求概率为：

$$p_n = 1 - \frac{\binom{365}{n} n!}{365^n}.$$

If $n \geq 23$, then $p_n \geq \frac{1}{2}$ ($p_{23} = 0.5073, p_{22} = 0.4757$).

Matlab程序： `1-nchoosek(365, 23)* factorial(23)/36523`

第二次作业

第二次作业（钟开来书）

P. 43-45: 第11, 12, 13, 14, 22, 25, 28题。

P. 70-71: 第2, 3, 5, 8, 12, 14, 16题。

作业：(1) “基础习题课”线下答疑：每天晚7点-8点（国庆假期10月1日、2日除外）。本周答疑地点：旧水304（以后地点再通知）。

(2) “基础习题课”计算机语言资料：

<https://cloud.tsinghua.edu.cn/d/9cd721673d644f3594d9/>

(3) 在作业第1页左上角醒目学号、姓名。

(4) 将作业扫描成单个、PDF文件（不要压缩、不要Word格式）。

预习内容：随机变量、分布、期望。