

概率论 笔记

任云玮

目录

1	样本空间与概率	2
1.1	概率模型	2
1.2	条件概率	3
1.3	全概率定理与 Bayes 准则	4

1 样本空间与概率

1.1 概率模型

1 定义

1. 对于一次实验，定义其可能产生的结果的全体为**样本空间** Ω .
2. 称一个集合 A 为**事件**，若它是样本空间 Ω 的一个子集.

评注 对于样本空间，在选取的时候需要注意结果需要是良定义的（无歧义的），同时需要实验的所有结果都在 Ω 中. 对于一个实验，求解某个事件的概率，最基础的步骤就是确定样本空间的组成.

2 定义 (概率律) 设 Ω 是一个样本空间，称定义在 Ω 中事件全体上的函数 P 为**概率律**，若它成立

1. 非负性. 对任意事件 A ， $P(A) \geq 0$.
2. 可加性. 对任意不相交的 A 和 B ，成立 $P(A \cup B) = P(A) + P(B)$. 或者更一般的，对于两两互不相交的 $\{A_n\}_{n=1}^{\infty}$ ，成立 $P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots$.
3. 归一性. $P(\Omega) = 1$.

评注 需要注意，概率律指的是特定的某一类函数，并非所有定义在样本空间上的函数都是概率律. 对于概率律，显然成立 $P(\emptyset) = 0$. 另外，一般在讨论概率律的时候，不区分只包含一个结果的事件和该结果本身.

3 定理 (概率律的性质) 给定概率律 P ，事件 A, B, C ，则

1. 若 $A \subset B$ ，则 $P(A) \leq P(B)$.
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
3. $P(A \cup B) \leq P(A) + P(B)$.
4. $P(A \cup B \cup C) = P(A) + P(A^c \cap B) + P(A^c \cap B^c \cap C)$.

评注 这些式子的证明都是 trivial 的，其中 [3.] 至少可以推广至有限个事件. 对于 [4.]，它实际上演示了一个将重合的事件拆分成不相交事件的方法.

4 引理 (Bonferroni 不等式) 设有事件 A_1, A_2, \dots, A_n ，则成立

$$P\left(\bigcap_{i=1}^n A_i\right) \geq \sum_{i=1}^n P(A_i) - (n-1).$$

证明 对 n 施归纳法, 利用定理 3 [2.] 即可.

5 定理 (容斥原理) 设 A_1, A_2, \dots, A_n 为样本空间 Ω 的事件, 则

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{\substack{I \subset \{1, \dots, n\} \\ |I|=k}} P(A_I).$$

6 定理 (连续概率) 设有事件序列 $\{A_n\}_{n=1}^\infty$, 成立 $A_n \subset A_{n+1}$. 令 $A = \bigcup_{n=1}^\infty A_n$, 则 $P(A) = \lim_{n \rightarrow \infty} P(A_n)$.

证明 考虑将 $P(\bigcup_{n=1}^\infty A_n)$ 拆分成级数的形式. 定义 $B_0 = \emptyset$, $B_n = A_n - A_{n-1}$, 则只需证明 $\bigcup_{k=1}^\infty B_k = A$, 再利用 定义 2 拆分 L.H.S 即可. ■

评注 这一定理意味着, 当事件单调增时, 极限运算与 P 可换序. 对于“单调减”的事件序列, 把并换成交, 可以有类似的结论.

1.2 条件概率

7 定义 (条件概率) 对于一个样本空间 Ω 、概率律 P 以及概率不为零的事件 $B \subset \Omega$, 对于事件 $A \subset \Omega$, 定义在 B 发生下 A 发生的概率为

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

评注 首先需要注意, $P(A|B)$ 是一个关于事件 A 的函数, 用更加符合一般函数的记法, 它相当于 $P_B(A)$, 即 $P(A|B)$ 和 $P(A|C)$ 是完全不同的两个函数. 同时, 我们可以验证, 若固定 B , 按照上述定义, $P(A|B)$ 确实是一个概率律, 即它满足定义 2. 另外, 此定理的另一个同样十分常用的记法为

$$P(A \cap B) = P(B)P(A|B).$$

它相当于将两个原本并列的事件安排了一个先后关系.

8 定理 (乘法法则) $P(\bigcap_{i=1}^n A_i) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|\bigcap_{i=1}^{n-1} A_i)$.

评注 此定理可以直接利用条件概率的定义证明. 这一定理描述了如果事件 A 发生的充要条件为事件 A_1, \dots, A_n 都发生, 那么可以将 A 看作 A_i 逐个依次发生的结果, 反之亦然.

9 习题 (Monty Hall) TODO

1.3 全概率定理与 Bayes 准则

10 定理 (全概率定理) 设 A_1, A_2, \dots, A_n 形成了样本空间的一个划分且 $P(A_i) > 0$, 则对于任意事件 B , 成立

$$P(B) = \sum_{i=1}^n P(A_i \cap B) = \sum_{i=1}^n P(A_i)P(B|A_i)$$

评注 证明是显然的. 这一定理常被用来求解事件的无条件概率, 其关键点在于确定恰当的划分 $\{A_i\}$.

