概率，随机变量和随机过程（第一章）

右武卫大将军

2019 年 3 月 29 日

1　引言

概率论用来研究相继发生或同时发生的大量现象的平均特性。人们观测到，在许多领域中，当观测次数增加时，某些量的平均会趋向于一个常数；即使平均是对试验前特定的任何子序列进行，其值扔保持不变。概率论的母的就是用事件的概率来描述和预测这些平均值。事件 A 的概率是赋予这一事件的一个数 $P(A)$，它可以解释为：

如果实验重复进行 n 次，事件 A 发生 n_A 次，则当 n 足够大时，A 发生的相对概率 n_A/n 以高度的确定性接近 $P(A)$:

$$P(A) \approx n_A/n \tag{1}$$

这种解释是不精确的。术语“以高度的把握性”，“接近”，“足够大”的含义都不明确。但是这种不精确是无法避免的，概率论只能以不准确的形式和物理现象相联系。

2　定义

2.1　公理化定义

概率论中有三个基础性的公理：

1. 任一事件 A 的概率 $P(A)$ 是赋予此事件的一个非负实数

$$P(A) \leq 0$$

2. 必然事件的概率等于 1

$$P(S) = 1$$

3. 如果两个事件 A 和 B 是互斥的，则：

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

2.2　相对频率定义

相对频率方法是基于下述定义：一事件的概率 $P(A)$ 是极限

$$P(A) = \lim_{n \leftarrow \infty} \frac{n_A}{n} \tag{2}$$

式中 n_A 是 A 是发生次数，n 是试验次数。

2.3 古典定义

对概率的古典定义为：一个事件的概率 $P(A)$ 可以不经过实际试验而先验的确定：

$$P(A) = \frac{N_A}{N} \tag{3}$$

其中 N 是可能结果的总数，N_A 是事件 A 的结果数。这种粗糙的定义忽略了每一个结果发生的可能性是不同的。这种有意地忽略是由不充分推理原理引出的，即：当没有先验知识时，我们只能假定事件 A_i 具有等概率性。即使规定了各个结果是等可能的，但同样存在一些问题：

1. 所谓的“等可能”也就是“等概率”的意义，但这一定义本身就是解决概率的定义问题。
2. 所谓的“等可能”适用的问题极为有限。
3. 这种“等可能”本身也是基于大量生活经验总结的，也并不是所谓的基于逻辑。
4. 当可能的结果无穷多的时候，这种定义就失效了。

3 概率与归纳

4 因果性和随机性

本质上，所有的物理现象或者社会现象都是确定性的，因果的，但前提是能够观测到所有的相关因素，由于这是无法实现的条件，所以，试验结果最终对观测者来说就是随机的。对于这一争议的回答是：物理学家并不关心什么是真正的，只关心什么是能够观测到的。