



独立性检验创设, 概率与统计交汇

■江苏省兴化市第一中学 刘 林

概率与统计专题相关的知识点错综复杂,又环环相扣,相关考题通常不是孤立地考查一个知识点,而是对多个知识的综合考查。而基于独立性检验的创新情境问题,是新高考数学试卷中一个比较常见的题型,由此实现概率与统计的合理交汇与融合,使得各个独立的知识点加以合理联系,综合考查同学们的“四基”。

一、与统计的交汇

例 1 2022 年是共青团建团 100 周年,某校组织“学团史,知团情,感团恩”知识测试,现从该校随机抽取了 100 名学生,并将他们的测试成绩按 $[50, 60)$, $[60, 70)$, $[70, 80)$, $[80, 90)$, $[90, 100]$, 分为 5 组,得到如图 1 所示的频率分布直方图。

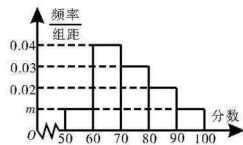


图 1

(1)求图中 m 的值,并估计这 100 名学生测试成绩的平均数;(同一组数据用该组数据所在区间的中点值作代表)

(2)规定测试成绩不低于 80 分为“优秀”,请将表 1 所示的 2×2 列联表(单位:人)补充完整,求出 x, y, z ,并判断是否有 95% 的把握认为“测试成绩是否优秀与性别有关”。

表 1

	优秀	非优秀	合计
女生	x	30	$x+30$
男生	y	z	55
合计	$x+y$	$30+z$	100

$$\text{附: } K^2 = \frac{(a+b+c+d)(ad+bc)^2}{(a+b)(c+d)(a+c)(b+d)}.$$

表 2

$P(K^2 > k_0)$	0.1	0.05	0.01
k_0	2.706	3.841	6.635

解析: (1)由题意得 $10 \times (m + 0.04 + 0.03 + 0.02 + m) = 1$, 解得 $m = 0.005$ 。

估计这 100 名学生测试成绩的平均数 = $(55 \times 0.005 + 65 \times 0.04 + 75 \times 0.03 + 85 \times 0.02 + 95 \times 0.005) \times 10 = 73$ 。

(2)在抽取的 100 名学生中,成绩优秀的学生人数为 $100 \times 0.25 = 25$,由此可得完整的 2×2 列联表,如表 3 所示:

表 3

	优秀	非优秀	合计
女生	15	30	45
男生	10	45	55
合计	25	75	100

$$\text{可得 } K^2 = \frac{100 \times (15 \times 45 - 30 \times 10)^2}{45 \times 55 \times 25 \times 75} =$$

$\frac{100}{33} \approx 3.03 < 3.841$, 所以没有 95% 的把握认为“测试成绩是否优秀与性别有关”。

点评:以实际应用为创新情境,设置独立性检验为问题背景,结合统计中的频率分布直方图、样本估计等知识,综合考查概率与统计中的相关应用问题。

二、与概率的交汇

例 2 经常有人说“数学学不好,物理也很难学好”,这话听着好像很有道理的样子,那么真实情况的确是这样吗?为此,某校数学兴趣小组收集了 500 名同学的数学成绩和物理成绩,记单科成绩在平均分之上为优秀,整理数据形成如图 2 所示的统计扇形图。

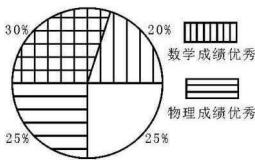


图 2

(1)根据图 2 完成表 4 所示的 2×2 列联表,并判断能否有 95% 的把握认为数学成绩与物理成绩有关;

表 4

	物理优秀	物理不优秀	合计
数学优秀			
数学不优秀			
合计			

(2) 视频率为概率,从全校数学成绩优秀的学生中随机抽取 3 人,记抽取到的 3 人中物理成绩优秀的人数为随机变量 X ,求 X 的分布列与期望。

$$\text{附: } K^2 = \frac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}.$$

表 5

$P(K^2 > k_0)$	0.1	0.05	0.025	0.01
k_0	2.706	3.841	5.024	6.635

解析:(1)由题意, 2×2 列联表如表 6:

表 6

	物理优秀	物理不优秀	合计
数学优秀	150	100	250
数学不优秀	125	125	250
合计	275	225	500

$$\text{可得 } K^2 = \frac{500(150 \times 125 - 100 \times 125)^2}{250 \times 250 \times 275 \times 225}$$

$\approx 5.051 > 3.841$, 所以有 95% 的把握认为数学成绩与物理成绩有关。

(2) 由于 X 的可能取值为 0, 1, 2, 3, 可得

$$P(X=0) = \frac{8}{125} = 0.064; P(X=1) = \frac{36}{125} =$$

$$0.288; P(X=2) = \frac{54}{125} = 0.432; P(X=3) =$$

$$\frac{27}{125} = 0.216.$$

所以 X 的分布列为表 7:

表 7

X	0	1	2	3
P	0.064	0.288	0.432	0.216

$$\text{所以 } E(X) = 0 \times 0.064 + 1 \times 0.288 + 2 \times 0.432 + 3 \times 0.216 = 1.8.$$

点评:以实际应用为创新情境,设置独立性检验为问题背景,结合概率中的重复独立

试验、随机变量的分布列与期望等知识,综合考查概率与统计中的相关应用问题。

三、与其他知识的交汇

例 3

2022 年 5 月以来,世界多个国家报告了猴痘病例,非洲地区猴痘地方性流行国家较多。我国目前为止尚无猴痘病例报告。我国作为为人民健康负责任的国家,对可能出现的猴痘病毒防控提前做出部署。同时国家卫生健康委员会同国家中医药管理局制定了《猴痘诊疗指南(2022 年版)》。此《指南》中指出:①猴痘病人潜伏期 5~21 天;②以往接种过天花疫苗者对猴痘病毒存在一定程度的交叉保护力。据此,援非中国医疗队针对援助的某非洲国家制定了猴痘病毒防控措施之一是要求与猴痘病毒确诊患者的密切接触者集中医学观察 21 天。在医学观察期结束后发现密切接触者中未接种过天花疫苗者感染病毒的比例较大。对该国家 200 个接种与未接种天花疫苗的密切接触者样本医学观察结束后,统计了感染病毒情况,得到表 8 所示的列联表:

