粗议一套数学试题

苏 淳　　江 涛

前两天看到一套 2025 年青岛市拔尖人才选拔数学测试题, 其中概率统计试题占比之高引起我们的关注. 仅在此粗浅地做点议论, 不足为评.

一共 19 道试题, 概率统计试题就有 7 道. 其中, 8 道单选题中有 2 道概率统计题, 3 道多选题中竟也有 2 道概率统计题, 3 道填空题中有 1 道概率统计题, 5 道解答题中有 2 道概率统计题.

在多选题和解答题中各有 1 道统计题, 其余 5 道为概率题.

概率统计试题占比如此之高, 可以反映出人才选拔中对概率统计知识的重视程度, 抑或是把对概率统计试题的处理能力视为工作能力的一项指标, 也未可知.

解答题中的概率题

18. (15 分)

甲乙二人参加射击选拔赛, 规则如下: 抽签决定首次射击方, 两人轮流射击, 各次射击结果相互独立. 每次射击时, 射中目标者得 1 分, 对方得 0 分; 射不中目标者得 0 分, 对方得 1 分. 得分领先 2 分者胜出, 选拔赛结束. 将结束时的射击总次数记作 n $(n \geq 2)$. 设每次射击时, 甲射中目标的概率为 $\dfrac{3}{5}$, 乙射中目标的概率为 p $(0 < p < 1)$. 若射击两次, 甲胜出的概率为 $\dfrac{2}{5}$.

(1) 求 p;

(2) 求射击两次, 甲得分 X 的分布律及均值;

(3) 以 $P_n(A)$ 记甲胜出的概率, 证明: $\dfrac{2}{5} \leq \displaystyle\sum_{i=2}^{n} P_i(A) < \dfrac{3}{4}$.

解: (1)若射击两次, 甲胜出, 根据题意, 甲两次都得分; 但两次是两个人各射击一次, 依靠抽签决定首次射击方, 每次射击时, 射中目标者得 1 分, 对方得 0 分; 射不中目标者得 0 分, 对方得 1 分. 所以甲射击时射中了目标, 乙射击时未射中目标; 两人各以 $\frac{1}{2}$ 的概率射击第一枪. 所以由全概率公式知

$$\frac{1}{2} \cdot \frac{3}{5} \cdot (1-p) + \frac{1}{2} \cdot (1-p) \cdot \frac{3}{5} = \frac{2}{5},$$

即
$$\frac{3}{5}(1-p)=\frac{2}{5} \quad \Rightarrow \quad p=\frac{1}{3}.$$

(2) 射击两次, 一人一次, 所以谁打第一枪无关紧要, 易知甲得分 X 的取值集合为 $\{0,1,2\}$, 容易求得

$$P(X=0)=\frac{2}{5}\cdot\frac{1}{3}=\frac{2}{15}, \quad \text{(甲未射中, 送乙一分; 乙射中了, 夺走一分)};$$
$$P(X=2)=\frac{2}{5}, \quad \text{(根据题意知)};$$
$$P(X=1)=1-P(X=0)-P(X=2)=1-\frac{2}{15}-\frac{2}{5}=\frac{7}{15}.$$

若直接计算, 则

$$P(X=1)=\frac{3}{5}\cdot\frac{1}{3}+\frac{2}{3}\cdot\frac{2}{5}=\frac{7}{15} \quad \text{(或两人皆中, 各得一分; 或两人都未中, 互相成就一分)}.$$

总之, 都有

X	0	1	2
P	$\frac{2}{15}$	$\frac{7}{15}$	$\frac{2}{5}$

(3) 根据题意, $P_2(A)=\frac{2}{5}$, 所以

$$\sum_{i=2}^{n} P_i(A) \geq P_2(A)=\frac{2}{5},$$

这就是不等式左端.

