

示性函数与两道高考题的解答

苏 淳 江 涛

我们曾经在文 [1] 中介绍过示性函数, 并在多处运用过它, 今后我们还会不断地运用它来解答问题.

设 S 为某个集合, $A \subset S$ 是它的一个子集, 对 $s \in S$, 令

$$I_A(s) = \begin{cases} 1, & \text{如果 } s \in A; \\ 0, & \text{如果 } s \notin A, \end{cases}$$

那么 $I_A(s)$ 就从数值上告诉我们, 元素 s 是否属于子集 A , 故将之称为子集 A 的示性函数.

在概率问题中, S 就是全事件 (通常把全事件写作 Ω), 而 A 则是我们所关心的随机事件. 在这里, I_A 是一个只取 0 或 1 两个值的随机变量. 并且当 $I_A = 1$ 时表示事件 A 发生; 而当 $I_A = 0$ 时表示事件 A 不发生. 易知 $E I_A = P(A)$.

示性函数在概率论中运用得很多, 最重要的一种运用就是帮助我们计算随机变量的期望.

例 1. 一个箱子里有 5 个球, 分别以 1 – 5 标号, 若有放回地取球三次, 记至少被取出一次的球的个数为 X . 求 $E X$.

这是 2025 年新高考一卷第 14 题, 原题是填空题, 只要答案, 不要过程. 但是要有答案, 就必须有办法算出答案. 传统的解法是先求出 X 的分布律, 再求期望. 易知, X 的取值集合是 $\{1, 2, 3\}$. 要计算 X 取它们的概率需要运用排列组合知识, 容易出错.

我们来用示性函数计算 $E X$. 令

$$I_k = \begin{cases} 1, & \text{如果 } k \text{ 号球至少被取到一次,} \\ 0, & \text{如果 } k \text{ 号球三次都没有被取到,} \end{cases} \quad k = 1, 2, \dots, 5.$$

容易看出 $I_1 + I_2 + \dots + I_5$ 就是至少被取出一次的球的个数, 意即

$$X = I_1 + I_2 + \dots + I_5.$$

所以

$$EX = EI_1 + EI_2 + \cdots + EI_5.$$

而对于每个 $k \in \{1, 2, \cdots, 5\}$, 都有

$$P(I_k = 0) = \left(\frac{4}{5}\right)^3, \quad P(I_k = 1) = 1 - \left(\frac{4}{5}\right)^3, \quad EI_k = 1 - \left(\frac{4}{5}\right)^3.$$

所以

$$EX = EI_1 + EI_2 + \cdots + EI_5 = 5 \left(1 - \left(\frac{4}{5}\right)^3\right) = \frac{61}{25}.$$

这个解答过程写起来慢, 但是思维过程却非常简单, 一眼就能看看到底, 对于解答填空题尤其有效.

无独有偶, 2025 年中科大强基计划数学试题中也有一道类似的填空题:

例 2. n 个球放进 m 个盒子, 则装有球的盒子的个数的期望是 ____.

虽然与例 1 相比, 一个是取球, 一个是放球, 其实解决思路完全相同. 这个题目的答案就是:

$$EX = m \left(1 - \left(\frac{m-1}{m}\right)^n\right).$$

再看 2025 年北京高考数学试卷中的概率解答题.

例 3. 有一道选择题只考察一个知识点, 甲乙两校各有 100 人参加考试. 甲校有 80 人答对该题, 乙校有 75 人答对该题. 用频率估计概率.

(1) 从甲校随机抽取一人, 求该人答对该题的概率;

(2) 从甲乙两校各随机抽取一人, 以 X 记他们中答对该题的人数, 求恰有一人答对该题的概率, 并求 X 的期望;

(3) 甲校掌握该知识点的人 100% 能答对该题, 而乙校掌握该知识点的人 85% 能答对该题. 未掌握该知识点的人则都是从四个选择项里随机选择一个. 设甲校学生掌握该知识点的概率是 p_1 , 乙校学生掌握该知识点的概率是 p_2 , 试比较 p_1 与 p_2 的大小.

本大题三个小题分别考核不同的知识点. 第 (1) 小题最为简单.

(1) 由于甲校 100 名参考学生中有 80 人答对题目, 用频率估计概率, 可知从甲校随机抽取一人, 该人答对该题的概率是 0.8.

为解第 (2) 小题, 可以引入示性函数. 令

$$I_1 = \begin{cases} 1, & \text{如果抽出的甲校学生答对题目,} \\ 0, & \text{如果抽出的甲校学生未答对题目;} \end{cases}$$

$$I_2 = \begin{cases} 1, & \text{如果抽出的乙校学生答对题目,} \\ 0, & \text{如果抽出的乙校学生未答对题目.} \end{cases}$$

那么 $X = I_1 + I_2$ 就是所抽出的人中答对题目的人数.

易知, I_1 与 I_2 相互独立, 它们的分布律分别为

$$\begin{pmatrix} 0 & 1 \\ 0.2 & 0.8 \end{pmatrix} \quad \text{与} \quad \begin{pmatrix} 0 & 1 \\ 0.25 & 0.75 \end{pmatrix}.$$

故知

$$\begin{aligned} P(X=1) &= P(I_1 + I_2 = 1) = P(I_1 = 1, I_2 = 0) + P(I_1 = 0, I_2 = 1) \\ &= 0.8 \times 0.25 + 0.2 \times 0.75 = 0.2 + 0.15 = 0.35. \end{aligned}$$

而

$$EX = EI_1 + EI_2 = P(I_1 = 1) + P(I_2 = 1) = 0.8 + 0.75 = 1.55.$$

至于第 (3) 小题, 可根据全概率公式, 由题中条件, 列出两个等式

$$80\% = p_1 \cdot 100\% + (1 - p_1) \cdot 25\%, \quad \text{①}$$

$$75\% = p_2 \cdot 85\% + (1 - p_2) \cdot 25\%. \quad \text{②}$$

从 ① 式解得 $p_1 = 11/15$; 从 ② 式解得 $p_2 = 5/6$. 所以 $p_1 < p_2$.

由此看来, 尽管甲校学生掌握该知识点的情况不如乙校, 可是每个掌握该知识点的人都能 100% 能答对题目, 使得全校的正确率反而超过乙校.

在解第 (2) 小题的过程中所引入的两个示性函数 I_1 与 I_2 给概率 $P(X=1)$ 和期望 EX 的计算带来一定程度的方便. 而如果我们要求我们计算 X 的方差的话, 那么好处就更加明显了. 事实上, 对于只取 0 和 1 两个值的随机变量 ξ , 我们有

$$D\xi = E\xi^2 - (E\xi)^2 = P(\xi=1) - P^2(\xi=1) = P(\xi=1) \cdot P(\xi=0).$$

所以

$$DI_1 = 0.8 \times 0.2 = 0.16; \quad DI_2 = 0.75 \times 0.25 = 0.1875.$$

而由 I_1 与 I_2 的独立性, 即知

$$DX = D(I_1 + I_2) = DI_1 + DI_2 = 0.16 + 0.1875 = 0.3475.$$

参考文献

- [1] 苏淳: 示性函数与抽屉原理,