表 8

接种天花疫苗 与否人数	感染猴痘 病毒	未感染猴痘 病毒
未接种	30	60
已接种	20	90

(1) 是否有 99% 的把握认为密切接触者感染猴痘病毒与未接种天花疫苗有关。

(2) 以样本中结束医学观察的密切接触者感染猴痘病毒的频率估计概率。现从该国所有结束医学观察的密切接触者中随机抽取 4 人进行感染猴痘病毒人数统计,求其中至多有 2 人感染猴痘病毒的概率。

(3) 该国现有一个中风险村庄,当地政府决定对村庄内所有住户进行排查。在排查期间,发现一户 3 口之家与确诊患者有过密切接触,这种情况下医护人员要对其家庭成员逐一进行猴痘病毒检测。每名成员进行检测后即告知结果,若检测结果呈阳性,则该家庭被确定为“感染高危家庭”。假设该家庭每个成员检测呈阳性的概率均为 $p(0 < p < 1)$ 且



高考中概率统计问题的分析与展望

■河北省廊坊市第一中学 杨明

通过对近几年高考试卷的研究,发现概率统计模块的考查更注重在数学知识交汇处命题,强调数学知识的综合性与应用性。解答题中对于概率统计的考查,新高考数学试题朝着“重视基础、强调综合、体现应用与着力创新”等特点的命题方向发展。

一、强调综合

“综合性”主要体现在概率统计模块的知识与其他数学知识之间的综合与应用,特别是与函数、方程、不等式、数列等知识的交汇,同时与其他学科及数学思想方法的综合。

例 1 数据显示,中国直播购物规模近几年保持高速增长态势,而直播购物中的商品质量问题逐渐成为人们关注的重点。已知某顾客在直播电商处购买了 $n(n \in \mathbf{N}^*)$ 件商品。

(1)若 $n=10$,且买到的商品中恰好有 2 件不合格品,该顾客等可能地依次对商品进行相互独立。记该家庭至少检测了 2 名成员才能确定为“感染高危家庭”的概率为 $f(p)$ 。试问:当 p 为何值时, $f(p)$ 最大?

$$\text{附: } K^2 = \frac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(a+c)(b+d)}.$$

表 9

$P(K^2 > k_0)$	0.1	0.05	0.025	0.01
k_0	2.706	3.841	5.024	6.635

解析: (1) 依题意知, $K^2 = \frac{200 \times (30 \times 90 - 60 \times 20)^2}{90 \times 110 \times 50 \times 150} \approx 6.061 < 6.635$, 所以没有 99% 的把握认为密切接触者感染猴痘病毒与未接种天花疫苗有关。

(2)由题意得,该地区每名密切接触者感染病毒的概率为 $p = \frac{30+20}{200} = \frac{1}{4}$, 设随机抽取的 4 人中至多有 2 人感染病毒为事件 A, 则 $P(A) = 1 - C_4^3 \cdot \left(\frac{1}{4}\right)^3 \cdot \frac{3}{4} - \left(\frac{1}{4}\right)^4 = \frac{243}{256}$.

行检查。求顾客检查的前 4 件商品中不合格品件数 X 的分布列。

(2)抽检中发现直播电商产品不合格率为 0.2。若顾客购买的 n 件商品中,至少有 2 件合格产品的概率不小于 0.998 4,求 n 的最小值。

解析: (1)由题意可知, X 的所有可能取值为 0, 1, 2。

$$\text{则 } P(X=0) = \frac{C_2^0 C_8^4}{C_{10}^4} = \frac{1}{3}; P(X=1) = \frac{C_2^1 C_8^3}{C_{10}^4} = \frac{8}{15}; P(X=2) = \frac{C_2^2 C_8^2}{C_{10}^4} = \frac{2}{15}.$$

所以 X 的分布列为表 1。

表 1

X	0	1	5
P	$\frac{1}{3}$	$\frac{8}{15}$	$\frac{2}{15}$

(2)记“顾客购买的 n 件商品中,至少有

(3)设事件 B = 检测了 2 名成员确定为“感染高危家庭”,事件 C = 检测了 3 名成员确定为“感染高危家庭”,则 $P(B) = (1-p)p$, $P(C) = (1-p)^2 p$, 所以 $f(p) = (1-p)p + (1-p)^2 p = p(1-p)(2-p)$, 设 $x = 1-p$, 则 $p = 1-x$, 所以 $y = (1-x)x(1+x) = x - x^3$, 求导得 $y' = 1 - 3x^2$, 令 $y' = 0$, 得 $x = \frac{\sqrt{3}}{3}$, 此时 $p = 1 - \frac{\sqrt{3}}{3}$, $f(p)$ 取得最大值。

点评:以实际应用为创新情境,设置独立性检验为问题背景,结合概率中的重复独立试验,函数与导数的应用等知识,综合考查概率与统计中的相关应用问题。

基于独立性检验创设的概率与统计专题,由于独立性检验的介入,使得问题的起点相对来说比较低,同时又可以多层面设计,起伏比较大,为问题的设置与能力的考查创设很好的基础,备受各方关注,成为新高考数学试卷中概率与统计专题部分的一个特色与亮点。

(责任编辑 王福华)