为证不等式右端, 需要研究数列 $\{P_n(A)\}$. 在 n 次射击中, 甲乙共得 n 分, 但若二者相差两分, 则二人得分的奇偶性相同, 故它们的和是偶数, 这就表明 $P_{2m-1}(A)=0$. 从而只需考察 $P_{2m}(A)$. 在偶数次射击中, 由于是交替射击, 两人射击次数相等, 所以谁先打第一枪不影响得分概率, 故无需区分, 关键的是在前 $2m-2$ 次竞技中, 在每相连两次射击中甲乙二人各得 1 分, 而最后两次射击中, 甲连得 2 分. 故可由第**(2)**小题结论知

$$P_{2m}(A)=P^{m-1}(X=1)\cdot P(X=2)=\left(\frac{7}{15}\right)^{m-1}\cdot\frac{2}{5}.$$

所以对任何正整数 $n \geq 2$, 都有

$$\sum_{i=2}^{n} P_i(A) < \sum_{m=1}^{\infty} P_{2m}(A)=\sum_{m=1}^{\infty}\left(\frac{7}{15}\right)^{m-1}\cdot\frac{2}{5}=\frac{3}{4}.$$

评述: 本题在得分规则上挺绕人, 在谁先开打上也采用选择机制, 给应试者设置障碍, 需要应试者细心.

填空题中的概率题

14. (15 分)

设有 n 个袋子, 每袋有 a 个白球和 b 个黑球. 先从第一个袋中随机摸出一球, 记下颜色后把它放入第二个袋中, 再从第二个袋中随机摸出一球, 记下颜色后把它放入第三个袋中, 依次进行, 最后从第 n 个袋中随机摸出一球, 记下颜色. 把在这 n 个球中的白球数目记作 Y_n. 试求 EY_n.

答案: $EY_n = \dfrac{na}{a+b}$.

附注: 原题是填空题, 只要答案, 不要解题过程, 更不讲究解题方法. 我们还是要来讲一讲解题方法. 设

$$I_k = \begin{cases} 1, & \text{如果从第} k \text{ 个袋子摸出的是白球}, \\ 0, & \text{如果从第} k \text{ 个袋子摸出的是黑球}, \end{cases} \qquad k = 1, 2, \cdots, n.$$

那么 $I_1 + I_2 + \cdots + I_n$ 当然就是从 n 个袋中所摸出的白球数目, 意即 $I_1 + I_2 + \cdots + I_n = Y_n$. 所以

$$EY_n = E(I_1 + I_2 + \cdots + I_n) = EI_1 + EI_2 + \cdots + EI_n.$$

现在关键是要求出 $EI_k = P(I_k = 1),\ k = 1, 2, \cdots, n$.

由于 $P(I_1 = 1) = \dfrac{a}{a+b}$, 所以

$$\begin{aligned} P(I_2 = 1) &= P(I_2 = 1 | I_1 = 1) \cdot P(I_1 = 1) + P(I_2 = 1 | I_1 = 0) \cdot P(I_1 = 0) \\ &= \frac{a}{a+b} \cdot \frac{a+1}{a+b+1} + \frac{b}{a+b} \cdot \frac{a}{a+b+1} = \frac{a}{a+b}. \end{aligned}$$

由此并结合归纳法, 即知对一切 $k = 1, 2, \cdots, n$ 都有

$$P(I_k = 1) = \frac{a}{a+b}.$$

所以

$$EY_n = \sum_{k=1}^{n} EI_k = \sum_{k=1}^{n} P(I_k = 1) = \frac{na}{a+b}.$$

评述: 这道题再一次显示了掌握示性函数这一重要工具的必要性.