11 定理 (Bayes 准则) 设 A_1, \dots, A_n 形成样本空间的一个划分且 $P(A_i) > 0$. 则对任意 B , 若 $P(B) > 0$, 则成立

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_{j=1}^n P(A_j)P(B|A_j)}.$$

证明

$$P(B)P(A_i|B) = P(B \cap A_i) = P(A_i)P(B|A_i) \Rightarrow P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)}.$$

将 $P(B)$ 用全概率公式展开即可完成证明. ■

评注 就数学角度而言, Bayes 公式将条件概率 $P(A|B)$ 和 $P(B|A)$ 联系了起来. 而从实际含义的角度而言, Bayes 表示如果我们知道事件 A_i 导致事件 B 的概率, 同时知道 B 确实发生了, 那么反过来可以推断事件 A_i 发生的概率.

12 定理 (Murphy) 设 $\{A_i\}$ 为一列事件, 满足 $A_n \subset A_{n+1}$, 定义 $A = \bigcup_{i=1}^n A_n$, 若 $P(A|A_n^c) \geq \varepsilon > 0$, 则 $P(A) = 1$.

证明 根据全概率公式, 可知

$$P(A) = P(A_n)P(A|A_n) + (1 - P(A_n))P(A|A_n^c) \geq P(A_n)P(A|A_n) + \varepsilon(1 - P(A_n)).$$

根据定理 6, 可知当 $n \rightarrow \infty$ 时, $P(A_n) \rightarrow P(A)$, $P(A|A_n) \rightarrow 1$, 所以

$$P(A) \geq P(A) + \varepsilon - \varepsilon P(A) \rightarrow P(A) \geq 1.$$

而根据概率律的定义, $P(A) \leq 1$, 从而 $P(A) = 1$. ■

评注 这一定理，通俗地说，意味着“只要一件事情有可能发生，则它必定会发生”。将 A_n 看作这件事在前 n 天发生这一事件，则 $P(A|A_n^c)$ 意味着在前 n 天没有发生这件事的前提下，这件事发生的概率，它大于一个大于零的常数，它对应了这件事“可能会发生”这一条件，而 $P(A) = 1$ 则表明这件事终究会发生。

13 引理 设 A, B, C 是样本空间 Ω 上的事件. 记 $P_B = P(*|B)$ ，它仍是样本空间 Ω 上的一个概率律，所以可以继续定义 $P_B(*|C)$. 成立

$$P_B(A|C) = P(A|B \cap C).$$

证明

$$\begin{aligned} P_B(A|C) &= \frac{P_B(A \cap C)}{P_B(C)} = \frac{P(A \cap C|B)}{P(C|B)} \\ &= \frac{P(A \cap B \cap C)}{P(B)} \frac{P(B)}{P(B \cap C)} = \frac{P(A \cap (B \cap C))}{P(B \cap C)} = P(A|B \cap C). \quad \blacksquare \end{aligned}$$

评注 实际上这一引理描述的是嵌套的条件概率就是要求两个条件都满足情况下的条件概率.

14 定理 (条件概率的全概率公式) 设 C_1, \dots, C_n 构成了样本空间的一个划分. 设 A 和 B 为两个事件且满足 $P(B \cap C_i) > 0$ ，则成立

$$P(A|B) = \sum_{i=1}^n P(C_i|B)P(A|B \cap C_i).$$

证明 记 $P_B(A) = P(A|B)$ 看作新的概率律，对它施全概率公式并利用引理 13，有

$$P(A|B) = \sum_{i=1}^n P_B(C_i)P_B(A|C_i) = \sum_{i=1}^n P(C_i|B)P(A|B \cap C_i). \quad \blacksquare$$

15 定义 (暗示) 对于事件 A 与 B ，设 $P(A) > 0$ ， $P(B) > 0$ ，称事件 B 暗示事件 A ，若 $P(A|B) > P(A)$ ，称事件 B 不暗示事件 A ，若 $P(A|B) < P(A)$.

16 命题 事件 A 暗示事件 B 的充要条件为事件 B 暗示事件 A . 若 $P(B^c) > 0$ ，则 B 暗示 A 的充要条件为 B^c 不暗示 A .

证明 根据 Bayes 公式，有 $P(B)/P(A) = P(B|A)/P(A|B)$ ，所以前者成立. 对于后者，由于 $(B^c)^c = B$ 成立，所以只需要证明一边即可. 利用 Bayes，得

$$P(A|B^c) - P(A) = \frac{P(A)}{P(B^c)}(P(B^c|A) - P(B^c)) = \frac{P(A)}{P(B^c)}(P(B) - P(B|A)).$$

根据前一条定理， $P(B) - P(B|A) < 0$ 成立，所以 $P(A|B^c) < P(A)$ 成立. \blacksquare