多选题中的概率题

10. 已知随机变量 $X \sim B(n, p)$, 其中 $0 < p < 1$, 则

A. 若 $EX = \dfrac{2}{3},\ DX = \dfrac{1}{3}$, 则 $p = \dfrac{1}{2}$;

B. 若 $P(X = 3) > P(X = 4)$, 则 $0 < p < \dfrac{4}{n+1}$;

C. 若 $n=6$, 则 $P(X=k) \le P(X=3),\ k=0,1,2,\cdots,6$;

D. 若 $0<p<\dfrac{1}{2}$, 则 X 取值为奇数的概率小于 $\dfrac{1}{2}$.

答案: 正确的断言有 B 和 D.

分析与判断:

A. 由于 $X \sim B(n,p)$, 所以 $EX=np,\ DX=np(1-p)$, 似乎应当有

$$1-p=\frac{DX}{EX}=\frac{\frac{1}{3}}{\frac{2}{3}}=\frac{1}{2},$$

故有人据此求得 $p=\dfrac{1}{2}$. 但这样一来, 还应该有

$$EX=\frac{n}{2}=\frac{2}{3} \Rightarrow n=\frac{4}{3}$$

此与 n 应该是正整数的性质相矛盾. 事实上, 无论是 EX 还是 DX, 都应该与 n 有关, 所以本小题就是一种无厘头的设置, 当然是错误的断言.

B. 由 $P(X=3)>P(X=4)$, 可得

$$C_n^3p^3(1-p)^{n-3}>C_n^4p^4(1-p)^{n-4}$$
$$C_n^3(1-p)>C_n^4p \quad \Rightarrow \quad \frac{1-p}{n-3}>\frac{p}{4}.$$

由此解得

$$P<\frac{4}{n+1}.$$

可见断言 B 正确.

C. 若 $n=6$, 那么由 B 可知, 若 $P(X=3)>P(X=4)$, 则 $0<p<\dfrac{4}{7}$. 然而在 C 的条件里对 p 的值没有任何界定, 故知该断言未必都能成立. 所以断言 C 不正确.

D. 设 $0<p<\dfrac{1}{2}$, 我们记 $q=1-p$,

$$P_{\text{奇}}=C_n^1pq^{n-1}+C_n^3p^3q^{n-3}+\cdots,$$
$$P_{\text{偶}}=q^n+C_n^2p^2q^{n-2}+C_n^4p^4q^{n-4}+\cdots,$$

则 $P_{\text{奇}}$ 与 $P_{\text{偶}}$ 分别是 X 取值为奇数的概率和取值为偶数的概率. 由于 $0<p<\frac{1}{2}<q$, 所以

$$P_{\text{偶}}-P_{\text{奇}}=(q-p)^n>0, \quad \Rightarrow \quad P_{\text{偶}}>P_{\text{奇}},$$

既然

$$P_{\text{偶}}+P_{\text{奇}}=(q+p)^n=1,$$

所以 $P_{偶} > \frac{1}{2} > P_{奇}$, 亦即 X 取值为奇数的概率小于 $\dfrac{1}{2}$. 所以断言 D 正确.

评述: 供作判断的四个断言里面都有骨头, 需要细嚼慢咽方能知其滋味, 颇有嚼头.

单选题中的概率题

2. 某校举行游泳和乒乓球比赛, 某学生只能参加其中一项竞赛, 他参加游泳和乒乓球比赛的概率分别为 0.4 和 0.6. 他在游泳和乒乓球比赛中获得冠军的概率分别是 0.3 和 0.7. 今知他获得冠军, 则他参加游泳比赛的概率是

$$\text{A. } \frac{5}{9}, \qquad \text{B. } \frac{4}{9}, \qquad \text{C. } \frac{2}{9}, \qquad \text{D. } \frac{1}{9}.$$

答案: 选 C.

解: 以 A 记该生参加游泳比赛的事件, 那么 $\overline{A}$ 就是该生参加乒乓球比赛的事件. 再以 G 记他夺冠的事件. 今知

$$P(A) = 0.4, \quad P(\overline{A}) = 0.6, \quad P(G|A) = 0.3, \quad P(G|\overline{A}) = 0.7.$$

于是由贝叶斯公式, 知

$$\begin{aligned} P(A|G) &= \frac{P(A)P(G|A)}{P(A)P(G|A) + P(\overline{A})P(G|\overline{A})} \\ &= \frac{0.4 \times 0.3}{0.4 \times 0.3 + 0.6 \times 0.7} = \frac{2}{9}. \end{aligned}$$

评述: 这道题说到底是一道计算题, 算出的结果肯定只有一个, 所以该选谁一清二楚.

本题是全卷唯一涉及贝叶斯公式的题目. 这类题目当是下一段重点关注的内容.

5. 设随机变量 $X \sim \mathcal{N}(0,1)$; $Y \sim \mathcal{N}(1,2^2)$, 则

$$\begin{aligned} &A.\ \ P(Y \ge 1) \ge P(Y \ge 0), \qquad && B.\ \ P(X \ge 2) \le P(X \le 1), \\ &C.\ \ P(X \ge t) \ge P(Y \ge t), \qquad && D.\ \ P(X \le t) \le P(Y \le t). \end{aligned}$$

答案: 选 B.

解: 这是一道单选题, 只有一个答案是对的, 所以只要看到正确的答案, 就不用管其他的备选答案了.

注意到 $X \sim \mathcal{N}(0,1)$, 知 $P(X \le 0) = \dfrac{1}{2}$. 故有

$$P(X < 2) > P(X \le 1) > P(X \le 0) = \frac{1}{2}$$

从而

$$P(X \geq 2) = 1 - P(X < 2) < \frac{1}{2} < P(X \leq 1).$$

所以选 B.

分析: 本题作为单选题, 我们的分析解答自然没问题, 但是如果把这道题放在多选题里, 那么难度就陡然上升了. 其余三个断言都值得仔细探讨, 都应该弄清楚它们到底是对是错, 如果是错, 错在哪里?

先看 A. 由于 $(Y \geq 1) \subset (Y \geq 0)$, 所以 $P(Y \geq 1) \leq P(Y \geq 0)$, 故知 A 错. 这里考核的是事件之间的包含关系, 谁包含谁, 不要看反了.

再看 C. 不等式 $P(X \geq t) \geq P(Y \geq t)$ 中没有明确 t 的具体值, 这意味着该式要对所有的实数 t 都成立. 可事实不然. 由于 $X \sim \mathcal{N}(0,1)$; $Y \sim \mathcal{N}(1,2^2)$, 所以 $P(Y \geq 1) = \frac{1}{2}$, 而 $P(X \geq 1) < \frac{1}{2}$, 可见 $P(X \geq 1) < P(Y \geq 1)$. 所以断言 C 不能对一切实数 t 都成立. 故知 C 错.

最后看 D. 该断言的毛病与 C 类似. 由于 $X \sim \mathcal{N}(0,1)$; $Y \sim \mathcal{N}(1,2^2)$, 所以 $P(X \leq 0) = \frac{1}{2}$, 可是 $P(Y \leq 0) < \frac{1}{2}$, 可见 $P(X \leq 0) > P(Y \leq 0)$. 所以断言 D 不能对一切实数 t 都成立. 故知 D 错.

评述: 本题中的四个断言, 除了 A 是考核事件的包含关系的之外, B,C,D 都考核的是同一个知识点, 即对于正态分布 $\mathcal{N}(a,\sigma^2)$ 而言, 均值 a 就是中位数点. 只要抓住这一点, B,C,D 就迎刃而解.

至此, 我们已对该卷中的 5 道概率题目都做了解答和分析. 这是一套极其重视概率统计知识的试题. 里面涉及了全概率公式和贝叶斯公式. 这一对公式任何时候都应该是关注的重点, 贝叶斯公式今后更应成为关注的重点. 本套试题还考核了二项分布 $B(n,p)$ 和正态分布 $\mathcal{N}(a,\sigma^2)$ 的有关知识, 值得肯定.